Evaluating Added Value of Augmented Reality to Assist Aeronautical Maintenance Workers — Experimentation on On-Field Use Case

Quentin Loizeau[1*], Florence Danglade[1], Fakhreddine Ababsa[1] and Frédéric Merienne[1]

[1]LISPEN EA 7515, Arts et Métiers, Institut Image, Chalon-sur-Saône, 71110 France
quentin.loizeau@ensam.eu, florence.danglade@ensam.eu,
fakhreddine.ababsa@ensam.eu, frederic.merienne@ensam.eu

Abstract. Augmented Reality (AR) technology facilitates interactions with information and understanding of complex situations. Aeronautical Maintenance combines complexity induced by the variety of products and constraints associated to aeronautic sector and the environment of maintenance. AR tools seem well indicated to solve constraints of productivity and quality on the aeronautical maintenance activities by simplifying data interactions for the workers. However, few evaluations of AR have been done in real processes due to the difficulty of integrating the technology without proper tools for deployment and assessing the results. This paper proposes a method to select suitable criteria for AR evaluation in industrial environment and to deploy AR solutions suited to assist maintenance workers. These are used to set up on-field experiments that demonstrate benefits of AR on process and user point of view for different profiles of workers. Further work will consist on using these elements to extend results to AR evaluation on the whole aeronautical maintenance process. A classification of maintenance activities linked to workers specific needs will lead to prediction of the value that augmented reality would bring to each activity.

Keywords: Augmented Reality, Aeronautical Maintenance, On-field Evaluation, Criteria, Use case selection, Added value.

1 Introduction

Aeronautics is a demanding sector as the security of millions of passengers is at stake every day [1]. It requires rigorously following procedures in order to reach the right quality standards. Aircrafts are made of advanced products and equipment optimized to assure essential advanced technological function while responding at best to numerous constraints.

Aeronautical maintenance deals with both the complexity and the variety of equipment and the quality requirements to fulfill. For workers, it heavily impacts maintenance manuals which are overloaded by all the information needed to conduct each maintenance task on each configuration of equipment while applying the right procedures according to regulation structure [2].

This industrial activity is constrained by an ever expanding aeronautical sector [1] and the renewing of its workforce [3] that increases the need of recruiting and training new operators. However, beginners cannot assimilate fast enough all the knowledge that retiring experimented workers learned in more than thirty years to work efficiently with current tools and instruction manuals.

The support of work information needs to evolve to ease the tasks of researching and understanding the right information. The goal is for the new operators to be operational at the best level possible and to accompany them as they improve their skills. This aims to impact global productivity of beginners from the beginning.

Augmented reality (or AR) appears to be a well suited response to the current needs in the aeronautical maintenance activity. AR technology enhances the perception of a situation by superposing digital information onto the real world, at the right place and at the right time [4]. A user equipped with well suited AR has access to a better and faster understanding of situations.

It is becoming an unavoidable tool for the future of various professional fields such as culture, entertainment, medical, retail or industry [5]. Pertinent use cases of augmented reality are demonstrated on a wide range of specific conditions and there is a will to extend AR applications to more common and widespread tasks [6].

2 Scientific and Technical Problem

AR technology can simplify the access to useful information and instructions by providing elements filtered according to the on-going tasks. It eases understanding by displaying data and models of equipment, tools and manipulations directly onto the physical work environment [7]. The digital support also makes interactions between users and data more efficient [8].

This explains why it seems interesting to implement AR tools to help workers on aeronautical maintenance activities. However, there are only few returns on actual measurement of what this technology could bring concretely to this industrial activity.

It is needed to evaluate the added value of the use of AR to assist operators in aeronautical maintenance in real working conditions in order to determine more precisely the benefits of using it and support further deployment.

First it is necessary to have a clear vision not only of criteria for evaluation of augmented reality but also of aeronautical maintenance environment to select or adapt right criteria while taking the context of usage into account. Then it will be possible to apply this on real use-cases, using augmented reality tools developed for this purpose, to experiment with workers on assembly subtasks and reach the goal of the study.

Thus, the three questions we aim to answer in this paper are:

- Q1: How to evaluate benefits brought by AR on aeronautical maintenance tasks?
- Q2: How to deploy an AR tool for assistance on aeronautical maintenance tasks?
- Q3: What is the added value of AR on aeronautical maintenance tasks?

This paper is structured as follows. Next part, section 3, details the evaluation criteria for AR applications through state-of-the-art study, followed by the selection criteria suited for our subject according to aeronautical maintenance and on-field constraints. Section 4 describes the methodology used to deploy AR to assist operators on maintenance use-cases in relation to current support of information and working instructions. Following, section 5 presents the implementation of an experiment conducted to evaluate the added value of AR comparing to current practices on on-field tasks, the results obtained and the related analysis. Finally, section 6 ends the paper with a conclusion on the work and an opening to further research on the subject.

3 State of the Art – AR Evaluation Criteria for Aeronautical Maintenance

Augmented Reality (AR) is defined by Azuma et al. [5] as a technology which combines in interactive way virtual elements with the real world both spatially, in three dimensions, and temporally, in real time. This technology brings digital data to the users directly in context, permitting him to focus more on the physical tasks he works on and to have a better understanding of situations [7].

Augmented reality relies on software capable of recognizing and tracking elements of the real environment (2D, 3D, plan, and geo-localization) to place and maintain virtual elements in the right physical position [9]. This AR software works on different types of hardware that can acquire data needed for localization and display digital elements directly (projectors, see-trough glasses) or indirectly (screen) [10].

In the fields of industry, AR can be used in the whole life of a product, on design, planning, manufacturing, inspection and maintenance activities [11]. Reviewing recent studies on the subject, Quandt et al. [12] identified general requirements for the development of industrial applications. Among those tasks are design of products and factories, training of operators, assistance to production, support of logistics and remote maintenance. Requirements concerns all elements impacted by applications, from cost and security in integration to set-up, accuracy and reliability during usage.

