Nationaal Lucht- en Ruimtevaartlaboratorium

National Aerospace Laboratory NLR

Executive summary

The Applicability of Eye Movements as an Indicator of Situation Awareness in a Flight Simulator Experiment

Problem area
This article discusses the applicability of eye movements as a means to assess Situation Awareness (SA) in a flight simulator setting. To ascertain a sufficient level of understanding of the system pilots are trained to visually scan their instruments in a structured manner as to assess the status of the aircraft and it's functioning. Therefore, understanding the visual scanning behaviour of pilots, insight may be gained into their visual attention which, in turn, may provide insight into their level of Situation Awareness.

Description of work
A scenario was designed that would hamper SA by introducing a malfunction (a fuel leak). It was expected that the pilots would spend more time on the displays on which the malfunction could be found (Electronic Centralised Aircraft Monitoring display; ECAM). This finding would reflect the perceptual level of SA (level 1). Eye track measures served as dependent variables and consisted of fix rates on the displays, dwell time on the displays and scanning entropy. The Crew Awareness Rating Scale (CARS) was used to compare the outcomes of the eye tracking

Report no.
NLR-TP-2009-702

Author(s)
G.K. van de Merwe
H. van Dijk
G.D.R. Zon

Report classification
UNCLASSIFIED

Date
February 2010

Knowledge area(s)
Training, Simulatie en Operator Performance
Cockpit

Descriptor(s)
Eye tracking
Situation Awareness

This report is based on a paper to be published in Aviation, Space and Environmental Medicine.

measurements with the pilot's subjective interpretation of their SA.

Results and conclusions

It was found that in the period before the malfunction (pre-period), the most time was spent by the crew on the NAV display and the PFD. After the malfunction was introduced, the ECAM display was the prime object of attention. Interestingly, the time spent on the NAV display and the PFD was traded against the time spent on the ECAM display whilst the focus on the 'rest' of the viewing area remained the same. This finding was further confirmed by the negative relationship between the search period and the time spent on the ECAM display. That is, crews that spent more time on the ECAM display were able to solve the problem quicker. Fix rates showed similar results compared to the dwell time on the different displays. Both measures showed the change in viewing behaviour from the NAV display and PFD to ECAM display in the same manner. The entropy analyses showed an increase in the randomness of the viewing pattern in the post-period when the malfunction had to be found. When comparing the successful crews with the unsuccessful ones the latter were much less structured in their cockpit scanning behaviour compared to the successful ones. This means that their cockpit scanning pattern was much more predictable than the patterns of the unsuccessful ones. This seems to suggest a higher level 3 SA for the successful crews. If this is true than it can be inferred that they also had a better comprehension of the situation (level 2 SA) spurring them to more efficiently scan the cockpit instruments in search of the malfunction. The CARS outcomes showed an increase in SA during the post-period. It was hypothesised that SA would be reduced because the pilots were dealing with an unknown malfunction and were searching for a solution. Furthermore when comparing the successful crews with the unsuccessful ones the former reported a much greater improvement of their SA compared to the unsuccessful participants. In other words, because they were successful in effectively scanning the right displays they were able to build up a good mental picture and able to solve the problem.

Applicability

This study was able to demonstrate the applicability of eye-movement analyses as an indicator of SA. A better understanding of the perceived SA of pilots was made possible by objectively studying the pilot's search patterns in relation to information acquisition (fix rates and dwell time; level 1 SA), and the use of this information to effectively steer new information acquisition activities (entropy, level 3 SA).

Nationaal Lucht- en Ruimtevaartlaboratorium, National Aerospace Laboratory NLR

Anthony Fokkerweg 2, 1059 CM Amsterdam,
P.O. Box 90502, 1006 BM Amsterdam, The Netherlands
Telephone +31 20 511 31 13, Fax +31 20 511 32 10, Web site: www.nlr.nl

Nationaal Lucht- en Ruimtevaartlaboratorium

National Aerospace Laboratory NLR

NLR-TP-2009-702

The Applicability of Eye Movements as an Indicator of Situation Awareness in a Flight Simulator Experiment

G.K. van de Merwe, H. van Dijk and G.D.R. Zon

This report is based on a paper to be published in Aviation, Space and Environmental Medicine.

This publication has been refereed by the Advisory Committee AIR TRANSPORT.

