Eye-data metrics to characterize tower controllers’ visual attention in a Multiple Remote Tower Exercise

Christoph Moehlenbrink & Anne Papenfuss
German Aerospace Center (DLR), Institute of Flight Guidance

Remote Tower Centers, where Air Traffic Services are provided independently of local airport control towers, promise novel opportunities like tailored staffing concepts or more efficient work organization. In a joint approach the German Aerospace Center (DLR) and the German Air Navigation Service Provider Deutsche Flugsicherung GmbH (DFS) evaluated the concept for Multiple Remote Tower Control (RTC) in a realtime simulation exercise, addressing human factor relevant aspects such as workload, situation awareness, visual attention and acceptance.

Within this paper we focus on eye-data analysis. 16 air traffic controllers from DFS participated in the high-fidelity human-in-the-loop simulation study for multiple remote tower. Different eye-data metrics are discussed, that can characterize the visual attention of tower controllers in this new work environment.

Keywords Remote Tower Operations, Validation, Eye-tracking, ATCO performance, monitoring behaviour,

I. INTRODUCTION

Controllers are highly qualified people, recruited and trained to provide air traffic control for large airports with densely packed air traffic. However, also small airports exist, where air traffic control service must be provided for only a few aircraft movements a day. During periods with low traffic volume those towers are often manned by a single controller. While a lot of research addresses the problem of controllers’ high workload situations, in which their visual attention is overstressed, it is seldom discussed that controllers at single manned towers are rather faced with the problem to stay focused for only a few aircraft movements a day. In these towers, the visual attention capacity is not the major problem, but rather low job satisfaction due to low traffic numbers.

The motivation for multiple remote tower control is mainly motivated by economic reasons. To provide staff and infrastructure for those small airports is very cost intensive (Oehme et al., 2012). Though, the concept for multiple remote tower also bares the potential to organize air traffic control in a different way, in order to eliminate single manned towers and to increase job satisfaction. But how many small airports can be assigned to a single controller? Although controllers are skilled in multitasking, their monitoring capacity is limited by the visual system and therefore by a limited visual attention capacity.

A. Multiple Remote Tower Control

The first validation exercise researching into Multiple Remote Tower control was realized within the DLR-project RAiCe (Remote Air Traffic Control Center) (Moehlenbrink et al. 2012). Three different work organizations were considered when two airports are controlled from a remote tower center. Within that study three conditions of work organization were realized. First, a team of two controllers (Tower Controller and Coordinator) provided air traffic control to both airports. In a second condition, one single controller provided air traffic control to both airports. In addition, a baseline condition was realized, in which each controller was responsible for one of the two airports. A major result of this study was that workload increased compared to the baseline condition (different traffic load), however it was also found that augmented vision aspects can reduce controllers workload, under high traffic conditions for the (1) team and (2) single controller condition, but not for the (3) baseline condition. The results were further discussed with respect to redundancy and the four-eye principle within the team condition.

Within the German aeronautical research program (LuFo IV) the project VICTOR (Virtual Control Tower Research Studies) led by Deutsche Flugsicherung researched into the concept that three airports are remotely controlled from a remote tower center (Oehme et al. 2012). The concept included a new master controller working position, who was responsible for ATC at several airports, during time periods with low traffic volume.

Within SESAR, research on Multiple Remote Tower is addressed within the operational focus area 6.3.1 Remote Tower. There are ongoing validation activities by LFV & SAAB as well as by DLR & DFS.

B. Visual attention of tower controllers

From the literature it can be found that the work of tower controllers is more bottom-up (data driven) in comparison to enroute or approach controllers which can be described as rather top-down (Dittmann et al. 2000) Therefore, tower controllers’ visual monitoring of the airport and its surrounding control zone can be seen as a critical subtask for performing their control task successfully.

Tower controllers cannot always anticipate future situations and ensure that there will be no conflict for the next 5 minutes. Therefore, they must monitor whether the processes on the airport are in line with their clearances. Those clearances often bare the risk for conflicts, e. g. if they allow a pilot to taxi to holding point RWY 28, there is the potential that the AC does not hold short of RWY 28. In these situations tower controllers often must react within seconds in order to ensure safe operations. It can be therefore said that tower controllers visual monitoring is guided by direct or indirect triggers that indicate the potential for a critical situation. A list of such triggers has been provided by Moehlenbrink et al. (2010).

