The Evaluation of Pilot's First Fixation and Response Time to Different Design of Alerting Messages

Wen-Chin Li[1(✉)], Jiaqi Cao[2], Jr-Hung Lin[3], Graham Braithwaite[1], and Matthew Greaves[1]

[1] Safety and Accident Investigation Center, Cranfield University, Cranfield, UK
wenchin.li@cranfield.ac.uk
[2] Commercial Aircraft Corporation of China (COMAC), Shanghai, China
[3] Office of Institutional Research, National Central University, Taoyuan, Taiwan

Abstract. Current research investigates the limitation of current cockpit design for Crew Alerting System (CAS) and Quick Reference Handbook (QRH), and proposes a potential design solution that might enhance crew performance. Using eye tracking device, an experiment was conducted based on Flight Simulator Software. Objective eye movement data as long as subjective feedback from participants were collected to evaluate the design. 24 participants experienced 4 scenarios with both traditional design and integrated design. Results demonstrated that participants never made error in locating the instructions by integrated design. It is due to the obvious fact that the instructions of integrated design will present itself, hence decrease the chance of executing the wrong NNC by eliminating the degree of choices. On the other hand, 75% of the participants experienced at least one error in finding the correct instruction, which is probably due to the limitation of QRH itself due to options too close in formatting which causes clicking the wrong page accidently. Furthermore, a trend was identified by eye movement patterns for longer fixation duration, smaller saccade amplitude, and less time fixated on instruction area occur in integrated design. Current research found that integrated design of CAS and QRH is superior in acquiring accurate solutions for emergent situations and processing the information presented compared with traditional design.

Keywords: Cockpit design · Crew Alerting System · Eye movement · Human-computer interaction · Quick Reference Handbook

1 Introduction

Commercial aircrafts equipped with Engine indication and Crew Alerting System (EICAS) are able to monitor real-time aircraft health and provide Crew Alerting System (CAS) messages which indicate abnormal situations to flight crews. With the support of Quick Reference Handbook (QRH), in which contains Non-Normal Checklist (NNC), flight crews will manage to find solutions for associated situations. However, there is no link or coordination between these two systems. The integrated CAS and QRH is very intuitive to fit the principle of human-centered design. It is

D. Harris (Ed.): EPCE 2017, Part I, LNAI 10275, pp. 21–31, 2017.
DOI: 10.1007/978-3-319-58472-0_2

obvious that once flight crews are alerted by CAS of an emergency, they would like to know how to deal with it immediately; hence the systems present the checklists in QRH to provide pilot relevant information. Kirsh (1995) proposed that contracting the number of decisions users must make is an indication of a good design. The task of handling emergency situations can be simplified as a single cognitive process, using cues from external environment to search for a specific procedure. If new design is able to reduce search space and time, crews will have more time solving the problems in external world. It is possible to make representations of QRH more active in order to help crews see what is most relevant to deciding what to do next (Hollan et al. 2000). One rule in Proximity Compatibility Principle (Wickens and Carswell 1995) in terms of spatial proximity can be used for this integration, which suggests two pieces of information that need to be integrated on a cluttered display should be placed in close spatial proximity. Electronic Checklist (ECL) developed for Boeing 777 is a close to ideal model of this integrated design.

QRH is designed with the intention of allowing flight crews to minimize the need for a lot of effortful analysis when time may be limited and workload is high (Burian 2004), as long as the expectation that they will correctly interpret the cues available to them and find the appropriate checklist to solve problems (Burian 2006). However, it is very important for QRH designers to be aware of how human will behave in emergent situations. Currently there are two types of QRH, paper or electronic form. Nowadays, paper QRH is less frequently used due to the large amount of effort needed to retrieve the correct NNC and the placement of electronic QRH in modern glass cockpit. However, Civil Aviation Authority (2006) requires that paper QRH should be still onboard of aircraft in case of system failure in CAP 676, which is the guideline on design, presentation and use of paper form NNC. NNC are used when the aircraft is experiencing one or more system failures (Boorman 2001), which will be displayed automatically upon detection of the related alert, so ECL's provide a means to accomplish checklists with a reduction of searching time and a reduction in the possibility of crew error (Federal Aviation Administration 1996). A slight downside of Boeing's ECL is that the NNC is presented in another display beneath EICAS display and a few steps still need to be done before reaching the checklist (Fig. 1).

