Measuring the Situation Awareness of Tower Controllers by Using Eye Movement Analysis

Guo Xinyao[1], Liu Yawei[1,*], Si Qingmin[1], Niu Linqing[1] and Jin Huibin[2]

[1]*School of Civil Aviation, Zhengzhou University of Aeronautics, Zhengzhou 450046, China*
[2]*The General Aviation College of CAUC, Civil Aviation University of China, Tianjin 300300, China*

_____

Abstract

The lack of situation awareness (SA) is the main reason for tower controllers' human errors. How to evaluate the SA of tower controllers is an important problem for controlling the risk of air traffic control system. The current method of measuring the SA of air traffic controllers by using eye movement analysis focuses on radar controllers. To verify the applicability of eye movement analysis in measuring the SA of tower controllers, an experiment platform was built by using the eye tracker and tower control simulation software. The comparison of the two group features including SA, task load, and operation performance, were conducted by the nearest neighbor index (NNI). Results show that the subjects whose fixation feature is dispersed distribution have high SA and improve operation performance under a low task load. The Pearson correlation coefficient between the NNI and SA reveals a significant positive moderate association (r=0.602, p<0.01). Eye tracking can measure the SA of tower controllers, and the NNI can be an auxiliary eye index for SA measurement in dynamic, uncontrolled environments. The conclusion of this study can provide a new reference index for the real-time quantitative assessment of tower controllers' situation awareness level.

Keywords: Tower controllers, Situation awareness, Eye movement, Nearest neighbor index, Fixation mode

_____

1. Introduction

Air traffic controllers (ATCos) play the core role of guaranteeing the safe, orderly, and efficient operation of air traffic control (ATC) systems. With the rapid increase of air traffic flow and the continuous improvement of the level of ATC automation, controllers must possess good situation awareness (SA) to make right decisions in a continuously enhanced dynamic system of human–machine information interaction. This condition is especially true for tower controllers, who are primarily responsible for the command and coordination work in the aircraft take-off and landing phases, which are the highest risk periods in flights where more than half of aircraft accidents occur [1, 2].

These two periods are known as the "black thirteen minutes" in the civil aviation, where good SA is of great significance in ensuring sound decision making and ATC safety. SA refers to operators' cognitive processing of target information in an evolving situation; it includes three levels, namely, perception, comprehension, and prediction [3]. SA errors by operators result in wrong cognition and decision making [4], leading to unsafe civil aviation incidents and even accidents [5, 6]. According to Aviation Safety Reporting System statistics, 38% of the unsafe incidents (including crew and ATCos) that occurred from January 1986 to May 1992 can be attributed to SA errors [7]. Jones and Endsley [8] analyzed the National Transportation Safety Board survey of 24 civil aviation accident investigation reports from 1989 to 1992 and found that up to 88% of the accidents caused by human errors were due to SA errors.

Therefore, to prevent human errors from ATCos and ensure ATC safety, an objective and sensitive approach for effectively measuring the SA of tower controllers is needed.

Considering the limitations of the current SA measurement technology, this study uses eye tracking experiments to discover new indicators for SA measurement, through a combination of qualitative and quantitative analysis, providing certain experimental theoretical support for tower controllers' SA measurement and improvement.

2. State of the Art

The method of SA measurement varies according to the research orientation and field, and many researchers adopt different classifications [9-12]. The current SA measurement methods can generally be divided into four categories: subjective techniques (e.g., self-rating and observer-rating techniques), performance measures (e.g., global, external task, and imbedded task measures), memory probe measures (e.g., posttest, freeze probe techniques, and real-time probe techniques), and physiological techniques (e.g., P300, eye tracker). These categories each have their respective shortcomings. The obvious flaw of subjective techniques is that each person's understanding of a situation is not the same, and therefore the consistency of the assessment between subjects cannot be guaranteed [13]. The biggest issue with performance measures is that good SA may be a necessary condition, but it is not a sufficient condition for good performance; thus, such measures may not reflect SA [14]. Memory probe measures are the most consistent with Endsley's definition of SA because they evaluate operators' SA through their self-report of the contents memorized. However, memory probe measures also have some