An overview of eye movements measures has been provided by Wickens & McCauley (2008). Beside the
definition of fixation, saccade and dwell times, they introduce the concept of area of interest (AOI) which stands for an area within which all individual fixations are considered by the researcher as functionally equivalent. Tower controllers’ visual attention is often analyzed using pre-defined AOIs such as out-of-the-tower view, radar, flight strips etc. Aggregated data for number of fixation, or total dwell time for a whole simulation run are often reported (Pinska 2006, Moehlenbrink et al. 2010). Those aggregated measures are used to discuss e.g. tower controllers head down times (time looking at instruments). This metric is seen critical, because, if controllers look too much on their instruments, they might miss critical events that are only detectable by the out-of-the-tower view.

Although, these aggregated measures are helpful to report how much attention tower controllers spend on different AOIs, these aggregated measures are not very helpful to establish a relationship between controllers visual attention and their monitoring task of certain events. Therefore, it has been suggested to use event-related data analysis metrics, to describe controllers’ visual attention (Moehlenbrink et al. 2010) on task relevant AOIs. Within that paper it was analyzed how often tower controllers monitored TakeOffs and Landings for arriving and departing air traffic.

In another research paper on tower control it has been addressed that no empirical research is needed to conclude that monitoring four displays for conflicts is less critical than monitoring eight displays for safety critical events (Moehlenbrink & Papenfuss 2011, compare Moray & Inagaki 2000). To evaluate whether one controller can control two airports it is therefore of interest, whether controllers can develop efficient attention allocation strategies that allow for monitoring of two airports in a safe manner. The mean first passage time, another eye-data metric suggested by Wickens could be used to quantify, how long an ATCO didn’t look at the other airport, when he was responsible for two airports at a time (Wickens & McCarley, 2008).

C. Functional representation of the work environment

In order to analyze, if there are major differences in controllers overt visual attention, if they control the same amount of traffic at one airport or at two airports, it is of interest to structure the work environment into functional units: Instead of distinguishing the two airports, within this study we will only consider, if the controller was looking at flight strips, radar, outside view, radio instrument or weather data. We will not be interested whether these instruments belong to one airport or the other.

D. Research question

Within this paper, it will be discussed in which way eye-data metrics can provide valuable insight how controllers visual attention is affected, when s/he is not only responsible for one airport, but for two airports. For the eye-data analysis we are mostly interested in identifying characteristic differences for different working conditions:

A) When a single controller works two airports in the condition multiple baseline

B) When a single controller works two airports in the condition multiple advanced (automatic tracking & callsigns)

C) When a single controller works one airport (baseline)

We will look at aggregated metrics like fixation duration, dwelltimes, but we will also introduce moving average time windows (MAW), which have been suggested by Anders (2001) to analyze visual attention over time. We suggest to use the MAW to plot controllers maximal visual attention for a certain time interval, to identify characteristic behavior of tower controllers. Further, an event-related analysis is completed for landings and take offs. It will be discussed in what way, the different metrics can be used to characterize, how the new work environment can affect controllers visual attention.

II. METHOD

A high-fidelity human-in-the-loop simulation was completed in the TowerLab of the Institute of Flight Guidance, DLR in Braunschweig.

A. Sample

Sixteen ATCOs from the Deutsche Flugsicherung participated in the study, two of them were female. All of them had a valid ATCO license and they were between XX to YY years old.

B. Simulation environment

The study was conducted using a high-fidelity simulator setup, simulating a 180° degree tower view of Braunschweig (EDVE) and Erfurt (EDDE) airport. The different experimental conditions were realized by two different set-ups, a single remote tower work environment and a multiple remote tower work environment. The latter included an approach radar and a pan-tilt zoom camera for each airport, electronic flight strips (left side EDDE, right side EDVE) and a coupled radio frequency. Weather data were integrated into the displays of the outside view. For the single remote tower set-up (baseline condition), was comparable but with instruments only for one airport. Within the multiple advanced conditions, it was simulated that ATCOs could use automatic tracking of aircraft and labels were overlaid in the outside view.