Customer	European Commission
Contract number	----
Owner	National Aerospace Laboratory NLR
Division NLR	Air Transport
Distribution	Unlimited
Classification of title	Unclassified
	February 2010

Approved by:

Author	Reviewer	Managing department
27-1-2010	M. de Leeuw 28-01-2010	15-02-'10

Summary

This article discusses the applicability of eye movements as a means to assess Situation Awareness (SA) in a flight simulator setting. A flight simulator experiment was set up that was specifically designed to manipulate SA in order to better understand its implications for eye movements. In this scenario SA was hampered by introducing a system malfunction in the form of a fuel leak that ultimately resulted in a fuel imbalance. Twelve airline pilots operating in teams of a Captain and a First Officer participated in the experiment. Area's of Interest (AoI) were defined throughout the cockpit, namely: the Navigation display (NAV), Primary Flight Display (PFD), Electronic Centralised Aircraft Monitoring display (ECAM) and the 'rest' of the cockpit (including the outside view). This study was able to demonstrate the applicability of eye-movement analyses as an indicator of SA. When comparing the successful crews with the unsuccessful ones the latter were much less structured in their cockpit scanning behaviour compared to the successful ones. A deeper understanding of the perceived SA of pilots was made possible by objectively studying the pilot's search patterns in relation to information acquisition (fix rates and dwell time; level 1 SA), and the use of this information to effectively steer new information acquisition activities (entropy, level 3 SA).

Contents

Abbreviations

AoI	Area of Interest
ASL	Applied Science Laboratory
CARS	Crew Awareness Rating Scale
ECAM	Electronic Centralised Aircraft Monitor
EPOG	Eye Point Of Gaze
GRACE	Generic Research Aircraft Cockpit Environment
HILAS	Human Integration in the Lifecycle of Aviation Systems
NAV	Navigation display
PF	Pilot Flying
PFD	Primary Flight Display
PNF	Pilot Not Flying
SA	Situation Awareness

1 Introduction

This article discusses the applicability of eye movements as a means to assess Situation Awareness (SA) in a flight simulator setting. For a pilot to have SA a certain level of understanding of his environment is required, including the parameters presented by the cockpit automation. This understanding is essential in forming the basis for subsequent decision making and performance in such a complex, dynamic and information-rich environment (Endsley 1995a). To ascertain a sufficient level of understanding of the system pilots are trained to visually scan their instruments in a structured manner as to assess the status of the aircraft and its functioning. Therefore, understanding the visual scanning behaviour of pilots insight may be gained into their visual attention which, in turn, may provide insight into their level of Situation Awareness.

A frequently used definition of Situation Awareness is ‘the perception of 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’ (Endsley, 1988, 1995b). The first part of the definition describes the perceptual aspect of information processing, the second part describes the mental activities that process the perceived information and the third part describes the resulting (mental) activity. Clearly the higher levels of SA (mental processing) depend on the lower levels (perception of information). A pilot may notice a difference in fuel levels between the left and the right Primary Flight Display (PFD), comprehends that this imbalance may mean a faulty pump and subsequently understands that this may cause a shift in the centre of gravity of the aircraft which will degrade its performance. The differences between these three stages are important since it points to differences in perceptual and cognitive processes ranging from low-level perceptual processes to (level 1) to higher level cognitive processes (level 2 and 3; Wickens, 2008). Since the majority of information in a cockpit is presented visually it is interesting to study visual information processing in the form of eye movements to gain insight into the perceptual qualities that underpin SA.

Eye movements are an indicator of visual attention (see Rayner 1998 for a review). An important benefit of studying eye movements is its non-intrusive continuous character. Eye activity is captured without interrupting the activity, is measured continuously during the activity and is measured in an objective manner. Two types of visual attention can be described: covert and overt visual attention (Styles, 1997; Johnson and Proctor, 2004). Covert visual attention is the movement of attention that can occur without movements of the eyes (also referred to as the attentional spotlight). For example a pilot can focus his gaze on the Integrated Control Panel whilst altering the settings of the auto pilot but focus his attention on the PFD to see if the changes have any effect. Overt visual attention is the actual movement of the gaze to a certain point in space. For example after setting the autopilot the pilot focuses his gaze on the PFD to confirm the changes. This means that attention can move without the accompanying eye movements. This may seem to make the usability of eye movements as a means to study visual attention limited. However, covert and overt attention are often aligned in space making eye movements a useful means to assess visual attention during search (Zelinsky, 2008).

Although humans can look in one direction and attend to another, previous research has shown that before a saccadic movement occurs to a certain location, attention is directed to this location first (see Johnson and Proctor 2003 for an overview). That is, covert attention frequently scans the visual field for interesting objects. Once such an object is found, overt attention is shifted to this location. Therefore one can be reasonably confident that when the eyes focus on a certain location, attention is also focussed at this specific location (see Shinar, 2008, for a discussion).

An important part of Endsley's SA definition is the perception of information in the environment. Perceiving relevant information in the environment is a crucial first step in the establishment of SA. With modern means of capturing eye movements, the eyes make for a convenient measurement of visual attention compared to using apparatus for measuring neural activity or indirect measurements such as reaction times or questionnaires.

