

# Future Ability Requirements for Operators in Aviation regarding Monitoring

Catrin Hasse, Carmen Bruder, Dietrich Grasshoff and Hinnerk Eißfeldt

*Keywords: automation, monitoring, personnel selection, eye tracking, future ATM*

## Zusammenfassung

Mit zunehmender Automatisierung von Mensch-Maschine-Systemen wird die Überwachung automatischer Prozesse und Systeme zur zentralen Herausforderung für zukünftige Luftfahrtoperatoren. Die Abteilung Luft- und Raumfahrtpsychologie des Deutschen Zentrums für Luft- und Raumfahrt (D.L.R.) ist in der Personalauswahl von Piloten und Fluglotsen tätig. Im Rahmen des Forschungsprojektes Aviator 2030 stellt sich die Frage nach der Identifikation von Luftfahrtoperatoren, die sich durch ein adäquates Überwachungsverhalten auszeichnen. Als adäquat wird ein Überwachungsverhalten bezeichnet, wenn es dem Operateur ermöglicht, bei Bedarf die Kontrolle über das System selbst zu übernehmen.

Mittels eines dafür entwickelten Simulationsprogramms werden Zusammenhänge zwischen Parametern des Überwachungsverhaltens und der manuellen Systemsteuerung untersucht. Mittels Blickbewegungserfassung wird die Überwachungsleistung gemessen. Es werden erste Ergebnisse von 30 Bewerbern für Luftfahrtberufe vorgestellt. Die Einbeziehung von Blickdaten bei der Personalauswahl ist ein innovativer Ansatz, der neue Möglichkeiten für die zukünftige Eignungsdiagnostik eröffnet.

## Abstract

The research project Aviator 2030 at the German Aerospace Centre (DLR) focuses on finding an optimal fit between technical innovations and human operators in future aviation. Due to the increase in automation, accurate and efficient monitoring poses a great challenge to future operators in aviation. As the DLR's Department of Aviation and Space Psychology is responsible for the selection of pilots and air traffic controllers, our objective for the selection of future personnel is to identify operators monitoring appropriately (OMA). OMA are assumed to monitor in such a way as to enable them to resume control if automation fails.

We developed a simulation tool that helps identify future monitoring requirements for pilots and controllers. In our study, we combine eye movement parameters and manual performance data in order to reveal the predictive power of eye movements, standing for monitoring performance, for manual control. Preliminary results of an experiment with job candidates are presented. Personnel selection based on eye movements is innovative and will establish new approaches for assessing selection profiles in future air traffic management (ATM) scenarios.

## Aviator 2030

Technical innovations in aviation and improvements in ATM, aircraft systems and organizational structures are great challenges facing aviation in the 21st century. 'Aviator 2030' deals with changes that will concern pilots and air traffic controllers in the future, with the objective of adapting selection profiles to suit future ability requirements.

Workshops with experienced pilots and air traffic controllers were conducted in order to reveal their expectations regarding future tasks, roles and responsibilities. Summing up workshop results, monitoring and teamwork in a highly automated workplace pose challenges to future aircraft operators (Bruder, Jörn & Eißfeldt, 2008). As a consequence, research should focus on

the ability to monitor as one of the new core competencies necessary for success as an airline pilot or air traffic controller.

We developed a simulation tools that represent future workplaces in aviation. Experiments with subjects operating in these simulated future workplaces help identify appropriate monitoring behaviour in highly automated environments (Hasse, Bruder, Grasshoff & Eißfeldt, 2009). Use of this tool may, in the long term, make it possible to select candidates on the basis of pre-defined monitoring behaviour.

## **Monitoring Requirements**

As the automation of future workplaces in the field of aviation is assumed to increase monitoring ability, automated systems will be one of the new core competencies relevant for success as an airline pilot or air traffic controller. As the DLR's Department of Aviation and Space Psychology is responsible for the selection of pilots and air traffic controllers, we aim to adjust selection profiles to fit with future developments in aviation. In the long run, we aim to develop a test that helps to identify job applicants that meet the criteria for a good monitoring operator. In other words, we aim to identify operators monitoring appropriately (in short, OMA). We assume OMA will be able to manually control the system when necessary by relying on the system understanding that they have built up whilst monitoring the automated system.