shortcomings. For example, in freeze probe techniques, the simulation task stops at a random time point and all task-related information are cleared (e.g., a blank screen). During the freeze period, the subjects are required to answer questions related to the task. The hypothesis of this approach is that the subject has already stored all the pertinent information in his/her working memory, but the fact is that he/she needs the information presented by the display to maintain SA. Furthermore, this method causes great interference in the experiment, affecting the objectivity of data collection [15]. Compared with the former three SA methods, physiological techniques are rarely used, probably because the capability of such measures to directly touch the high level of cognitive processes, such as SA, is still unknown. For eye tracking cases, several studies show that the "look-but-not-see" phenomenon exists in visual attention; thus, whether or not the person has already processed the object he/she sees is unclear [16]. Owing to the uncertainty, the use of eye movement to measure SA needs further investigation. As a typical physiological measurement method, eye tracking also has the advantages of collecting data without bias and being non-invasive to the experiment. Furthermore, it can be applied in uncontrolled environments where the other three methods will not work. The eye also provides important access to external information and has a close connection to the brain's attention and information-processing mechanisms [17-20]. Thus, it provides a new perspective for studying SA and is favored by many SA researchers. For example, in the field of pilot SA research, Merwe et al. [21] used fixation rates and dwell times as "Level 1 SA" indicators and scanning entropy as a "Level 3 SA" indicator to measure pilot SA; Yu et al. [22] conducted an eye movement otherness research on military pilots with different SA levels, including pupil sizes and average fixation duration; and through a fighter aircraft simulation experiment, Liu and Su [23] proved that the superiority of eye tracking in measuring pilots' SA lies in having more extensive differences in sensitivity. In the research on controllers' SA, Hyun et al. [24] studied the SA differences between novice and expert controllers by using eye movement measures and a thinking-aloud experiment. Hauland [25] measured the individual and team SA of controllers through their time allocation strategy in different areas of interest (AOIs).

Eye movement systems have numerous indexes, among which the geometry of the scan path is considered the most valuable index for studying the human cognitive process in complex human–machine interaction systems. Existing scholars introduced the nearest neighbor index (NNI) into recognition of subjects' fixation pattern and found that the NNI is highly sensitive to workload changes [26, 27]. NNI was soon adopted by many scholars [28-33], because it considers all the fixation points of a given situation and does not need to prioritize the AOI, thus greatly increasing the availability of eye movement in the real environment. For instance, Nocera et al. [29] used the NNI to calculate the random distribution of fixation points at the flight deck to characterize the pilot's visual glance strategy and explore its relationship with the workload. Fidopiastis et al. [30] characterized the unmanned operator's eye movement model with the NNI and studied the effects of automation and workload on eye movement patterns. Moore and Gugerty [31] introduced the NNI to analyze the aggregation and dispersion degree of the controller's fixation points on the radar interface and studied its relationship with SA and performance, arriving at the conclusion that how "operators

allocate attention (distributed widely or narrowly) affects their SA, as well as their task performance".

To measure the SA of tower controllers, understanding the function of SA in information processing is necessary and should be prioritized. Currently, the most widely accepted definition of SA is given by Endsley [3]: "SA is the perception of the 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". Specifically, for tower control work, tower controllers first apperceive diversified information from the current environment (e.g., radar display terminal, flight strip, and air–ground communication) by visual and auditory means and then perform preliminary information filtering and synthesis. Next, on the basis of their analysis of current information in their short-time working memory and in association with the tower control rules and experience in their long-term memory system, the tower controllers can understand the current traffic situation and predict the future status. Finally, tower controllers make decisions and act based on the preceding steps. These steps comprise the information processing (Figure 1) that a tower controller must experience, in which the importance of good SA has been proven by several studies [34, 35]. Furthermore, in the Next Generation Air Traffic Management System (NextGen), the interaction between the controllers and the ATC system is frequent, "so there is a need to assess how NextGen tools and concepts of operation will impact operator's SA" [36]. The current measurement techniques for SA are also far from satisfying the requirement of performing real-time, easy-to-conduct, and accurate measurements in real situations (out of the lab).

In summary, using eye movement analysis to measure ATCos' SA requires further exploration, and the NNI is a valuable eye movement index. Our objectives are to explore the applicability of eye movement analysis in measuring tower controllers' SA through a simulated tower control experiment and investigate whether the NNI can be used as an eye movement indicator for SA measurement.

The remainder of this study is organized as follows: The third section describes the participants and training, the apparatus, and the experimental task and design. The fourth section is the data processing (including the calculation of NNI, NASA-TLX, 3D-SART, and operation performance) and the analysis and discussion of the results (including the comparison analysis of SA, task load, and operation performance, and the correlation analysis between NNI and SA). The last part summarizes this article and gives relevant conclusions.

3. Methodology

3.1 Participants and training
Twenty senior students who majored in Air Traffic Management in the Civil Aviation University of China were selected as participants, including 16 males and 4 females, with an average age of 21.96 years and a standard deviation of 1.32. None of them has color weakness or color-blind, and their eyesight exceeds 1.0. All participants are familiar with all the details of tower control and can complete control tasks independently.

To ensure that each participant could use the simulated tower control software skillfully in the experiment, training was provided for a week, followed by testing.

