An Analysis of Language Impact on Augmented Reality Order Picking Training

Nela Murauer
BMW Group
Max-Diamand-Straße 5
80788 München
nela.murauer@bmw.de

Florian Müller
TU Darmstadt
Hochschulstraße 10
64289 Darmstadt
mueller@tk.tu-darmstadt.de

Sebastian Günther
TU Darmstadt
Hochschulstraße 10
64289 Darmstadt
guenther@tk.tu-darmstadt.de

Dominik Schön
TU Darmstadt
Hochschulstraße 10
64289 Darmstadt
schoen@tk.tu-darmstadt.de

Nerina Pflanz
BMW Group
Max-Diamand-Straße 5
80788 München
nerina.pflanz@bmw.de

Markus Funk
TU Darmstadt
Hochschulstraße 10
64289 Darmstadt
funk@tk.tu-darmstadt.de

ABSTRACT

Order picking is a difficult and cognitively demanding task. Traditionally textual instructions are helping new workers to learn different picking routines. However, the textual instructions are sometimes not written in the workers' native languages. In the area of Industry 4.0, where digital functions are finding their way into manufacturing processes, language-independent instructions are possible. Through a user study with 15 participants, we compare textual feedback in the workers' native language, textual feedback that is written in an unknown foreign language, and visual Augmented Reality (AR) feedback. We found that AR feedback is significantly faster and leads to a lower perceived cognitive load.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality** • **Human-centered computing → Empirical studies in HCI**

KEYWORDS

Order Picking; Augmented Reality; Assistive Systems

PETRA '18, June 26–29, 2018, Corfu, Greece

ACM ISBN 978-1-4503-6390-7/18/06…$15.00
https://doi.org/10.1145/3197768.3201570

1 INTRODUCTION

Manufacturing companies are currently undergoing a dynamic transformation from traditional flow production to a smart factory. Designing this change automotive logistics focus on the implementation of Augmented Reality (AR) in order picking processes, robotics, autonomous forklifts and autonomous tugger trains. In particular, AR-technology offers advantages to series production and training. Today foremen explain every process step to novice employees and accompany the worker for several days. This training procedure is time consuming and causes high personnel costs for the foremen. AR-training scenarios in head mounted displays offer the opportunity to guide the trainee with the aid of indoor navigation and the visualization of work instructions. Especially at order picking workstations, the guidance through the process is more important than the training effect. As orders vary in their amount and their composition of components or products, each order picking process must be supported by additional information. There are different visualization methods for order picking processes. *Pick-by-Paper* means picking with the aid of a paper-based picking list, *Pick-by-Voice* guides the worker with auditory navigation, *Pick-by-Light* points out the next target to an employee with a little light fixed below the box and *Pick-by-Vision* supports the picker by augmented additional information in head mounted displays. Especially *Pick-by-Voice* and *Pick-by-Vision* are suitable for enabling unskilled workers, such as novice employees or temporary staff, to perform at the workstation with the aid of step-by-step instructions. Compared to *Pick-by-Voice,* AR-assistance can display different information at the same time and, for example, overlay the path to the correct bin with arrows similar to a head-up display design in cars.

As team characteristics are multi-cultural and thus multi-lingual, we observed foreman training in different languages. Additionally, in some production plants the working language is not the local language. For this reason, we would like to address these language-barriers integrating different ways of information visualization in our AR design concept. Afterwards, we will analyze the impact of language on training efforts in our study.

2 RELATED WORK AND BACKGROUND

2.1 Augmented Reality Training

A number of authors have given their attention to the application of AR-supported training in the environment of production, dexterity and manual labor. This section introduces some of their works as an indicator of the current state of knowledge on AR. Reyes et al. aimed to replace the trainer for inexperienced students by an app, which teaches the handling of machinery [1]. The tool developed and implemented by [1] delivered good results since students accepted the AR instructions positively. This information compliments studies by Wiedenmaier and Wiedenmaier et al. [2, 3] finding shorter task completion times (TCT), better assembly quality, and lower human workload comparing HMD to paper-based instructions for the process of assembling a car door. Two separate and independent studies by [4] assessed the usability of virtual training in automotive manufacturing. Langley et al. observed error reductions in tasks performance after virtual introduction and training [4]. Participants were again for the most part positive concerning the overall use of the virtual training systems for assembly operation training.

