Evaluation of Augmented Reality in Industry

Tone Lise Dahl, Manuel Oliveira, Emrah Arica

Evaluation of Augmented Reality in Industry

Tone Lise Dahl[1[0000-0003-2524-9091]], Manuel Oliveira[2[0000-0003-3687-1118]] and Emrah Arica[3[0000-0002-2054-9514]]

SINTEF, Trondheim, Norway
{tone.lise.dahl,manuel.oliveira, emrah.arica}@sintef.no

Abstract. Augmented Reality (AR) is seen as a key technology for the development of smart manufacturing. Despite the many possibilities and affordances of this emerging technology, it is a fairly new technology in industry without widespread adoption. Research indicates that there are many affordances of using this technology in industry as well as some challenges. However, there seems to be a lack of research on how to evaluate AR in industry. Based on literature and a case study, we propose guidelines for evaluating AR based on identified dimensions.

Keywords: Industrial Augmented Reality, Evaluation, Guidelines

1 Introduction

Augmented Reality (AR) refers to a real-world environment enhanced with computer-generated information such as sound, video or graphics. AR can be delivered by using smartphones, tablets, head-mounted-displays (HMDs) and projectors [12] [14]. The last few years, augmented reality technology (AR) has arisen in manufacturing [9] and is considered one of the nine enabling technologies that will power the transformation supported by Industry 4.0 initiative [4][10][13]. Although industrial augmented reality (IAR) is one of the key pillars of the industry 4.0 enablers, adoption levels remain low in industry [8]. There also seems to be a lack of research about how to assess IAR in real industrial environments, namely the manufacturing shop floor. The lack of literature on the subject is the genesis for the creation of practical guidelines to evaluate IAR by distilling the lessons learnt from the evaluation studies conducted in the HUMAN research project over a period of two years.

This paper is structured with a theoretical background about AR and evaluation of IAR, followed by a description of the HUMAN project and lessons learnt from the evaluation activities focused on the use of an IAR solution to cognitively augment the worker on the shop floor. The paper then distils the lessons into a set of guidelines to assist in the evaluation of IAR in real manufacturing work environments.

2 Theoretical Background

In this section, we present three categories based on literature to consider when evaluating industrial augmented reality: lab or industrial environment, safety and operability challenges and usability and interface.

2.1 Evaluating AR in industry

Lab or industrial environment

Evaluating AR in industry can be conducted in a lab or in a real industrial environment. When searching for literature, most of the studies we found was conducted in a lab and there seems to be few studies conducted on the shop floor. A survey shows that most of the technical studies have been carried out only in laboratory settings, without implementing the AR system in a real context [2]. Most implementations of AR for assembly seems to be carried out in laboratory settings, and AR studies targeting ergonomics issues have previously been entirely carried out in laboratory settings, with real case applications still lacking in the literature. Laboratory studies typically represent a proof-of-concept of the AR applicability and should be seen as the first step for the development of a scalable solution [2].While evaluations conducted outside the lab considers the state in which the current system is, evaluations in real environments is important because the system should ultimately be used in the real industrial context [5].

Safety and operability challenges

When evaluating AR in a real industrial context, there seems to be some safety and operability challenges that might present a major obstacle, such as working in construction sites [3]. Changes in worker´s workload, time of the day, and location of the room are examples of unquantifiable factors that might skew the findings in the process. In addition, gathering data on an active job site for a pilot study presents potential safety and financial risks for the contractor when using unproven and untested technologies. Also, tracking the environment and accurately displaying the content requires well-lit areas and is very sensitive to heavy shadows [3].

Usability and interface

Usability is an essential variable of all Machine-Interface (HMI) when complexity and number of functions increase. If a technology is not regarded as easy to use and easy to understand, it will not enhance the user´s performances and process at hand and will not be accepted by the users [6]. User interface is described as "does not require special knowledge to use, user friendly and easy to use, with minimum interactions required", and has been identified as one important feature for the future adoption of AR [11]. A consequence of evaluating AR with operators is that they might spend more time to complete the task in total, especially in the first steps, because they don't have previous experience with the technology and therefore might need time to understand instructions given in the AR-system [9]. When assessing usability, it is important that usability is measured relative to the chosen group of users within the specified scope of tasks [6].

The System Usability Scale (SUS) has been used to measure usability [6], others have added questions to a questionnaire based on NASATLX [9] and formulated a hypothesis and conducted a post-experiment questionnaire to evaluate ease of use, satisfaction level, and intuitiveness for each instruction mode [12]. The ultimate desired state of an IAR system that prove its usability, is if the system was deployed to the end-users and is being used without requiring constant presence of the developers. It is also important evaluate if the AR system is developed with input of industrial partners. This is an important factor as it usually introduces a level of reality to a prototype [5].