3.2 Apparatus

The experimental platform consists of the simulated tower control software and the Eye Tribe Tracker. The interface of the simulated tower control software includes four functional areas: airspace surveillance radar, airport surveillance radar, progress strip, and command send bar. The experimental requirements of a simulated tower control can be satisfied. The size of the experimental computer's screen is 14 in, and the resolution is 1366×768 pixels, with an average brightness of 22.95 cd/m². The accuracy of the Eye Tribe Tracker is 0.5°, the sampling rate is 60 Hz, and the test range is 45–75 cm.
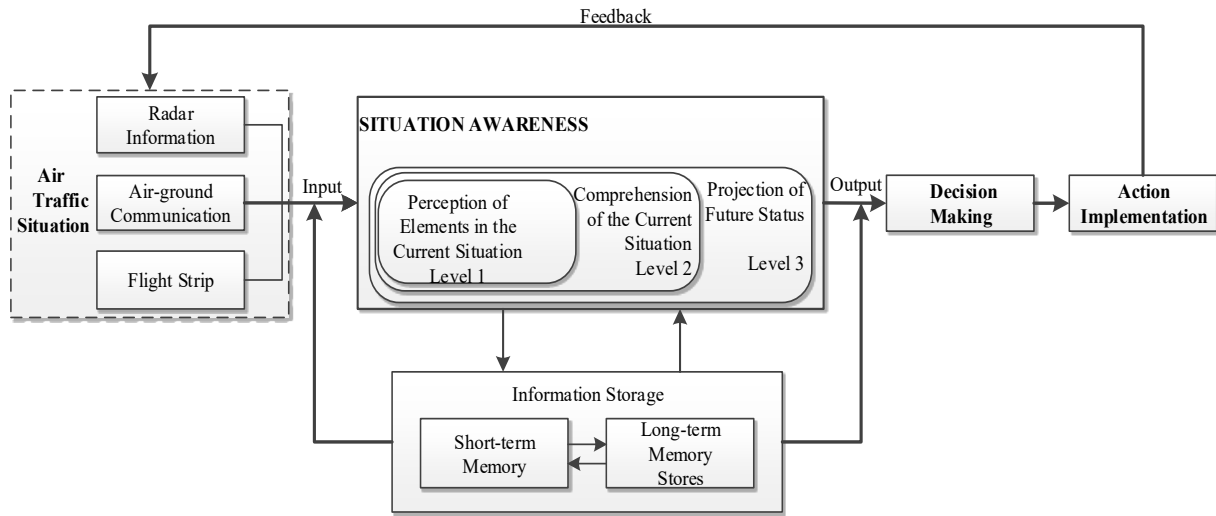

Fig. 1. Information processing of the SA by tower controllers

3.3 Task

In this experiment, with the safety interval satisfied, the participants attempted to reduce the conflict and the flight delay by controlling the entry and departure flights, including issuing landing commands for the entry flights and then directing them into the appropriate parking apron after landing; allocating the appropriate runway for departing flights and then issuing take-off commands; and transfer control to approach controllers.

3.4 Experimental design

The difficulty of simulated tower control, an independent variable in our experiment, was divided into three levels: simple, medium, and difficult. The numbers of aircraft under these three levels of task difficulty were 7, 12, and 16, respectively. The experiments for each participant lasted for 30 min.

The experimental dependent variables included fixation features, workload, SA, and operation performance. The fixation features were classified using the NNI. The workload and SA were assessed by the National Aeronautics and Space Administration Task Load Index (NASA-TLX) and the Three-Dimensional Situation Awareness Rating Technique (3D-SART) scale, respectively. Operation performance refers to the number of operation errors.

To eliminate the errors caused by the test sequence and fatigue, the experiment was conducted in accordance with the Latin square design. Before the experiment, the participants were guided on how to fill the NASA-TLX and 3D-SART scales. The detailed procedure was as follows:

1) The experimenters ran the simulated tower control software and set the control task. Next, the participants sat in front of the experimental computer screen and adjusted the seats and their postures to calibrate the Eye Tribe Tracker.

2) The individuals started the experiment, with the Eye Tribe Tracker recording the eye movement data synchronously. To ensure sufficient sampling of eye movement data, the participants had to keep their heads in the same position.

3) Finally, the participants completed the NASA-TLX and 3D-SART scales.

4. Results Analysis and Discussion

4.1 Data Processing

4.1.1 NNI algorithm

Widely used in the fields of geospatial analysis and pattern classification, NNI is a type of clustering algorithm based on the distance between points within the region. In eye movement analysis, the coordinate distance of the fixation points on the experimental interface is first calculated according to the NNI algorithm and then the NNI value can be obtained as follows:

$$NNI = \frac{D(NN)}{D(ran)} \tag{1}$$

where $D(NN)$ represents the average minimum distance between the fixation points, and $D(ran)$ represents the average distance between the randomly distributed fixation points. Specifically,

$$D(NN) = \sum_{i=1}^{n} \left[\frac{Min(d_{ij})}{n} \right], 1 \leq j \leq n, j \neq i \tag{2}$$

where $Min(d_{ij})$ is the distance between fixation points i and j, and n is the number of fixation points within the interface. In addition,

$$D(ran) = 0.5\sqrt{\frac{A}{n}} \tag{3}$$

where A is the area of the measured interface area, and n is the number of fixation points.

