

The Investigation of Visual Attention and Workload by Experts and Novices in the Cockpit

*Wen-Chin Li¹, Fa-Chung Chiu², Ying-shin Kuo³, Ka-Jay Wu⁴

¹ Air Transport Department, Cranfield University, United Kingdom

^{2,3&4} Psychology Department, National Defense University,

Beitou District, Taipei City 112, Taiwan

Tel: +886-2-28962979 ext. 11, Fax: +886-2-28964426

Wen-Chin Li: w.li.2002@cranfield.ac.uk

Abstract. Under high pressure of flight mission and dynamic aircraft maneuvers in the tactic missions, pilot faces additional difficulties and increased mental workload. Workload could increase the error of flight operation, decrease efficiency of pilot's decision-making. Experts had significantly shorter dwells, more total fixations, more aim point and airspeed fixations and fewer altimeter fixations than novices, experts were also found to have better defined eye-scanning patterns. This research applies the eye-tracking technology for analyzing visual attention, emWave-2 for measuring physiological coherence, and NSAS-TLX for investigating subjective cognitive efforts. The participants of this research consisted of 41 fighter pilots. The present study is applying new technology to understand the pilots' workload and visual attention in the cockpit for conducting a simulated air-to-air tactic operation. There is a raising need for further research regarding mental workload and stress management program for real-time flight operations.

Keywords: Cognitive Processes, Eye Movement, Visual Attention, Workload

1. Introduction

The natural limitation of human being's cognitive processes and the vast number of tasks are reasons for increasing critical workload levels for military pilots. Under high pressure of tactic missions and dynamic aircraft maneuvers, the pilot faces additional difficulties and increased workload during hostile environmental (Ahlstrom, 2003). Workload can negatively affect operator performance and increase the probability of operational hazards. Peter, Jennifer and Joey (2001) found that experts had significantly shorter dwells, more total fixations, more aim point and airspeed fixations and fewer altimeter fixations than novices. Experts were also found to have better defined eye-scanning patterns. The visual scanning difference between novice and experts are correlated with better performance by experts, as experts should have shorter dwell and more fixations than novices on all the instruments (Bellenkes, Wickens, & Kramer, 1997).

Fox, Merwin, Marsh, McConkie, & Kramer (1996) found that experts are more likely to use peripheral vision to process a broader range of visual cues than novices do. The pattern of acquisition of cue-based information provides an opportunity to assess the application of distinct cognitive skills. Eye movements are useful to reveal the diagnostic information that enables the development of appropriate strategies which efficiently target a particular feature of the performance for completing a task (Rayner, 1998). However, deliberate missions result in higher workload, such as controlling aircraft for tactic maneuver is a stressful operation which needs high situation awareness to make risk assessment, might increase pilot's mental workload (Schipani, 2003). Pilots have to make in-flight decisions not only about the management of the systems in the cockpit, but also about the operating

state of that airspace in dynamic situations, the workload may be increased dramatically during abnormal situations and system failures (Weiner, 1989).

The NASA Task Load Index (TLX) is a popular technique for measuring subjective mental workload. It relies on a multidimensional construct to derive an overall workload score based on a weighted average of ratings on six subscales (Hart & Staveland, 1988). The scales range from 0 to 100 and were divided into 20 increments. By analyzing the six weighted scores to understand which dimension has higher degree of workload. Aside from using self-reported subjective workload ratings such as NASA-TLX for evaluating operator's workload, pupillary response has also been proposed as an index of the amount of cognitive processing (Beatty, 1982). Eye movement measurement offers deep insights into human-machine interaction and the mental processes of pilots. Measurements based on different aspects of ocular behavior, such as the number of fixations, dwell time, and the dilation of pupil, have been used to reveal the status of mental workload. There are existing diversity findings of eye movement in the literature. Beatty (1982) reviewed a large amount of experimental data and concluded that increasing in pupil size correlate with increasing in workload. However, some experiments have not detected a relationship between task difficulty and pupil dilation (Lin, Zhang, & Watson, 2003). There was evidence that increasing in workload could increase dwell time and the frequency of long fixations (Van Orden, Limbert, Makeig, & Jung, 2001). Athènes, Averty, Puechmorel, Delahaye and Collet (2002) found that workload could increase the error of operation, and the decreased fixation duration appeared to predict upcoming errors in the auditory task (Tsai, Viirre, Strychacz, Chase, & Jung, 2007).