2.2 Augmented Reality Order Picking

Some studies address order picking supported by AR-technologies. Guo et al. [5] compared in their order picking study card-mounted displays, head mounted displays (HMD), Pick-by-Light and Pick-by-Paper with 12 participants. The subjective workload, measured by NASA-TLX [6], was the lowest using a HMD. The task completion time was significantly shorter and the error rate was significantly lower subjecting to AR-instruction compared to the other methods. Odenthal et al. recruited 48 participants for a study comparing the delivery of assembly information by head-mounted or table-mounted Augmented Vision Systems [7]. For [7] AR increased the error detection capability significantly, but they concluded that error detection time increased as well. Kampmeier et al. [8] analyzed in their laboratory test the difference between paper-based and HMD-based assembly support. To estimate the impact of wearing the HMD on the worker, they added a third scenario, where the participants wore a HMD, but information was provided paper-based. After ophthalmological examination they concluded that using a HMD does not cause eye impairment. The resulting evidence shows that the workload is not significantly higher using a HMD. Deduced results were that quality and quantity of workers' performance renders HMD support more useful in order picking process rather than assembly tasks. Funk et al. [9] recommended a cart-mounted projector system. In a comprehensive study, they found that a cart-mounted system is faster and leads to less cognitive effort than traditional picking methods. Further, Funk et al. [10] advised a similar camera-projector system that is worn on the worker's head to assist during order picking tasks. Reif et al. [11] and Schwerdtfeger et al. [12] proposed the implementation of an attention funnel visualization [13] based on [14] for head-mounted displays to show workers the path to the target bin. Theis et al. [15] compared in their study with 60 participants different types of HMDs with a monitor. The performance of the group, which assembled with the aid of HMDs, was significantly lower in comparison to the screen-supported group. Tümler [16] shares the view that paper-based order picking is faster and thus preferable to HMD. After Tümler, estimation of the subjective cognitive workload is equally higher for HMD than paper-based tools.

An overview about Augmented Reality and Virtual Reality systems for manufacturing environments is presented by Buettner et al. [17].

Results are inconclusive at this point as the range of time, error rate and cognitive workload is not transferable from one use-case to another. We assume thus that neither the results nor the advantages and disadvantages are context independent. Hence, AR- support requires testing in the targeted work environment.

3 SYSTEM

To evaluate the use of both textual instructions and Augmented Reality instructions for order picking, we created three order picking assistance systems: a textual German system, a textual Finnish system, and an Augmented Reality system. All systems were implemented in Unity using a Microsoft HoloLens. In the following, we describe the three order picking systems in detail.

3.1 Textual Feedback: German

As a baseline condition, we implement a textual order picking system, which gives textual picking instructions that are written in German language. We choose German as a language for this textual baseline order picking system as German is the work-language in our factory.

The system shows the compartment to pick from and the quantity to pick in a text overlay that is at a fixed position in the worker's field of view. Using this fixed position, we ensure that wherever the worker is currently looking at, the picking instruction is visible and in the center of the worker's field of view. Figure 1(left) shows an example of a German picking instruction.

To place the previously picked object, the system also shows a textual instruction for the worker. Again, the place instruction is placed at a fixed position in the worker's field of view. Figure 1 (right) shows an example of a German placing instruction.

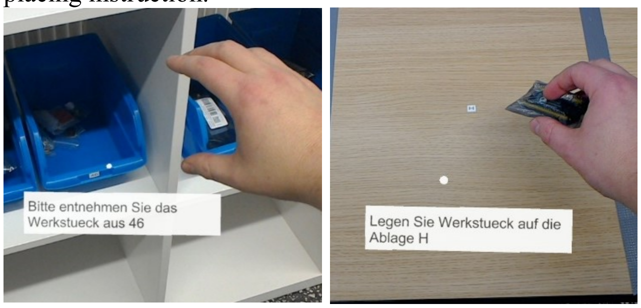

Figure 1: The text-based German picking instructing telling the participant to pick a part from a bin (left) and to place the previously picked part (right).