The question remains: which parameters of monitoring behaviour help us to identify OMA that fit future man-machine systems in aviation? This study aims to find criteria for the identification of appropriate future operators based on their monitoring behaviour.

## **Devising a normative model of monitoring behaviour (Assumptions)**

Based on the empirical background (for a detailed explanation see Hasse et al., 2009) we devised a normative model, which describes the monitoring behaviour of OMA. According to models of adequate and efficient monitoring behaviour (Niessen & Eyferth, 2001; Whitfield & Jackson, 1982; Wickens, Helleberg, Goh, Xu & Horrey, 2001) as well as differences between experts and novices (Underwood, Chapman, Brocklehurst, Underwood & Crundall, 2003) it can be stated that OMA show target oriented attention allocation in general as well as during monitoring phases, i.e. orientation phase, anticipation phase, operation phase, and debriefing phase. Whereas the first assumption requires the operator to adapt attention allocation to the specific requirements of a given situation in general, the second assumption focuses the allocation of attention in phases. The operator is required to demonstrate flexibility in:

- anticipating system operations (during anticipation phases)
- detecting relevant system operations (during operating phases)
- controlling system performance afterwards (during debriefing phases)

We further assume that “good monitoring” is associated with an accurate manual system handling in case of automation failure, and therefore aim to connect monitoring behaviour with manual control behaviour. We assume that this link reflects differences in understanding of the underlying principles of the automatic system. The question we need to answer is: which monitoring criteria are most important with regard to identifying OMA that have the ability to manually control the system?

In order to answer this we aim to test the following hypotheses:

35. Keeping an overview of the overall automated system during an entire monitoring run is related to an accurate manual control of the system in case of necessity.
36. Anticipating and detecting important automated system operations in time as well as controlling them is related to an accurate manual control in case of necessity.

## Method

With the objective of identifying monitoring parameters that have an impact on manual system handling, we developed a simulation tool that allows for the assessment of monitoring performance.

## Simulation Tool

We developed a simulation tool called “Self Separation Airspace” (SSAS) that is a simplified and dynamic simulation of the basic requirements for future flight operators (for further description of this tool, see Hasse et al., 2009). In short, the operators’ task is to control the traffic flow between two airports (see Fig. 1). The operator either monitors an automatic process or controls the traffic manually, allowing us to collect data on their performance of both types of task separately. With the objective of varying complexity and dynamics of the automatic system, we developed four scenarios reflecting different degrees of difficulty (ibid.).

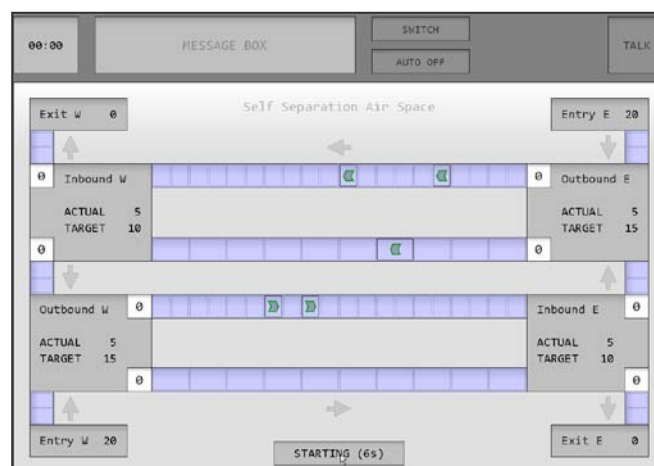


Fig.1: Simulation tool: Self Separation Air Space (SSAS)

## Measurements

Monitoring performance was measured by recording eye movements. We used relative fixation counts<sup>1</sup> and mean fixation durations based on predefined areas of interest. Fixation counts can be used as a measure of expectations and assumptions of the person (Rötting, 2001), where important objects are likely to be fixated upon more often than less important ones (Göbel, 1999). Fixation durations can be used as measure of information processing duration (Inhoff & Radach, 1998, S. 37, cit. in Rötting, 2001). Accordingly, processing difficulty as well as personal strain is reflected in the fixation duration (Rayner, 1978, 1982, Balota et al., 1985).