NNI is affected by the visual search strategy of the participants. Thus, it can characterize the distributed randomness of the fixation points. When the NNI value is less than 1, the average minimum distance between the fixation points is less than the average random distance, indicating that the fixation points follow a clustered distribution; when the NNI value is equal to 1, the fixation points follow a random distribution; when the NNI value is more than 1, the fixation points follow a regularly dispersed distribution. The possible distribution state of the fixation points is shown in Figure 2.
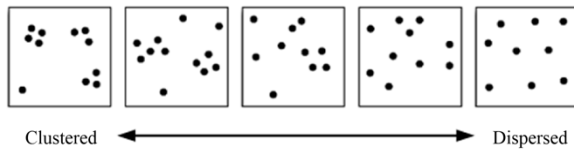

Fig. 2. Distribution state of the fixation points

The eye movement data sampling rate of two participants were less than 50%, which had no value for the next analysis. Finally, the eye movement data of 18 participants were selected for NNI analysis.

NNI analysis was conducted using A Simple Tool for Examining Fixations (ASTEF), an algorithm routine based on Matlab [37, 38]. The eye movement data from the 18 participants in each group were imported into the ASTEF analysis program and then the NNI values were acquired.

Fig. 3. NNI analysis interface of ASTEF

4.1.2 NASA-TLX
NASA-TLX is a subjective scale developed by NASA to evaluate workload including six dimensions: mental, physical, temporal demands, frustration, effort, and performance [39].

4.1.3 3D-SART calculation
3D-SART is used for evaluating the SA of pilots. 3D-SART requires participants to perform self-assessment through questionnaires at the end of the experiment. It interferes minimally with the experiment and is tested with sensitivity and construct validity [40]. The 3D-SART scores are calculated using the following formula:

$$SA = U - (D - S) \tag{4}$$

where U represents situational understanding (Understand), D represents the demand for attention resource (Demand), and S represents the supply of attention resource (Supply).

4.1.4 Operation performance
The number of operation errors was determined by video playback, and the errors were classified on the basis of Endsley's SA research [7, 8].

4.2 Analysis and discussion
According to their different fixation features, the participants were divided into two groups: the dispersed group whose fixation feature is dispersed distribution (NNI>1) and the clustered group whose fixation feature is clustered distribution (NNI<1). The details are shown in Figure 4.

Using the independent sample t-test to determine the difference significance of the two groups' NNI under three different difficulty experiments, we found that the differences were all significant ($p<0.05$). Therefore, the grouping method is reasonable.

Furthermore, the descriptive statistics and difference significance test for the NNI, workload, SA, and operation errors of two groups were conducted using the SPSS. The descriptive statistics are presented in Table 1.
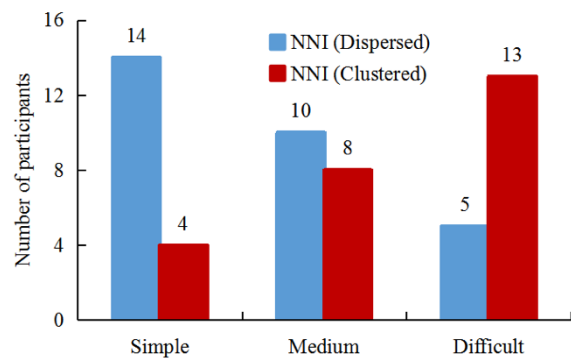

Fig. 4. Number of participants in different groups

Table. 1. Descriptive statistics of data results (Mean SD)

Measures	Simple Experiment		Medium Experiment		Difficult Experiment	
	Clustered	Dispersed	Clustered	Dispersed	Clustered	Dispersed
NASA-TLX	50.47 (5.00)	42.95 (2.42)	59.48 (3.50)	54.94 (3.07)	73.53 (2.44	68.33 (3.62)
SA	85 (4.76)	85.50 (0.98)	74.5 (1.20)	81.20 (4.69)	61.85 (1.72)	74.40 (7.54)
Operation Errors	4 (0.82)	2.86 (1.03)	5.38 (1.30)	4.10 (1.29)	9.46 (1.76)	6.60 (2.51)

4.2.1 SA of tower controllers
The influence of experimental difficulty on the SA of the different groups was studied, as shown in Figure 5.

In the independent sample t-test for the SA of the different groups, we found no significant difference in the simple experiment ($p=0.83>0.05$), a very significant difference ($p=0.001<0.01$) in the medium experiment, and a significant difference ($p=0.02<0.05$) in the difficult experiment. Under the same experimental difficulty, the dispersed group exhibited a higher SA than the clustered group. This result is consistent with the findings of Moore and Gugerty [31] that participants whose attention distribution is more extensive could consider the information of each functional area in the simulated tower control interface and have a better SA. Regardless of the group, the participants' SA decreased as the experimental difficulty increased because increasing the experimental difficulty greatly increased the demand for attention resource (Demand), while the situational understanding