C. Traffic Scenarios

From an operation perspective, the requirement for the traffic scenario was 6-a/c-at-a-time for the one hour simulation scenarios. For the multiple remote work environment, two different scenarios were realized, varying the traffic distribution. In one scenario traffic was equally distributed among both airports (3:3) in the other scenario, there were 5 movements in EDVE and only one or two movements at EDDE. The scenarios for the different conditions were comparable in traffic load, but different callsigns were realized. Each controller completed three scenarios (baseline, multiple, multiple advanced). Half of the sample was assigned to the 5:1 traffic distribution, the other half to the 3:3 traffic distribution.

D. Experimental Setup

A schematic representation of the multiple remote tower work environment is depicted in Figure 1 and the corresponding 3-D environment for the eye data recording in Figure 2. The top row represents the panorama view of Braunschweig airport (5 screens, planes 2-6). The second row depicts Erfurt airport (5 screens, planes 7-11). The approach radar for Erfurt and Braunschweig are included in plane 12. The bottom row depicts the radio touch input interface for Braunschweig and Erfurt (plane 14). On the right, the user interface for the zoom cameras was placed (plane 15). Plane 13 includes the electronic flight strip bay and maps.
Eye data were recorded with the head-mounted eye tracking system iView-X combined with an optical head tracker from ART (SMI 2007; A.R.-Tracking 2009). The DLR Software EyeTA (Eye-tracking Analyser) was used for eye-data processing. The tool allows for the analysis for defined areas of interest (AOI) within the simulation environment and for event-related eye data analysis. In Figure 2 the pre-defined planes for the eye-data recording are depicted which are directly related to the different information sources represented in Figure 1.

### E. Experimental Design

A 3*2 factor design with the within subject factor work environment (single, multiple, multiple advanced) and the between subject factor traffic distribution (5:1, 3:3, 6:0) was completed. The traffic scenarios varied in positioning and timing of arriving and departing aircraft. The order in which ATCOs completed the conditions was varied and ATCOs were matched by age to the conditions (5:1) and (3:3) to make the subsamples comparable. Prior to the experimental traffic scenarios, ATCOs completed a training procedure to become familiarized with the work environment, the provided technical equipment and the multiple remote tower work environment. In Table 1 it is depicted that 14 out of 16 eye-data sets were available for each condition (total n=42). In addition, for each eye data set a validity metric was calculated referring to the percentage of eye data unequal to zero or minus one.

<table>
<thead>
<tr>
<th>TABLE I. AVAILABLE EYE-DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of data-set</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>Validity (Min, Max)</td>
</tr>
<tr>
<td>Summary</td>
</tr>
</tbody>
</table>

### F. Procedure

For each day, two ATCOs were invited. ATCOs completed sessions of 1 hour on 2 succeeding days. The first day started with a briefing, followed by the training (about 30 min). ATCOs completed the scenarios in parallel, one controlling traffic for the single remote tower, the other for the multiple remote tower. Only for the multiple baseline condition the ATCOs completed their scenarios sequential in order to have the second ATCO in the role of an expert observer. The instantaneous self-assessment scale was used in two minute intervals for online workload assessment. After each scenario ATCOs had to fill out post-run questionnaires asking for situation awareness, workload and addressing safety aspects. A half-structured interview was then used to get ATCOs feedback for specific traffic situations that occurred during each scenario. On the second day, a final questionnaire was used followed by a final debriefing and a feedback session.

### G. Task

The ATCOs were asked to control one or rather two airports at a time while procedures, rules and regulations correspond to the existing regulations for tower control today.

### III. RESULTS

#### A. Dwelltimes and Fixations

**Dwelltimes.** In all three conditions ATCOs showed highest dwelltimes of visual attention on the flight strips, followed by the panorama (comparable to out-of-the-tower-view) followed by the radar. The mean values and standard deviations of the percentaged dwelltimes for the different AOIs are summarized in Table 02.