A flight simulator experiment was set up that was specifically designed to manipulate SA in order to better understand its implications for eye movements. A scenario was designed that would hamper SA by introducing a malfunction (a fuel leak). The pilots were only aware of the fact that 'something' was malfunctioning and should be discovered. The aim of this scenario was to trigger visual search activity of the pilots when they were trying to discover the issue. It was expected that a shift in focus on the cockpit displays would occur once the pilots were aware of a malfunction. It was expected that the pilots would spend more time on the displays on which the malfunction could be found. This finding would reflect the perceptual level of SA (level 1). Furthermore assessing the order in which the displays were viewed provided insight into the search strategies that were used. These activities may reflect higher order SA indicating comprehension of the situation (level 2) and strategic search behaviour (level 3). It was expected that visual search strategies may become less structured since the pilots were searching for the malfunction.

2 Method

The current experiment was part of a larger study into the applicability of relevant Human Factors tools in the design, evaluation and operation of aviation systems (Human Integration into the Lifecycle of Aviation Systems; HILAS; Zon and Van Dijk, 2009).

2.1 Participants

Twelve airline pilots operating in teams of a Captain and a First Officer participated in the experiment. All pilots were active and qualified to fly an Airbus A320. Each pair of pilots acted as Pilot Flying (PF) and Pilot Not Flying (PNF) respectively. Their average age was 38 years (SD = 6.0 years) and with an average of 7450 hours of flight experience (SD = 4850) on different types of aircraft.

2.2 Technical set-up

The simulations were run in the Generic Research Aircraft Cockpit Environment (GRACE) simulation facility configured as an Airbus A320 (Heesbeen, Ruigrok and Hoekstra, 2006).

Before commencing the experiment ample training runs were performed to familiarize the pilots with the simulation facility. For this experiment Applied Science Laboratories (ASL) 6000 eye trackers were used. The eye trackers use infrared optics and a camera to track eye movements. An optical head tracker is used to track head movements. The combination of eye- and head movements results in the Eye Point of Gaze (EPOG). This information was used to identify the location of the pilot’s visual focus.

Area’s of Interest (AoI) were defined throughout the cockpit, namely: the Navigation display (NAV), Primary Flight Display (PFD), Electronic Centralised Aircraft Monitoring display (ECAM) and the ‘rest’ of the cockpit (including the outside view).

2.3 Scenario description

The scenario consisted of a trip from London Heathrow to Amsterdam Schiphol and lasted about 25 minutes. The scenario started during cruise and ended after touchdown at Schiphol. In this scenario SA was hampered by introducing a system malfunction in the form of a fuel leak that ultimately resulted in a fuel imbalance. This malfunction was depicted on the ECAM display in the bottom-left corner. At first the pilots were warned about an ‘undefined system error’, after about 5 minutes the warning changed to ‘fuel imbalance’ (see Figure 1).

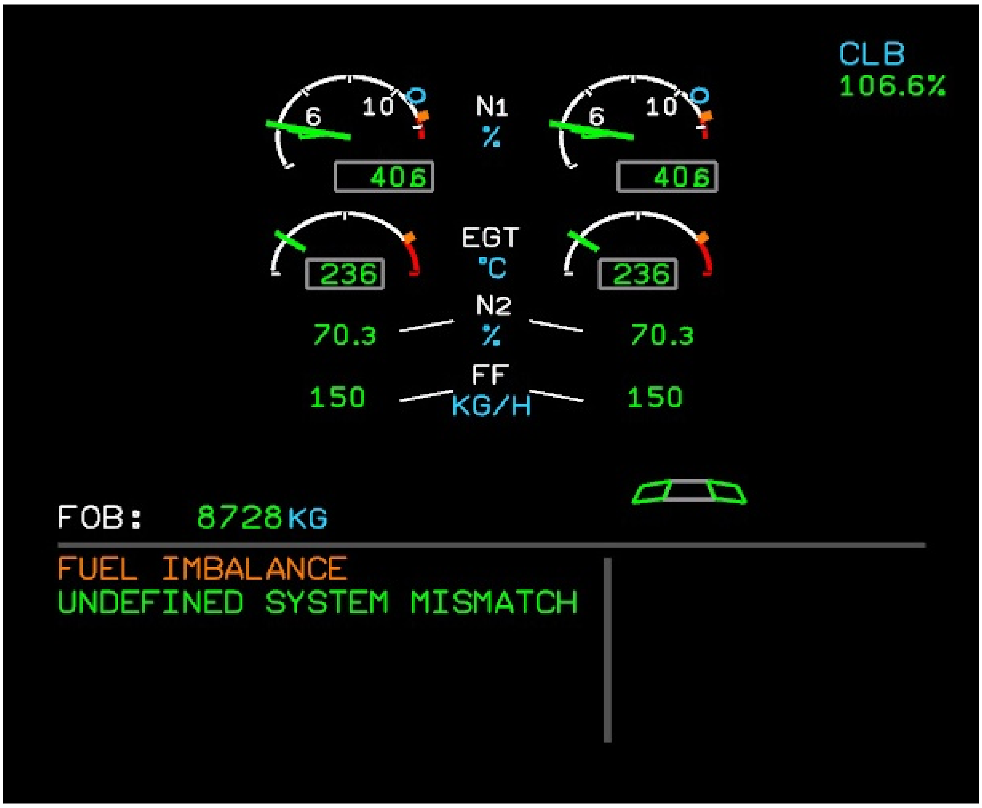

Figure 1. The ECAM display with the fuel indicator (FOB)

