Please have a Look here: Successful Guidance of Air Traffic Controller's Attention

Oliver Ohneiser; Hejar Gürlük; Malte-Levin Jauer
German Aerospace Center (DLR),
Institute of Flight Guidance,
Lilienthalplatz 7, 38108 Braunschweig, Germany
Oliver.Ohneiser@DLR.de

Ádám Szöllősi; Dóra Balló
HungaroControl Zrt.,
Igló u. 33-35,
1185 Budapest, Hungary

***Abstract*—Keeping the operator's attention on the right spot of the situation data display is one of the key factors to successfully guide air traffic. However, this becomes particularly difficult with complex and dense traffic situations displayed on larger screens. This paper describes our developed prototypic attention guidance (AG) system for air traffic controllers (ATCO). This system uses eye-tracking as an input for the ATCO's current attention. Different attention guidance was implemented for specific air traffic control (ATC) events such as handover and conflict alerts. For those events, different visual cues are presented step-wise within various levels of escalation in case the ATCO did not pay attention to ATC events. The AG system was tested in human-in-the-loop validation trials with five ATCOs. The simulated Hungarian Flight-Centric airspace was chosen as a test-case. The validation trials revealed promising results for the Solution controller working position (CWP), which was equipped with AG functionality. ATCOs reported less workload and improved situation awareness with the Solution CWP than without AG support. Increased acceptance and confidence with the Solution system were also reported. ATCOs felt strongly supported by our robust and smoothly interacting attention guidance system encouraging further development of our prototype towards operational use.**

Keywords*—Air Traffic Controller; Attention Guidance; Controller Working Position; Human Machine Interface; Eye-Tracking; Visual Cues; Flight-Centric Air Traffic Control

I. Introduction

Air traffic controllers are scanning their situation data display continuously in order to keep up situation awareness and handle all relevant events safely and timely with respect to urgency and importance. Thus, ATCOs have to determine and prioritize ATC events and plan their controller tasks accordingly. As ATC events are mostly connected to aircraft on the radar screen, there are relevant screen sections that need to be looked at and paid attention to.

This task becomes more difficult with more complex and dense traffic, particularly when displayed on large format monitors. Current human machine interfaces (HMI) often provide single types of alerts to warn the ATCO. These mostly visual alerts can be over-looked or may be too intrusive with regards to the aforementioned challenges.

The feedback cycle providing the ATCO's reaction to the HMI only starts with the conflict resolving. In conclusion, there is a demand for smooth and non-intrusive guiding of ATCO's attention to the relevant HMI spots.

In our prototypic AG system, this is done via an eye-tracking system determining which area the ATCO currently looks at. The corresponding data feeds a trigger algorithm for relevant ATC events. It is then used to potentially escalate visual cues on the HMI in different steps if the ATCO does not pay attention to the relevant spots.

This paper reviews related work on the topic of attention and its guidance in chapter II. Chapter III outlines the concept for an Attention Guidance (AG) prototype at a Controller Working Position (CWP) and its implementation within the given Flight-Centric Air Traffic Control (FC-ATC) use case. The study setup to reveal benefits and drawbacks of the AG prototype is outlined in Chapter IV. Chapter V presents the results of the AG Human-in-the-Loop study. Those results are discussed in Chapter VI. Chapter VII summarizes, concludes, and gives an outlook on future work.

II. Related Work

A. *Operator Attention*

According to Broadbent's filter metaphor [1], the attention of a human operator represents a filter to the environment. This filter reduces irrelevant input to focus on relevant – potentially multiple – input streams [2]. This selective filter avoids overload of the human's brain [1]. Many aspects such as task demands, operator's situation understanding, different channels and senses as well as the related perceptual limits influence the effectivity of the filter [2]. Operator's attention is connected to distinguishing, remembering, reacting in a certain amount of time, perceiving, and conceiving [3].

The operator's attention and thus filtering is a prerequisite for proper situation awareness seen as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" following Endsley [4]. As the situation awareness itself is limited, meaning that better understanding of some elements decreases understanding of others, this factor is especially important in complex environments [5].

Air traffic control can be seen as such a complex environment where visual attention plays a key role for proper situation awareness. There are six tasks that ATCOs need to fulfill [2] related to visual attention. (1) Scanning and orienting in a goal-directed and undirected way. (2) Supervisory control to assure that aircraft characteristics stay within a required range.

(3) ATCOs need to notice unexpected events like emergencies when monitoring the radar screen. (4) ATCOs are searching for specific aircraft to issue clearances. (5) ATCOs read all relevant information such as radar labels shown on the situation data display. (6) After a clearance has been given to an aircraft there must be a follow-up check if the current flight behavior is conform to the desired behavior. The above knowledge about visual attention and situation awareness needs to be taken into account when attempting to guide operator's attention.

B. Guiding Attention

In the SESAR2020 exploratory research project Mitigating Negative Impacts of Monitoring high levels of Automation (MINIMA), some eye-tracking based guidance of attention has already been done [6]. A semi-transparent circle was used to highlight aircraft on the radar screen that have not been looked at for a certain amount of time. However, this highlighting of inattention did not take the relevance and necessity to pay attention to this specific aircraft into account. A study on attention in a supervisory task with large displays revealed that movements were better for notifications than (animated) color [7]. Faster perception of notified aircraft was achieved by background opacity, concentric circles, and pulsating boxes. However, the intrusiveness differed on the use of color [7].