3.2 Textual Feedback: Finish

We are interested in assessing the effect of the language the instructions are written in on the worker's performance. Therefore, we argue to present the picking instruction in an unknown foreign language and assess the effect of the foreign language on the dependent variables. As a result, we chose Finnish as a language for the textual control condition as in our factory Finnish is a language that not only no one has experience speaking in but also no one has pre-knowledge of. For the Finnish textual instructions, we use exactly the same representation than in the German textual instruction, i.e., a textual overlay that is displayed on a fixed position in the worker's field of view. Figure 2 (left) shows an example of a Finnish picking instruction and Figure 2 (right) shows a Finnish placing instruction.

Figure 2: The text-based Finnish picking instructing telling the participant to pick a part from a bin (left) and to place the previously picked part (right).

3.3 Augmented Reality Arrow Visualization

Based on related work and our previous research activities, we build a 3D arrow visualization which shows the worker where to pick parts from and where to place them. The arrow is displayed at a dynamic position in the worker's field of view and is always facing the position to pick from or the position to place the currently picked part. The amount of items to pick and items to place is also displayed using a text label. Once the correct amount is picked or placed, the arrow starts pointing to the next target. If the last action is performed, the arrow is not shown anymore and the current picking round is finished. Figure 3 (left) shows an example of the Arrow picking instruction and Figure 3 (right) shows a placing instruction.

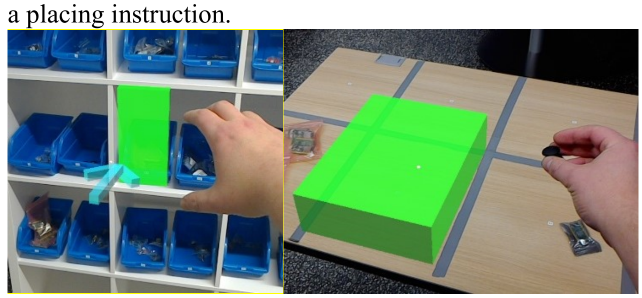

Figure 3: The Augmented Reality picking instructing telling the participant to pick a part from a bin using green rectangles (left) and to place the previously picked part (right).

4 EVALUATION

We scientifically evaluated the previously proposed order picking systems. In the following the design, procedure, participants and the results of the study are presented.

4.1 Design

We designed the study according to a repeated measures design with the used order picking system as the only independent variable consisting of three levels: German Text, Finnish Text, and Augmented Reality Guidance. As dependent variables, we measured the Task Completion Time (TCT), the number of errors that were made (ER), and the perceived cognitive load using the Raw NASA-TLX questionnaire (RTLX) [6].

4.2 Task and Study Setup

To compare the previously introduced order picking visualizations for the Microsoft HoloLens, we designed a study environment consisting of a picking area and a placing area. The picking area consists of two IKEA KALLAX shelves containing 48 traditional picking bins. We labeled the picking bins with a number from 1 to 48 to being able to identify them better. As items to pick and place, we use electronic boards with different shapes and sizes. Thereby one picking bin held one type of electronic board. For the placement area, we used a standard table (160cm x 80cm), which we divided into 10 equal placement areas. We used tape to optically mark the different placement positions. We labeled the different placement positions with letters from A to J.

We designed the picking and placing tasks to pick one item from each picking bin. This results in equally complicated picking tasks for each conditions. However, the order of the bins and the order of the placement areas were changed. Both picking area and placing area are depicted in Figure 4. For controlling the instructions that are shown on the Microsoft HoloLens device, we created a backend interface for the experimenter to control the content that is shown on the head-mounted display. The backend interface is also implemented in Unity3D and communicates with the Microsoft HoloLens using the HoloToolkit SharingService. Figure 5 shows the backend interface, which contains two buttons for indicating if a picking or placing task has been performed correctly or not. Also, the backend interface contains a routine to connect to the HoloLens application and to choose the order picking visualization according to the chosen condition. The experimenter uses this backend interface to control all aspects of the user study.