The fuel leak was introduced after a pause in the scenario. During the break the pilots were informed that once they commenced their flight, a system malfunction had occurred. Their task was to investigate what this malfunction was. Once the fuel amount was below a certain threshold, a warning message was depicted on the ECAM display. The fuel leak message on the ECAM display indicated a significant reduction in the amount of fuel in one of the fuel tanks. The flow meters on the ECAM display were only visible if this specific page was displayed. The auditory warning that followed the significant reduction in fuel quantity was also displayed on the warning and caution part of the ECAM screen and was always visible irrespective of the ECAM page.

For the post-experimental data analyses, the time it took the crew to discover the fuel leak after the warning was displayed was named the ‘post-period’. In the analysis, this post-period was compared to a reference period (the ‘pre-period’) of equal length as the post-period and took place immediately before the onset of the fuel leak. A graphical representation of the timeline is depicted in Figure 2.

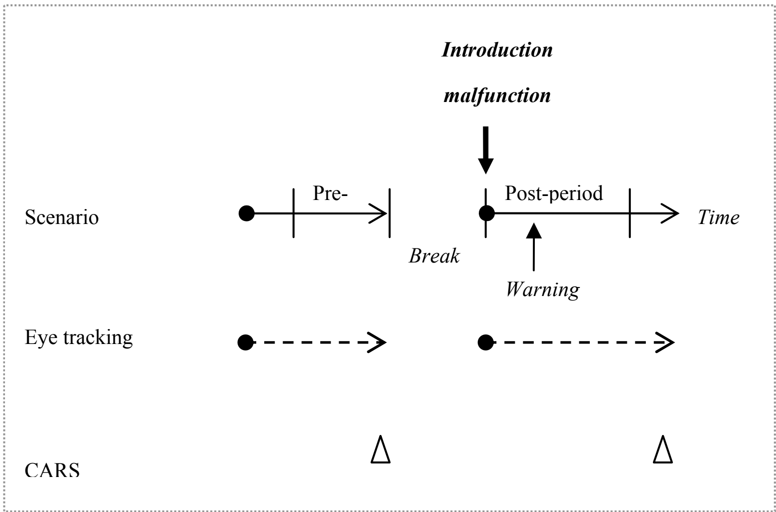

Figure 2. Illustration of the timeline and measurement periods for the scenario

2.4 Independent and dependent variables

The scenario was developed specifically to manipulate SA. Therefore SA served as an independent variable. Eye track measures served as dependent variables and consisted of fix rates on the displays, dwell time on the displays and scanning entropy. These variables were measured continuously throughout the scenario. The crew's SA was also measured subjectively through the CARS questionnaire which was administered directly after the pre-period and the post-period (see also Figure 2).

The number of fixations per minute is called 'fix rate'. Eye activity is considered to be a fixation when the eye remains focussed on a specific point. Based on a literature review by Jacob and Karn (2003) a fixation is defined as any eye movement on the display that is less than one degree of visual arc for a minimum time period of 150 ms. Fixations are interesting to study since it is considered that visual information is only acquired during fixations and not during intermediate saccades (Rayner, 1998). Therefore fixations on a particular display provides insight into the visual workload it generates (Fitts, Jones and Milton, 1950).

The total fixation time within a specific AoI is called dwell time. The ratio between dwell time on a specific AoI and the total dwell time on all AoI was used as a dependent variable. Dwell time is interesting to investigate because it reflects the importance of that display to the pilot (Jacob and Karn, 2003).

The randomness of the eye activity is called entropy (Harris, Glover and Spady, 1986). In the present experiment the normalised stratified entropy was used as a dependent variable which varies between 0 (scan patterns completely predictable) and 1 (scan pattern completely random).

Entropy is interesting to assess since it indicates the level of strategic visual search activity (or lack thereof) of the pilot (Zhang, Smith and Witt, 2006). A high entropy level indicates low visual workload and vice versa.

In addition, the Crew Awareness Rating Scale (CARS; McGuinness and Foy, 2000) was used to compare the outcomes of the eye tracking measurements with the pilot's subjective interpretation of their SA. The CARS is a post-trial questionnaire based on Endsley's definition of SA (1995) and focuses on how well the pilot could identify, comprehend, predict and decide in the given scenario. The use of a subjective rating scale is interesting as it provides insight into the pilot's perceived SA and provides a basis for comparison with the eye track measurements.

3 Results

An α of 5% was used as a significance value and Cohen's d as a measure of effect size.

3.1 Fix rate

A significant interaction effect was found between the pre-period, post-period and display type ($F(3,8) = 11.434$, $p < .01$, $\eta^2_p = .811$). The effects of the fixation rates are depicted in the graph below. It is shown that in the pre-period the largest number of fixations were on the NAV display and the PFD. Once the crew was informed that a malfunction had occurred the ECAM display became the primary focus of attention at the cost of the number of fixations on the NAV and PFD (see Figure 3).