Peripheral cues should be presented next to the target stimulus if ATCO's view is anywhere else in order to enable proper cue recognition. Those cues need to be salient enough to be recognized [2] especially if they are far away from the current gaze focus. However, cues should not mask other information. They should also not be displayed too often or too long as this might distract ATCOs. Reliable cues can be achieved via integration of multiple sensor data and a model-based AG system. If cues are reliable, exogenous cues are preferred over endogenous cues as they are processed faster by humans [8] [9]. These aspects were considered for the AG concept and their implementation is described in the following section.

III. Attention Guidance Concept and Implementation

A. Basic Assumptions for the AG Concept

The display spot, where (a) the ATCO's gaze targets at respectively (b) the mouse-cursor on the radar display (when moved), are assumed as the ATCO's current area of visual attention. Thus, this area is determined (a) via an infrared contact-free eye-tracker mounted at the bottom of a 40 inch screen respectively (b) via mouse data positions. An assistance system of the FC-ATC environment calculates the next relevant ATC events – so called triggers. If there are multiple unsolved ATC events, the AG system prioritizes them with respect to urgency and importance. In most cases, the ATC event with highest priority requires the ATCO's attention. If the current and preceding area of attention of the ATCO does not match the expected area of attention, the AG system checks if the minimum look-ahead time of some seconds or minutes depending on the ATC event type has passed. If so, the AG system raises the escalation level [10] and raises it even further if the ATCO still does not pay attention.

B. Escalation Levels and De-Escalation for ATC Events

The AG system comprises four different escalation levels (0 to 3) if there is a relevant ATC event. The current AG implementation includes short-term conflict alerts (STCA), medium-term conflict alerts (MTCA; with and without right of way) as well as handover events. Basic level 0 identifies the state without additional visual cues. As short-term conflicts are very important, they do not have an escalation level 0, but are directly escalated to level 1. In general, there is a rectangle frame around the ATC event affected aircraft radar label in escalation level 1. The frame is accompanied by a round semi-transparent flashlight effect in level 2. This flash-light effect obtains a colored "glowing" circular frame in escalation level 3. The higher the level, the more salient the visual cue (compare figure 1).

Figure 1. Screenshots of the three escalation levels' visual cues for a handover event.

If the controller notices all visual cues related to an ATC event, the escalation level will immediately be set to zero and the cue disappears. In case of a handover event it is just this aircraft that needs to be noticed. In case of an aircraft conflict all involved aircraft need to be noticed. There are two options for an ATCO to notice a visual cue. The primary option is to look at the aircraft radar label or head symbol. The eye-tracker will then detect the ATCO's gaze and initiate the de-escalation. The eye-tracker has a certain detection range of a few centimeters, so that multiple aircraft can be noticed at a time. This bigger detection range allows for more body movement of the controller that otherwise would result in less robust recognition. The secondary option is to move the mouse cursor over the radar label to indicate the noticing of the event. For more details on the AG concept, please refer to [11].

C. ATC Event Resolution

Even if the ATC event was noticed, the AG system will remain active in the background until the event was actually resolved. The resolution is carried out by the ATCO usually by issuing clearances into the aircraft label or by coordinating with other ATCOs. If the event resolution takes longer than a certain ATC event type dependent threshold time and the controller again did not notice the involved aircraft for a certain amount of time, another escalation through the levels will follow. The non-resolution can have two reasons. Either, the ATCO forgot or did not actively notice the event at all. However, from a safety perspective, this does not matter a lot due to the re-escalation.

D. Use Case: Flight-Centric Air Traffic Control

SESAR2020 (Single European Sky Air Traffic Management Research Programme) foresaw the integration and validation of an Attention Guidance prototype as part of PJ.16-04-03 within the Flight-Centric ATC environment of PJ.10-01b.

The Flight-Centric ATC concept focuses on ATCO's responsibility for a number of aircraft instead of geographic airspace sectors. This concept has been researched at DLR in cooperation with the German ANSP DFS Deutsche Flugsicherung GmbH since 2008 [12]. The general feasibility has been proven for the upper airspace area [13]. Furthermore, assignment strategies have been analyzed [14] so that the incoming traffic is balanced between CWPs in current FC-ATC software. ATCO support tools for conflict detection and planning are another essential factor to enable the FC-ATC concept [15].

It has to be noted that the Flight-Centric ATC part just served as a use case to integrate and present the Attention Guidance functionality. Thus, results reported in this paper focus on the comparison of a CWP equipped with an AG system compared to a CWP without an AG system and do not consider benefits or drawbacks from the FC-ATC environment. All ATC events, related times, and the radar appearance of basic escalation level 0 are part of FC-ATC.

IV. Human-in-the-Loop Study with Attention Guidance Prototype

A. *Validation Setup with Software, Hardware, and Simulation Exercise Staff*

The Human-in-the-Loop Study to evaluate the DLR Attention Guidance prototype took place on January 17th, 2019 at HungaroControl premises in Budapest, Hungary.

The final software setup consisted of the FC-ATC software (provided by the DLR pilot assistance department) on the one hand and the AG software (provided by the DLR controller assistance department) on the other hand. The FC-ATC part included the traffic simulation itself, the aircraft assignment, automatic conflict solving options, communication infrastructure, and finally the situation data display.

The FC-ATC software was connected to the AG system, which is the logical core part of the presented concept. From a software engineering perspective, the FC-ATC software can be seen as the "model" and the "view", whereas the AG system is the "controller" governing the current appearance of the HMI.