Palmarini et al. [13] classified the different area identified for application of AR in industry. They detect high interest for AR application on aeronautics and maintenance activities such as assembly, repair, inspection and training.

3.1 Evaluation of AR Technology

Evaluation of AR on Technology Side. One focus on AR evaluation is on the comparisons on the technology itself. Through a Systematic Literature Review, Palmarini et al. [14] identified a set of main characteristics that are compared in most studies including hardware, development platform, tracking technics and interaction methods.

Baumeister et al. [15] compared the possibility offered by hardware for procedural tasks by evaluating the mean response time to AR indications. Renner and Pfeiffer

[16] compared visualization methods for guidance purpose by measuring mean time for action and mean head movement during tasks.

Evaluation of AR on Usage Side. The evaluation against other tools in controlled environment is another research topic [17]. Webel et al. [18] compared AR to video for training. Rios et al. [19] compared AR to paper instruction for troubleshooting. The same way, Syberfelt [20] evaluate AR guidance for assembly. Fiorentino et al [21] testing AR on large screen against paper instructions, distinguish the time on tasks when AR can be useful or not. Blaga et al. [22] used a virtual environment to measure the precision of gestures with freehand gesture interactions in terms of reaction time, action time and accuracy.

In most cases, these experiments are conducted in laboratory conditions and on specific tasks where the conditions make it easier to identify and measure comparison criteria to highlight differences between solutions.

Evaluation of AR on User Side. The third important element working with AR is the user. The tools used impact the way of work on the task and thus the level of engagement needed for the work and felt by the user through the task. This concept of cognitive load [23] can be evaluated with quantitative or qualitative observations.

For quantitative observations, direct solutions are to measure physiological values of the users before and during the task like brain activation [24] or to conduct dual task measures by adding another task to do beside the original task [24]. The evolution between the two situations indicates the cognitive load induced by the task.

For qualitative observations, Hornbaek et al. [25] evaluated different solutions and concluded that standard questionnaires are better than homemade ones. Rubio et al. [26] compared subjective methods for evaluation of mental workload and identified the NASA-TLX for prediction of user's performance on tasks.

Hart et al. [27] defined the NASA-TLX (task load index) through years of research on workload evaluation. The users need to assess the six elements defining workload on a rating scale after the task. The mean workload score of the task can be used for predictions of performance, a low value is linked to better performance.

Brooke et al. [28] worked on usability to define a questionnaire focused on the user feeling of the tools during the task with ten questions rated on a five-points Likert scale. This score can be used to compare the usability of different kind of solutions. Bangor et al. [29] analyzed hundreds of studies using the SUS and linked SUS score to acceptability rating of a system through percentile rank comparison.

The validity of both subjective questionnaires NASA-TLX and SUS to predict objective impacts on performance has been confirmed through compilation of hundreds of studies in retrospectives more than twenty years after their construction [30-31].

3.2 Constraints of Aeronautical Maintenance

The objective of aeronautical maintenance is to respond to maintenance requirements adapted for aeronautics sector through processes like MSG-3 [32]. This reliability-oriented process is built to ensure maintaining safety and reliability levels of the aircraft or restore them to optimum level after deterioration and obtain required data to improve design at minimal cost.

To fulfill these objectives, maintenance activities encompass numerous procedures punctuating the life of all aeronautic equipment and depending on lifetime, flight cycles and direct observations. Some circumstances occur only a few times in the life of the equipment, which leads to low frequency of product going through some maintenance tasks. Modifications creates new configuration of equipment and thus different procedures to follow, impacting the complexity of maintenance.

Needs in Aeronautical Maintenance. Martinetti et al. [33] state that it requires gathering and using a substantial amount of information on standard procedures, specific to the various tasks needed and to each equipment to conduct the maintenance tasks. Each product being able to go through maintenance process at different moment of his life, every product can be different from the other, which makes them even harder to work on than in production where every component is the same.

The needs are to identify the products, to collect the right information necessary to conduct the task, to interpret documentation and understand instructions and eventually to check the result and validate the proper execution of the task. These needs are not fulfilled yet with current tools, but AR technologies seem to bring some solutions.

Constraints and Evaluation. Constraints apply on the global activity with the growth of the aviation industry worldwide [3] which speeds up the need to find solutions to maintain quality and increase productivity on tasks along with improving efficiency of operator training in various working conditions. Moreover, variability of type and configurations of maintained equipment makes it impossible memorize information related to each task and the requirements of aviation industry make it necessary to be certain that the right information have been used.

Various indicators are used to evaluate aeronautical maintenance at different scales. On a large scale, the most important factor is reliability, measured through indicators such as mean time before failure. However, these elements pilot maintenance on a global level and don't connect to tasks conducted during maintenance, which are monitored through other criteria. Some criteria related to operator safety are considered in these evaluations.

The major criterion related to the maintenance activities is the quality level. Improper execution of maintenance instructions can affect the product which must be discarded or redo part of the maintenance cycle. Parts can even be damaged during task which is considered as non-quality. This leads to significant financial impacts associated to the cost and time for repairing or repurchasing parts and impact the overall maintenance cycle through time required to remake the maintenance on the parts concerned and delay on the following maintenance steps.

The second equally important criterion related to the maintenance activity is the time required to complete the maintenance tasks. Equipment maintenance times are contractually defined between the owners of the equipment and the maintenance facility. It is necessary to respect at best these times to avoid penalties. The delays that can accumulate with each maintenance step and impact the delivery date.

All these requirements will be considered during the selection and the evaluation of AR impact in aeronautical maintenance.

Evaluation Criteria for Aeronautical Maintenance Activities. The activities of aeronautical maintenance are evaluated in terms of productivity. From this point of view, the main interest for evaluation is to observe evolution of performance due to the use of augmented reality. Criteria used should highlight direct impact of AR on quality indicators and time of realization for tasks. However, this would neglect other advantages of AR on operators.