(Understand) and the supply of attention resource (Supply) remained limited. However, the average SA value of the dispersed group decreased slowly because dispersed attention can help participants improve their understanding and cognition of the entire scene, ensuring that no potentially important information is neglected. According to the SA "leapfrog" view of Liu et al. [41], in the information search process, ordinary controllers tend to search the whole objects, including useful and other redundant information, to predict and project. By contrast, experienced controllers (those with high SA levels) only need to search for critical information to fulfill the task, presenting a "leapfrog" information search pattern and a dispersed distribution fixation feature. This capability is especially important for controllers who control multiple aircraft simultaneously. While ordinary controllers are still busy processing large amounts of information, experienced controllers can understand the whole project and make a quick prediction with skip scanning. Thus, the dispersed group outperforms the clustered group in the multi-threaded process.
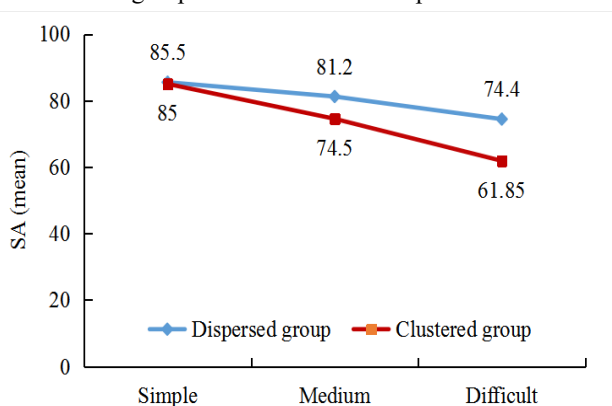

Fig. 5. Influence of the experimental difficulty on the SA of different groups

4.2.2 Workload of tower controllers

Variance homogeneity analysis was performed on the NASA-TLX scores under the three groups of experiments, and the significance was 0.77 ($p>0.05$). The three groups were then analyzed by using one-way analysis of variance. The results show significant differences among workloads of the three groups ($p<0.05$), indicating that the design of the experiment system is reasonable and reliable. We likewise investigated the impact of experimental difficulty on the workload scores, as shown in Figure 6.
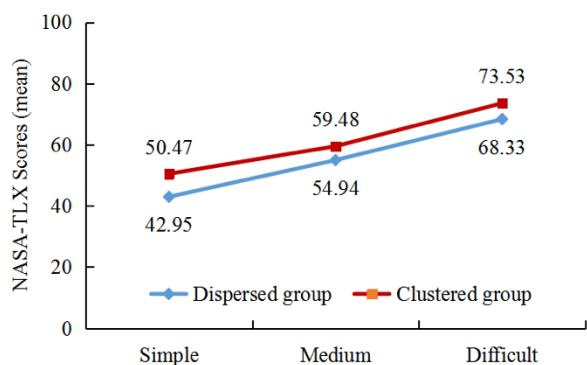

Fig. 6. Comparison of workload between different groups

Regardless of the participants' fixation feature, their workload increased with the experimental difficulty.

Interestingly, under the same experimental difficulty, the dispersed group obtained lower workload scores than did the clustered group. Considering the previous SA analysis, this finding may be ascribed to the differences in information search approaches. Compared to the global information search pattern used by the clustered group, the key information search pattern used by the dispersed group had less information. Thus, the dispersed group bore a smaller task load than did the clustered group.

4.2.3 Performance of tower controllers

Furthermore, we studied the operation errors of different groups in terms of experimental difficulty.
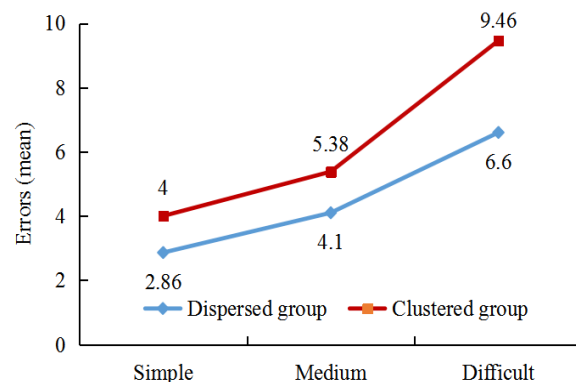

Fig. 7. Comparison of errors between different groups

Figure 7 shows that the average number of operation errors increases with the difficulty of the experiment. For the same experimental difficulty, the average number of operation errors of the dispersed group is less than that of the clustered group. The independent sample t-test for the different groups' number of errors in the simple, medium, and difficult experiments showed a nearly significant difference between the dispersed and clustered groups (the respective p values of the simple, medium, and difficult experiments were 0.06, 0.05, and 0.01). Specifically, the participants whose fixation feature is dispersed distribution (NNI>1) had a better operation performance, which is inconsistent with the conclusion of Moore and Gugerty [31] that "Errors also increased as NNI increased towards 1, which indicates that the spatial distribution of fixations became less aggregated and more randomly dispersed". The possible reasons are as follows: (1) the interfaces of these two experiments were not the same, that is, Kristin's was based on the radar control interface, whereas ours was based on the simulated tower control interface. (2) The different control tasks in these two experiments led to different fixation strategies. (3) The participants were not the same (those of Kristin's are radar controllers, whereas ours are tower control students).

Drawing lessons from Endsley's classification of SA errors, we also attempted to classify the errors made by the participants. Table 2 lists the SA errors in the simulated tower control experiment in detail. Notably, the SA errors here were not equal to the accidents.