Workload has impact to cognitive processing and creating psychological and physical stress. The deleterious effects of stress on human beings' performance are pervasive, as stress may result in physiological changes such as emotional reactions such as fear, anxiety, frustration, and cognitive effects such as narrowed attention, longer reaction time to peripheral cues and decreased vigilance, and degraded problem solving (McCraty, Atkinson, Tomasino, & Bradley, 2009). Training programs, often called stress management programs (SMPs), have been developed, primarily by clinical psychologists, for this expressed purpose. Among the most common procedures are progressive muscle relaxation, meditation, biofeedback, and cognitive-behavioral skills training. Human beings cognitive processes and performance have been the subjects of much concentration in research and practice including situation awareness, decision-making, mental workload, and operational performance. However, there are lots of argument for finding strong empirical evidence and lacking scientific status for military pilots. Therefore, this research applied the eye-tracking technology, emWave-2 and NASA-TLX to investigate objective and subjective cognitive effort involved in attending to information presented in stressful situation for military pilots. By examining pilots' eye movement's patterns and performance compared with pilots' physical coherence levels, and subjective stress levels, it is hope to discover the appropriate interventions in flight operations for improving the training effectiveness and aviation safety.

2. Method

2.1 Participants

There were 41 fighter pilots participated in this research. The flying hours is between 335 and 3200 hours, the rank is between first lieutenant and colonel, the age is between 25 and 45 years old.

2.2 Apparatus

2.2.1 Eye-tracker

Eye movement data were collected by an head-mounted ASL (Applied Science Laboratory) Mobile Eye with 76 gram in weight. When combined with an optional head-tracking device and eye/head integration software, the eye tracker can also measure pilot's eye line of gaze with respect to stationary surfaces in the environment. It is designed to be durable under a variety of active applications and its light weight is suitable to detect the eye movement when pilot operates the aircraft under active and dynamic flight tasks.

2.2.2 Flight Simulators

This study used fighter simulator, a dynamic high fidelity trainer that replicates actual aircraft performance, navigation and weapon systems. This simulator provides a realistic representation of the flight management system. The instructors can supervise the participated pilot's performance and the instrument data from three screens. In instrument flight task, the integrate control panel (ICP) 、 data entry display 、 horizontal situation display 、 head-up display (HUD) , Left and Right Multi-display provide critical information for tactic mission. Pilots cross check those instrument to maintain the speed 、 altitude 、 heading and position. Therefore, this study set those six gauges as the area of interests (AOIs) to analyze the eye movement data. The scenario of flight simulator is Air-to-Air Combat.

2.2.3 NASA-TLX

The NASA Task Load Index (TLX) is a popular technique for measuring subjective mental workload. It relies on a multidimensional construct to derive an overall workload score based on a weighted average of ratings on six subscales: (1) Mental demand: How much mental demand and perceptual activities you would use; (2) Physical demand: How much is the degree of physical demand; (3) Temporal demand: How much is the degree of time pressure; (4) Performance: How do you feel about the flying time and the performance in flight? (5) Effort: How much difficult do you think? (6) Frustration: How much frustration and disappointment do you feel (Hart & Staveland,1988).

2.2.4 EmWave-2

Physiological coherence is measured by HRV (heart rate variability) analysis, which reflect heart-brain interactions and autonomous nervous systems (ANS) dynamics and is reflected in a heart rhythm pattern. The Coherence Ratio bar chart is divided into three levels: Low, Medium and High coherence. The positive emotions such as appreciation and compassion, as opposed to negative emotions such as anxiety, anger, and fear, are reflected in a heart rhythm pattern that is more coherent. The coherent state has been correlated with a general sense of well-being, and improvements in cognitive, social, and

physical performance. Heart rhythm coherence is indeed associated with significant improvements in cognitive performance. EmWave-2 is used in many areas including military operational personnel and Olympic athlete, it has a significant improvement on stress management (McCraty, Atkinson, Tomasino, & Bradley, 2009).

2.3 Procedure

Participants were asked to perform Air-to-Air Combat using the flight simulator, the procedure included: (1) an orientation to the experiment (10 minutes); (2) eye-tracker calibration in the cockpit of flight simulator (5-10 minutes); (3) operate Air-to-Air Combat task on flight simulator for collecting eye movement data by eye-tracker, and coherence level by emWave-2 (3-5 minutes); (4) rate NASA TLX (10-15minutes). Each session was conducted by an eye-tracker operator and a flight instructor. The instructor evaluates pilot's performance base on the figures of flight simulator, and the simulator control panels record the time for terminating target and weapon consumption. Eye movement patterns, video, verbal protocol data were collected for further analysis.