The normative model postulates that OMA keep an overview of system operations during an entire monitoring run; in this experiment during the automatic mode of one scenario (hypothesis 1). Moreover, OMA are expected to anticipate, detect and control automated operations in time; in this experiment reflected by different operations performed by automation within the automatic mode of one scenario (hypothesis 2). As for testing the first hypothesis, we defined scenario specific areas of interest (AOIs), that is areas on the simulation screen that we expect to be pre-conditional for keeping an overview of system behaviour. As for testing the second hypothesis, we determined AOIs that help to anticipate and detect system operations as well as to debrief them. As anticipation, detection and debriefing of system operations are only possible within definite time frames within a scenario, we cut every scenario into time frames.

<sup>1</sup> We used relative fixation counts (relative fixation number in relation to all fixations in the same time frame on defined AOIs) in order to account for general individual differences in fixation counts.

Every time frame stands for a monitoring phase and is characterised by AOIs being conditional for monitoring adequately in this phase, e.g. anticipating a system operation adequately. Hence, this model leads us to expect eye movements on areas of interest that are generally relevant for a specific scenario as well as for monitoring phases within specific time frames.

Regarding the performance of a test subject during the manual phase of each scenario, we used the mean deviation of actual values from target values of in- and outbounds. To avoid the possible impact of a general ability on manual performance when controlling a system, we corrected manual performance by deducting the baseline measurement when both parameters are significantly correlated.

## Experimental Procedure

Participants were first given a questionnaire measuring their trust in automation, and the instruction for the following experiment. Participants were informed they would work on four scenarios, each consisting of two phases; first an automation phase and then a manual phase. Referring to the automation phase of each scenario, participants were instructed to monitor the automation with the objective of understanding the rule-based dynamics of the given scenario. Referring to the hand control phase, participants were instructed to manually control the same system they had seen during the automated phase. Eye movement parameters were recorded by the Eyegaze Analysis System. Participants seated in front of a 19-inch LCD computer display with a distance of approximately 60 cm. After a calibration (15 s) phase, the persons were presented with the four scenarios, each of a five minute duration. There was a smooth transition between the automatic mode and manual mode within each scenario but pauses were placed between each scenario. The four scenarios were presented in a fixed order for every subject, beginning with the easiest (scenario 1), finishing with the most complex (scenario 4). After each scenario, subjects evaluated their difficulty as well as reported the strategies of automation they identified.

## Test Subjects

Our experiments were conducted with 90 candidates for DFS (Deutsche Flugsicherung GmbH) and DLH (Deutsche Lufthansa AG). This enables us to compare our experimental data about monitoring in future man-machine systems with abilities measured in personnel selection tests. Preliminary results reported in this article refer to a sample size of the first 30 participants (29 male, 1 female; candidates of both DFS and DLH are equally represented).

## Results

Preliminary analyses refer to two scenarios, scenario 2 and 3, with scenario 3 being the more complex one. The predictive validity of fixation counts on relevant areas of interest (reflecting monitoring performance) for manual performance data was estimated by correlation analyses.

Referring to hypothesis 1, there is a significant negative correlation between relative fixation counts on scenario specific areas of interest (AOIs) and subsequent manual performance ( $r = -.403$ ;  $p = .03$ )<sup>2</sup>. The greater the proportion of fixations that fall on the predetermined areas, the smaller the mean deviation between actual and target values and the better the performance when actively controlling the simulation. As for scenario 3, no significant correlation between relative fixation on scenario relevant AOIs and manual performance can be established. However, when absolute fixation counts are taken into account, the number of fixations on scenario specific AOIs correlated significantly with manual performance ( $r \approx -.4$ ;  $p \approx .03$ ).

