

How to Measure Monitoring Performance of Pilots and Air Traffic Controllers

Catrin Hasse, Dietrich Grasshoff & Carmen Bruder
German Aerospace Center
Department of Aviation and Space Psychology
catrin.hasse@dlr.de

Abstract

In prior research on the future of aviation it was established that operators will have to work with highly automated systems. Increasing automation will require operators monitoring appropriately (OMA). OMA are expected to demonstrate the use of distinctly different monitoring phases (orientation, anticipation, detection, and recheck). Within these phases, they must grasp in time the relevant information that would enable them to take control should automation fail. The presented study aims at finding appropriate measurements for the identification of OMA on the basis of eye tracking. In order to do this, a normative model of adequate monitoring behavior was designed including the definition of areas of interest. We tested 90 participants who had to monitor a dynamic automatic process, and then take control. In order to decide on suitable eye tracking parameters it was asked which parameters are significantly related to manual control performance. The results show that the suitability of parameters depends on the specific phase of the monitoring process. Gaze durations allow for differentiating between high and low performing subjects during orientation phases. In contrast, relative fixation counts are suitable for predicting monitoring performance during detection phases. In general, the results support the assumption that eye tracking parameters are appropriate for identifying OMA.

Keywords: automation, human monitoring performance, eye tracking, aviation, future ATM

1 Introduction

Technical innovations in aviation and improvements in ATM, aircraft systems and organizational structures are great challenges facing aviation in the 21st century. According to prior research on the future of aviation, such as the Single European Sky ATM Research (SESAR) Program, operators will have to work with highly automated systems.

As the automation of future workplaces in the field of aviation is assumed to increase, the ability to monitor automated systems will be one of the new core competencies relevant for success as an airline pilot or air traffic controller.

Workshops were conducted with experienced pilots and air traffic controllers in order to gather their expectations about their future tasks, roles and responsibilities [Bruder et al. 2008]. Findings from the workshop debriefings suggest a new crucial requirement for humans operating in man-machine settings: ‘operational monitoring’. Operational monitoring includes using one’s senses to follow up meaningful information from various sources (e.g. an automated system) without direct need for action. It involves being prepared to fully take over the handling of a system at any time, for example in the case of malfunction [Eißfeldt et al. 2009].

Thus, the increase in automation requires operators monitoring appropriately (OMA). OMA are assumed to monitor in such a way as to enable them to detect system errors in time, and to take control if automation fails. As Wickens et al. [1998] noted, increased automation might affect system performance due to the new skills that are required of operators, and that human operators might not have been adequately selected and trained for these changes. In order to prepare future training and selection processes, the question therefore arises: How can OMA be identified?

According to models of adequate and efficient monitoring behavior [e.g. Niessen and Eyferth 2001; Whitfield and Jackson 1982], and the differences between experts and novices [Underwood et al. 2003] it can be stated that appropriate monitoring behavior consists of target-oriented attention allocation both in general and during monitoring phases, i.e. orientation phase, anticipation phase, detection phase, and recheck phase. Based on this background, a normative model can be devised which describes the monitoring behavior of OMA. Whereas the first assumption requires the operator to adapt attention allocation to the specific requirements of a given situation, the second assumption focuses on the allocation of attention in phases. The operator is required to demonstrate flexibility in:

- orienting towards an approaching situation in general (during the initial orientation phase),
- anticipating system operations (during anticipation phases),
- detecting relevant system operations (during detection phases) and
- controlling system performance (during recheck phases).

We have so far derived criteria for effective monitoring behavior but what is missing is a suitable way of measuring monitoring performance.

A variety of psychophysiological and imaging studies support the idea that eye movements offer an appropriate means for measuring the efficient and timely acquisition of visual informa-

Copyright © 2012 by the Association for Computing Machinery, Inc. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions Dept, ACM Inc., fax +1 (212) 869-0481 or e-mail permissions@acm.org.

ETRA 2012, Santa Barbara, CA, March 28 – 30, 2012.
© 2012 ACM 978-1-4503-1225-7/12/0003 \$10.00

tion. For example, shifts in attention are usually reflected in the fixations [Findlay and Gilchrist 2003]: fixation counts can be used as a measure of the expectations and assumptions of the person [Rötting 2001], fixation durations reflect information processing duration [Inhoff and Radach 1998], and the total gaze duration per AOI is a measure of the difficulty of recording the information viewed [Rötting 2001].

However, in order to utilize eye movement measurements in the context of monitoring, it is necessary to first advance our understanding of the manner in which human monitoring is reflected by eye movements. This raises the question of how OMA typically guide their eye movements. In particular, which scanning profile enables the operator to assume control when necessary? The aim is to identify suitable eye tracking parameters which record the monitoring process and, at the same time, are related to manual control.