In Table 2, Level 1 SA error refers to the error caused by missing or misconceiving information, such as failing to monitor the requested information from the aircraft, which may cause the aircraft to fly in circles, and forgetting to transfer control after take-off because of a lapse in memory. Level 2 SA error corresponds to misunderstanding of information, including the change in instructions and the

irrational allocation of the runway (the controller directs the aircraft to a far runway while the nearby runway is idle). Level 3 SA error refers to the incorrect projection of the future situation of the system, including runway conflict and short distance without satisfying the safe interval between take-off and landing of aircraft. The different SA level errors that account for the total errors of each group are presented in Figure 8.

Level 1 SA error	Aircraft go around
	Flight delay
	Forget to transfer control after taking off
	Aircraft fails to slip to the terminal after landing
Level 2 SA error	Change the command
	Irrational allocation of the runway
Level 3 SA error	Runway conflict
	Distance between taking off and landing aircraft less than safe interval
	Others

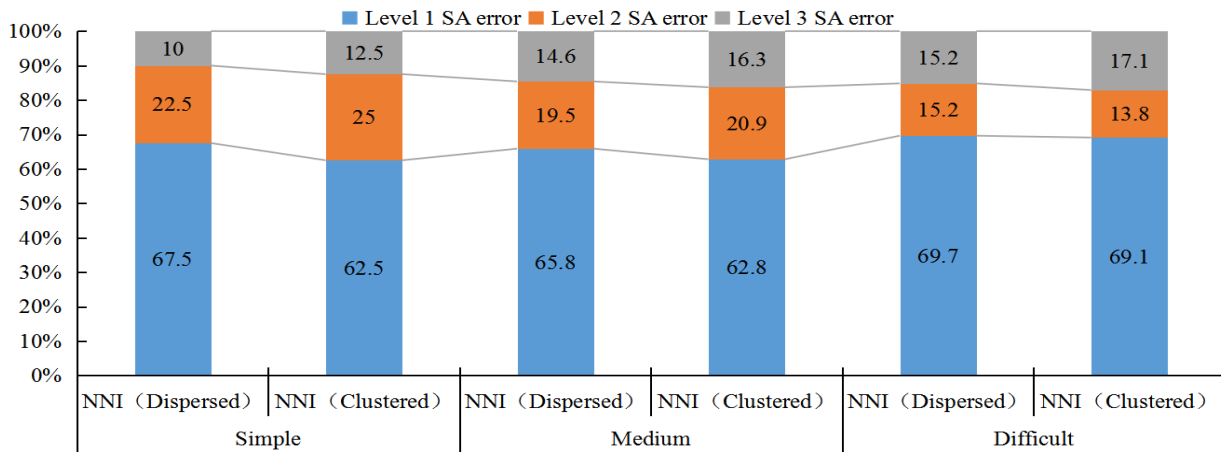**Table 2.** Taxonomy of Situation Awareness Errors

Fig. 8. Distribution of different SA level errors between different groups

Figure 8 indicates that irrespective of the fixation features, Level 1 SA errors account for more than half (60%–70%) of the entire operation's errors, and Level 3 SA errors account for the smallest proportion. This conclusion is basically consistent with the research of Jones and Endsley [8], who found that 72.4% of the total number of errors was Level 1 SA errors, 17.2% were Level 2, and 10.4% were Level 3. Most of the SA errors occurred in the acquisition and perception of information, and only a small part of the SA errors was caused by the incorrect understanding of information and the deviation of projection.

Moreover, the frequencies of Level 2 and Level 3 SA errors made by the dispersed group were lower than that of the clustered group. In other words, the dispersed group has a better situational understanding and projection capability. Although the proportion of Level 1 SA errors made by the dispersed group was higher than that of the clustered group, the Level 1 SA errors made by the dispersed group consisted mainly of flight delay and failure to slip to the terminal after landing, which rarely cause devastating consequences. Meanwhile, the Level 1 SA errors made by the clustered group consisted mainly of aircraft flying in circles and forgetting to transfer control after take-off, which may cause catastrophic accidents. The proportions of aircraft flying in circles and forgetting to transfer control after take-off errors that account for the Level 1 SA errors made by the dispersed group versus that made by the clustered group were as follows: 37% versus 50% in the simple experiment, 40.7% versus 55.6% in the medium experiment, and 30.4% versus 45.9% in the difficult experiment. Thus, in the simulated tower control experiment, the dispersed group participants performed better than the clustered group participants.

4.2.4 Correlation analysis between NNI and SA
Pearson correlation analysis was performed to explore the relationship between the NNI and SA, and the partial correlation analysis of controlling the workload variable was conducted. We found that without considering the impact of workload variable, the Pearson correlation coefficient between the NNI and SA is 0.602. Obviously, it is a moderate positive correlation ($0.5 < r < 0.8$), and the significance level is high ($p < 0.01$). When considering the impact of the workload variable, the Pearson correlation coefficient becomes slightly larger (0.626). Similar to the former case, it is also a moderate positive correlation, and the significance level is high.