The visual cues of the escalation levels are shown on the ATCO display as the handover example in figure 1 points out. Other ATC events are displayed in a very similar but differently colored way. Details can be found in [11].

Five Flight-Centric ATC CWPs with height movable chairs for the eye-tracking calibration were available as demonstrated in figure 2. The ATCOs communicated with five simulation pilots. The traffic – respectively its coordination – was automated. The simulation crew was available in the background in case of upcoming questions.

The average age of the five participating HungaroControl ATCOs was 33.2 years (standard deviation, SD: 10.1 years), with an average job experience as a controller of 7.4 years (SD: 10 years). All of them were en-route area control center (ACC) controllers for Hungary (LHCC flight information region, FIR, Budapest).

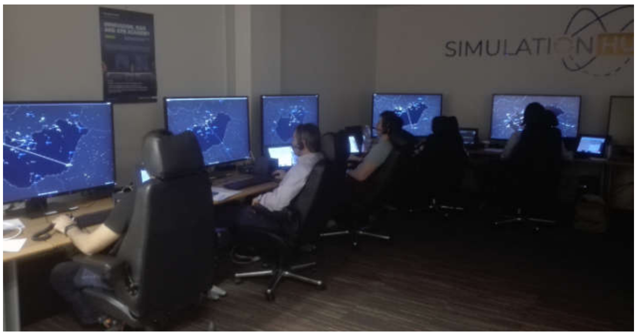

Figure 2. SESAR2020 PJ.16-04-03 Attention Guidance Human-in-the-Loop validation exercise EXE-16.04-TRL4-TVALP-310 with DLR AG prototype and five ATCOs at HungaroControl Simulation Hub in Budapest, Hungary.

B. *Simulation Run Conditions*

For the simulation runs two different conditions were designed in order to compare effects on human performance. One condition depicted a Baseline Flight Centric CWP without AG functionality, whereas the Solution CWP represented the other experimental condition encompassing the Flight Centric CWP with AG functionality. Both conditions incorporated a high density Flight-Centric ATC traffic scenario in order to raise the probability of provoking higher numbers of aircraft conflicts or missed aircraft handovers. For a traffic setup time and familiarization, both simulation runs started with a 7.5 minutes initialization phase running in double simulation speed. Afterwards, the runs were directly continued in real time (announced by the simulation staff) until one hour simulation run time was completed.

As all five ATCOs changed their CWP for the second run, they had to handle different assigned aircraft of the complete Hungarian airspace and thus different air traffic situations in both runs to avoid a scenario learning effect. A restart of the display was necessary for ATCOs 3 and 4 in the Solution condition. This restart lasted 23 respectively 27 seconds, but should not have significantly affected any of the results compared to the simulation duration of 3,000 seconds.

C. *Organizational Preparation of Simulation Runs*

The ATCOs already received a pre-briefing document some weeks in advance comprising AG prototype functionalities and the schedule. The trials started with a briefing explaining the purpose of the study and an overview of the AG system functionalities including escalation levels and visual cues. Furthermore, participant agreement sheets needed to be signed and a demographics questionnaire to be filled. Afterwards, a short training run familiarized the ATCOs with the eye-tracking calibration process and the further visual cues in case of ATC event escalation. As the ATCOs were already trained with the Flight-Centric ATC CWP the days before, the training concentrated on the appearance of visual cues and how to let them disappear.

D. *Eye-Tracking Calibration Process*

The eye-tracking calibration was done for each controller before participating in the Solution simulation run. First, a comfortable and technically appropriate seating position regarding chair height and distance to the screen needed to be found.

After that, the ATCO's gaze had to be fixed on four specific spots on the screen for some seconds in order to let the eye-tracker's infrared sensors learn the pupils' positions.

Finally, aircraft (radar labels) that are recognized as being noticed by the ATCO on the radar display were highlighted in yellow. This feature for eye-tracking recognition demonstration was switched off for the validation runs and just served for calibration transparency reasons.

E. AG Data Acquisition Activities: Simulation Runs, Questionnaires, and Debriefing

During the following first simulation runs two ATCOs worked with the Solution CWP and three ATCOs with the Baseline CWP. ATCOs filled out the common PJ.16-04 human performance questionnaire after the run and before the subsequent break. Then, eye-tracking calibration was executed for the other ATCOs; this time, three of them worked with the Solution and two ATCOs with the Baseline CWP. This order was chosen to balance the run sequence. The second round of questionnaires again included the human performance questionnaire. In addition, the tailor-made Attention Guidance parts were filled out by the ATCOs.

The questionnaire items, statements, scales, and other details are explained in results Section V.A. The final group debriefing followed a semi-structured interview method. Log files of eye-tracking and mouse data of ATCOs as well as ATC event data were recorded during the Solution run. The complete validation exercise lasted slightly more than 3.5 hours in the afternoon.

V. Results of Attention Guidance Validation Exercise

This section presents the results with respect to human performance questionnaires, AG log files, and more general statements. Values for questionnaire ratings and times are reported as arithmetic averages with standard deviations SD:

$$SD = \sqrt{\frac{\sum(x-\bar{x})^2}{(n-1)}}$$

where n is the number of values, $\bar{x}$ is the arithmetic average, and x is the rating value. The above equation is used, as values are a random sample of the population.