Fig. 1. Position of EICAS display and NNC

The application of eye-tracking in the study of flight simulation is promising as it provides direct feedback, which could diagnose potential factors that impact upon pilot attention and situation awareness on the flight deck (Robinski and Stein 2013). As suggested by Jones and Endsley (1996), over 75% of pilot errors are caused by perceptual failures, so it is interesting to study visual information processing in the form of eye movements and gain insight into the perceptual qualities. By applying eye tracking device to evaluate the relation between human and design, that eye movement, including gaze, fixation and saccade, is controlled by ongoing cognitive processing, so that it is possible to analyze human behavior by examining eye movement. This assumption has been validated by many previous researchers in reading (Carroll and Slowiaczek 1986; Just and Carpenter 1980; Rayner and Pollatsek 1992; Rayner et al. 1989), cognitive tasks (Ahlstrom and Friedman-Berg 2006; Salvucci and Goldberg 2000), information processing tasks (Rayner 1998), scanning behavior (Allsop and Gray 2014), interface evaluation (Goldberg and Kotval 1999), and Human Computer Interaction (Yu et al. 2014). Pilots have to manually go through the QRH and identify the relevant checklists for current situation then act accordingly. It might increase respond time for searching suitable information hence degrade crew performance for emergent events. By examining the visual characteristic of the integration CAS and QRH, it might speed up pilots' reaction time under urgent situations. Therefore, this research will evaluate the effectiveness of proposed new cockpit design and whether or not it will impact to pilots' performance accordingly. However, Yu et al. (2014) argued that most eye tracking experiments are performed in the laboratory and restrict subjects' head and body motion, which differs from the naturalistic setting and limits the application. In this research, a portable glass-like eye tracker is introduced to counter these effects, which allows subjects perform as usual.

2 Method

2.1 Participants

The study involved twenty-four participants (ages range from 21 to 50 years old, M = 27.5, SD = 6.9) consisting of airlines pilots, PPL pilots, Engineers in aviation industry and Professionals in aviation domain (flight experience between 0 and 3,000 h, M = 154.1, SD = 608.3). All participants are reported normal or corrected vision with contact lens. As data was gathered from human participants, a research proposal was submitted to the Cranfield University Research Ethics System (CURES) for ethical approval. Ethical approval was granted for the research prior to starting the experiment by the CURES (CURES ID: 1773).

2.2 Apparatus

2.2.1 Simulator Set Up

The experiment was conducted in a controlled environment (lighting condition and disturb-avoidance), which contained one laptop with Microsoft Windows OS (Operating System) for Microsoft Flight Simulator X (FSX), one joystick and one control lever to assist FSX, one laptop with MAC OS to record eye movement data, one display screen and one laptop for playing the videos of emergent scenarios, and one IPAD to serve as electronic version of Boeing QRH (Fig. 2a).

Fig. 2. (a) Image of Flight Simulator, (b) Pupil Pro eye tracker

2.2.2 Eye Tracking Device

To capture eye movement data while participants were watching the videos of emergent scenarios, a Pupil Labs “Pupil Pro” eye tracking device was used, which carry one Eye Camera with 640 × 480 at 60 fps resolution and one World Camera with 1280 × 720 at 60 fps resolution. Unlike conventional eye tracking devices, the Pupil Pro is a glasses-like eye tracker which is very portable and easy to be configured to multiple test environments. The data can be record and process in a computer (an Apple Mac Book Pro Laptop was used for this experiment) via USB interface. This device has two cameras, which will be synchronized after calibration. The ‘World Camera’ is mounted on the right top of glasses showing the orientation and view of the wearer’s head. The World Camera used in this experiment is a high speed camera with 100-degree diagonal lens. A second camera, the Eye Camera, is mounted offset right and low which is adjustable to suit different wearer’s facial layout and track their pupil data accordingly (Fig. 2b).

2.3 Research Design

To design the emergent scenarios for experiment, FSX software was used to record 4 flight segments of the same flight path with fixed cockpit view by 2 different design layouts for each scenario which are current design and integrated design. In current design, instructions for each emergent event are located in QRH by electronic form of IPAD. In integrated design, the same instructions are listed directly on EICAS display under the CAS message (Fig. 3a). The instructions in both electronic versions of QRH and EICAS display are modified to the same context in this experiment. These 4 scenarios as following.