---

<sup>2</sup> A negative relation is expected because the SSAS performance parameter is reversed. The smaller the deviation between actual and target, the better the performance in the manual phase.

Referring to hypothesis 2, results concerning the link between fixations during monitoring phases and manual performance seem to be dependent on the scenario. Regarding scenario 2, the relative frequency of fixations during all *operating phases*<sup>3</sup> on related AOIs shows a significant correlation of medium strength with manual performance ( $r = -.411$ ;  $p = .027$ ). The higher the proportion of fixations on AOIs, the better the performance in the simulation (see Fig. 2). Regarding scenario 3, the relative frequency of fixations on AOIs in all *debriefing phases* shows a significant negative correlation of medium strength with manual performance ( $r \approx -.437$ ;  $p \approx .018$ ).

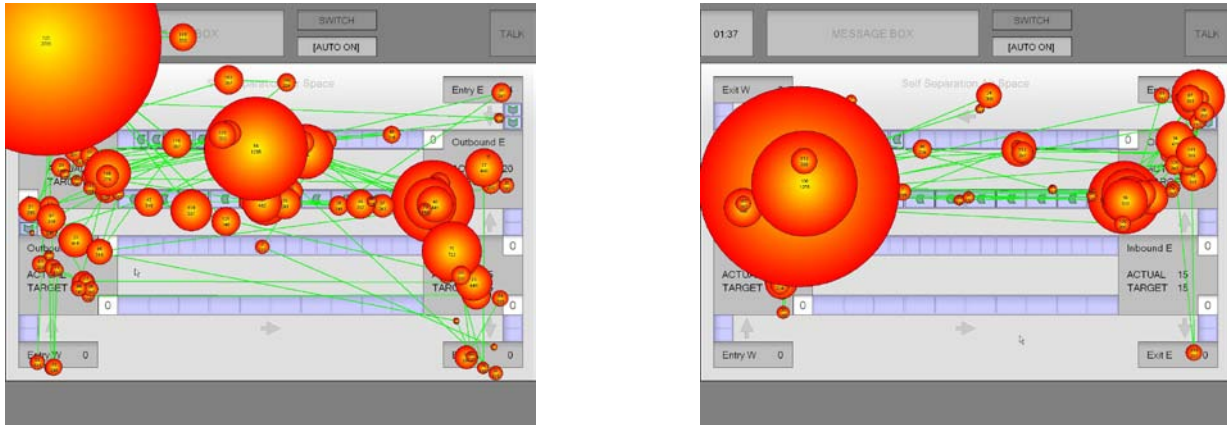


Fig. 2: Distributions of fixations as scan paths during operating phase of scenario 2. Amount of circles within an area indicate fixation number, circle size indicates fixation duration. A participant with poor manual performance and inadequate attention allocation (left) is compared to a participant with good manual performance and adequate attention allocation (right).

## Discussion (status quo and further steps)

Regarding the core competencies of future operators in aviation, we expect future operators to be able to monitor automated systems appropriately and to manually control systems where necessary on the basis of their monitoring behaviour. It was questioned, *which* parameters of monitoring behaviour were crucial with regard to the ability to manually control the system after monitoring and whether or not it would, in the long term, be possible, to select candidates on the basis of predefined monitoring behaviour.

Regarding fixations on areas of interest that are generally relevant, it could be verified for scenario 2 that suitable operators direct their attention to relevant areas during monitoring scenarios. Taking total fixations instead of relative fixation durations, this could also be verified for scenario 3. Further analyses will concentrate on scenarios 1 and 4.

Fixation number during anticipation, operating and debriefing phases within scenarios 2 and 3 seem to be dependent on scenario difficulty. Regarding scenario 2, less complex than scenario 3, fixations on related AOIs during all *operating phases* were related to manual performance. By contrast, fixations on AOIs in all *debriefing phases* were related to manual performance in scenario 3. Thus, when estimating the impact of eye movements during automation on manual system handling afterwards, the scenario's difficulty has to be taken into account. On this note, it can be speculated that the more complex the operations done by automation are, the more difficult it is for the human operator to anticipate operations or to detect them in time.