2 Method

An empirical study was undertaken to test the normative model of monitoring behavior, i.e. its postulated monitoring phases and their relationship to manual control. A simulation tool was developed to assess both monitoring performance and manual control performance. Diverse eye movement parameters were applied to record monitoring performance in order to decide on their suitability for identifying OMA.

2.1 Simulation equipment

A simulation tool called “Self Separation Airspace” (SSAS) was developed which allows the performance assessment of monitoring behavior and of manual control (Figure 1). As the tool is a simplified and abstract simulation of the basic requirements of future flight operators, test subjects need no prior experience as a pilots or air traffic controllers.

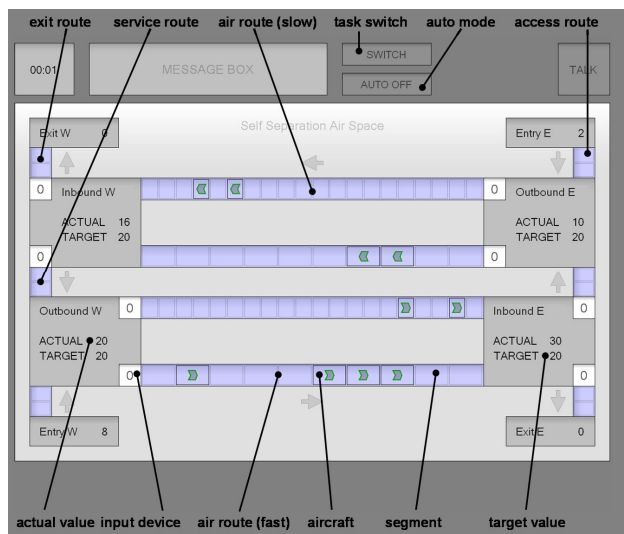


Figure 1 SSAS Simulation

The traffic flow simulation can be controlled either automatically or manually using input devices. The task of both the automated system and the human operator is to bring all actual values into agreement with target values [for further information, see Hasse et al. 2009].

2.2 Eye tracking equipment

Eye movements are recorded by the Eyegaze Analysis System manufactured by L. C. T.. The system was combined with the simulation tool SSAS, i.e. both systems use the same timestamp. NYAN software, developed by Interactive Minds, was used to manage the raw data. Subjects were seated in front of a 19-inch LCD computer display at a distance of approximately 60 cm.

2.3 Test Subjects

Experiments were conducted in conjunction with the regular selection process at the German Aerospace Center’s Department of Space and Aviation Psychology. The department is responsible for the selection of ab-initio pilot or air traffic controller trainees, or “ready entries”, as they are referred to in aviation. The sample consisted of 90 potential future aviation operators ranging in age from 17 to 26 years. 82% were male. Participation was not dependent on success in the selection process to ensure that potentially suitable candidates as well as less suitable candidates could be measured. Candidates received 20 € for participating and were assured that their performance in the experiment would not affect the selection process.

2.4 Procedure

Participants were first given the instructions for the experiment. They were informed they would work on a traffic flow simulation consisting of two mode; an automatic mode and a manual mode. For the automatic mode, participants were instructed to monitor the automated processes with the objective of understanding the rule-based dynamics of the simulation. Participants were informed that the system controls the flow of traffic in an ideal way. Their task was to grasp, in time, all relevant input devices in order to be prepared to fully take over the handling of the system. Referring to the manual control phase, they were required to manually control the system in terms of the rules and dynamics that they learned from monitoring the scenario during the automation phase. After a trial run, they were presented with the simulation, each mode for five minutes. Both modes started off with an orientation phase of 12 seconds where the display was frozen. The orientation phase should allow the participant to build up a picture of the simulation before it starts flowing dynamically.

2.5 Measurements

Eye tracking data and manual performance were the dependent variables.

Relative fixation counts (fixation number for defined AOIs in relation to all fixations in the same time frame, to account for individual differences in fixation counts), mean fixation and total gaze durations on predefined AOIs were measured. AOIs were determined which directly represent system operations done by the automation: focusing on these AOIs at the right time should help to keep an overview of system operations during an entire monitoring run, to orient towards system operations, and to anticipate, detect and control them.

As orientation, anticipation, detection and control are only possible within certain time frames during a scenario, every scenario was cut into sections. Each time frame stands for a monitoring phase and is characterized by AOIs that are necessary for monitoring adequately during this phase, e.g. effectively anticipating a system operation. Hence, this model leads us to expect

eye movements to focus on areas of interest that are generally relevant for the simulation as well as for monitoring phases within specific time frames.

The simulation allowed for differentiation between the orientation phase at the beginning of each scenario, anticipation phases (time frames in which important system operations can be anticipated), detection phases (time frames in which important system operations takes place and should be detected), and re-check phases (time frame in which the completion of system operation takes place and should be carefully controlled).