Jetter et al. [34] identified what could be KPI (Key Performance Indicators) on the integration of AR systems in automotive maintenance through subjective analyses and questionnaires. They extracted two main KPI for the usefulness of AR which are reduction of time & error and perceived ease of use of AR solutions.

4 Methodology – For Implementation and Evaluation of AR Application on Maintenance Use Cases

The methodology driving our work on the use of augmented reality on aeronautical maintenance tasks is synthetized in figure 1. Experimentations on aeronautical maintenance task are conducted to obtain results on the added value of AR. The setup of the experimentations requires selecting the evaluation criteria in accordance to the task (left side of the figure). It also requires having deployed a use case in augmented reality in the same environment as current practices (right side of the figure).

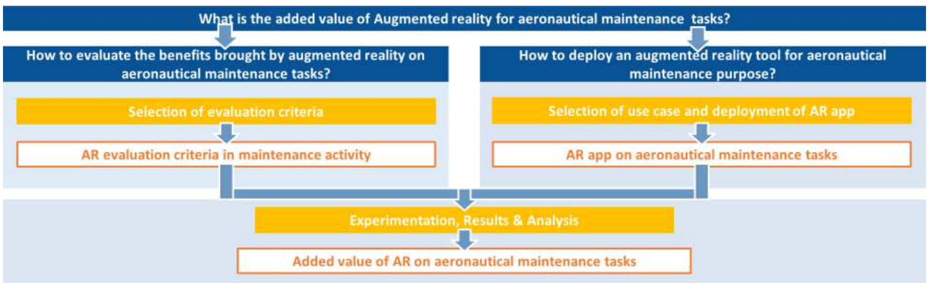

Fig. 1. Methodology for evaluating the added value of augmented reality on aeronautical maintenance tasks

4.1 Selection of Evaluation Criteria

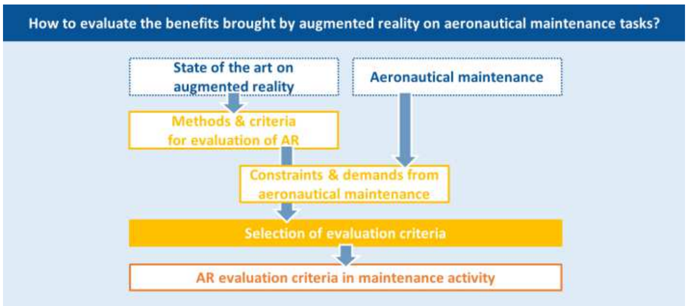

Fig. 2. Method for selecting criteria for AR evaluation in aeronautical maintenance

As synthetized in figure 2, by applying constraints of the industrial activity to the list of criteria established to evaluate augmented reality we are able to extract the criteria best suited for our study. The impact on the maintenance process will be assessed through the comparison of instruction understanding time, task action time and global time needed for task. To complete the observation, we compare the contribution of AR to the operators work with the two questionnaires, NASA-TLX for cognitive load measurement and SUS for usability of the AR solution.

Proposition for AR Evaluation. Table 1 summarizes observation methods (col.2) for evaluation of AR on different points of view (col.1). These are applicable and pertinent to compare AR to other solutions in specific and controlled conditions (col.3) but not all suitable to be used on-field, on real maintenance tasks. However, the needs and constraints of aeronautical maintenance require selecting criteria that can be observed in industrial conditions and relevant for deployment of AR in this activity.

Subject	**Observation methods**	**Setup constraint**	**Ref**
AR technology	Response time to indications	Laboratory only	[15]
AR technology	Head movement during task	Laboratory only	[16]
AR technology	Precision of hand gesture	Laboratory only	[22]
Tools comparison	Time to complete a task	No constraint	[20]
Tools comparison	Quality of the result of the task	No constraint	[19]
Impact on user	Dual task	Laboratory only	[15]
Impact on user	Physiological observation	Laboratory only	[24]
Impact on user	Questionnaires	No constraint	[16]

Table 1. Possibilities for evaluation augmented reality

Selection in Accordance to Constraints. AR facilitating understanding and access to right information, it is suited to address the needs on maintenance tasks and reach time gains as well as reducing quality issues. There is a need to measure feedback on current job with evaluation criteria related to current activities and way of measuring productivity in terms of quality and time taken to realize the task. On-field it is impossible to use intrusive measures constraining the operators such has in laboratory.

The solution is to use criteria and apply method suited for the workshop environment, which are, referring to table 1, questionnaires for the impact on user and tools comparisons through mean time to complete a task and mean quality of the result of the task. This allows gathering results both on process side through productivity indicators and user appropriation side with cognitive load and usability questionnaires.

Summary of the Selected Criteria. According to the indicators identified in previous sub steps, the structure most suitable for classification of criteria is based on the notion of usability described with the three elements of the ISO 9241-11 norm [35]:

- Effectiveness is the ability of users to complete tasks using the system and correspond to the selection of use cases where AR is applicable with potential benefit.
- Efficiency concerns the level of resource consumed in performing tasks. It is linked to the performance indicators on maintenance tasks which includes measurement of task execution time and assessment of the error rate during these tasks.
- Satisfaction embraces subjective feedbacks on the use of the system through measurement of user acceptance & cognitive load with SUS and TLX questionnaires.

The vision by mean of the usability notion highlights two elements essential for evaluation that are the impact on the maintenance operation through efficiency criteria and the acceptance of AR by the users measured with satisfaction indicators.

Time Measurement Values. AR have different impact on steps dealing with the understanding of information and steps of work on parts, thus the duration of each subtask has been divided between “Understanding Time” (or T_U) and “Action Time” (or T_A). T_U corresponds to the time used by the participant to research, understand and translate the instructions from the support to the real parts. T_A corresponds to the time used by the participant to execute the instructions on the equipment.