Figure 4: The setup that was used to conduct the user study. In the background there is the picking area containing of 48 picking pins. In the front is the placement area consisting of 10 placement spots

4.3 Procedure

We followed the same standardized procedure for every participant in our user study to ensure equal starting conditions. At first, we explained the participant the course of the study and gave a general introduction on order picking. After filing a consent form and informing the participants about the data that is being collected in the study, we collected the demographics and asked about previous experience with Augmented Reality. Then, we gave a general introduction to the Microsoft HoloLens and made participants familiar with mounting the HoloLens. As interacting with content on the HoloLens was not required for this study, we did not include any user interaction in the tutorial and in the application. As we were only interested in the effect of the visual instructions on the dependent variables, we excluded the picking detection and placing detection from the study and used a Wizard of Oz (WoZ) to forward the instructions once a pick or a place was done by the worker. Using this WoZ approach, we assured that the pick and place detection always worked perfectly.

We gave the participants some time to get familiar with the device. Once the participants indicated that they were ready to start the study, we started with the first picking system. To avoid ordering effects of the picking systems, we counterbalanced the order of the conditions using a Balanced Latin Square. Each picking condition was done with a task consisting of 48 items to pick and place. During the study, a researcher was observing the participant. The researcher also counted if the participants made an error. After the task was done, we repeated the procedure for the other conditions.

To recreate a realistic setting that would be found in an industrial order picking scenario, a researcher was there at all times if the participant had questions regarding the study. However, the researcher did not answer any questions about the picking task during the study. We deliberately designed the study in this way as this represents the reality in our company.

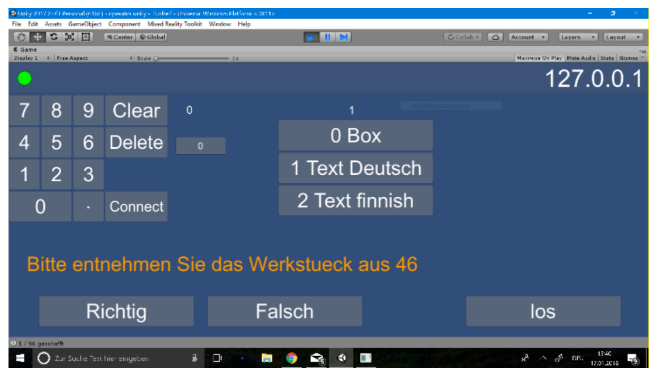

Figure 5: The backend which is used by the experimenter to connect to the HoloLens application and to control the shown order picking visualization. The interface on the bottom of the application enables the experimenter to log errors.

4.4 Participants

For our study, we invited 15 participants (3 female, 12 male) who were between 22 and 33 years old (M = 27.4 years, SD = 3.02 years). The participants were researchers with different backgrounds and students with different majors. Eight of the 15 participants had previous experience with the Microsoft HoloLens. Five of the 15 participants considered themselves as Augmented Reality experts. None of the participants was familiar with the positions of the items and the order picking tasks that were used in this study. The study took approximately 30 minutes per participant. All participants volunteered to take part in the study and did not receive a compensation.

4.5 Results

We statistically analyzed the Task Completion Time, the number of errors, and the perceived cognitive workload using the Raw NASA-TLX questionnaire. Mauchly's test of sphericity did not indicate a violation for all dependent variables. We used a Bonferroni correction for all post-hoc tests.

We statistically compared the Task Completion Time across the three order picking systems. The Augmented Reality system was the fastest (M=3.31s, SD=0.57s), followed by the textual instructions in German (M=4.25s, SD=0.97s) and the textual instructions in Finnish (M=4.37s, SD=0.92s). A one-way repeated measures ANOVA revealed a statistically significant difference between the conditions $F(2,28) = 11.674$, $p < 0.001$. A post-hoc test revealed a significant difference between the Augmented Reality condition and both textual conditions ($p<0.05$). The effect size estimate revealed a large effect ($\eta2 = 0.455$). Figure 9 depicts the average TCT graphically.