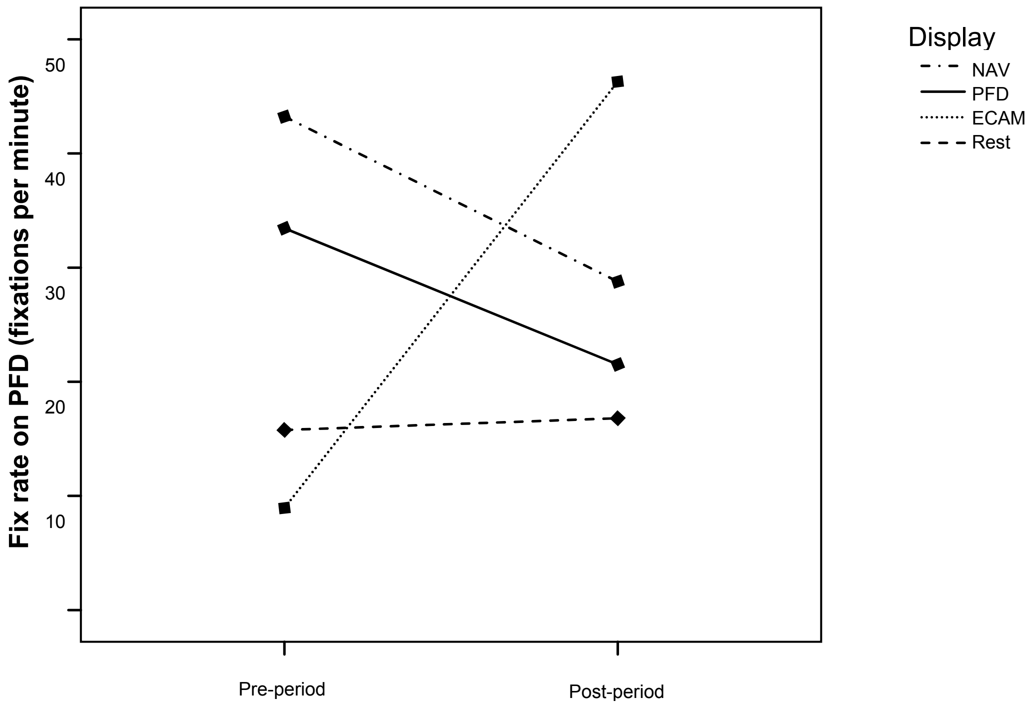

Figure 3. The fixation rates on the displays in the pre- and post period

The relationship between the time it took for the crew to discover the fuel leak on the ECAM display and the amount of time spent looking at the different displays was also analysed. It was assumed that the error could only be discovered when the ECAM display was viewed and the right page was shown. A significant correlation was found between the number of fixations per minute on the ECAM display and the duration of the search period ($r(11) = -.742, p < .01$). This means that the higher the number of fixations per minute on the ECAM, the quicker the discrepancy was found. Further negative correlations were found between the NAV fix rate and the ECAM fix rate ($r(11) = -.803, p < .01$), the rest of the areas of interest and the ECAM display ($r(11) = -.643, p < .05$) and a trend between the PFD and the NAV display ($r(11) = -.552, p = .078$).

3.2 Dwell time

A significant interaction effect was found between the pre-period, post-period and display type ($F(3,8) = 11.322, p < .01, \eta^2_p = .809$). This is depicted in the graph below in which in the pre-period most time was spent on the NAV display and the PFD (see Figure 4). Once the crew was informed that a malfunction had occurred the ECAM display became the primary focus of attention at the cost of dwell time on the NAV and PFD.