A. Subjective Controller Ratings on Human Performance and the AG concept

Five questionnaire parts mainly with Likert scales [16] have been answered by the ATCOs two times (after each simulation run). The questionnaires were extracted from the EUROCONTROL Human Performance repository [17]. In addition, a more general AG questionnaire was answered just once after finishing both simulation runs to compare the Solution system with the Baseline system.

1) Situation Awareness

First, ATCOs filled in the China Lakes Scale [18] and rated their situation awareness on a decision tree from low (1) to high (10). Average situation awareness score was 8.25 (SD: 0.96) for the Baseline system, but 9.0 (SD: 0.82) for the Solution system.

Secondly, ATCOs rated five items about - traffic understanding, aircraft messages, coordination, and identification during the run from bad (1) to good (7). The awareness score was 5.84 (SD: 1.06) for the Baseline system and 5.72 (SD: 1.19) for the Solution system.

2) Workload

First, ATCOs filled in the Bedford Workload Scale [19] from easy (1) to hard (10) for peak and average workload experienced during the run. The peak workload was 5.4 (SD: 1.34) for the Baseline system and 4.8 (SD: 1.34) for the Solution system, respectively. The average workload was 3.4 (SD: 1.14) for the Baseline system and 3.4 (SD: 0.55) for the Solution system, respectively, as shown in figure 3.

Secondly, ATCOs rated an ATC Workload Scale with nine items from easy (1) until hard (10) for multitasking, planning, decision making, team awareness, information processing, attention direction, problem solving, memory management, and maintaining awareness as experienced during the run. The average workload score was 5.16 (SD: 0.71) for the Baseline system and 3.96 (SD: 0.5) for the Solution system, respectively (see figure 3).

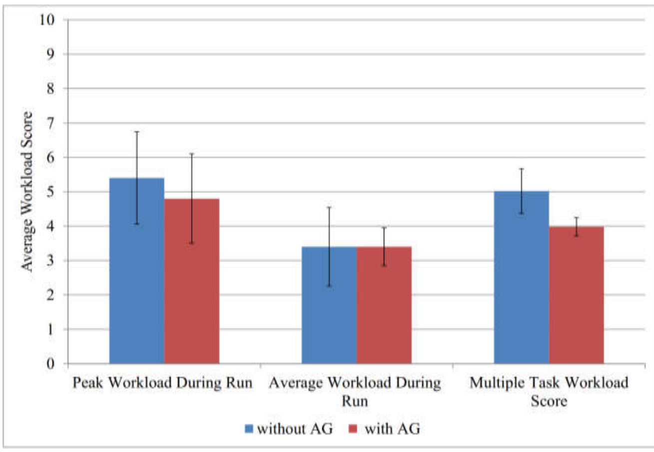

Figure 3. ATCO workload ratings for Baseline without AG (blue) and Solution with AG (red) with standard deviations as black vertical lines.

3) Usability and Controlling Tasks

ATCOs had to rate 19 different statements regarding usability and controlling tasks about air traffic control functionalities on a scale from "strongly disagree" (1) to "strongly agree" (5) or even "not applicable". The average for the Baseline system was 3.86 (SD: 0.69), for the Solution system 4.0 (SD: 0.53).

The five statements about separation determination, coordination of in- and outbound traffic, regular scanning cycle, conflict detection, and separation assurance had a 0.4 points better value for Solution than Baseline in average. The statement that ATCOs were able to rapidly prioritize alerts was even rated 1.25 points better in Solution than Baseline.

4) User Acceptance

The adapted Controller Acceptance Rating Scale (CARS [20]) delivered an average value between bad (1) and good (10).

ISSN 0770-1268

For the five ATCOs it was at 2.8 (SD: 1.64) for the Baseline system, but 7.4 (SD: 3.24) for the Solution system (figure 4).

5) User Confidence

Four statements on confidence with a scale from "completely disagree" (1) to "completely agree" (10) had to be rated. Statements asked if the tool supports work, if ATCOs feel adequately trained, if information is suitable for their tasks, and if an overall confidence is given.

User confidence was at an average of 4.5 (SD: 3.3) for the Baseline system and 5.95 (SD: 2.95) for the Solution system, respectively, as presented in figure 4. One ATCO noted that his score for the Solution system was low as he could not split the FC-ATC and AG part for his rating.

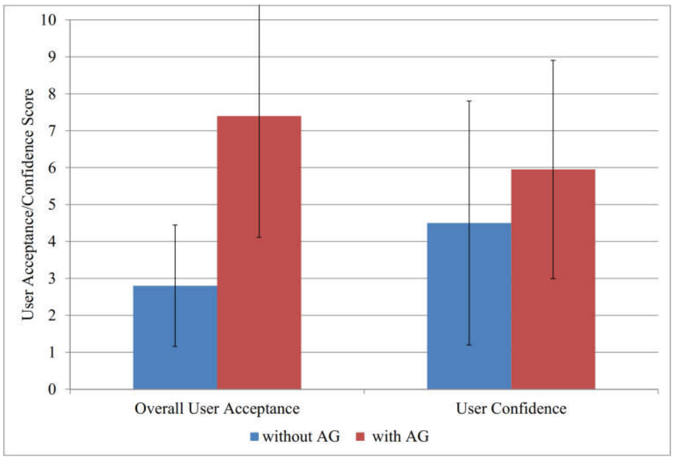

Figure 4. ATCO acceptance and confidence for Baseline without AG (blue) and Solution with AG (red) with standard deviations as black vertical lines.