2.3.1 Left Engine Fire at 46 s

After the aircraft takes off from Heathrow airport, it takes a gentle turn heading 137 and initiates a climb to follow the flight path to the waypoint of GURLU with autopilot

(a) (b)

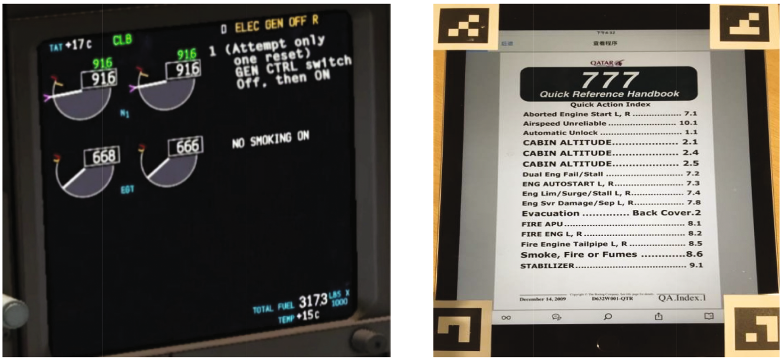

Fig. 3. (a) Integrated design of CAS and QRH, (b) Electric form of QRH

engaged. When it reaches 4,200 feet at 254 knots, red warning CAS message, FIRE ENG L, appears on the EICAS display with an audio alert and a red warning light.

2.3.2 Cabin Altitude Warning Occurs at 61 s

The aircraft is in climb phase, heading 137, on the flight path to the waypoint of GURLU with autopilot engaged. When it reaches 14,720 feet with airspeed of 320 knots, red warning CAS message, CABIN ALTITUDE, appears on the EICAS display with an audio alert and a red warning light. The emergency event is caused by left cabin door malfunction.

2.3.3 Right Engine is Shut Down at 26 s

The aircraft is in cruise phase, heading 137, airspeed 318 knots, altitude 10,000 feet, on the flight path to the waypoint of GURLU with autopilot engaged. After a sudden engine noise decreased, engine indication of N1 and EGT drop dramatically, then amber caution CAS message, ENG FAIL R, appears on the EICAS display with an audio alert and an amber caution light.

2.3.4 Right Engine Generator Fails at 40 s

The aircraft is in climb phase, heading 137, on the flight path to the waypoint of GURLU with autopilot engaged. When it reaches 3,800 feet with airspeed of 251 knots, firstly, amber caution CAS message, FUEL PUMP CENTER R, appears on the EICAS display, shortly after, amber caution CAS message, ELEC GEN OFF R, appears on the EICAS display with NO audio alert and NO caution light.

2.4 Data Collection and Analysis Processes

All participants were provided an instruction sheet of the experiment design for briefing in the same way. Then a detailed demonstration of the task with 2 screenshot pictures of examples of emergent scenarios were performed, which would assist participants acquire the basic knowledge regarding current design and integrated design, the color of the CAS message, what to do after see the CAS message by call out, then find the instruction for the CAS message in QRH or under the CAS message; read the instructions, then go back monitoring PFD until the scenario is ended. Before calibrate eye tracker, participant was asked to sit comfortably in order to maintain the same seating position during the whole experiment process. After that, the world camera view was checked via asking the participant to look at the screen and look at the IPAD, to ensure the world camera would capture all the components. Afterwards, the eye camera view was also manipulated through adjusting the eye camera mount to keep the right pupil of the participant within the correct view (Fig. 4). Eleven points of calibration which cover different space of the outside views were conducted. Following the calibration, the participant performed the experiment by monitoring 8 developing scenarios in the random sequence. Random table was used to reduce practise effect and random error due to individual differences (Mitchell and Jolley 2012).

Fig. 4. Eye camera adjustment view and manual calibration process

All participants' eye movement data are analyzed in the same 30-seconds of time period, which starts from the emergency event occurs and ends 30 s after. This time period contains the most significant eye movement data which can reflect to cognitive processes of participants interacted with non-normal flight operations. The data analysis process is divided into 4 sectors: (1) define Area of Interests (AOIs); (2) select time period; (3) export and organize data; and (4) analyze data. AOIs are the areas which contain critical eye movement information for analysis. In this case, there are three defined AOIs including, (1) CAS message area on EICAS; (2) Instructions on QRH; and (3) Instructions under CAS. There are two independent variables in this research. The first one is the design layout of EICAS, whether it is traditional design which instructions for emergent situations should be found in QRH, or integrated design which instructions are integrated with the CAS message. The second independent variable is four different scenarios. The dependent variables are time of first fixation placed on CAS message (T1), time of first fixation placed on instructions (T2), Task Completion Time (T3), percentage of fixation on instruction AOI, mean fixation duration (MFD) on instruction AOI, mean saccade amplitude (MSA) on instruction AOI, number of fixation on different displays and perceived mental Workload (PMW). For each data, test of normality is conducted, which determine whether the sample can represent general population. T1, T2, T3, percentage of fixation, MFD, MSA and PMW will be analyzed against two design layouts of EICAS and four flight scenarios, using two-way repeated measure Analysis of Variance (ANOVA).