<table>
<thead>
<tr>
<th>TABLE II. PERCENTAGED DWELLTIMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area of Interest</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>flight strips</td>
</tr>
<tr>
<td>panorama</td>
</tr>
<tr>
<td>radar</td>
</tr>
<tr>
<td>radio</td>
</tr>
<tr>
<td>weather</td>
</tr>
<tr>
<td>other</td>
</tr>
</tbody>
</table>

A 3*6 factor repeated measurement ANOVA with the factors (condition and AOI) revealed a significant interaction effect COND*AOI ($F_{(10,156)}=5,29, p=.25$). Post-hoc tests reveal that within the single baseline condition ATCOs spent significant less attention on the flight strips compared to both multiple conditions (p<.01). For the radar, controllers spent...
sign, less attention on the radar in the multiple advanced condition compared to the single baseline condition. There is no difference of ATCOs visual attention on the panorama. The results for the percentaged dwelltimes are depicted in Figure 3.

![Figure 3. Percentaged dwelltimes for all three conditions](image)

Number of Fixations. For the 60 min lasting simulation sessions mean number of fixations are presented in Table 3. There are no significant differences between conditions. While the dwelltimes for flight strips showed significant differences, this is not represented by the number of fixations for flight strips. Also in contrast to the trend that the dwelltime for the strips increases. On the other hand, the number of fixations for the panorama is lowest for the single baseline condition. The mean values and standard deviations are depicted in Table 3.

<table>
<thead>
<tr>
<th>Area of interest</th>
<th>Single</th>
<th>Multiple Ba</th>
<th>Multiple Ad</th>
</tr>
</thead>
<tbody>
<tr>
<td>flight strips</td>
<td>1016 (325)</td>
<td>1055 (408)</td>
<td>1126 (408)</td>
</tr>
<tr>
<td>panorama</td>
<td>1397 (425)</td>
<td>1523 (354)</td>
<td>1558 (175)</td>
</tr>
<tr>
<td>radar</td>
<td>769 (180)</td>
<td>622 (123)</td>
<td>555 (115)</td>
</tr>
<tr>
<td>radio</td>
<td>107 (48)</td>
<td>114 (34)</td>
<td>119 (34)</td>
</tr>
<tr>
<td>weather</td>
<td>65 (21)</td>
<td>139 (49)</td>
<td>143 (22)</td>
</tr>
<tr>
<td>other</td>
<td>716 (254)</td>
<td>662 (181)</td>
<td>548 (175)</td>
</tr>
</tbody>
</table>

Mean Fixation Duration. The mean fixation durations are of main interest for the AOIs panorama, flight strips and radar. The data are represented in Table 4.

<table>
<thead>
<tr>
<th>Size of time interval</th>
<th>AOI</th>
<th>Single</th>
<th>Multiple Ba</th>
<th>Multiple Ad</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1s,10s]</td>
<td>flight strips</td>
<td>0.98 (0.25)</td>
<td>1.25 (0.38)</td>
<td>1.22 (0.38)</td>
</tr>
<tr>
<td></td>
<td>panorama</td>
<td>0.57 (0.23)</td>
<td>0.42 (0.07)</td>
<td>0.46 (0.07)</td>
</tr>
<tr>
<td></td>
<td>radar</td>
<td>0.88 (0.14)</td>
<td>0.87 (0.16)</td>
<td>0.85 (0.18)</td>
</tr>
</tbody>
</table>

Due to the fact that the differences in dwelltimes for flight strips could not be explained by the mean number of fixations it becomes apparent that for the mean fixation duration on the flight strips an interaction effect AOI*COND (F(4,72)=3.55; f²=.16) was found. For the single baseline mean fixation duration on the flight strips are shorter than for the multiple conditions (p<.01). For the AOIs panorama and radar, there are no sign. differences in fixation durations.

B. Moving average time windows

Visual attention over time. While dwelltimes, number and duration of fixations are often reported to describe ATCOs visual attention for complete simulation runs, these metrics are not very informative for ATCOs visual attention over time. Calculating moving average time windows (MAWs) is one way to respect for ATCOs visual attention within smaller time intervals. For the data presented in Table 5 and 6 we calculated moving average time windows for 10 (or rather 30 [s] intervals for each simulation run. In a second step, we averaged the MAWs over participants (n=39) per condition (3). For these three data sets (mean(SiBa), mean(MuBa), mean(MuAd)) we then identified for every time interval which AOI had the maximum value. This way it can be identified for how many time intervals ATCOs main visual attention was on the panorama, the radar, or the flight strips.