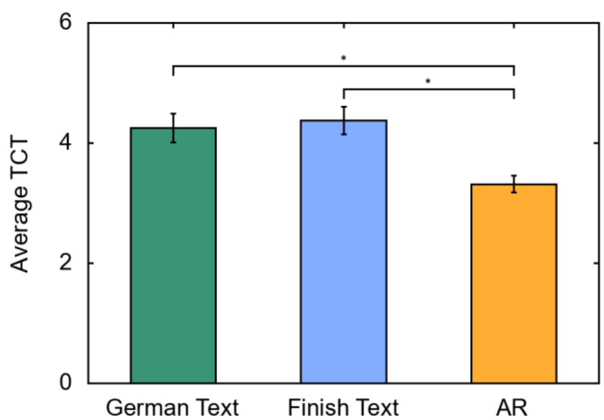

Figure 6: The average Task Completion Time (TCT) that was needed for completing one pick in our user study. The error bars show the Standard Error (SE). Bars that are marked with a * indicate a statistically significant difference.

Considering the errors that were being made by the participants, all order picking systems led to the same amount of errors (M=0.07, SD=0.258). This is equivalent to 1 error per condition. The different errors were made by different participants. Hence, a one-way repeated measures ANOVA did not reveal any significant difference ($p>0.05$). Figure 10 shows a diagram depicting the average errors.

Regarding the perceived cognitive load that was measured using the Raw Nasa-TLX (RTLX) questionnaire, the Augmented Reality order picking system led to the lowest RTLX score (M=19.13, SD=10.01), followed by the German textual order picking system (M=28.71, SD=12.33) and the Finnish textual order picking system (M=29.92, SD=11.83). A one-way repeated measures ANOVA revealed a significant difference between the conditions, $F(2,28) = 9.022$, $p = 0.001$. A post-hoc test revealed a significant difference between the Augmented Reality condition and both textual conditions ($p<0.05$). The effect size estimate revealed a large effect ($\eta2 = 0.486$). The results for the average RTLX score are also shown in Figure 11.

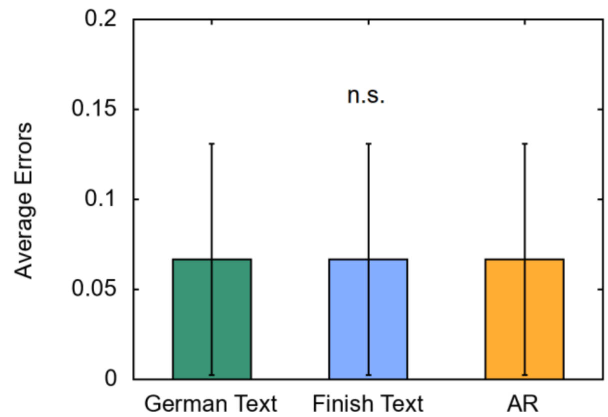

Figure 7: The average errors that were made while completing one pick in our user study. The error bars show the Standard Error (SE). Bars that are marked with a * indicate a statistically significant difference.

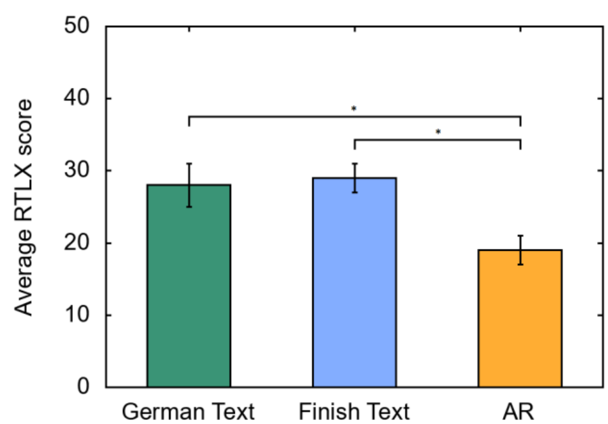

Figure 8: The average Raw NASA-TLX (RTLX) score that was scored for each condition in our user study. The error bars show the Standard Error (SE). Bars that are marked with a * indicate a statistically significant difference.