3. Results and Discussion

3.1 Sample Characteristics

Participants consisted of 41 fighter pilots from R.O.C. air force. The qualifications for flying experience were not combat ready, combat ready, 2-aircraft team leader, 4-aircraft team leader, daytime back seat instructor, night back seat instructor, and training instructor; flying hours between 335 and 3200 hours; the age of participants between 25 and 45 years old (Table 1).

Table1 Demographical Variables of Participants (*N=41*)

Variables	Group	Frequencies (%)
Age	25 – 30	14(34.1%)
	31 – 35	11(26.8%)
	36 – 40	9(22%)
	41 – 45	7(17.1%)
Rank	First lieutenant	1(2.4%)
	Captain	17(41.5%)
	Major	9(22%)
	Lieutenant Colonel	13(31.7%)
	Above colonel	1(2.4%)
Qualification	Not combat ready	2(4.9%)
	Combat ready	12(29.3%)
	Two aircraft team leader	4(9.8%)
	Four aircraft team leader	9(22%)
	Daytime back seat instructor	3(7.3%)
	Night back seat instructor	0(0%)
	Training instructors	11(26.8%)
Flying hours	Under 1000	17(41.5%)
	Above 1000	24(58.5%)

3.2 The Different Workload level by NASA-TLX and Coherence Scores

There were two groups of pilots by operational qualifications, junior pilots including not combat ready and combat ready pilots; and senior pilots including team leaders, back-seat instructors, and training instructor. The workload scores of NASA-TLX and emWave-2 (coherence) during Air-to-Air Combat by t-test shown as table 2. There were two dimensions of NASA-TLX, Performance ($p < .05$) and Frustration ($p < .05$), with significant difference between junior and senior pilots. Also, senior pilots showed significant higher coherence levels than junior pilots by emWave-2 ($p < .001$).

Results show that senior pilots have higher performance ($M=8.80$ vs $M=5.45$) and coherence levels ($M=27.71$ vs $M=15.22$), and lower frustration than junior pilots ($M=7.41$ vs $M=10.45$). The coherent state has been correlated with a general sense of well-being, and improvements in cognitive, social, and physical performance. Senior pilots have more experience of tactic training, competent of system integrations for flight operations, and mature maneuver skills than junior pilots. The findings of high performance and low frustration by NASA-TLX, and high physical coherence of heart rhythm by emWave-2 for senior pilots in this study, have coincided with both laboratory and natural settings (Tiller, McCraty, & Atkinson, 1996). There is abundant evidence that emotions alter the activity of the body's physiological systems and that beyond their pleasant subjective feeling, heartfelt positive emotions and attitudes provide a number of benefits that enhance physiological, psychological, and social functioning (McCraty, Atkinson, Tiller, Rein, & Watkins, 1995). Zakowski, Hall and Baum (1992) concluded that stress management is a valuable weapon not only against temporary interference with performance attributable to stress, but also against certain types of disease and physical illnesses and against the progression of diseases already established. Furthermore, stress management training could bring individuals' ability to diminish stress and increase their adaptation based on a stressful situation (Linden, 2004). The different approaches between objective measurement of emWave-2 and subjective rating scale of NASA-TLX show senior pilots have better performance and physical coherence for coping with stress. There is a raising need for developing stress management interventions for military pilots for improving aviation safety.

Table 2 The Different Workload level by NASA-TLX and emWave-2

<i>Dimensions</i>	<i>Means (SD)</i>		<i>t</i>	<i>D</i>	<i>p</i>
	Seniors	Juniors			
Mental Demand	10.13 (7.48)	11.23 (9.05)	.55	.13	.59
Physical Demand	15.64 (9.95)	16.32 (8.44)	.33	.07	.74
Temporal Demand	12.70 (8.18)	10.20 (8.13)	1.20	.31	.23
Performance	8.80 (7.21)	5.45 (6.03)	2.25	.50	.03
Effort	11.45 (6.86)	13.55 (6.94)	1.37	.30	.18
Frustration	7.41 (6.34)	10.45 (6.22)	2.18	.48	.03
Coherence	27.71 (5.53)	15.22 (7.80)	8.47	1.85	.001