<table>
<thead>
<tr>
<th>Table V. Maximum Visual Attention within 10[s] Time Window</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAW</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>[10s,1s]</td>
</tr>
<tr>
<td>[10s,1s]</td>
</tr>
<tr>
<td>[10s,1s]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table VI. Maximum Visual Attention within 30[s] Time Window</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAW</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>[30s,1s]</td>
</tr>
<tr>
<td>[30s,1s]</td>
</tr>
<tr>
<td>[30s,1s]</td>
</tr>
</tbody>
</table>

Size of time interval. In Table 5 and 6 two different time intervals were chosen. It becomes clear that the size of the time interval must be defined with respect to the task of interest. The data show that within the single baseline condition, 909 time intervals could be identified in which the averaged major visual attention was on the AOI Panorama. For the multiple baseline condition only 330 and for the multiple advanced only 399 such intervals could be found. In addition, it can be seen from a comparison of Table 5 and 6 that the numbers with maximum visual attention on the AOI Panorama and radar decrease, when the length of the interval is increased from 10 to 30 [s]. On the other hand, the number of time intervals with maximum visual attention on the flight strips increases.

Maximum visual attention. In Figure 4 2*3 MAW plots corresponding with Table 5 and 6 are depicted. The first three graphs report the MAW for the [1s, 10s] interval and the second three graphs for the [1s, 30 s] interval. These graphs can be read as a cursor for controllers’ major visual attention over time. It does not mean that controllers’ looked exclusively on that AOI, but that for a defined time window of 10 or rather 30 s their major visual attention was on a specific AOI. Grey bars stand for major visual attention on the flight strips, green bars for the panorama and black bars for the radar. Looking at Table 5 and 6 or comparing the single with the multiple conditions clearly show that on average there are
AOI max  Moving average time window [1s: 10s]  panorama=green  radar=black  flight strips=grey

Multiple Baseline

Multiple Advanced

Single Baseline

------------------------------------------------------------------------------------------------------------------------------------------------------------

AOI max  Moving average time window [1s: 30s]  panorama=green  radar=black  flight strips=grey

Multiple Baseline

Multiple Advanced

Single Baseline
more time intervals with maximal attention on the panorama or radar within the single baseline condition. For the multiple condition there are more intervals with maximal attention on the flight strips. The height of the bars (0-1) represents the proportional visual attention on the AOI. It will be discussed in the next section, in what way this eye data measure can be of interest to characterize controllers monitoring behaviour and how it might help to better define when controllers might be overloaded based on their visual attention capacity.

IV. DISCUSSION

In this paper it has been addressed in what way moving average time windows might provide valuable analysis to evaluate whether tower controllers are still able to show a sufficient monitoring behaviour when controlling multiple airports. It was demonstrated how the moving average time window for the maximum AOI can nicely visualize which AOI received controllers’ maximum visual attention for a predefined time interval. In order to evaluate controllers’ monitoring behaviour it would be possible to quantify how much visual attention controllers’ have left to monitor the processes of the airport using the out-of-the-tower view over time.

A. Multiple Remote Tower: Dwelltimes and Fixations

In the literature, tower controllers head down times are seen critical, as during these time periods, controllers do not have the chance to identify unexpected events that might occur within their control zone (Hilburn 2004). It can be reported by the aggregated measure percentaged dwelltime that controllers spent less attention on the flight strips in the single baseline condition compared to the multiple conditions. This effect is not explained by a higher number of fixations on the flight strips, but by shorter fixations on the flight strips within the single baseline condition in comparison to the multiple baseline condition. This effect might be explained by the fact that in the multiple conditions controllers coordination task is higher, so they have to keep in mind, which airport the flight strip belongs to.