3 HUMAN Project

The HUman MANufacturing (HUMAN) project was an European H2020 project aimed to develop a platform that is contextually aware of the factory and the human operator to support the operator in performing their tasks with the desired quality, whilst ensuring their well-being. The development of the solutions was permeated with co-creation principles, engaging with different stakeholders in manufacturing companies from inception to deployment, where evaluation played a crucial role in collating feedback and validation. An outcome of the HUMAN project was a fully integrated industrial AR system that aims to increase manufacturing productivity whilst providing support for quality assurance (KIT-AR).

3.1 KIT-AR

Three distinct end-user organizations contributed to the co-creation process that led to the development of KIT-AR: Airbus, Comau and Royo Group. As shown in Figure 1, KIT-AR technology consists of four interactive system modules.

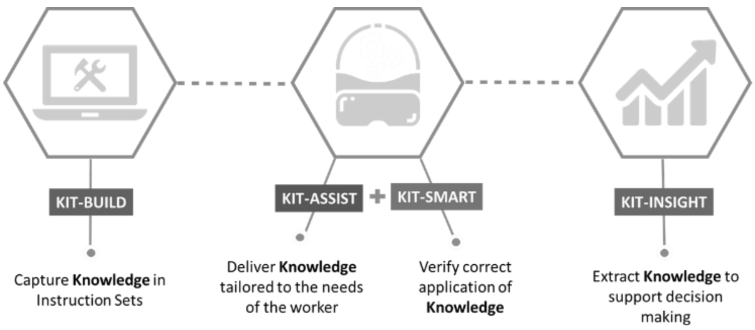

Figure 1.KIT-AR system

3.2 The evaluation studies

The evaluation studies conducted at the three end-users, had the following lifecycle stages: data collection, technical testing, UX testing formative evaluation study and a final study. For the purpose of training the reasoning of the KIT-SMART modules, data was gathered from multiple sensors and the KIT-AR solution. This data gathering happened at shop floor level with the operators. It usually came together with the technical testing. The particular instance of the KIT-AR solution was tested with operators performing their everyday tasks. Preliminary tuning of KIT-AR occurred during all interactions with initial feedback being obtained from operators. Normally, technical testing went together with data collection in the same visit. A particular session to assess the UX of the KIT-AR solution, used a mix of quantitative and qualitative methods and tools. The aim was to improve iteratively the user interface until one can focus on the utility of the solution. With a formative evaluation study, the KIT-AR service was evaluated from all aspects, including the perception of utility. In-depth discussion to expand the initial responses was carried out to understand the underlying motivation and explore potential alternatives to inform the subsequent development cycle. At this stage, the key assumptions of the final study were validated, namely the KPIs. The final stage of evaluation corresponded to the assessment of utility of the KIT-AR solution. The chosen KPIs were measured to determine the impact assessment.

3.3 Lessons learnt

As a result of the evaluation process in each of the three end-user organizations, several lessons were distilled and are succinctly summarized in this section.

Operators availability and scheduling

A fundamental reality of conducting evaluation studies in an industrial environment is that production will not change due to evaluation. The traditional approach is to reduce the impact in production and setup an offline area to setup the experiment for evaluation. However, the human resources are limited, and adequate planning is necessary to ensure that sufficient operators are made available for the evaluation study, with the company incurring costs for the operator’s time taking part. Another shortcoming of offline evaluation is the difficulty to infer transfer of the results into production. When considering evaluation in production, scheduling becomes even more stringent as the evaluation study cannot determinately impact on the production output as measured by the KPIs. In addition, the nature of the task needs to be able to accommodate the overhead of the evaluation activity to avoid introducing unnecessary stress to the operator or risk exceeding the tolerance of the operator and thereby introducing a negative bias in the evaluation results. The access to the operator population is severely limited on two accounts: motivation of operators and the production planning. Although operators are interested in new ways of doing work more efficiently, they are less interested in taking part of evaluation studies. This has an impact on operators' approach of using novel tools that are not easily integrated into existing work practices without an adequate change management process. The production planning imposes constraints on the availability of operators, thus making quantitative studies extremely difficult to carry out. Consequently, one approach used in HUMAN was to conduct qualitative studies with real operators and complement the population of potential participants with students or workers with different roles. This approach requires an experimental design that attempts to characterize the differences between the different subject populations, thus increasing the credibility of combining the evaluation results from both populations. Ultimately, one must plan for small number participants and consider the fusion of evaluation results from operators and additional participants of different type. It is clear that the company cost of the operator's time, and the impact of using the AR solution, needs to be considered in the evaluation study.