The mean deviation of actual values from target values was used to measure the performance of a test subject during the manual phase of each scenario.

3 Results

The relationship between eye gaze data and performance data was examined. In addition, on the basis of manual performance, data groups with distinct positive or negative results were identified and compared with regard their to eye movement parameters.

As expected, there is a significant negative correlation between relative fixation counts on scenario-specific AOIs and subsequent manual performance ($r = -.214$; $p < .05$). This negative relationship is expected because the SSAS performance parameter is reversed: the greater the proportion of fixations that fall on the predetermined areas, the smaller the mean deviation between actual and target values. This corresponds to better performance when actively controlling the simulation.

In addition, the relative frequency of fixations on relevant AOIs during all detection phases shows a significant correlation of medium strength with manual performance ($r = -.296$; $p < .01$): the higher the proportion of fixations on AOIs, the better the performance in the simulation.

Concerning orientation phases, absolute gaze durations correlate significantly with manual performance ($r = -.332$; $p < .01$). The longer a person gazed at relevant areas during orientation phases,

the smaller the mean deviation between actual and target values and the better the performance when actively controlling.

To get a deeper understanding of the link between monitoring and manual performance, participants were divided into three groups according to their manual performance (high performers, moderate performers and low performers). One-way analysis of variance (ANOVA) was used to compare manual performance (between-subjects factor) and eye movement behavior (within-subjects factor). By way of a post-hoc Tukey test, differences which are responsible for the outcome were analyzed. Analysis of variance reveals a significant difference between groups concerning the absolute fixation count ($F(2,86) = 3,669$; $p < .05$). A post-hoc Tukey test showed that high performers demonstrated a significantly higher fixation count than low performers ($p < .05$). That is, participants with better manual performance looked at relevant areas more often than participants with poor manual abilities (compare Figures 3 and 4). The results of an ANOVA reveal a significant difference between the means of the high and low performing groups with regard to their absolute gaze durations on AOIs during the orientation phase ($F(2,61) = 6,945$; $p < .005$). That is, participants with poor manual performance show inadequate attention allocation during the orientation phase, whereas participants with good manual performance show adequate attention allocation.

4 Discussion

Generally, our results support the assumption that suitable operators direct their attention to relevant areas during monitoring scenarios. Monitoring performance explains independent portions of variance. It could be verified that well performing operators direct their attention to relevant areas. That is, relative fixation counts on scenario relevant AOIs were significantly related to manual performance. Moreover, it could be shown that relative fixation counts on related AOIs during detection phases were related to manual performance. When absolute gaze durations were taken into account, the time spent on AOIs correlated significantly with manual performance. The longer a person gazed at relevant areas during the orientation phase, the better the performance when actively controlling the system.

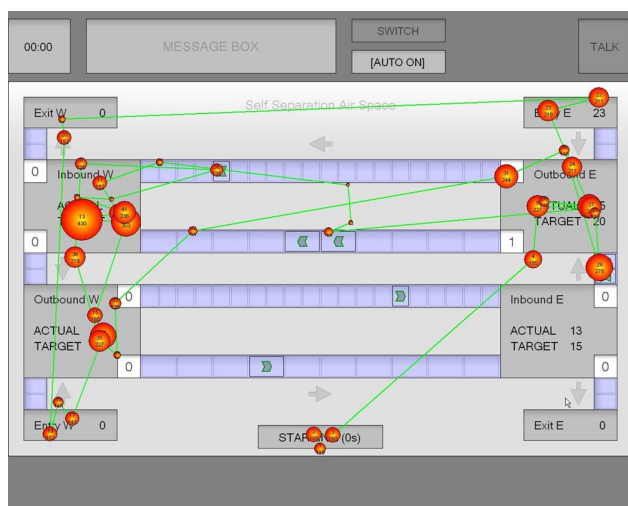


Figure 3 Participant with good manual performance and adequate attention allocation: Distribution of fixations as scan paths during orientation phase of scenario 2.

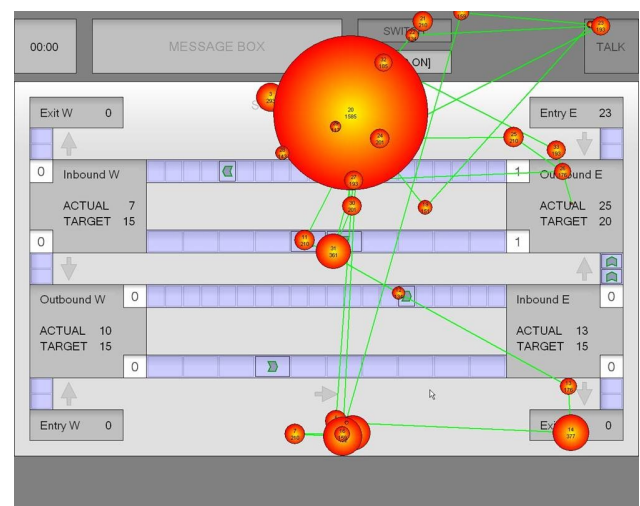


Figure 4 Participant with poor manual performance and inadequate attention allocation: distribution of fixations as scan paths during orientation phase of scenario 2.