6) Tailor-made Attention Guidance Rating

The ATCOs rated six statements about the AG concept and the hardware setup on a scale from "I do not agree at all" (1) to "I totally agree" (5). The statements were about if the display of AG escalation levels for STCA/MTCA/handover are understood easily, if the AG logic is transparent, if the eye-tracking works reliably, and if the radar screen is sufficiently large. ATCOs gave 30 single rating scores with an average of 4.73 (SD: 0.25).

7) Summarized and Normed AG Rating

Parts 1 to 6 comprise nine different questionnaire sections with 47 rating items. For better readability and comparability, all scales have been normed (scale from 1 to 10) and some scales have been inverted so that a higher score is always better than a lower score in figure 5.

Hence, the Usability and Controlling Tasks part, as well as the tailor-made AG part have been normalized by multiplying with 2 to enlarge the rating scale from 1 to 5 up to 1 to 10. The second situation awareness part with a scale from 1 to 7 has been multiplied with a factor of 10/7. The workload ratings have been inverted (10 to 1; 9 to 2; …; 1 to 10) as lower workload scores indicate better results.

This inverted rating was combined with the other workload parts into a single so called "Relax Score". Furthermore, both situation awareness parts as well as the usability and controlling part were combined to a single score as well as another single score for the user acceptance and confidence parts as figure 5 shows.

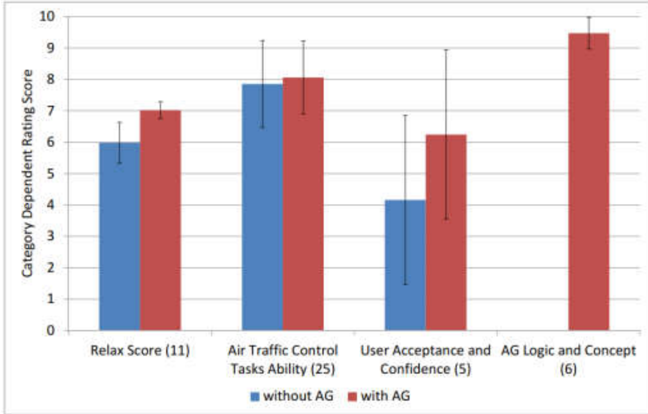

Figure 5. Combined questionnaire parts with number of single rating items in brackets for both conditions: Baseline without AG (blue) and Solution with AG (red) with standard deviations as black vertical lines.

B. Log File Results on ATC Events and Aircraft Noticing

1) Number of Escalated ATC Events

114 ATC events in total have been escalated in the five simulation runs of 50 analyzed net minutes duration per each of the five ATCOs. This means 22.8 escalated ATC events per ATCO (SD: 2.9). 108 of the escalated ATC events were noticed by the ATCOs detected via eye-tracker, just three ATCOs noticed two escalated handovers, each detected via mouse-over functionality.

95 of those ATC events were handovers (83.3%), 15 medium-term conflict alerts (13.2%; thereof 3 with right of way and 12 without right of way), and four short-term conflict alerts (3.5%). Just one ATCO had no STCA during the high-density Flight-Centric ATC scenario; however this ATCO had the most MTCAs. 75 ATC events were escalated until level 1 (65.8%), 30 events until level 2 (26.3%), and only 9 events until the highest escalation level 3 (7.9%) as shown in figure 6.

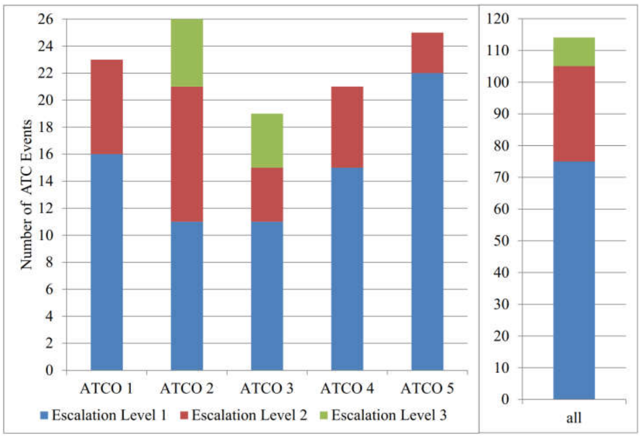

Figure 6. Number of ATC events escalated to level 1/2/3 (blue/red/green) until detection per ATCO and over all ATCOs.

Regarding just handover escalations, 62.1% (59) were noticed in level 1, 28.4% (27) in level 2, and 9.5% (9) only in level 3. 16 conflict alerts (both medium- and short-term) have been noticed in level 1, the other three in level 2. Just two ATCOs experienced ATC events that were escalated up to the highest level 3 (see ATCO 2 and 3 in figure 6).

All escalated ATC events have been resolved during the simulation time except of 10 handovers and 6 medium-term conflict alerts that only appeared during the last minutes. Hence, there would have been time to solve these conflicts if the simulation would have continued.

2) Visual Escalation Cue Noticing Times

In average it took an ATCO 8.3 seconds to notice an escalated ATC event (SD: 10.2s). For all types of conflicts – so without handovers – the average time was 3.1 seconds (SD: 3.0s). From the highest escalation level, a general ATC event was escalated to (1, 2 or 3), it took the ATCO only 3.6 seconds (SD: 5.6s) to notice the visual cue. The noticing time for conflict alerts from the highest escalation level for all ATCOs was only 2.3 seconds (SD: 1.8s). As one ATCO needed more than twice as much time to notice escalated ATC events compared to the average of the others, this heavily influences the above reported average times (see figure 7).