Total Time Gain calculation. The “Time Gain” (or T_G) quantity has been defined to evaluate the added value of AR on the performance on the subtasks. With T_{AR} the time recorded “with AR” and T_{CS} the time recorded “without AR”. T_G consists on the comparison of time requested for each phase calculated with the following formula:

$$T_G = (T_{CS} - T_{AR}) / T_{CS} \quad (1)$$

Understanding/Action ratio calculation. The quantity “U/A ratio” (or $R_{U/A}$) has been defined to evaluate the repartition of time needed to research and understand the instructions T_U and time needed to execute actions T_A in regard of the total time T_T needed to complete the task or subtask. It is calculated with the following formula:

$$R_{U/A} = T_A / T_T \text{ with } T_T = T_A + T_U \quad (2)$$

4.2 Selection of Maintenance Use Cases for AR Evaluation

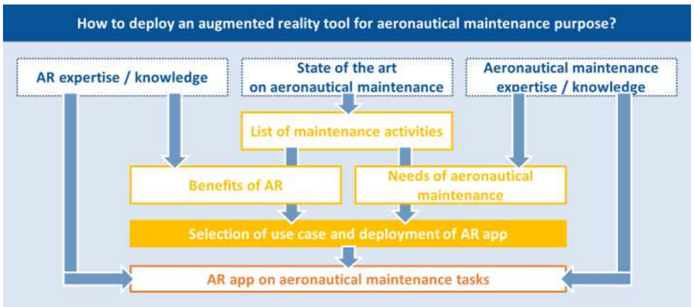

Fig. 3. Method for selecting and deploying AR apps to assist maintenance task

Details on Aeronautical Maintenance. First element of deployment method (fig 3.) is to work on maintenance activities. These activities occur at different times in the lifecycle of equipment and can be light unscheduled intervention, for replacement of sub-equipment, or heavy scheduled intervention requiring specific resources and skills such as overhauls or modifications. The second type identified as depot-level is more suited for AR tools due to the complexity of tasks and the variety of equipment covered. This level of maintenance can only be performed by accredited Maintenance, Repair & Overhaul workshops (or MROs) and concerns important repairs and a wide variety of tasks (fig. 4.) leading to the complete overhaul of equipment.

Maintenance Activities	Description of the activities	Instructions Quantity	Instructions Complexity	Instructions Variability	Hands needed	Mobility needed
Preliminary Inspection	Identifying the global state of the on-coming equipment	HIGH	LOW	HIGH	NO	YES
Disassembly	Separating each part of the equipment	HIGH	MID	MID	YES	YES
Cleaning	Cleaning each re-usable part, removing treatments when needed	LOW	LOW	LOW	YES	NO
Testing	Controlling the state of individual parts	MID	MID	MID	YES	NO
Repair	Repairing parts that need/can be repaired	MID	HIGH	MID	YES	NO
Treatments	Applying treatments on part that need to be treated again	MID	MID	MID	YES	YES
Sub-assembly	Assembling the sub-components of the equipment	HIGH	HIGH	HIGH	YES	NO
Final Assembly	Assembling the sub-components together	MID	MID	MID	YES	YES
Final Inspection	Verifying the conformity of the final state of the equipment	HIGH	LOW	MID	NO	YES

Fig. 4. Tasks conducted during overhaul maintenance process

Core elements of these tasks are different and call on different skills and resources but there also are similarities in the global action sequence related to each task. Every task requires referring to standard instructions describing the elements needed to conduct the task (tools, grease …), the task itself and the way to validate the execution.

The focus on this study is on the use of AR content in the shop and the help it gives to the operators for MRO tasks. It does not explore yet the authoring part and the management of maintenance data, which are also impacted by the use of AR technology, even if the use of digital and identified content should bring gains to these types of tasks, especially in terms of content updates.

Selection of Activities. For a same task, complexity varies according to the type of equipment and subtasks. It is essential to work with the experts to select the right use-cases. Demonstration of AR on an operation well known to users helps them extrapolate to their daily activities. It allows gathering feedback from the field, identifying specific complex tasks and selecting the use cases in accordance to the needs.

Assembly tasks on sub-equipment before final assembly were selected as they consist on multiples subtasks and thus requires important amount of information to search, understand and translate into actions for each step.

4.3 Implementation of AR Application in Aeronautical Maintenance

Once the use case is selected, it is necessary to observe the current work environment to identify which AR solution (fig. 5.) is the more pertinent and choose one well suited depending on the constraints detailed in previous chapter.

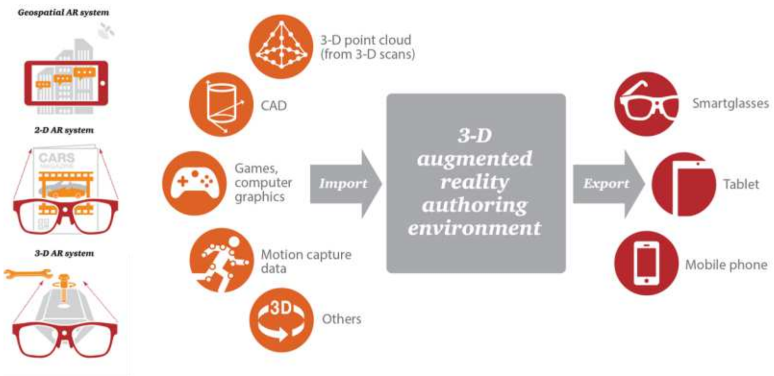

Fig. 5. AR solutions (hardware and software possibilities) [10]

Software Selection. Concerning the software, workshop conditions and aeronautics environment advise against the use of markers recognition. It would require development of specific tooling and procedures to install it. Maintenance tasks consist on working on mechanical parts which qualifies 3D-model based recognition.