4.6 Qualitative Results

Through a semi-structured interview at the end of each condition, we also collected valuable qualitative feedback. In general, participants were mentioning the comfort of wearing a HoloLens. They stated that "at the end of the study wearing the HoloLens was becoming a little bit painful" (P3) and that "the HoloLens was getting a little uncomfortable with increasing time" (P4). Another participant stated that "the field of view of the HoloLens is a problem to see things in the arrow condition" (P7). That was also noticed by another participant by stating that for understanding the arrow visualization you "have to take a step back to grasp the whole picture" (P9). Also for the arrow feedback condition, some participants reported that "I completely lost the connection to what I was picking, as at some point I was only paying attention to the visual picking instructions" (P4, P10, P11, P14). However, they also enjoyed the comfort of the arrow based instructions as they stated that "the arrow-based instructions were really intuitive" (P5, P7, P15). They particularly liked that "for the arrow-based feedback [they] don't have to understand the feedback and can start right away" (P6, P8, P11). One participant even stated that while during picking, "[he] could turn off [his] head" (P12). For the textual instructions, participants stated that "It didn't make a difference for me if the feedback was displayed in German or in Finnish as I was only paying attention to the numbers and letters indicating the picking and placement locations" (P3, P4, P5, P9, P11, P12, P15).

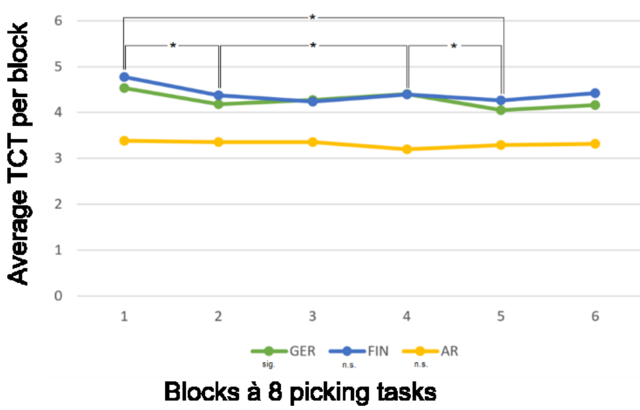

Figure 9: Learn effect for the groups using German, Finnish and AR

In order to assess a learn effect for the conditions, we divided the task into 6 sub groups consisting of 8 picking tasks per group. While we could not find a significant learn effect for the groups using AR and Finnish, there was a significant difference ($p>0.05$) between the blocks 1 and 2, 2 and 4, 4 and 5, and 1 and 5 (as depicted in Figure 9).

5 DISCUSSION

Through the conducted user study, we made several main findings that are summarized and discussed in this section.
We found that the arrow visualization led to significantly faster picking results than both textual conditions. We assume that this is due to the effort the workers have to put in to read the instructions first. The qualitative feedback revealed that with the arrow visualization, the participants could start with the picking right away. The processing time the participants needed for using the AR arrow visualization might have been less than processing a textual instruction.
Further, we found that the arrow visualization leads to significantly less cognitive effort based on the measured RTLX score than both textual instructions. This finding was also confirmed by the qualitative feedback that we received by the participants. We assume that this is also due to the reduced effort that participants need to understand the AR arrow visualization in comparison to a textual picking instruction. Based on these two findings of the performed user study, we can see a preference towards the arrow-based order picking visualization.
Another interesting finding of this study is that participants did not seem to pay attention about the language that is used for describing the picking instructions. This finding is mostly based on the qualitative feedback that we received by the participants.

6 CONCLUSION

AR visualizations support novice employees learning the order picking process irrespective of their native language background. It is an advantage of AR-supported that the foreman is relieved of the task to train incoming workers with a low level of experience.
Guidance by symbols rather than textual information also introduces an element to the training process, which renders the trainee more autonomous and enables unskilled workers to take up work not only quickly but also at a variety of workstations. This flexibility in job rotation is an enrichment for the worker and contributes to task variation for the single employee. Unskilled workers can become part of the staff quickly. In particular, production plants with high worker fluctuation will profit from the decreased training times under supervision and assistance of the foremen. Furthermore, the role of language in the training phase becomes less important using visual guidance to avoid language-induced communication problems. How far-reaching the AR-technologies and visual guidance should be, remains to be discussed. A healthy balance between automatization and autonomous behavior needs to be defined and upheld as decoupling of task comprehension and task execution inevitably comes in focus.