3 Results and Discussions

3.1 Response Time of First Fixation on CAS Message

Response time is analyzed based on the information process of participants during the task performance, which starts from noticing the emergent event (time of first fixation on CAS message, T1) to locating solution for the event (time of first fixation on instructions, T2), then finally time of completing the task (T3). Mauchly's test shows that the assumption of sphericity is met for the effect of scenario, $\chi^2(5) = 4.34$, $p > 0.05$. For the interaction between design and scenario, the assumption of sphericity is violated, $\chi2(5) = 11.06$, $p < 0.05$, therefore degrees of freedom were corrected by using Huynh-Feldt estimated of sphericity ($\varepsilon = 0.82$) (Field 2013). The result indicated no significant main effect of design layouts, F (1, 23) = 0.105, p = 0.748, $\eta_p^2 = 0.005$, suggesting no significant difference between two design layouts on the time of first fixation on CAS message. However, there is a significant main effect of scenario, F (3, 69) = 1.639, $p < 0.001$, $\eta_p^2 = 0.214$. Further post-hoc comparisons showed that participant's first fixation of scenario 3 (M = 0.57) is significant faster than that of scenario 2 (M = 0.93) and scenario 4 (M = 0.96). There is no significant interaction between the design and scenario, F (2.76, 63.40) = 0.192, p = 0.887.

3.2 Response Time of First Fixation on Instruction

There is a significant main effect of design, F (1, 23) = 143.00, $p < 0.001$, $\eta_p^2 = 0.861$. The time of first fixation on instruction AOI of integrated design (M = 6.02) is significant less than that of the traditional design (M = 17.14). There is a significant main effect of scenario, F (2.16, 49.63) = 47.81, p < 0.001, $\eta_p^2 = 0.491$. Further post-hoc comparisons showed that participant's first fixation of scenario 1 (M = 7.18) is significant less than scenario 3 (M = 14.64) and scenario 4 (M = 15.87). Participant's first fixation of scenario 2 (M = 8.62, SD = 7.55) is significant less than scenario 3 (M = 14.64) and scenario 4 (M = 15.87). The effect suggests that time of first fixation on instruction is different among 4 scenarios. There is a significant interaction between designs and scenarios, F (3, 69) = 22.11, $p < 0.001$, $\eta_p^2 = 0.243$. The Fig. 5a shows the average T2 of different scenarios for each design. Generally speaking, participants spent less time finding the instructions for emergent events (T2) in integrated design than in traditional design. As the level of scenarios increases, average T2 increases as well, which might suggest that the instructions for emergent events in latter scenarios are more difficult to be found than former ones. The reason for this might be that Engine Generator Fail and Engine Shut Down are less experienced or trained by crews than Engine Fire and Cabin Altitude, hence it takes flight crews more time to locate the instruction and find the solutions to solve the emergent situations.

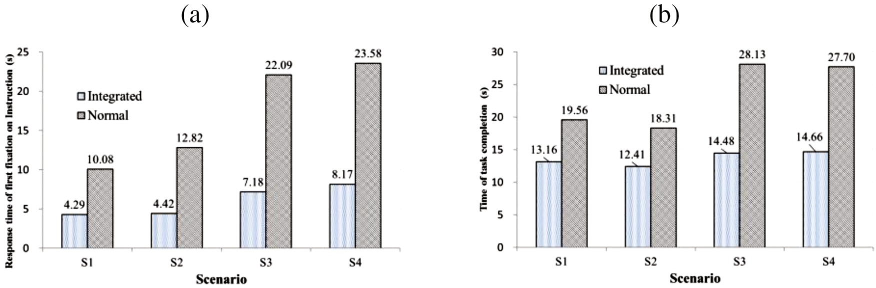

Fig. 5. (a) T2 of 4 scenarios in 2 design layouts, (b) T3 of 4 scenarios in 2 design layouts