Utility

Traditionally, evaluation studies are conducted in a session-based manner in well-controlled conditions and with hands-on support from researchers. The real assessment of the utility of an AR solution is when the corresponding AR device is left over a period of time within easy access of the operators, who then have the option to either use the solution or not. Consequently, the operator will only use the solution if it deemed of value, with clear benefits in terms of productivity and/or quality by means of doing less mistakes. In addition to monitoring of the usage of an AR solution, one can complement with relevant tools such as diaries [1].

Lack of control

A key distinguishing aspect of field evaluation versus a laboratory setting is the almost excruciating lack of control on the work environment. A manufacturing company will not modify the work environment for the benefit of the study, in particular when considering the study in production and not an offline setting. As part of the evaluation methodology adopted in HUMAN, extensive technical testing was carried out to validate the usage of technology considering the constraints of the work environment, such as:

- the production environment had significant impact on the use of wireless connectivity, which enforced edge computing;
- the HoloLens device (version 1) is highly susceptible to the amount of luminosity that had an impact on the rendering of holograms and consequently requiring redesign of the instruction sets in some cases;
- The seasonal ambient temperature of the work environment may have an impact on the correct functioning of the IAR device.

In any case, the potential acceptance of any deviations to the work environment requires careful negotiation with different stakeholders in the manufacturing company beyond the expected operator and managers, extending to include social actors such as trade unions. Under no conditions should an evaluation study compromise the health and safety regulation of the work environment.

Quality of user interface and content

The quality of the user interface has significant impact on the evaluation results and consequently, it is highly relevant to have appropriate tools to assess the usability with the purpose of ensuring the focus is on the utility of the solution. The expectations of operators in a production environment are high and there is less tolerance for poor usability, thus the recommendation is to iteratively improve the solution until a usability tool such as SUS [7] indicate that acceptable UX has been achieved. However, even then, one should design the effectiveness of the interface design, possibly experimenting different variants of the user interface. In addition to the quality of the user interface, one needs to consider the quality of the content of the AR solution being used. Consequently, it is recommended that the instructions are created by a product engineer or individual with appropriate expertise and experience. An alternative is to involve experienced operators to co-create the instruction sets. However, irrespective of how content is created, it is recommended to conduct heuristic evaluation on the quality of the content.

Onboarding

The introduction of a new technology, such as AR, requires learning of the solution to enable operators to build the appropriate mental model. This onboarding must be considered when conducting the evaluation to eliminate the learning effect as it will distort the results otherwise and possibly introduce bias in the analysis and interpretation of the collated data.

4 Evaluation Guidelines

Based on the lessons learnt in the HUMAN project, combined with the limited knowledge reported in literature concerning IAR evaluation in real work environments, a set of four guidelines for evaluating IAR is proposed in the following subsections:

Experimental Design

A fundamental premise of evaluation of IAR in real manufacturing work environments is the fact that the experimental design must accommodate the constraints of the industrial environment in terms of:

- **Technical.** The characteristics of the work environment may negatively impact the correct operation of the IAR solution. *Consequently, it is necessary to accommodate adequate testing and piloting of the experimental design.*
- **Cost.** There is always a cost to the company to run the evaluation with potential risk to loss of productivity, exacerbating the cost further. *It is necessary to negotiate with the company the financial implications and possibly modify the experimental design to accommodate the cost impact.*
- **Participant population.** The workers are an extremely scarce resource that impose constraints on number of participants affecting the experimental design. *Start the discussion with production managers early to identify the actual number of participants available. Consider the fusion of qualitative results of studies involving workers with quantitative results from studies involving non-workers.*
- **Production.** The production planning dictates when evaluation activities can take place and the health/safety regulations cannot be compromised. *Consider the need to conduct the evaluation study over a long enough period to accommodate the production schedule.*
- **Return on Investment (ROI).** There is a cost of conducting evaluation on the shop floor. *Understand the needs of the company and manage their expectations*

concerning the outcome of the evaluation. Strive to associate evaluation goals to the business as it contributes to an understanding of potential ROI.

Consequently, it is necessary to plan in advance for the evaluation and negotiate with the relevant stakeholders, from management to team leaders.

Content Quality

An IAR solution provides digital overlay of knowledge to workers and traditionally assessing the impact of the quality of the content is disregarded when conducting evaluation studies. *Consider the following two measures:*

- *Co-design the instructions with the workers and relevant stakeholders. This reduces the impact of poorly designed content;*
- *Conduct heuristic evaluation of the content quality with trainers/experienced worker. This measure may contribute to the improvement of the quality of the content or to understand the limitations of the final instance of the content.*

Interface and usability

It is fundamental to ensure that the User eXperience (UX) is well designed so usability is not a barrier to utility. *Consequently, one should consider the following measures:*