The Diota [36] software solution has been selected as it uses a 3D model-based recognition which is able to track mechanical parts with accuracy to overlay 2D or 3D data on job cards. It avoids setup constraints imposed by other types of AR recognition and make the solution pertinent for an industrial use in aeronautics. The authoring solution allows using existing 3D models of parts from design for creation of static or animated job cards which facilitate the deployment of applications.

Hardware Selection. Hardware takes many forms from computers or tablets to unusual ones as glasses or projectors, each having advantages and drawbacks. The choice is impacted by the mobility needed, the possibilities of manipulations by hand, the available software and the readiness of hardware for use in industrial environment.

The AR software selected is usable on hardware running Windows 10 Operating System like PC, tablet, HoloLens V1 and a specific Diota projector. The projector has been rejected because the way of displaying information was not relevant to the use case and its volume. The HoloLens have small field of view, low autonomy and requires practicing time for interactions. The tablet requires manipulating the device with one or both hands, which is problematic when it is needed to interact with parts.

A PC equipped with a touch screen and an industrial mobile camera was selected and installed on a standard mobile working desk used in the workshop.

4.4 Implementation Process for AR Application

The last step is to create the content for the AR app itself using the data coming from current process, AR development tools and the feedback of maintenance operators. Current documentation provides information and key elements to consider. AR development tools helps organize each step around reference parts which link the virtual and physical world together. Existing 3D models from design are used for visual instructions and set up into the work environment in 3D to make AR use natural.

Reviews and Validation with the Users. The review of early version of the application with maintenance operators permit to detect mistakes due to wrong interpretation of current documentation or identifying missing information needed for the task. It also helps verifying the correct disposition of elements. The creation process continues by applying changes and iterating with the operators until the final version fixes the AR app content for industrial use. It can be introduced into the maintenance process to train the operators, conduct experiments and evaluate AR.

5 Experiment – Evaluation of AR on Selected Use Cases

The experiment was conducted on an assembly task divided into eight subtasks of similar complexity to observe the impact of AR on this kind of task in industrial conditions. The participants completed the task into the workflow of the factory in two conditions, one with the current paper supports used by workers and one with an AR support available on a mobile workstation. The completion time was recorded and the feedback of participants has been assessed through questionnaires right after the task.

5.1 Implementation

The use case has been selected according to the methodology detailed in previous section. It is on assembly, a task type involving many steps that can be enhanced with current AR technology. Environment study and interview with the maintenance operators conducted to a selection of eight complex subtasks, on which a need for a different instruction format has been identified such as installing parts or axles in position, tighten bolts and protect areas with grease (fig. 6.).

The AR application was deployed with the industrial AR software solution Diota V2.3.0 connected to the 3D authoring software CatiaComposer R2017X. The hardware is a PC running Windows 10 Operating System equipped with a 27” touch screen and an industrial HD camera installed on a standard mobile desk. This mobile workstation can be moved to be installed on the assembly line to test the “with AR” condition in the real environment without disturbing the workflow.

The AR instructions were provided through an industrial AR application. For each subtasks the main part is tracked in 3D and an AR job card overlay models of other parts in position, reference numbers of parts and the standardized textual instructions.

Fig. 6. Pictures of the equipment, AR application and workstation used during experimentation

5.2 Measurement Protocol

9 participants were recruited among the operators of the workshop to take part in the study. 6 participants completed the eight subtasks in both conditions (“without AR” & “with AR”) and 3 participants completed them only in the “with AR” condition before filling the two qualitative questionnaires. None of them add previous knowledge of AR technology. Participants were divided into three groups (Table 2) according to their knowledge on the task and on maintenance.

Groups	**Knowledge on task**	**Knowledge on job**	**N° of participants**
Beginner	Null	Less than 3 years	2
Confirmed	Null	More than 3 years	5
Expert	Significant	More than 3 years	2

Table 2. Repartition of participants according to their knowledge on the task and the job

Two conditions were evaluated on the assembly line in the real working area. For the “without AR” condition, there were few adaptations compared to the current process. For the “with AR” condition, a mobile workstation (Fig. 6) has been built to bring AR into the assembly line.

The “without AR” instructions consisted in standard working documents containing the information need to do the assembly. It was composed of overview figure, standardized textual instructions and details pictures for each subtask.

Procedure. A demo application not related to the observed task was presented to explain the possibilities of AR and the way it works. The experiment conditions, the questionnaires and the measures were presented to participants before volunteering.

Before the experiment, the observer prepares parts needed for the assembly, the current support for the task, the AR application, questionnaires and a support for time measures. Participants are reminded of the experiment conditions. In the "with AR" conditions, the participant goes through a brief overview of the controls. The mobile workstation is installed into the workspace and calibrated for recognition of the parts.

The experiment begins when the observer, the participant, the support and the equipment are ready. The participant navigates through the support to find the information and realize the eight subtasks of the assembly. The observer writes down the understanding and action time needed by the participant to complete each subtask.

After the experiment, the participant fills in the NASA-TLX questionnaire to assess the cognitive load induced on him by doing the task with the associated support. He also fills in the SUS questionnaire to evaluate the usability of the AR support.

5.3 Results

The results of the experiment are of two types, quantitative for the evaluation of the impact of AR on user's performance doing the task and qualitative for the cognitive load induced on the users by the use of each instruction support and subjective evaluation of the AR application usability.

The results are divided by experiment conditions ("with AR" and "without AR") and by profile of participant ("Beginner", "Confirmed" or "Expert") for each subtask.

Quantitative Performance Measurement – Process Side

The results are synthetized in table 3 and presented in figure 7 for "Total Time Gain" (T_{G-T}) and in figure 8 for $R_{U/A}$ calculated with formula (1) and (2) from chapter 4.1.