3.3 Time of Task Completion

The result of within-subject AVOVA for T3 suggests that there is a significant main effect of design, F (1, 23) = 175.40, $p < 0.001$, $\eta_p^2 = 0.884$. The completion time on integrated design (M = 13.68) is significant less than the traditional design (M = 23.42). There is a significant main effect of scenario, F (2.17, 50.00) = 19.66, $p < 0.001$, $\eta_p^2 = 0.461$. The effect suggests that time of completing the task is different among 4 scenarios. Further post-hoc comparisons showed that participant's first fixation of scenario 1 (M = 16.36) is significant less than scenario 3 (M = 21.30) and scenario 4 (M = 21.18). Participant's first fixation of scenario 2 (M = 15.36) is significant less than scenario 3 (M = 21.30) and scenario 4 (M = 21.18). Furthermore, there is a significant interaction between designs and scenarios, F (3, 69) = 8.97, $p < 0.001$, $\eta_p^2 = 0.281$. It seems that participants spent less time completing the task (T3) in integrated design than in traditional one. It might provide preliminary evidence that integrated design is more efficient for participants to deal with emergency. As the level of scenarios increases, average T3 increases in general, however, there is a slight decrease of average T3 in Scenario 2. These results suggest that the tasks in latter scenarios (4 & 3) are more complicated than former ones (2 & 1), with Scenario 2 to be the easiest one (Fig. 5b). The reason for this might be that the solutions for Cabin Altitude and Engine Fire are curter than others, as the situation is very urgent and demand quick response from flight crews.

The time of first fixation on instruction is a key aspect for the proposed integrated design, because it defines how efficient it is for users to locate the solution, which would provide more time for them to execute the instructions or gain extra time to conduct decision-making and problem-solving processes, in terms of improving SA. From the results of Within-Subjects ANOVA and T-test for T2 and, Fig. 5a, it is obvious that no matter what emergent scenario it is, participants spent less time finding the intended instruction and the difference between them can extend to more than 10 s. In emergent situations, every second is valuable, as it might lead to very different consequences, hence integrated design could be considered as a better design than traditional one. The significant main effect for scenarios might be contributed to the differences in contents of each instruction, in terms of facilitating perception, comprehension and projection for future status. The integrated design in this research is based on Proximity Compatibility Principle (Wickens and Carswell 1995) to design a warning system. There is of little use of alerting design if it is not efficient at disengaging attention from current task (Dehais et al. 2011). Thus, time of first fixation to be placed on CAS message is a very important factor to examine whether the alerts in the design is salient enough. Human operators' visual attention would be influenced by features in the operational environment (Carmichael et al. 2010), hence the result of their visual behavior will become more valid in evaluating the effectiveness of the design. As for the significant main effect of different scenarios, it might be the reason that different scenario has different form in presenting the alert, from color of the text, sound effect, to the content of the alert itself, so it might affect the time for participants to identify. Future study could examine this effect.

Task Completion Time is another key factor for evaluating a design, as the reason similar to T2, it defines how quickly users can achieve their goals. The results derived from T3 reveal similar effect, as participants took longer time finishing the task in current design. The significant main effect for scenario of completing the task suggests that the scenario settings are different, which is an indication that different emergent events pre-programmed are not common, thus provide different levels for this experiment. In integrated design, participants never made error in locating the instructions. It is due to the obvious fact that the instructions of integrated design will present itself, hence decrease the chance of executing the wrong NNC by eliminating the degree of choices.

4 Conclusion

Restructuring CAS message and QRH can facilitate cognitive function which will be able to reduce the cost of visual search by integrated them as an oracle to advise crews on what they must do next, or where particular information is to be found. Icons, objects, texts and emergent structure in CAS and QRH are not incidental to crews' cognition, but part of their thinking process to achieve cognitive goals, and space is an invaluable resource that can be exploited and managed (Hollan et al. 2000). A fundamental way to achieve this is by re-organizing and restructuring messages presentation to facilitate perception, hence making it easier to find relevant items (Kirsh 1995). The integrated design in this research is based on Proximity Compatibility Principle (Wickens and Carswell 1995) to design a warning system. Based on the analysis of collected eye movement data, it can be concluded that integrated design is no difference than current design in alert detecting time, which means the integrated design is as good as the current design as the warning system in current one meets regulations. Furthermore, the integrated design is significant quicker than current design on both finding the solutions and task completion time.