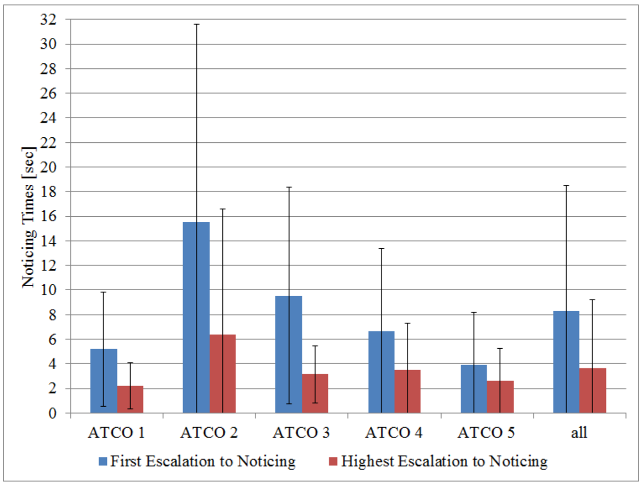

Figure 7. Average time needed to notice an escalated ATC event per ATCO and over all ATCOs (blue left columns). Average time needed to notice an ATC event after its highest escalation level per ATCO and over all ATCOs (red right columns); with positive/negative standard deviation (black lines).

C. AG Concept Results and General Comments

1) Tailor-made Attention Guidance Questionnaire

ATCOs were also asked to complete eleven open questions comparing the Baseline and the Solution once after finishing both simulation runs. Four ATCOs preferred the Solution system over the Baseline system (one ATCO gave no answer on this).

This paper does not analyze the FC-ATC part that served as a Baseline system. Hence, only the following summarized view on the ATCO's questionnaire statements 1 to 5 is given for completeness reasons: Information filtering and situation awareness were poor for the time with all potential subsequent effects on safety.

Some meaningful and mostly positive answers about the Solution system (questionnaire statements 6 to 10) are paraphrased in the following. When using the Solution system, ATCOs liked best that incoming and outgoing traffic as well as conflicts were highlighted if not scanned or being forgotten. In addition, the general AG idea was liked. The system was rated as simple to use and really helped to find blind spots.

The majority of ATCOs did not mention any disadvantage for the Solution system. Moreover, they only saw advantages as the AG system complements existing features. However, one ATCO reported that the system sometimes requires taking the focus away from an area or problem which needs to be focused on. Thus, it might be better suited only for training new ATCOs. Others especially wanted to work with the visual guidance tool (Solution system) in dense traffic situations or when they are more tired in order to draw the attention back. The only issue with the AG system might exist if somebody does not want to get eye-tracked. Some ATCOs wished that the AG system would already have been implemented in their CWP as it really provided them assistance.

2) Debriefing Comments

In the common verbal debriefing session all five ATCOs preferred the Solution to the Baseline. Some feedback sentences are paraphrased in the following. ATCOs found the visual cues to be non-intrusive. They also thought that eye-tracking worked really robust after calibration and thus the interaction was fine. The AG functionality really supported ATCOs to have a look at HMI spots that they would not have looked at this timely. The AG system was assessed as already ready to be used in operational life. Furthermore, most ATCOs would like to have AG in their CWP, in the near future, regardless of Flight Centric ATC environment.

ATCOs wanted to have individual settings for aircraft highlighting in case of handovers, i.e. a possibility to switch on and off respectively a different visual cue style. Some preferred to only highlight aircraft when they left their sector without being handed over to the next position or entered their sector without being handed over to them respectively aircraft being assumed or not assumed. Others also liked the highlighting before aircraft enter or exit their sector. The possibility to switch off the complete AG functionality existed, but was not explained to the ATCOs as exactly this functionality needed to be evaluated.

The mouse-over functionality was only used in seldom cases as the eye-tracking was very accurate. ATCOs also found that the debugging functionalities of the eye-tracking could be very useful for training purposes. The instructor could improve scan patterns of trainees better when visualizing the gazes. This could lead to a whole new methodology in ATCO training.

ATCOs felt well supported by the AG system and reported unanimously that AG provided additional benefit to their HMI working routines compared to working at the FC-ATC CWP without AG assistance. ATCOs reported they were fine with respect to the timing behavior of cues except of MTCAs that should only be escalated with eight minutes look-ahead time to avoid false alerts. ATCOs rarely observed intrusiveness if they intensively "worked" at a different spot.

Then it might be slightly disturbing if their peripheral attention was tracked to look somewhere else. Many ATCOs said there was nothing annoying at all with the AG functionality.

VI. Discussion of Attention Guidance Study Results

A. Human Performance Results Discussion

The human performance results showed ATCOs' preference for the Solution system with attention guidance functionality over the Baseline system. However, no statistical significance was tested due to the limited number of test subjects (n=5). Therefore, results showed tendencies in ATCO ratings, but have differences in mean values bigger than the corresponding standard deviations in some of the categories indicating a strong tendency.

The Solution system had a better score for overall situation awareness, but no great difference between compared systems in the second situation awareness questionnaire. The Solution system showed slightly better usability and support for controlling tasks than the Baseline system in average. This small difference in the rating might also be affected by the FC-ATC part which in general was not perceived to provide good situational awareness.