A significant moderate positive correlation seemingly always exists between the NNI and SA despite the difficulty of the experiment. Therefore, the NNI can be used as an auxiliary eye movement index to measure SA. However, evidence on whether the NNI can be treated as an independent indicator is lacking because the NNI is a physiological measure that does not measure SA directly. Following Endsley's three levels of SA, namely, perception, comprehension, and prediction, we easily found that eye movement reflects the perception level of SA. However, whether eye movement can represent comprehension and projection requires more theoretical and experimental exploration.

5. Conclusions

In this work, the SA of tower controllers were investigated by performing eye movement analysis based on the NNI. The following conclusions were drawn：

(1) Participants with a dispersed distribution fixation feature (NNI>1) have better SA and operation performance than those with a clustered distribution fixation feature (NNI<1). This conclusion shows that the attention characteristics of controllers are closely related to their operational performance, which can provide reference for controller training. The controller should be trained to use the "leapfrog" information search pattern, to enhance their ability to control global information and command multiple aircrafts at the same time. "Top-down" driven cognitive behavior is conducive to the improvement of SA of controllers.

(2) The correlation between the NNI and the controllers' SA is significantly moderate positive. The NNI fixation index is not only related to situational awareness in data, but also to the cognitive characteristics and information processing patterns of controllers. Therefore, the NNI can be used as an auxiliary eye movement index to measure SA. But whether NNI fixation index can be used as an independent eye movement index to measure SA level effectively still needs more experimental verification and theoretical support.

(3)

In the future, SA can be combined with other eye movement indices (e.g., fixation point, fixation time, and scanning). Analyzing the difference in eye movement features between the tower and radar controllers also makes sense. The results may provide some insights into SA measurement and ATC automation based on machine learning in the real situation for improved understanding.

References

1. Belcastro, C. M., Foster, J. V., "Aircraft loss-of-control accident analysis". In: *Proceedings of AIAA Guidance, Navigation, and Control Conference*, Toronto, Canada: AIAA, 2010, pp.26-42.
2. Li, F. H., Xie, Z. N. "Study of regular analysis and preventive measures of flight accident and accident proneness in the stage of China's civil aviation approach". *Journal of Civil Aviation Flight University of China*, 19(4), 2008, pp.3-7.
3. Endsley, M. R. "Toward a theory of situation awareness in dynamic systems". *Human Factors*, 37(1), 1995, pp.32-64.
4. Jensen, R. S. "Pilot judgment: training and evaluation". *The Journal of the Human Factors and Ergonomics Society*, 24(1), 1982, pp.61-73.
5. Härtel, C. E. J., Smith, K., Prince, C. "Defining aircrew coordination: Searching mishaps for meaning". In: *Proceedings of the 5th International Symposium on Aviation Psychology*, Columbus, USA: Ohio State University, 1989, pp.12-22.
6. Yan, S. H. "Classification and analysis of air traffic controllers' errors based on information processing model". *China Safety Science Journal*, 19(08), 2009, pp.121-125.
7. Endsley, M. R. "A taxonomy of situation awareness errors". *Human Factors in Aviation Operations*, 3(2), 1995, pp.287-292.
8. Jones, D. G., Endsley, M. R. "Sources of situation awareness errors in aviation". *Aviation, Space, and Environmental Medicine*, 67(6), 1996, pp.507-512.
9. Endsley, M. R. "Measurement of situation awareness in dynamic systems". *Human Factors*, 37(1), 1995, pp.65-84.
10. Salmon, P., Stanton, N., Walker, G. H., Green, D. "Situation awareness measurement: A review of applicability for C4i environments". *Applied Ergonomics*, 37(2), 2006, pp.225-238.
11. Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., Young, M. "Measuring situation awareness in complex systems: Comparison of measures study". *International Journal of Industrial Ergonomics*, 39(3), 2009, pp.490-500.
12. Yang, J. Z., Zhang, K. "Situation awareness: approaches, measures and applications". *Advances in Psychological Science*, 12(6), 2004, pp. 842-850.
13. Wang, L. L., Yang, Y. "On the situational evaluation awareness for the radar-receiver controllers". *Journal of Safety and Environment*, 19(02), 2019, pp.554-561.
14. Vidulich, M. A., Stratton, M., Crabtree, M., Wilson, G. "Performance-based and physiological measures of situational awareness". *Aviation, Space, and Environmental Medicine*, 65(5), 1994, pp.7-12.
15. Bao, J. P, "*Research on the Influence of Cockpit Noise on Pilot Situation Awareness*". Master thesis of Nanjing University of Aeronautics and Astronautics, China, 2019, pp.34-47.
16. Strayer, D. L., Cooper, J. M., Drews, F. A. "What do drivers fail to see when conversing on a cell phone?". In: *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, Denver, USA: SAGE, 2004, pp.2213-2217.
17. Deng, Z. "Theories, techniques and applied researches about eye-movement psychology". *Journal of Nanjing Normal University (Social Science Edition)*, 1(01), 2005, pp.90-95.
18. Kramer, A. F., McCarley, J. S. "Oculomotor behavior as a reflection of attention and memory processes: neural mechanisms and applications to human factors". *Theoretical Issues in Ergonomics Science*, 4(1-2), 2003, pp.21-55.
19. Liversedge, S. P., Findlay, J. M. "Saccadic eye movements and cognition". *Trends in Cognitive Sciences*, 4(1), 2000, pp.6-14.
20. Rayner, K. "Eye movements in reading and information processing: 20 years of research". *Psychological Bulletin*, 124(3), 1998, pp.372-422.
21. Merwe, K. V. D., Dijk, H. V., Zon, R. "Eye movements as an indicator of situation awareness in a flight simulator experiment". *The International Journal of Aviation Psychology*, 22(1), 2012, pp.78-95.
22. Yu, C. S., Wang, E. M. Y., Li, W. C., Braithwaite, G. "Pilots' visual scan patterns and situation awareness in flight operations". *Aviation, Space, and Environmental Medicine*, 85(7), 2014, pp.708-714.
23. Liu, Z. F., Su, H. "Exploring the individual and task difficulty difference of situation awareness in aviation working environment: the advantages of eye tracking method". *Chinese Journal of Applied Psychology*, 22(01), 2016, pp.12-25.
24. Hyun, S. H., Kim, K. T., Song, S. M., Yoon, J. S., Lee, J. I., Cho, S. M., Sohn, Y. W. "Effects of expertise and situation complexity on visual attention and action planning for air traffic control". In: *Proceedings of the Human Factors and Ergonomic Society 50th Annual Meeting*, San Francisco, USA: SAGE, 2006, pp.420-423.
25. Hauland, G. "Measuring Individual and team situation awareness during planning tasks in training of En route air traffic control". *International Journal of Aviation Psychology*, 18(3), 2008, pp.290-304.
26. Wei, K. K. "*Research on Fixation Characteristics of Tower Controller under Fatigue State*". Master thesis of Civil Aviation University of China, China, 2018, pp.15-21.
27. Clark, P. J., Evans, F. C. "On some aspects of spatial pattern in biological populations". *Science*, 121(3142), 1955, pp.397-398.
28. Dillard, M. B., Warm, J. S., Funke, G. J., Funke, M. E., Finomore, V. S., Jr., Matthews, G., Parasuraman, R. "The sustained attention to response task (SART) does not promote mindlessness during vigilance performance". *Human Factors*, 56(8), 2014, pp.1364-1379.