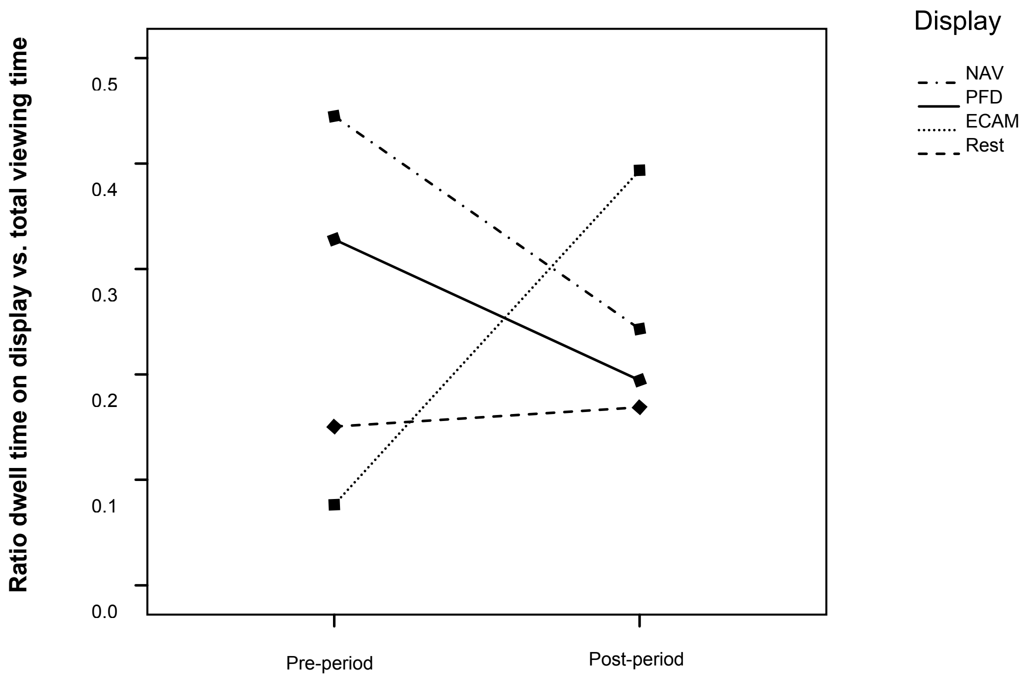

Figure 4. The ratio of dwell time on the displays vs. the total dwell time

A significant negative correlation was found between the discovery period and the amount of time spent on the ECAM display ($r(11) = -.763, p < .01$). This means that the more time the crew members spent on the ECAM display, the less time it took to discover the fuel leak. Significant negative correlations were found between the dwell time on the NAV display and the ECAM display ($r(11) = -.804, p < .01$), as well as the PFD and the ECAM display ($r(11) = -.631, p < .05$). These findings seem to indicate the previously mentioned trade-off between the dwell time on the NAV display and the PFD versus the dwell time on the ECAM display.

3.3 Entropy

A trend was found between the pre- and the post period that shows an increase in entropy in the post-period ($t(10) = -2.167$, $p = .06$, $d = -.65$). This means that the eye movements of the pilots become less systematic in their attempt recover the source of malfunction (see Table 1).

Table 1. The stratified normalized entropy for the pre- and post-period

Entropy	Mean	SD	SE mean
Pre-period	.79	.11	.04
Post-period	.89	.06	.02

3.4 CARS

The outcomes of the CARS questionnaire show a higher score in the post-period compared to the pre-period ($F(1,11) = 18.175$, $p = .001$, $\eta^2_p = .623$). This means that the pilots reported their SA to be higher in the period when they were searching for the malfunction compared to the period prior to the malfunction (see Table 2).

Table 2. CARS scores indicting a higher SA in the post-period

CARS	Mean	SD	SEM
Pre-period	.35	.10	.03
Post-period	.50	.11	.03

3.5 Successful vs. unsuccessful crews

Two of the seven crews were unable to pin-point the malfunction within the time-limits of the scenario. The data of these two crews were analysed and compared to the successful crews. No statistical analyses were performed because of the low number of participants. However, these results show some interesting insights into the differences in scanning behaviour.

In terms of dwell time, the successful crews showed a 45% increase on the ECAM display in the post-period compared to the pre-period versus only an 8% increase for the unsuccessful crews (see Figure 5). This seems to indicate that the crews that looked extensively at the ECAM display were in general more successful in allocating the malfunction.