Also ATCOs did not perceive significant differences in the average workload during the simulation runs. This seems to be logic as the scenarios were the same and ATCOs just had to handle different aircraft but roughly the same number of aircraft in their runs. However, the peak workload was much lower for the Solution system. Furthermore, the workload scale with nine questionnaire items showed a much lower score for the Solution system. Hence, workload seems to be lower with AG especially in dense and complex traffic operations. ATCOs may handle ATC events earlier without getting into stress situations, indicating a confirmation of the envisioned benefits of the AG system.

User acceptance and user confidence were rated much better for the Solution system, too. This can be a prerequisite when introducing this new system later on.

B. Log File Results Discussion

Almost two thirds of escalated ATC events have already been noticed in escalation level 1. Roughly another quarter was noticed in level 2. This shows the effectivity of visual cues to guide ATCO's attention, while still remaining non-intrusive as the controller statements show.

The evaluation of log files additionally confirms the suitability of the developed AG prototype for the improvement of situation awareness and timely handling of ATC events: All escalation levels did occur during the simulations and therefore helped the operator noticing ATC events he/she was previously unaware of.

Furthermore, the decreasing number of occurrences of higher escalation levels (see figure 6, right) and the reduced noticing time of events with higher salience (e.g. conflicts) undermines the suitability of the HMI design in terms of showing the most important information first.

As in general, the reaction times are within a one digit number of seconds and ATCOs reported being pointed to spots that they would not have realized so fast otherwise, demonstrates the efficiency potential of the AG system. However, some reaction times were also slower than ten seconds. Probably, ATCOs saw something in their peripheral view, however, finished their task at their current area of interest, but were aware of that they need to shift their attention afterwards. This fact can especially be assumed for less time-critical handover events. ATCOs also relied on the eye-tracking system for aircraft noticing rather than using the mouse-over functionality.

To sum up the time aspect, we took effective measures for the implementation of our AG prototype to ensure that ATC events with high priority are noticed by the controller in a timely manner to support safety.

C. Tailor-made Attention Guidance Questionnaire and Debriefing Comments Discussion

The results of the tailor-made AG ratings showed that the ATCOs had a clear understanding of the AG logic and that the system is viewed as being robust in this early stage of development. ATCOs' feedback on the AG functionality was almost completely positive. The only negative statements comprised the guidance of attention in situations where the ATCO likes to keep attention anywhere else and theoretical concerns on data privacy due to eye-tracking.

ATCOs did not feel patronized, but really felt supported individually by AG. The Solution system helped to put attention on important display areas that otherwise would have been looked at only later. ATCOs also experienced a well-working assistance as it was robust and non-intrusive. They even wanted to have it for their conventional CWP and formulated ideas for enhancements and further ATC events that could be included with respect to attention guidance. The ATCO statements stand for themselves and support the core AG validation result even more than just the subjective questionnaire ratings.

VII. Summary and Outlook

A. Summary of Attention Guidance Validation Trials

Our attention guidance prototype – based on prioritized ATC events and eye-tracking data of ATCOs – was successfully implemented and tested in the FC-ATC environment. The validation exercise of the AG prototype revealed very motivating results. ATCOs felt supported by the visual cues of escalated ATC events for handovers, medium- and short-term conflicts. As they were reminded of conflicts in case of non-resolution, AG may also serve as an additional safety net.

The event noticing times also depending on the escalation level were in the range of a radar update (few seconds). Even if not significant in all categories, relax score, ATC tasks ability (also comprising situation awareness), as well as user acceptance and confidence were higher using the Solution CWP with AG functionalities. ATCOs also rated the AG logic and concept very high.

The debriefing feedback was really encouraging. It hardly happens that ATCOs wish to have a new functionality – in their daily life CWP – to be noted that these were just first trials of a prototype. Furthermore, the used low budget eye-tracker and the few adaptations that would be necessary to integrate an Attention Guidance system into a CWP promise to deliver reasonable support for air traffic controllers.

B. *Outlook on Future Work*

Further ideas of which visual cues to escalate additionally or which aspects to be customizable have been developed. ATCOs uttered the idea to adjust some of the visual cues by their own. The personalization of AG settings as outlined in section V.C.2 is easily doable from a technical point of view, but needs further analysis with respect to a common CWP functionality basis.

Some ATCOs wanted to reduce the escalation time in advance of a medium-term conflict as some of them could be false alerts. ATCOs even wished to use the visual cues also for other ATC events. Escalation should be done in case of wrong Mode-S settings in the cockpit, for route adherence monitoring (RAM), approaching restricted areas, cleared flight level conformance alarm (CLAM) events, and if the current flight level is different to the exit flight level with the aircraft being close to the exit point.

The trigger algorithm could be adapted so that whenever the operator notices an important event, a “working time” is defined. Only after this working time has passed, the trigger logic will continue generating visual cues for the high-priority events. This could avoid unintended guiding of ATCO’s attention. One ATCO zoomed far out of his airspace looking on the whole European map to check whether all aircraft in the very center are detected as being noticed by him at the same time. ATCOs did not retry this during the simulation runs. However, it is of course one aspect to adapt the eye-tracking noticing area depending on the radar display zoom step.