29. Nocera, F. D., Camilli, M., Terenzi, M. "A random glance at the flight deck: Pilots' scanning strategies and the real-time assessment of mental workload". *Journal of Cognitive Engineering and Decision Making*, 1(3), 2007, pp.271-285.

30. Fidopiastis, C. M., Drexler, J., Barber, D., Cosenzo, K., Barnes, M., Chen, J. Y., Nicholson, D. "Impact of automation and task load on unmanned system operator's eye movement patterns". In: *Proceedings of the 5th International Conference on Foundations of Augmented Cognition*, San Diego, USA: Springer, 2009, pp.229-238.

31. Moore, K., Gugerty, L. "Development of a novel measure of situation awareness: The case for eye movement analysis". In: *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*, San Francisco, USA: SAGE, 2010, pp.1650-1654.

32. Jin, H. B., Liu, Y. W., Zhu, G. L. "Measuring situation awareness of tower controllers based on eye movement and performance analysis". *China Safety Science Journal*, 27(7), 2017, pp.65-70.

33. Moacdieh, N., Sarter, N. "Clutter in electronic medical records: Examining its performance and attentional costs using eye tracking". *Human Factors*, 57(4), 2015, pp.591-606.

34. Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., Nikolic, D., Manning, C. A. "Situation Awareness as a predictor of performance for En route air traffic controllers". *Air Traffic Control Quarterly*, 6(1), 1998. pp.1-20.

35. Endsley, M. R., Smolensky, M. W. "Situation awareness in air traffic control: The picture". *Human Factors in Air Traffic Control*, 1998, pp.115-154.

36. Chiappe, D., Vu, K. P. L., Strybel, T. "Situation awareness in the nextgen air traffic management system". *International Journal of Human Computer Interaction*, 28(2), 2012, pp.140-151.

37. Camilli, M., Nacchia, R., Terenzi, M., Di Nocera, F. "ASTEF: A simple tool for examining fixations". *Behavior Research Methods*, 40(2), 2008, pp.373-382.

38. Di Nocera, F., Capobianco, C., & Mastrangelo, S. "A Simple(r) tool for examining fixations". *Journal of Eye Movement Research*, 9(4), 2016, pp.1-6.

39. Bunce, D., Sisa, L. "Age differences in perceived workload across a short vigil". *Ergonomics*, 45(13), 2002, pp.949-960.

40. Jin H. B., Liu W. H., Chen J. "Effects of both situation awareness and gaze transition mode on conflict in tower control". *China Safety Science Journal*, 26(4), 2016, pp.72-77.

41. Liu, W., Mu, D. M., Zhuang, D. M., Yuan, X. G. "Study on situation awareness in aviation ergonomics". *Advanced Materials Research*, 328-330, 2011, pp.2414-2418.