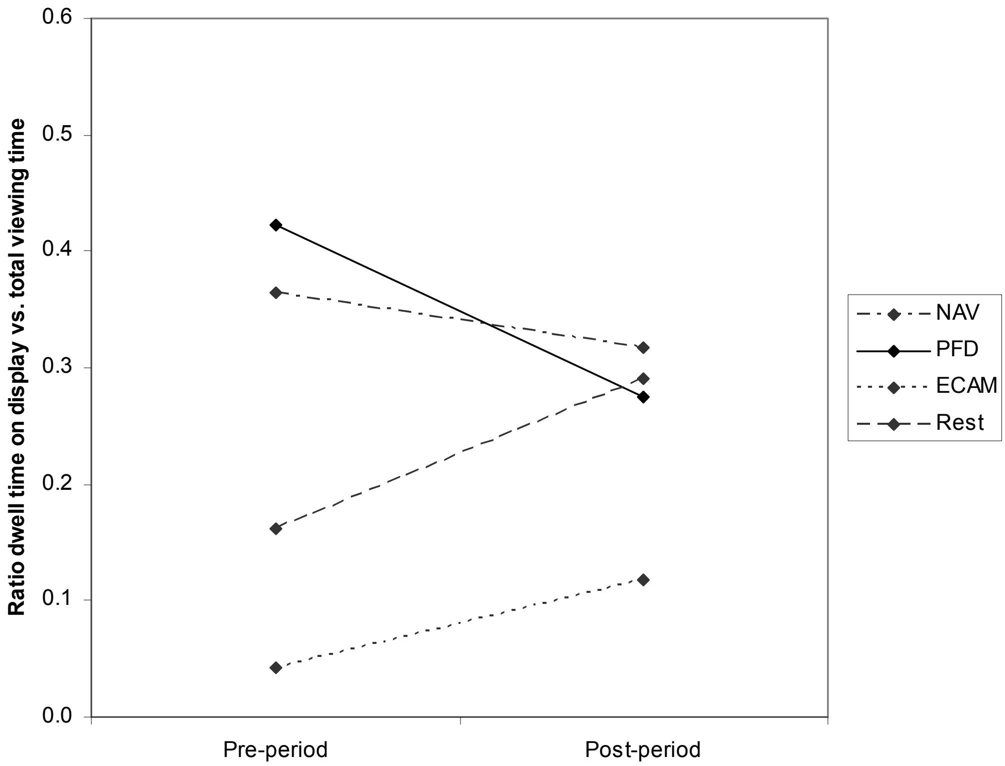

Figure 5. The dwell times on the display for the unsuccessful crews

The entropy scores of the unsuccessful crew increased around 28% in the post-period versus around only 5% in the successful crews. Apparently the crews that were unable to find the malfunction were scanning the cockpit in a more random manner than the crews that were able to find it.

Interestingly the increase in the CARS scores of the unsuccessful crews for the post-period was lower (+8%) than the increase for the successful crews (+51%) compared to the pre-period. That is, both the successful and the unsuccessful crews indicated that their SA scores improved but the successful crews indicated the largest improvement.

4 Discussion

The aim of this experiment was to investigate the potential for eye movement analysis as an indicator of SA. In particular it was hypothesized that because of the perceptive nature of eye movements these could be used to assess a pilot’s lower level SA (level 1) as well as strategic search behaviour (level 3).

A scenario was devised that specifically aimed to hamper SA as to assess the change in eye movements of the pilots. It was found that none of the pilots was aware of the sudden drop in fuel quantity as depicted on the ECAM display that was introduced after the scenario break. All pilots needed an auditory warning before they started investigating the display and search for the malfunction. After they received the warning their visual attention shifted from primarily the NAV display and the PFD to the ECAM display. It was found that in the pre-period, the most time was spent by the crew on the NAV display and the PFD. After the malfunction was introduced, the ECAM display was the prime object of attention. Interestingly, the time spent on the NAV display and the PFD was traded against the time spent on the ECAM display whilst the focus on the ‘rest’ of the viewing area remained the same. This finding was further confirmed by the negative relationship between the search period and the time spent on the ECAM display. Also, a negative relation between the NAV display and ECAM display and the PFD and the ECAM display showed that there was a trade-off between time spent viewing the NAV display and PFD versus the ECAM display. Fix rates showed similar results compared to the dwell time on the different displays. That is, here also a trade-off was found between the fixation rate on the NAV display, the PFD and the ECAM display. These results were further confirmed by the negative correlation between the ECAM display and the search period. This came at a cost of fixations on the remainder of areas of interest (NAV display, PFD and the rest).

Few differences were found in outcomes between the fix rate and the dwell time. Both measures showed the change in viewing behaviour from the NAV display and PFD to ECAM display in the same manner. Fix rate and dwell time were used to identify the difficulty in extracting information from the displays and the importance of the display respectively (Fitts, Jones and Milton, 1950; Jacob and Karn, 2003). As the contents of the displays did not change during the course of the experiment it is unlikely that pilots had more difficulty in extracting information from the displays thereby resulting in an increase in fix rate. It seems more likely that since the malfunction was displayed on the ECAM display this would elicit viewing behaviour towards this particular display. Therefore, the increase in fix rate on the ECAM display is more likely to be attributed to the change in relevance of the displays rather than a change in complexity. As a result, the similarity between the results of the dwell times and the fix rates seem to suggest that both measures of visual activity measure the same; i.e. relative importance of the display.

If fix rate and dwell time can be used to indicate the importance of the viewing area, as suggested above, they therefore provide similar insight into the pilot’s acquisition of information (i.e. level 1 SA). Accurate information acquisition is the first step in acquiring full SA and is crucial in laying the foundation for a good performance. When investigating the dwell times of the successful and unsuccessful crews this finding was further supported. That is, crews that spent limited time on the ECAM display were not able to find the source of the malfunction whereas the successful ones did.

The entropy analyses showed an increase in the randomness of the viewing pattern in the post-period when the malfunction had to be found. Interestingly a high entropy indicates a low mental workload (Harris, Glover and Spady, 1986). In our experiment this may have meant that the pilots were not exerting high levels of mental activity to find the source of the malfunction. This is a plausible explanation since the period of the flight in which the malfunction occurred was not very demanding (i.e. cruise section). Therefore the difference in entropy scores is more likely to indicate the pilot's strategic search behaviour to localise the malfunction rather than indicating their mental workload. As it was unclear for the pilots what exactly was the cause of the malfunction they found themselves in an ambiguous situation. This may have resulted in an increase in entropy in which pilots were searching in a less structured manner compared to their normal scanning pattern in the pre-period. Upon further analyses this finding seems to have been aggravated by the scanning patterns of the unsuccessful crews.