REFERENCES

1. Reyes, A.M., Villegas, O. O. V., Bojóquez, E.M., Sánchez, V.C., Nandayapa, M.: A Mobile Augmented Reality System to Support Machinery Operations in Scholar Environments,Computer Applications in Engineering Education 24 (6), November 2016 (pp. 967–981)

2. Wiedenmaier, S. (2004). Unterstützung manueller Montage durch Augmented Reality-Technologien, Aachen: Shaker.

3. Wiedenmaier, S., Oehme, O., Schmidt, L., & Luczak, H. (2003). Augmented reality (AR) for assembly processes: Design and experimental evaluation. International Journal of Human-Computer Interaction, 16(3), Pages: 497–514.

4. Langley, A., Lawson, G., Hermawati, S., D'Cruz, M., Apold, J., Arlt, F. and Mura, K.: Establishing the Usability of a Virtual Training System for Assembly Operations within the Automotive Industry Human Factors and Ergonomics in Manufacturing & Service Industries 26 (6), Pages: 667–679 (2016)

5. Guo, A., Raghu, S., Xie, X., Ismail, S., Luo, X., Simoneau, J.& Starner, T. (2014, September). A comparison of order picking assisted by head-up display (HUD), cart-mounted display (CMD), light, and paper pick list. In Proceedings of the 2014 ACM International Symposium on Wearable Computers (pp. 71-78). ACM.

6. Hart, S. G. "NASA-task load index (NASA-TLX); 20 years later." Proceedings of the human factors and ergonomics society annual meeting. Vol. 50. No. 9. Sage CA: Los Angeles, CA: Sage Publications, 2006.

7. Odenthal, B., Mayer, M. P., Kabuß, W., and Schlick, C. M.: A Comparative Study of Head-Mounted and Table-Mounted Augmented Vision Systems for Assembly Error Detection, Human Factors and Ergonomics in Manufacturing & Service Industries, 24 (1), Pages: 105–123 (2014)

8. Kampmeier, J., Cucera, A., Fitzsche, L., Brau, H., Duthweiler, M., Lang, G.K.: Eignung monokularer Augmented Reality – Technologien in der Automobilproduktion. In: Klinische Monatsblätter für Augenheilkunde, Georg Thieme Verlag, Stuttgart, 590-596 (2007)

9. Funk, M., Shirazi, A. S., Mayer, S., Lischke, L., & Schmidt, A. (2015, September). Pick from here!: an interactive mobile cart using in-situ projection for order picking. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 601-609). ACM.

10. Funk, M., Mayer, S., Nistor, M., & Schmidt, A. (2016, June). Mobile in-situ pick-by-vision: Order picking support using a projector helmet. In Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments (p. 45). ACM.

11. Reif, R., Günthner, W. A., Schwerdtfeger, B., & Klinker, G. (2009, February). Pick-by-vision comes on age: evaluation of an augmented reality supported picking system in a real storage environment. In Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (pp. 23-31). ACM.

12. Schwerdtfeger, B., Reif, R., Günthner, W. A., & Klinker, G. (2011). Pick-by-vision: there is something to pick at the end of the augmented tunnel. Virtual reality, 15(2-3), 213-223.

13. Schwerdtfeger, B., Klinker, G.: Supporting Order Picking with Augmented Reality. IEEE International Symposium on Mixed and Augmented Reality, 15-18 September, Cambridge (2008)

14. Biocca, F., Tang, A., Owen, C., & Xiao, F. (2006, April). Attention funnel: omnidirectional 3D cursor for mobile augmented reality platforms. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 1115-1122). ACM.

15. Theis, S., Pfendler, C., Alexander, Th., Mertens, A., Brandl, Ch., Schlick, Ch. M.: Head-Mounted Displays – Bedingungen des sicheren und beanspruchungsoptimalen Einsatzes. BauA Report, Project F 2288, Dortmund Berlin Dresden (2016)

16. Tümler, J.: Untersuchung zu nutzerbezogenen und technischen Aspekten beim Langzeiteinsatz mobile Augmented Reality Systeme in industrieller Anwendung. Fakultät für Informatik, Otto von Guericke Universität Magdeburg, Dissertation (2009)

17. Büttner, S., Mucha, H., Funk, M., Kosch, T., Aehnelt, M., Robert, S., & Röcker, C. (2017, June). The design space of augmented and virtual reality applications for assistive environments in manufacturing: a visual approach. In Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments (pp. 433-440). ACM.