	Nb of measures	T_G on understanding	T_G on action	T_G in total
All profiles	40	30%	-16%	9%
Beginner	16	25%	-11%	7%
Confirmed	16	39%	-27%	11%
Expert	8	15%	-38%	2%

Table 3. Results on time observation per subtask for all profiles of users

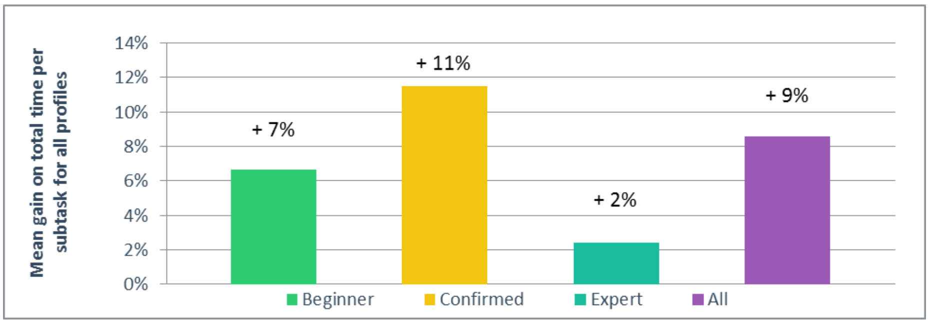

Fig. 7. Mean gain on total time per subtasks for all profiles of users

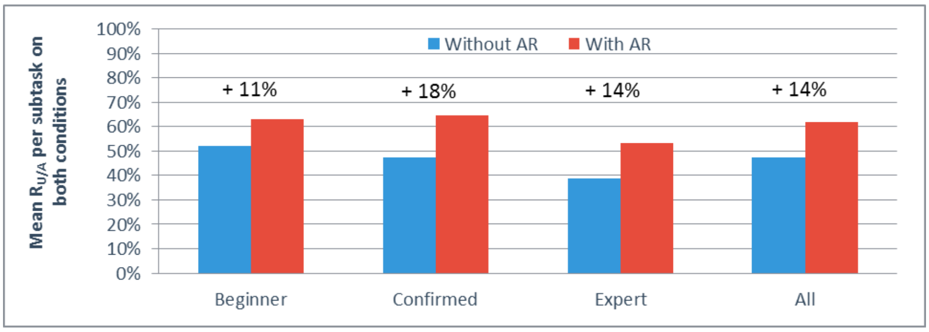

Fig. 8. Mean $R_{U/A}$ value per subtasks for all profiles of users in both conditions

On All Profiles. Considering all profiles of users, 5 out of the 9 subjects, on each of the eight subtasks we obtain 40 measures on both conditions. Per subtask, workers spend a mean value of 30% less time on the understanding phase and 16% more time on the action phase which leads to a gain of 9% on total time per subtask. $R_{U/A}$ shows a gain of 14 points.

On Each Profile. Considering beginner, confirmed and expert profiles separately on each of the eight subtasks we obtain respectively 16, 16 and 8 measures on both conditions. Observing mean value per subtask, workers spend less time on the understanding phase (25%, 39% and 15%) and more time on the action phase (16%, 27% and 38%) which leads to a gain on total time per subtask (7%, 11% and 2%). $R_{U/A}$ shows a gain for each profile of respectively 11 points, 17 points and 14 points.

Qualitative User Related Measurement – User Side. The results of NASA-TLX and SUS questionnaires are presented in figure 9 respectively left side and right side.

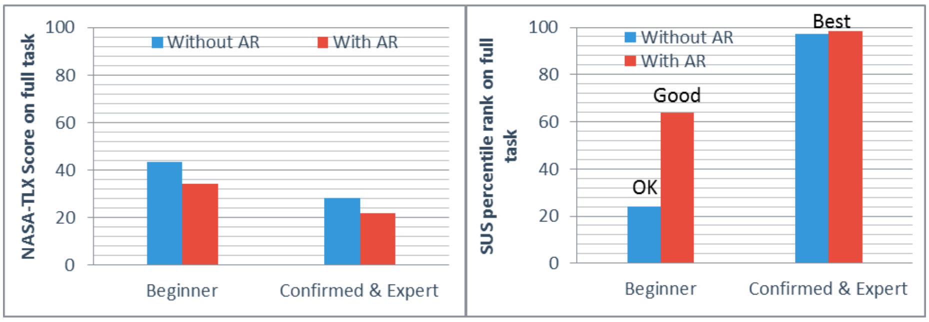

Fig. 9. Mean TLX score and SUS percentile rank in both conditions for each profile of user

5.4 Analysis

The distribution of workers profile for the experiment corresponds to global repartition of skills in the workshop. The aim with augmented reality is to assist workers in finding and understanding instructions to ease the execution of maintenance task. Depending on their knowledge, workers are not at the same level of ease on tasks. It is important to observe all the profiles to measure the differences between them and determine the impact of AR on workshop with different skills repartitions.

Quantitative Performance Measurements. Through total time gain $T_{G\text{-}T}$ and $R_{U/A}$, we observe a global positive impact of AR (fig. 7. and fig. 8.). The mean results on all users highlight a total time gain of 9% which is a high value in aeronautical processes.

Taking all profiles independently we also observe a gain on mean total time for subtasks for each profile. Thus, when the repartition of profiles evolves in the future, AR will continue to bring total time gain on the global workshop. Even with a small impact on expert users (2% $T_{G\text{-}T}$), there is no risk of losing time with AR.

Comparing beginner and confirmed profiles, the impact of AR is greater on confirmed profiles (respectively 7% and 11%) which was not expected. The hypothesis was that AR will benefit less to a user with more knowledge. This may be explained by the fact beginner are not familiar with the process and take more time to link instructions together whereas confirmed know instructions they are looking for and AR gives them an easier and intuitive access to these information. Upcoming beginners will benefit directly from AR and with practice they will become confirmed which will amplify the added value of AR. This could also promote versatility in workshops where workers could move between positions on many tasks.