It was questioned which parameters of monitoring behavior are crucial with regard to predicting monitoring performance, i.e. to detect errors and take control. Results indicate that the suitability of parameters depends on the specific phase of the monitoring process. Gaze durations allow for differentiating between high and low performing subjects during orientation phases. In contrast, relative fixation counts are suitable for predicting monitoring performance during detection phases. One reason for this finding could be that in the orientation phase, the simulation remains static. This allows for prolonged and reflected information processing to build up a picture of a certain traffic situation [Rötting 2001]. In contrast, the traffic flow changes dynamically during detection phases. In this case, it seems to be more efficient to fixate on relevant objects more frequently than to gaze at them for a longer time. By fixating on objects more frequently and thereby often jumping back and forth between objects, relationships between objects might be detected faster, and possible conflicts better identified.

Results indicate that orientation and detection phases are particularly suitable for identifying OMA. Analyses show substantial and robust effects. One reason could be that, for both of these phases, it can be clearly defined when people orient themselves and when they observe the operations of the automated system. In contrast, it is difficult to determine specific time frames in which anticipation and recheck have to occur.

However, relationships between monitoring behavior and manual control performance are substantial, but not very strong. The focus of this study was on monitoring parameters and their relationship to manual system handling. However, monitoring behavior, manual performance and the relationship between them might also be caused by other abilities aside from the ability to monitor appropriately.

In general, ability testing with dynamic simulations on the basis of eye movements is innovative and establishes new approaches for the assessment of future job requirements, such as operational monitoring. In this regard, the SSAS simulation is an adequate basis tool to investigate human performance in future ATM scenarios.

As part of the follow-up project, we are currently expanding SSAS to a test called MonT (Monitoring Test) by including a measurement of the ability to detect automation failures. This test could be used in the selection of future aviation personnel, such as pilots and air traffic controllers. Additional studies are planned to validate our model of adequate monitoring behavior with experienced controllers and pilots. Furthermore, a team version of MonT is going to be developed, rendering the assessment of team monitoring performance possible. Use of this tool may, in the long term, allow the selection of candidates on the basis of predefined monitoring behavior.

References

BRUDER, C., JÖRN, L., AND EIBFELDT 2008. Aviator 2030 - When pilots and air traffic controllers discuss their future. In *Proceedings of the 28th Conference of the European Association for Aviation Psychology*, A. Droog, and T. D'Oliveira, Eds., EAAP, 354-358.

EIBFELDT, H., GRASSHOFF, D., HASSE, C., HOERMANN, H.-J., SCHULZE KISSING, D., STERN, C., WENZEL, J., AND ZIERKE, O. 2009. *Aviator 2030 – Ability requirements in future ATM sys-*

tems II: Simulations and experiments. DLR Forschungsbericht 2009-28, DLR.

FINDLAY, J., AND GILCHRIST, I. 2003. *Active Vision*. Oxford University Press.

HASSE, C., BRUDER, C., GRASSHOFF, D., AND EIBFELDT, H. 2009. Future ability requirements for human operators in aviation. In *Engineering Psychology and Cognitive Ergonomics, 8th International Conference, EPCE 2009, Held as Part of HCI International 2009*, D. Harris, Ed., Springer, 537-546.

INHOFF, A. W., AND RADACH, R. 1998. Definition and computation of oculomotor measures in the study of cognitive processes. In *Eye guidance in reading, driving and scene perception*, G. Underwood, Ed., Elsevier, 29-53.

NIESSEN, C., AND EYFERTH, K. 2001. A model of the air traffic controllers' picture. *Safety Science* 73, 187-202.

RÖTTING, M. 2001. *Parametersystematik der Augen- und Blickbewegungen für arbeitswissenschaftliche Untersuchungen*. Shaker.

UNDERWOOD, G., CHAPMAN, P., BROCKLEHURST, N., UNDERWOOD, J., AND CRUNDALL, D. 2003. Visual Attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics* 46, 629-646.

WHITFIELD, D., AND JACKSON, A. 1982. The air traffic controller's picture as an example of mental models. In *Proceedings of the IFAC Conference on Analysis, Design, and Evaluation of Man-Machine Systems*, Johanssen, G., and Rijnsdorp, J. E., Eds., Pergamon Press, 45-52.

WICKENS, C. D., MAVOR, A. S., PARASURAMAN, R. & MCGEE, J. P. 1998. *The future of air traffic control: Human operators and automation*. National Academic Press.