References

Ahlstrom, U., Friedman-Berg, F.J.: Using eye movement activity as a correlate of cognitive workload. Int. J. Ind. Ergon. **36**(7), 623–636 (2006). doi:10.1016/j.ergon.2006.04.002

Allsop, J., Gray, R.: Flying under pressure: effects of anxiety on attention and gaze behavior in aviation. J. Appl. Res. Memory Cogn. **3**(2), 63–71 (2014). doi:10.1016/j.jarmac.2014.04.010

Boorman, D.: Today's electronic checklists reduce likelihood of crew errors and help prevent mishaps. ICAO J. **56**(1), 17–20 (2001)

Burian, B.K.: Emergency and abnormal checklist design factors influencing flight crew response: a case study. In: Paper presented at the Proceedings of the International Conference on Human-Computer Interaction in Aeronautics 2004 (2004)

Burian, B.K.: Aeronautical emergency and abnormal checklists: expectations and realities. in: Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting (2006)

Carmichael, A., Larson, A., Gire, E., Loschky, L., Rebello, N.S., Singh, C., Rebello, S.: How does visual attention differ between experts and novices on physics problems? In: Paper Presented at the AIP Conference Proceedings (2010)

Carroll, P., Slowiaczek, M.L.: Constraints on semantic priming in reading: a fixation time analysis. Mem. Cogn. **14**(6), 509–522 (1986)

Dehais, F., Causse, M., Tremblay, S.: Mitigation of conflicts with automation: use of cognitive countermeasures. Hum. Factors **53**(5), 448–460 (2011). doi:10.1177/0018720811418635

Federal Aviation Administration: Operational use & modification of electronic checklists. Fed. Aviat. Adm. Advisory Circular **120**(64), 1–9 (1996)

Field, A.: Discovering Statistics Using IBM SPSS Statistics: and Sex and Drugs and Rock 'n' Roll, 4th edn. Sage Publications Ltd (2013)

Goldberg, J.H., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. Int. J. Indus. Ergon. **24**(6), 631–645 (1999)

Hollan, J., Hutchins, E., Kirsh, D.: Distributed cognition: toward a new foundation for human-computer interaction research. ACM Trans. Comput. Hum. Interact. (TOCHI) **7**(2), 174–196 (2000)

Jones, D.G., Endsley, M.R.: Sources of situation awareness errors in aviation. Aviat. Space Environ. Med. **67**(6), 507–512 (1996)

Just, M.A., Carpenter, P.A.: A theory of reading: from eye fixations to comprehension. Psychol. Rev. **87**(4), 329 (1980)

Kirsh, D.: The intelligent use of space. Artif. Intell. **73**(1–2), 31–68 (1995)

Mitchell, M.L., Jolley, J.M.: Research Design Explained. Cengage Learning, Boston (2012)

Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychol. Bull. **124**(3), 372 (1998)

Rayner, K., Pollatsek, A.: Eye movements and scene perception. Can. J. Psychol. (Revue canadienne de psychologie) **46**(3), 342 (1992)

Rayner, K., Sereno, S.C., Morris, R.K., Schmauder, A.R., Clifton Jr., C.: Eye movements and on-line language comprehension processes. Lang. Cogn. Process. **4**(3–4), SI21–SI49 (1989)

Robinski, M., Stein, M.: Tracking visual scanning techniques in training simulation for helicopter landing. J. Eye Mov. Res. **6**(2), 1–17 (2013)

Salvucci, D.D., Goldberg, J.H.: Identifying fixations and saccades in eye-tracking protocols. In: Paper Presented at the Proceedings of the 2000 Symposium on Eye Tracking Research & Applications (2000)

UK Civil Aviation.: AP 676: A Guidance on the Design, Presentation and Use of Emergency and Abnormal Checklists. CAA, London (2006)

Wickens, C.D., Carswell, C.M.: The proximity compatibility principle: its psychological foundation and relevance to display design. Hum. Factors **37**(3), 473–494 (1995)

Yu, C., Wang, E.M., Li, W.-C., Braithwaite, G.: Pilots' visual scan patterns and situation awareness in flight operations. Aviat. Space Environ. Med. **85**(7), 708–714 (2014)

http://www.springer.com/978-3-319-58471-3

Engineering Psychology and Cognitive Ergonomics: Performance, Emotion and Situation Awareness
14th International Conference, EPCE 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part I

Harris, D. (Ed.)

2017, XXI, 419 p. 152 illus., Softcover
ISBN: 978-3-319-58471-3