The distinction between understanding and action phases highlights important elements. We observe a positive impact of AR on understanding time, 30% gain considering all profiles. This confirms that AR has an added value on task by facilitating the task of processing instructions for aeronautical maintenance. However, a negative impact of AR is seen on action time, 16% loss on total time considering all profiles. The hypothesis is that it is due to the use of a new unfamiliar device and a change of comportment where user tends to verify more information thanks to the closeness with instructions and action in AR. It could bring more benefits by anticipation on detection of errors before the end of the task thus reducing the needs of rework.

The added value of AR is also visible through the Understanding/Action ratio that is increased by 14 points considering all users. It means that on global workers spend less time processing instructions before each action with the assistance of augmented reality. The comparison of evolution of the ratio for the different profiles of users highlight that the impact is greater for confirmed than for beginners.

Qualitative User Related Measurements. SUS and TLX scores show positive returns on the use of AR (fig.9.). Considering all profiles, the usability of the AR support has been marked 7.5 points better than the current support. Transposing these results into percentile rank (PR) it means that the support for the task went from a good tool to an excellent tool (67.8 to 84.6 SUS PR). The impact of AR is also visible on cognitive load has it decreased between current support and AR support (34.2 to 26.8 TLX score). It means that less cognitive load is required to deal with the AR instructions, allowing the users to focus on tasks execution with higher efficiency.

To further analyze the impact of AR on users, we associated the measures on confirmed and experts. Those two profiles have knowledge and practice on the current support of instructions whereas beginners discover the current support of instructions. The usability scores given by confirmed and expert is similar on both conditions and corresponds to the best evaluation (over 85 points). For them the AR support is as usable as current support they learnt to use for years.

Beginners on the other hand gave lower score on both conditions. This is explained by the fact that they are less at ease with the maintenance process. However, we observe an important difference between current support that is just OK (SUS score of 57.5) and the AR application (SUS score of 72.5) which qualify the application usability as good when converting to percentile rank. For beginners there is a visible improvement on usability of the support with augmented reality.

This tendency is the same on TLX measures. Confirmed and experts tend to have lower cognitive load on the tasks than beginners and the impact of AR on the TLX score is more important for beginner. On the user point of view augmented reality seems to have a bigger impact for new workers discovering the maintenance process than for workers already familiar with the current support.

6 Conclusion and Perspectives

In this work we studied the impact of augmented reality technology on the execution aeronautical maintenance tasks from process and users' point of view.

To conduct this evaluation, we first had to answer the question of selecting right criteria. The goal was to identify criteria not only suitable to evaluate the use of AR to assist a user doing a task but also adapted to constraints of aeronautical maintenance environment. Among criteria and method to evaluate AR we selected those applicable in workshop without disturbing the execution of maintenance tasks. On the process we selected performance indicators related to tasks such as non-quality detected or overall task time but detailed to highlight specificities of tasks which should be impacted differently by AR. For users feedback we selected two questionnaires, one evaluating cognitive load felt by workers during the execution of the task and the other focusing on usability of the AR solution comparing to the current support.

To apply the selected criteria into an experiment, we had to deploy an AR solution covering the same information as current supports for the same task. We established a process based on knowledge on available AR technology (hardware and software), the situations where it is useful and the knowledge of experimented workers on aeronautical maintenance. This helped focusing on specific maintenance tasks and deploying an AR application with appropriated content according to instructions currently used by workers, as well as their feedback on the tasks.

With criteria and the AR application we were able to conduct experiment in real conditions, with AR in the maintenance process line without disturbing the workers. The experiment has been conducted on workshop population, gathering three different profiles, and on a long period of time. The results of the experiment bring answers on the added value of AR for the tasks. Considering all profiles of users, the experiment highlights a gain brought by augmented reality through each criterion. Comparing to current support, the AR application received a better usability score from the users which is echoes to a better ratio between understanding phases and action phases for each subtask. AR facilitates the access and the understanding of instructions thus the users can focus more easily on the action phase. A parallel can be made with the cognitive load score of the task that decrease and on the total time observed that is better with AR. This confirms the hypothesis that the RA adds something to the process point of view and this is visible from the user point of view.

The comparison between profiles highlights different impact of AR according to users. At first beginner will benefit more from the AR solution on their point of view (good usability & cognitive load reduction) which could ease their training. Then the observation that stands out is that the benefit will increase on productivity side with AR tools in the hand of confirmed workers familiar with maintenance process. And it does not interfere with the work of experts that are as at ease with AR as with the current supports. However, the AR application negative impact on the action phase should be reduced by better adaptation of AR content to needs.

Some limitations need to be addressed. One is the constraints on evaluation criteria which needed to be usable in workshop without disturbing the workers and easily deployable to conduct experiments on the maintenance lines. This is closer to real impact of AR on tasks, but it limits the observations. Another is the non-systematic selection of use cases based on the opportunities currently offered by the technology and on the users' experience of the workshop to link AR functionalities and complex tasks that need to be assisted by this. There were also constraints on experiments feasibility and frequency as aeronautics work on long cycles. The low availability of equipment on the selected task along with the small number of subjects in the workshop reduced the number of observations even on a long period of time.

The added value of AR has been identified on selected tasks and this result needs to be expanded to other tasks not yet addressed by this work. Improvement of the interactions between the user and the AR (data capture and processing) could impact other stages of the maintenance process. There is a need to be able to generalize the choice of activities. Further work will consist on using these elements to extend results to AR evaluation on the whole aeronautical maintenance process. A classification of maintenance activities linked to workers specific needs will lead to prediction of the value that augmented reality would bring to each activity.