3.3 The Percentage of Gaze in Region to Total Gaze for Different Area of Interests (AOIs)

The differences between senior and junior pilots for percent gaze in region to total gaze at different area of interests (AOIs) by t-test (table 3). The results show that senior pilots have significant more gaze on Left Multi-display than junior pilots during Air-to Air combat ($p < .01$). The result is difference compared with the research conducted by Peter, Jennifer and Joey (2001) which proposal that experts had significantly shorter dwells, more total fixations, more aim point and airspeed fixations and fewer altimeter fixations than novices. Experts should have shorter dwell and more fixations than novices on all the instruments. Pilots have better performance as experience pilots aware how to pay attention to certain instrument (Left Multi-display) at the critical timing for tactic maneuver.

Table 3 The percentage of gaze in region to total gaze for AOIs

<i>Dimensions</i>	<i>Means (SD)</i>		<i>t</i>	<i>D</i>	<i>p</i>
	Seniors	Juniors			
Head-up Display	53.28 (19.23)	50.78 (21.48)	.56	.12	.58
Integration Control Panel	1.35 (1.31)	1.76 (2.51)	.94	.20	.35
Data Entry Display	.18 (.38)	.12 (2.6)	.87	.03	.39
Right Multi-display	.67 (1.23)	.63 (1.50)	.13	.03	.90
Left Multi-display	3.37 (5.22)	.91 (1.85)	2.72	.63	.01
Horizontal Situation Display	.08 (.22)	.11 (.45)	.52	.08	.61

3.4 Average Pupil Diameter on AOIs

There were significant differences between senior pilots and junior pilots on the Average Pupil Diameter at Left Multi-display by t-test ($p < .05$). The data show that senior pilots' Average Pupil Diameter significant bigger than junior pilots at Left Multi-display ($M=73.69$ vs $M=55.60$) (Table 5). The results demonstrated that senior pilots' pupil diameter dilated significantly than junior pilots at Left Multi-display, it is for the collecting target's information from tactic radar. It is the focus of pilots' attention and getting situational awareness during Air-to-Air mission. Hilburn, Bakker, Pekela, & Parasuraman (1997) found an increase in pupil diameter when subjects performed well on the auditory task. The results also found senior pilots have better performance than junior pilots.

Table 5 Average pupil diameter on different AOIs

<i>Dimensions</i>	<i>Means (SD)</i>		<i>t</i>	<i>D</i>	<i>p</i>
	Seniors	Juniors			
Head-up Display	86.67 (11.42)	89.50 (10.92)	1.14	.25	.26
Integration Control Panel	81.28 (24.93)	83.79 (23.00)	.47	.10	.64
Data Entry Display	33.44 (42.14)	27.19 (40.56)	.68	.15	.50

Right Multi-display	46.91 (45.50)	43.69 (45.95)	.32	.07	.75
Left Multi-display	73.69 (36.42)	55.60 (44.59)	2.02	.53	.04
Horizontal Display	Situation 12.84 (3.24)	11.51 (29.66)	.20	.06	.85

4. Conclusion

This research applied subjective measurement by NASA-TLX, and objective measurement by eye-tracker and emWave-2 for approaching military pilots physical coherence, mental workload and eye movement pattern for developing future training interventions. The weakness of subjective measurement is the self-report scale might be affected by social desirability and introspective limits (Perugini & Banse, 2007). Therefore, the objective measurement of workload and attention distribution by emWave-2 and eye-tracker are to improve the defect of self-report approach. The emWave-2 found that the coherence scores have significant difference between expert pilots and novice pilots. The effect size by emWave-2 is 1.85 which is much bigger than .50 by self-report of NASA-TLX. Eye-tracker is a powerful research apparatus not only for objective measurement but also for investigating the cognitive processes for task performance (Rayner, 1998). The experienced pilots have paid more attention on Left Multi-display, and the pupil diameter larger than novice pilots. Rehder and Hoffman (2005) also found that operators would spend more time on the target which attracting attention for completing task. Expert pilots know how to pay attention at critical time for getting important information for tactic operations. Eye-tracking device has already been shown to be a useful measurement for attention allocation.