However, not all critical monitoring parameters during monitoring phases are significantly correlated with manual performance. Manual performance might also be caused by other abili-

<sup>3</sup> Note: temporal frame of reference for the eye movement parameters is here the summarisation of all operating phases during the automatic phase of scenario 2.

ties aside from monitoring behaviour. As our experiments were conducted with candidates, we will be able to compare our experimental data about monitoring in future human-machine systems with abilities measured in personnel selection tests. Beyond this, fundamental research on other future core abilities is planned, e.g. communication abilities within automated environments, shared monitoring between team-partners.

Ability testing with dynamic simulation based on eye movements is innovative and establishes new approaches assessing selection profiles. In this regard, SSAS is introduced as an appropriate base tool for investigating human performance in future ATM scenarios as well as the underlying ability requirements of human performance in future aviation.

## References

- Bruder, C., Jörn, L. & Eißfeldt, H.. (2008). *Aviator 2030 - When pilots and air traffic controller discuss their future*. In A. Droog & T. D'Oliveira (Hrsg.), *Proceedings of the 28th EAAP Conference, Vol. 2* (S. 354-358). Valencia: EAAP.
- Balota, D.A., Pollatsek, A. & Rayner, K. (1985). The interaction of Contextual Constraints and Parafoveal Information in Reading. *Cognitive Psychology*, 17, 364-390.
- Göbel, M. (1999). *Blickbewegungsregistrierung bei der Gestaltung von Fahrerarbeitsplätzen des ÖPNV*. In M. Rötting & K. Seifert (Hrsg.). *Blickbewegungen in der Mensch-Maschine-Systemtechnik* (S.118-132). Sinzheim: Pro Universitate Verlag.
- Hasse, C., Bruder, C., Grasshoff, D. & Eißfeldt, H. (2009). *Future Ability Requirements for Human Operators in Aviation*. In D. Harris (Hrsg.), *Engineering Psychology and Cognitive Ergonomics, 8<sup>th</sup> International Conference, EPCE 2009, Held as Part of HCI International 2009, San Diego, USA* (S. 537-546). Berlin, Heidelberg: Springer.
- Inhoff, A.W. & Radach, R. (1998). *Definition and Computation of Oculomotor Measures in the Study of Cognition*. In G. Underwood (Hrsg.), *Eye guidance in Reading and Scene Perception* (S. 29-53). Amsterdam: Elsevier.
- Niessen, C. & Eyferth, K. (2001). A model of the air traffic controllers' picture. *Safety Science*, 73, 187-202.
- Rayner, K. (1978). Eye Movements in Reading and Information Processing. *Psychological Bulletin*, 85 (3), 618-660.
- Rayner, K. (1982). *Visual Selection in Reading, Picture Perception and Visual Search – A Tutorial Review*. In H. Bouma & D.G. Bouwhuis (Hrsg.), *Attention and Performance X* (S. 67-96). Hillsdale: Lawrence Erlbaum.
- Rötting, M. (2001). *Parametersystematik der Augen- und Blickbewegungen für arbeitswissenschaftliche Untersuchungen*. Aachen: Shaker Verlag.
- Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J. & Crundall, D. (2003). Visual Attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics*, 46, 629-646.
- Whitfield, D. & Jackson, A. (1982). *The air traffic controller's picture as an example of mental models*. In Johannsen, G. & Rijnsdorp, J.E. (Hrsg.) *Proceedings of the IFAC Conference on Analysis, Design, and Evaluation of Man-Machine Systems* (S. 45-52). London: Pergamon Press.
- Wickens, C.D., Helleberg, J., Goh, J., Xu, X. & Horrey, B.(2001). *Pilot task management: Testing an attentional expected value model of visual scanning* (ARL-01-14/NASA-01-7). Savoy, IL: University of Illinois, Aviation Research Lab.