References

1. IATA Press Release No.: 62, https://www.iata.org/pressroom/pr/Pages/2018-10-24-02.aspx, last accessed 2019/06/14.
2. EASA Regulations, https://www.easa.europa.eu/regulations, last accessed 2019/06/14.
3. MRO Survey 2017: When Growth Outpaces Capacity, https://www.oliverwyman.com/our-expertise/insights/2017/apr/mro-survey-2017.html, last accessed 2019/06/14.
4. Azuma, R.T.: A survey of Augmented Reality. Presence, vol. 6, pp. 355-385 (1997).
5. Van Krevelen, D. W. F., Poelman, R.: A Survey of Augmented Reality Technologies, Applications and Limitations. International Journal of Virtual Reality 9(2), pp. 1-20 (2010).
6. Keynote Prof. Dr. Henry Fuchs - The XR Future - the coming utopia or a gamer's plaything, https://www.youtube.com/watch?v=lyX_0gcixf4, last accessed 2019/01/05.
7. Caudell, T., Mizell, D.: Augmented reality: an application of heads-up display technology to manual manufacturing processes. Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences (1992).
8. Mura, M. D., Dini, G., Failli, F.: An integrated environment based on augmented reality and sensing device for manual assembly workstations. Procedia CIRP 41, pp. 340–345 (2016).
9. Kin, S.J., Dey, A. K.: Simulated Augmented Reality Windshield Display as a Cognitive Mapping Aid for Elder Driver Navigation. CHI 2009 – Navigation, Boston (2009).
10. PWC: How will people create content for augmented reality?, https://usblogs.pwc.com/emerging-technology/how-will-people-create-content-for-augmented-reality/, last accessed 2019/06/14.
11. Fite-Georgel, P.: Is there a reality in Industrial Augmented Reality?. In: 10th IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2011 (2011).
12. Quandt, M., Knoke, B., Gorldt, C., Freitag, M., Thoben, K.D.: General Requirements for Industrial Augmented Reality Applications. Procedia CIRP 72, pp. 1130–1135 (2018).
13. Palmarini, R., Erkoyuncu, J. A., Roy, R.: An Innovative Process to Select Augmented Reality (AR) Technology for Maintenance. Procedia CIRP 59, pp. 23–28 (2017).
14. Palmarini, R., Erkoyuncu, J. A., Roy, R., Torabmostaedi, H.: A systematic review of augmented reality applications in maintenance. Robotics and Computer-Integrated Manufacturing, vol. 49, pp. 215-228 (2018).
15. Baumeister, J., Ssin, S. Y., ElSayed, N. A. M., Dorrian, J., Webb, D. P., Walsh, J. A., Simon, T. M., Irlitti, A., Smith, R. T., Kohler, M., Thomas, B. H.: Cognitive Cost of Using Augmented Reality Displays. IEEE Transactions on Visualization and Computer Graphics, vol. 23, pp. 2378-2388 (2017).
16. Renner, P., Pfeiffer, T.: Augmented Reality Assistance in the Central Field-of-View Outperforms Peripheral Displays for Order Picking: Results from a Virtual Reality Simulation

Study. Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017 (2017).
17. Werrlich, S., Eichstetter, E., Nitsche, K., Notni, G.: An Overview of Evaluations Using Augmented Reality for Assembly Training Tasks. International Journal of Computer and Information Engineering, vol. 11, pp. 1068-1074 (2017).
18. Webel, S., Bockholt, U., Engelke, T., Gavish, N., Olbrich, M., Preusche, C.: An augmented reality training platform for assembly and maintenance skills. Robotics and Autonomous Systems, vol. 61, pp. 398-403 (2013).
19. Rios, H., González, E., Rodriguez, C., Siller, H. R., Contero, M.: A mobile solution to enhance training and execution of troubleshooting techniques of the engine air bleed system on boeing 737. Procedia Computer Science (2013).
20. Syberfeldt, A., Danielsson, O., Holm, M., Wang, L.: Visual Assembling Guidance Using Augmented Reality. Procedia Manufacturing (2015).
21. Fiorentino, M., Uva, A. E., Gattullo, M., Debernardis, S., Monno, G.: Augmented reality on large screen for interactive maintenance instructions. Computers in Industry, vol. 65, pp. 270-278 (2014).
22. Blaga, A. D., Frutos-Pascual, M., Al-Kalbani, M., Williams, I.: Usability Analysis of an Off-the-Shelf Hand Posture Estimation Sensor for Freehand Physical Interaction in Egocentric Mixed Reality. Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017 (2017).
23. Paas, F. G. W. C.: Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. Journal of Educational Psychology, vol. 84, pp. 429-434 (1992).
24. Brunken, R., Plass, J. L., Leutner, D.: Direct measurement of cognitive load in multimedia learning. Educational psychologist, 38(1), pp. 53–61 (2003).
25. Hornbaek, K.: Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer Studies 64, pp. 79–102 (2006).
26. Rubio, S., Diaz, E., Martin, J., Puente, J. M.: Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. Applied psychology: an international review, 53(1), pp. 61–86 (2004).
27. Hart, S. G., Staveland, L. E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Amsterdam (1988).
28. Brooke, J.: SUS - A quick and dirty usability scale. (1996).
29. Bangor, A., Kortum, P., Miller, J.: Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. Journal of Usability Studies, vol. 4, pp. 114-123 (2009).
30. Hart, S. G.: NASA-task load index (NASA-TLX); 20 years later. (2006).
31. Brooke, J.: SUS: A Retrospective. Journal of Usability Studies, vol. 8, pp. 29-40 (2013).
32. Lugan, G.: L'élaboration de la maintenance aéronautique à travers la méthodologie MSG-3. Master Thesis (2011).
33. Martinetti, A., Rajabalinejad, M., Dongen, L. V.: Shaping the Future Maintenance Operations: Reflections on the Adoptions of Augmented Reality Through Problems and Opportunities. Procedia CIRP (2017).
34. Jetter, J., Eimecke, J., Rese, A.: Augmented reality tools for industrial applications: What are potential key performance indicators and who benefits?. Computers in Human Behavior, vol. 87, pp. 18-33 (2018).
35. ISO.: Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV) - Partie 11: Lignes directrices relatives à l'utilisabilité. ISO 9241-11 (1998).
36. Diota | Solutions 4.0 for Industry, http://diota.com/index.php/en/, last accessed 2019/06/14.