When comparing the successful crews with the unsuccessful ones the latter were much less structured in their cockpit scanning behaviour compared to the successful ones. This means that their cockpit scanning pattern was much more predictable than the patterns of the unsuccessful ones. This seems to suggest a higher level 3 SA for the successful crews. If this is true than it can be inferred that they also had a better comprehension of the situation (level 2 SA) spurring them to more efficiently scan the cockpit instruments in search of the malfunction.

The CARS outcomes showed an increase in SA during the post-period. It was hypothesised that SA would be reduced because the pilots were dealing with an unknown malfunction and were searching for a solution. In other words, since the pilots, for most of the time, did not know the cause of the malfunction, it was expected that their SA would be reduced for as long as they had not found the cause. However, apparently the malfunction prompted the pilots to more scrupulously search the various cockpit instruments in comparison to the pre-period. This may have improved the pilot's perception of their own SA. Furthermore when comparing the successful crews with the unsuccessful ones the former reported a much greater improvement of their SA compared to the unsuccessful participants. In other words, because they were successful in effectively scanning the right displays they were able to build up a good mental picture and able to solve the problem.

This study was able to demonstrate the applicability of eye-movement analyses as an indicator of SA. A better understanding of the perceived SA of pilots was made possible by objectively studying the pilot's search patterns in relation to information acquisition (fix rates and dwell time; level 1 SA), and the use of this information to effectively steer new information acquisition activities (entropy, level 3 SA).

5 Endnotes

The HILAS project is part of the 6th framework programme for aeronautics and space research, sponsored by the European Commission. The authors would like to thank the European Commission for sponsoring this research. Further, we would like to thank the HILAS flight deck technologies strand members for their contribution in the experiment.

6 References

Colvin, K., Dodhia, R.M., Belcher, S. and Dismukes, K. (2004). Scanning for Visual Traffic: An Eye Tracking Study. *Proceedings of the 12th International Symposium on Aviation Psychology,* 255-260.

Endsley, M.R. (1995a). Measurement of Situation Awareness in Dynamic Systems. *Human Factors, Vol. 37*(1), 65-84.

Endsley, M.R. (1995b). Towards a theory of Situation Awareness in Dynamic Systems. *Human Factors, Vol. 37*(1), 65-84.

Endsley, M.R. (1988). Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting,* 97-101.

Fitts, P.M., Jones, R.E. & Milton, J.L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review, 9*(2), 24-29.

Harris, R.L., Glover, B.L., & Spady, A.A. (1986). *Analytic techniques of pilot scanning behavior and their application* (NASA technical paper 2525). Hampton, VA: NASA Langley Research Center.

Heesbeen, B.W.W.M., Ruigrok, R.C.J. & Hoekstra, J.M. (2006). GRACE – a Versatile Simulator Architecture Making Simulation of Multiple Complex Aircraft Simple. *AIAA Modeling and Simulation Technologies Conference and Exhibit.*

Jacob, R.J.K. & Karn, K.S. (2003) Eye tracking in human computer interaction and usability research (section commentary), in J. Hyona, R. Radach, and H. Deubelin (Eds.) *The Mind's Eye Cognitive and Applied Aspects of Eye Movement Research*, Elsevier Science; Oxford.

Johnson A. & Proctor, R. W. (2004). *Attention: Theory and Practice.* Thousand Oaks: Sage Publications Inc.

McGuinness, B., & Foy, L. (2000). A subjective measure of SA: The Crew Awareness Rating Scale. In D. B. Kaber and M. R. Endsley (Eds). *Human performance, situation awareness and automation: user centered design for the new millennium.* Atlanta, GA: SA Technologies.

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin, Vol. 124*(3), 372-422.

Shinar, D. (2008). Looks Are (Almost) Everything: Where Drivers Look to Get Information. *Human Factors, Vol. 50*(3), 380-384.

Styles, E.A. (1997). *The Psychology of Attention.* East Sussex: Psychology Press Ltd.

Wickens, C.D. (2008). Situation Awareness: Review of Mica Endsley's 1995 Articles on Situation Awareness Theory and Measurement. *Human Factors, Vol. 50*(3), 397-403.

Zelinsky, G.J. (2008). A Theory of Eye Movements During Target Acquisition. *Psychological Review, Vol. 115*(4), 787-835.

Zhang, H., Smith, M.R.H. & Witt, G.J. (2006). Identification of real-time diagnostic measures of visual distraction with an automatic eye-tracking system. *Human Factors, 48*(4), 805-821.

Zon, R., Dijk, H. van (2009). A Selection of Human Factors Tools: Measuring HCI Aspects of Flight Deck Technologies. *Proceedings of the HCI International 2009,* 616-624.