- *Strive to improve the UX until the law of diminishing returns applies. This may involve changes to the development of the solution or content workarounds;*
- *Assess the impact of the different devices on the UX, both from the constraints on the interface and the ergonomics;*
- *Understand the impact of the UX on the outcome of the evaluation results by doing a small pilot and combining it with an heuristic evaluation by UX experts;*
- *It is fundamental to have onboarding of the operators in the use of IAR as they require a learning period to acquire the mental model of the system.*

Execution

The most well-planned evaluations are subject to the uncertainties of production realities and as such may disrupt the evaluation activities. *As such, it is necessary to be able to adapt, but most importantly use every opportunity to gather data and conduct analysis incrementally to best understand how to make necessary changes to the evaluation activities to mitigate the negative impact of disruption.*

Conclusion

The presented guidelines are not exhaustive, neither comprehensive, but are an initial approach that builds upon the experience of two years in conducting iterative evaluation studies using IAR. The aim is to provide insights on how to improve evaluation activities in real industrial work environments towards establishing a good practice of research on the shop floor.

Acknowledgements

The authors would like to thank the operators that took part of the evaluation activities using the KIT-AR solution. This work was partially funded by the Norwegian Research Council (#28026) and the H2020 HUMAN project (Grant Agreement nº: 723737).

References

1. Bolger, Niall & Davis, Angelina & Rafaeli, Eshkol. (2003). Diary Methods: Capturing Life as it is Lived. Annual review of psychology. 54. 579-616. 10.1146/annurev.psych.54.101601.145030.
2. Bottani, Eleonora & Vignali, Giuseppe. (2019). Augmented reality technology in the manufacturing industry: A review of the last decade. IIE Transactions. 51. 284-310.
3. Chalhoub, J., & Ayer, S.K. (2019). Exploring the performance of an augmented reality application for construction layout tasks. Multimedia Tools and Applications, 78, 35075 - 35098.
4. Davies, R. (2015). Industry 4.0: Digitalization for Productivity and Growth. European Parliamentary Research Service.Fite-Georgel,P. "Is there a reality in Industrial Augmented Reality?" 2011 10th IEEE International Symposium on Mixed and Augmented Reality, Basel, 2011, pp. 201-210.
5. Fite-Georgel,P. "Is there a reality in Industrial Augmented Reality?" 2011 10th IEEE International Symposium on Mixed and Augmented Reality, Basel, 2011, pp. 201-210.
6. Holm,M., Danielsson, O., Syberfeldt,A. Moore,P. & Wang,L. (2017). Adaptive instructions to novice shop-floor operators using Augmented Reality, Journal of Industrial and Production Engineering, 34:5, 362-374, DOI: 10.1080/21681015.2017.1320592olm et al, 2017
7. Lewis, J. & Sauro, J. (2009). The Factor Structure of the System Usability Scale. Proceedings of the 1st International Conference on Human Centered Design: Held as Part of HCI International. 5619. 94-103. 10.1007/978-3-642-02806-9_12.
8. Masood,T. & Egger,J.(2019). Adopting augmented reality in the age of industrial digitalisation,Computers in Industry,Volume 115,2020,103112, ISSN 0166-3615, https://doi.org/10.1016/j.compind.2019.07.002.
9. Mourtzis, D., Zogopoulos, V. & Xanthi, F. (2019). Augmented reality application to support the assembly of highly customized products and to adapt to production re-scheduling. Int J Adv Manuf Technol 105, 3899–3910.Https://doi.org/10.1007/s00170-019-03941-6
10. Romero, D., Stahre, J., Taisch,M. (2020).The Operator 4.0: Towards socially sustainable factories of the future, Computers & Industrial Engineering,Volume 139,106128, ISSN 0360-8352, https://doi.org/10.1016/j.cie.2019.106128.
11. Stoltz,H.,Giannikas,V., McFarlane,D., Strachan, J., Um,J., Srinivasan,R. (2017).Augmented Reality in Warehouse Operations: Opportunities and Barriers, IFAC-PapersOnLine, Volume 50, Issue 1, 2017, Pages 12979-12984,ISSN 2405-8963.
12. Uva, A.E., Gattullo, M., & Manghisi, V.M. (2018). Evaluating the effectiveness of spatial augmented reality in smart manufacturing: a solution for manual working stations. Int J Adv Manuf Technol 94, 509–521. Https://doi.org/10.1007/s00170-017-0846-4
13. Zolotová,I., Papcun,P., Kajáti,E, Miškuf,M & Mocnej,J. (2020).Smart and cognitive solutions for Operator 4.0: Laboratory H-CPPS case studies, Computers & Industrial Engineering,Volume 139,105471, ISSN 0360-8352, https://doi.org/10.1016/j.cie.2018.10.032.
14. XR association (n.y). XR PRIMER 2.0: A STARTER GUIDE FOR DEVELOPERS.