Workload is an increasingly salient factor in advanced complex environment, the potential for stress-induced error spans both high-technique cockpit and ground operation for aviation domain, and effects of stress on physiological reactions, cognitions, emotions, and social behavior are manifold. This research found that pupil dilation is a valid and reliable indicator of mental workload and increase in pupil diameter when pilots performed well on the cognitive process for getting situational awareness. More experience pilots have higher physical coherence, bigger pupil dilation and more gaze on Left Multi-display for Air-to-Air mission, and have better performance and lower frustration than less experienced pilots. Although the mechanism of human perception is in the same operation, the purpose of this study is to explore the correlation among pilots' performance, workload and eye movements. By examining pilots' eye movement's patterns and physical coherence levels compared with pilots' subjective stress levels, it is hope to discover the role of cognitive effort in flight operations for improving the training effectiveness and aviation safety.

5. References

- Ahlstrom, U. (2003). Current trends in the display of aviation weather. *Journal of Air Traffic Control*, 45(3), 14-21.
- Athènes, S., Averty, P., Puechmorel, S., Delahaye, D., & Collet, C. (2002). *ATC complexity and controller workload: Trying to bridge the gap*.

- Averty, P., Collet, C., Dittmar, A., Athénes, S., & Vernet-Maury, E. (2004). Mental workload in air traffic control: an index constructed from field tests. *Aviation, Space, and Environmental Medicine*, 75, 333–341.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91, 276–292.
- Bellenkes, A. H., Wickens, C.D. & Kramer, A.F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation Space and Environmental Medicine*, 68, 569-579.
- Endsley, M. R. (1997). *Level of automation: Integrating humans and automated systems*.
- Fox, J., Merwin, D., Marsh, R., McConkie, G., & Kramer, A. (1996). Information extraction during instrument flight: An evaluation of the validity of the eye-mind hypothesis. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*.
- Hart, S. G. & Staveland, L.E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.), *Human Mental Workload*.
- Hilburn, B. G., Bakker, M. W. P., Pekela, W. D., & Parasuraman, R. (1997). *The effect of free flight on air traffic controller mental workload, monitoring and system performance*.
- Lin, Y., Zhang, W. J., & Watson, L. G. (2003). Using eye movement parameters for evaluating human-machine interface frameworks under normal control operation and fault detection situations. *International Journal of Human Computer Studies*, 59, 837–873.
- Linden, W. (2004). *Stress management: From basic science to better practice*. Thousand Oaks, USA: Sage.
- McCraty, R., Atkinson, M., Tiller, W. A., Rein, G., & Watkins, A. D. (1995). The effects of emotions on short-term power spectrum analysis of heart rate variability. *The American Journal of Cardiology*, 76(14), 1089-1093.
- McCraty, R., Atkinson, M., Tomasino, B. A., & Bradley, D. (2009). The coherent heart heart-brain interactions, psychophysiological coherence, and the emergence of system-wide order. *Integral Review*, 5(2), 10-115.
- Peter, K., Jennifer, S., & Joey, H. (2001). Comparison of expert and novice scan behaviors during VFR flight. *The 11th International Symposium on Aviation Psychology*. Columbus: The Ohio University.
- Perugini, M., Banse, R. (2007). Personality, implicit self-concept and automaticity. *European Journal of Personality*, 21, 257–261.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3), 372.
- Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology*, 51, 1–41.
- Schipani, S. P. (2003). *An evaluation of operator workload, during partially-autonomous vehicle operations*: DTIC Document.
- Tiller, W. A., McCraty, R., & Atkinson, M. (1996). Cardiac coherence: A new, noninvasive measure of autonomic nervous system order. *Alternative Therapies in Health and Medicine*, 2, 52–65.
- Tsai, Y. F., Viirre, E., Strychacz, C., Chase, B., & Jung, T. P. (2007). Task performance and eye activity: predicting behavior relating to cognitive workload. *Aviation, Space, and Environmental Medicine*, 78(1), 176-185.
- Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T. P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(1), 111-121.
- Wickens, C. D. (2002). Situation awareness and workload in aviation. *Psychological Science*, 11(4), 128–133.
- Wiener, E. L (1989). *Human Factors of Advanced Technology (“Glass Cockpit”) Transport Aircraft*. (NASA Contractor Report 177528). Moffett Field, CA: NASA Ames Research Center.
- Zakowski, S., Hall, M. H., & Baum, A. (1992). Stress, stress management, and the immune system. *Applied and Preventive Psychology*, 1, 1-13.