The majority of ATCOs also wanted to have a visual cue if an aircraft was not looked at for a longer period of time. This should also be valid for all radar targets on the display that are not correlated with a flight plan. However, airspace regions should rather not be highlighted in general if there was a lack of attention. Although, this indicator could be tested in a further study after implementing the respective functionality.

The AG concept will be adapted to other laboratories and training CWP environments like (multiple remote) tower in the future. It will also be enhanced with additional auditory cues. Furthermore, a coupling of the ATCO authentication with the pre-defined user profile settings could ease the use. Despite all those ideas for further refinement (respectively enhancement) of the AG concept and its implementation, ATCOs found the AG system already ready for the next step towards operationalization.

Acknowledgment

The PJ.16-04 CWP HMI project also comprising the Attention Guidance activity (PJ.16-04-03) has received funding from the SESAR Joint Undertaking under the European Union’s grant agreement No. 734141.

References

[1] D. Broadbent, “Perception and Communication,” London: Pergamon Press, 1958.

[2] C.D. Wickens, ”Engineering Psychology and Human Performance,” 4th Edition, Psychology Press, 2015.

[3] W. James, “The principles of psychology,“ 2016, website cited May 14, 2019, URL: http://ebooks.adelaide.edu.au/j/james/william/principles/chapter11.html.

[4] M.R. Endsley, “Situation Awareness Global Assessment Technique (SAGAT),” in “Proceedings of the IEEE 1988 National Aerospace and Electronics Conference: NAECON,” 1988, p. 792.

[5] M.R. Endsley, “Designing for Situation Awareness in Complex Systems,” in “Proceedings of the 2nd International Workshop on Symbiosis of Humans, Artifacts and Environment,” Kyoto, Japan, 2001.

[6] O. Ohneiser, F. De Crescenzio, G. Di Flumeri, J. Kraemer, B. Berberian, S. Bagassi, N. Sciaraffa, P. Aricò, G. Borghini, and F. Babiloni, “Experimental Simulation Set-Up for Validating Out-Of-The-Loop Mitigation when Monitoring High Levels of Automation in Air Traffic Control,” in “International Journal of Aerospace and Mechanical Engineering,” 12 (4), 2018, pp. 307-318.

[7] J.-P. Imbert, H.M. Hodgetts, R. Parise, F. Vachon, F. Dehais, and S. Tremblay, “Attentional costs and failures in air traffic control notifications,” in “Ergonomics,” Taylor & Francis, 57 (12), 2014, pp. 1817-1832.

[8] J. Theeuwes and R. Godjin, “Parallel Allocation of Attention Prior to the Execution of Saccade Sequences,” in “Journal of Experimental Psychology, Human Perception and Performance,” 29, 2003, pp. 882-896.

[9] J. Jonides, “Voluntary versus automatic control over the mind’s eye’s movement,” in “Attention and Performance,” 9, 1987, pp. 187-203.

[10] H. Springborn, “Design and Assessment of Methods of Attention Guidance for the Sector-Less Air Traffic Management Controller Working Position,” Master’s Thesis, FH Joanneum, Graz, Austria, 2017.

[11] O. Ohneiser, M.-L. Jauer, H. Gürlük, and H. Springborn, “Attention Guidance Prototype for a Sectorless Air Traffic Management Controller Working Position,” German Aerospace Congress (DLRK), Deutsche Gesellschaft für Luft- und Raumfahrt - Lilienthal-Oberth e.V., 4.-6. Sep 2018, Friedrichshafen, Germany.

[12] B. Korn, C. Edinger, S. Tittel, D. Kügler, T. Pütz, O. Hassa, and B. Mohrhard, ”Sectorless ATM — A Concept to Increase En-Route Efficiency,” in “Proceedings of the 28th Digital Avionics Systems Conference (DASC) 2009”, Orlando, FL, USA, 2009.

[13] M. Biella, B. Birkmeier, B. Korn, C. Edinger, S. Tittel, and D. Kügler, ”Operational Feasibility of Sectorless ATM,” in “Proceedings of the International Conference of the European Aerospace Societies (CEAS) 2011”, Venice, Italy, 2011.

[14] A.R. Schmitt, C. Edinger, and B. Korn “Balancing Controller Workload Within a Sectorless ATM Concept,” CEAS Aeronautical Journal, 2, 2011, pp. 35-41.

[15] B. Birkmeier, “Feasibility Analysis of Sectorless and Partially Automated Air Traffic Management,” PhD dissertation as DLR Forschungsbericht 2015-12, ISSN 1434-8454, 2015.

[16] R. Likert, “A Technique for the Measurement of Attitudes,” in “Archives of Psychology 22,” No. 140, 1932, pp. 5-55.

[17] EUROCONTROL, SESAR JU electronic “Human Performance repository,” 2013, website cited May 14, 2019, URL: https://ext.eurocontrol.int/ehp/?q=Home.

[18] M.R. Endsley, “Measurement of Situation Awareness in Dynamic Systems,” Human Factors, 37, 1995, pp 65-84.

[19] A.H. Roscoe and G.A. Ellis, “A subjective rating scale for assessing pilot workload in flight: A decade of practical use,” RAE-TR-90019, Royal Aerospace Establishment Farnborough, United Kingdom, 1990.

[20] K.K. Lee, K. Kerns, R. Bone, and M. Nickelson, “The Development and Validation of the Controller Acceptance Rating Scale (CARS): Results of Empirical Research,” Proceedings of the 4th USA/Europe Air Traffic Management R&D Seminar, Santa Fe, NM, USA, 2001.