Although there is no dwelltime effect apparent for the AOI panorama, there is a radar effect showing that within the multiple advanced condition less attention is on the radar screen compared to the single baseline condition. This effect is in line with a former study by Papenfuss et al. (2010) which was also no effect was found for the out-of-the-tower view, but reduced visual attention on the radar. This reduced visual attention in the multiple advanced condition can be explained by the fact that within this condition controllers can get the callsign information directly from the panorama and therefore don’t need this information from the radar.

From a human factors perspective this result has to aspects. On the one hand, it is well liked by controllers to have the callsign integrated and close to the aircraft, but on the other hand no overlays in the panorama ensure that controllers look on the panorama to verify the system state of the aircraft (real time data), while they have to look on the radar to get the callsign information (background knowledge). If both information are integrated in the out-of-the-tower view, controllers might become blind for the aircraft state (e.g. check gear down) but just look for the aircraft label. This effect is although known as attentional tunneling, a phenomenon that describes that you don’t see information that is rather obvious, if you are focused due to specific information search (Wickens et al. 2009).

B. Functional Task Environment and moving average time windows

Areas of Interest. In this study the areas of interest were defined only as functional units, no matter if ATCOs had to control one airport or two airports. This way it was possible to directly compare ATCOs visual attention on the functional units, independent from their experimental condition.

In this paper not only dwelltimes aggregated over simulation runs were reported. Due to the fact that aggregated measures over simulation runs only report the proportion of visual attention, these measures are not very informative for controllers monitoring behaviour over time. It was demonstrated that moving average time windows provide a valuable visualization of controllers’ visual attention over time. The traffic scenarios within this simulation study were aiming at 6-a/c-at-a-time, independent of single or multiple tower control.

In order to evaluate novel work environments like multiple remote tower, moving average time windows can complement aggregated measures, such as dwelltimes and fixations. They allow for quantifying whether controllers are still able to perform a good monitoring task of the out-of-the-tower view. For evaluating the monitoring task it is of interest to understand controllers’ visual attention over time. We assume that the limited visual attention capacity of a controller is one limiting factor for how much traffic, or how many airports a controller can be responsible for. Therefore having long periods with time intervals with maximum visual attention on the flight strips, this might indicate that controllers are not ahead of traffic anymore and are engaged by other activities such as planning, coordination or strip marking.

As head-down times bear a great potential for missing critical events, it might be defined how much visual attention resources controllers have left to monitor the out-of-the-tower view. The MAWs in Figure 4, show that only in certain time periods, controllers maximum visual attention was on the panorama. On average there were more time intervals in the single baseline condition which fulfill the criteria maximum visual attention on the panorama then in the multiple conditions.

Setting the parameters. Choosing larger moving average time windows (e.g. the 30 s time interval) is rather liberal as it allows for longer time periods of head down times, not looking at the outside world at all.

In contrast, a smaller time interval (e.g. 10 seconds) is rather conservative as it would expect that for every 10 second time interval within a simulation controllers’ maximal visual attention is on the out-of-the-tower-view.

V. CONCLUSION

It will be of interest for future concepts in air traffic control, like the concept for multiple remote tower, how to ensure that controllers’ are able to perform their monitoring and control task sufficiently in order to guarantee for safety. In this paper, it was introduced in what way an analysis of the visual attention of controllers can provide valuable access to characterize controllers’ monitoring performance. It might be of interest for future research to define normative models that capture the expectations of visual scanning behaviour for tower
controllers. From our understanding the approach of moving average time windows might provide a robust approach for defining measures and parameters that are stable over inter-individual differences.

ACKNOWLEDGMENT

This research was supported by the SESAR Program co-financed by the EU and EUROCONTROL. The data were collected within the SESAR LOT 1 activity, EXE-06.08.04-VP-641. The opinions expressed in this paper reflect the authors’ view only. The SESAR Joint Undertaking is not liable for the use of any of the information included herein.

The authors thank Rainer Kaufhold, Paul Distelkamp and Hans Hedde from DFS for their operational perspective on the project. The authors are also indebted to the 16 ATCOs from DFS for participating in this simulation study. We would also like to thank those people whose effort was necessary to make this a successful validation exercise. We would like to mention Sebastian Schier, Felix Timmermann, Tim Rambau, Markus Schmidt, Ewa Strubelt, Anna Mavinga and the 14 pseudo pilots.

REFERENCES