[EN-A-022] Analysis of Positive and Negative Effects of Salience on the ATC Task Performance

+H. Yoshida*, H. Aoyama**, D. Karikawa***, S. Inoue**, T. Kanno*, K. Furuta*

*Graduate School of Engineering
The University of Tokyo
Tokyo, Japan
[h-yoshida | kanno | furuta]@[cse | sys | rerc].t.u-tokyo.ac.jp

**ATM Department
Electronic Navigation Research Institute (ENRI)
Tokyo, Japan
[aoyama | s-inoue]@mpat.go.jp

***Graduate School of Engineering
Tohoku University
Sendai, Japan
daisuke.karikawa@tohoku.ac.jp

Abstract: This paper reports effects of salience on the air traffic control (ATC) task performance and proposes screen design criteria considering the effects. The objective of this research is to develop screen design criteria considering human perception to reduce controllers' workloads and improve task performance. Among several kinds of perceptions, we focused on visual salience in this research since it is one of the user interface elements causing a high-impact perception in accurate and efficient ATC tasks. We carried out an experiment with ten participants in order to clarify effects of the salience on the ATC task performance. The experimental task was a multiple task including giving heading instructions as the main tasks and handing operations as the sub tasks. Experimental conditions provided four patterns of screen designs with (a) no, (b) small, (c) middle, and (d) large gap of salience between important airplanes and the others. We controlled salience using the color salience model that we developed in our previous works. Results of the experiment showed (1) larger salience gaps among displayed information improved novices' instruction timing to the airplanes, (2) larger salience gaps according to the importance in main tasks degrades the performance of sub tasks, (3) larger salience gaps among displayed information slightly improved novices' situation awareness (SA). Based on these results, we proposed the screen design criteria considering salience of displayed information.

Keywords: salience, performance of the ATC task, situation awareness, screen design criteria

1. INTRODUCTION

The role of air traffic control (ATC) recently becomes more important with the increasing demand of air traffic. In ATC tasks, the controllers are keeping and facilitating safe and smooth flight for airplanes. They accurately understand many kinds of dynamic information using radar screen and so on. If there are risky airplanes for safe flight, the controllers give instructions such as change of the altitude or direction to the airplanes using communication channels in order to avoid the risk. ATC tasks are multiple and complex tasks that require understanding many kinds of dynamic information and executing different tasks in parallel. Moreover, they are crucial tasks, because a small mistake may lead to major accidents. Thus, ATC tasks are very difficult task of a high cognitive demand [2]. To reduce such high cognitive workload of the controllers and improve safety and reliability of the entire ATC tasks, the screens need to be designed considering not only usability but also perceptual and cognitive characteristics. Especially, visual salience that affects visual attention has to be considered in the screen design to avoid overlooking and misunderstanding important information. However, predicting and controlling effects of salience are difficult even for expert screen designers, since the salience is a factor related to human perception and psychology.

Here, if salience of displayed information is quantified, the designers can objectively evaluate propriety of the screens in terms of salience corresponding to the importance of the information. Moreover, they can also make improvement plans easily. For example, they can specify problems in designing screen displays where the order of salience value does not match that of the importance of information by simple comparison. As for improvement plans, the designers can adjust properties related to salience by referring to the salience value. Assuming such usage, we have developed the salience model which calculates salience value of information on a screen in previous works. The model calculates integrated salience of each object considering effects of its color, size, and position (distance from point of regard).

In order to apply salience to actual screen design, screen design criteria considered relationship between salience value and users' task performance is necessary. Therefore, the objective of this study is to clarify effects of salience on the ATC task performance and to develop the screen design criteria considering the salience effects. In this study, we carried out an experiments with novice participants and the experimental screens were designed using the salience model that we have developed. There are many existing researches of the screen design [3][4][5]. Some of them focus on the ATC tasks and assume a single task as the experimental ATC tasks to simplify their problems and experimental design. Whereas, an aspect of multiple task cannot help considering since it actually results in high cognitive demands for controllers in the ATC tasks. Taking it into account, we especially focus on the aspect of multiple tasks in the ATC tasks and clarify the effects of salience to the tasks. This paper describes an experiment, its results, and the screen design criteria for the ATC tasks from the results.

2. METHOD

We carried out an experiments with participants in order to clarify the relationship between salience and the ATC task performance. Participants executed a multiple task containing giving instructions to airplanes and hand in and hand off operations in parallel on the ATC simulator screens [1].

2.1 Participants

Ten people from 20s to 30s of age participated in this experiment. All of them had little knowledge and experience about the ATC tasks. They usually use computers and used to operating computers with a mouse.

2.2 Experimental Task

The experimental task was a multiple task of simplified actual ATC tasks. It included a giving instruction task as the main task and a handing airplane task as the sub task. The giving instruction task was a task of timely giving heading instructions to airplanes just following an air traffic control plan (ATC plan). The plan was preliminary defined before starting the task and including targeted airplanes, timing and contents of the instructions. The handing airplane task was a task of executing hand in and hand off operations to airplanes passing through boundaries of sectors. Here, a hand in operation is an operation for receiving a new airplane from a neighboring sector to one's own sector. A hand off operation is an operation for handing an airplane from one's own sector to a neighboring sector. The operations have to be executed before the airplanes go through the boundaries. As the experimental task, participants gave instructions to airplanes just as the ATC plans and executed handing in and handing off airplanes in parallel.

2.3 Scenario

Two fictional scenarios imitating situations of air traffic were prepared for this experiment. Each scenario had a different ATC plan. In both scenarios, we supposed that all airplanes were cruising at their own altitude and that the ATC plans included only instructions on heading for simplicity. There were about ten airplanes displayed on the screens, and three or four of them were the airplanes to be controlled that were not able to keep separations without ATC instructions. The ATC plans of this experiment

Table 1 The Conditions for Colors and Salience in Each Screen Design.

Expression of airplane information \ Screen design		(a)no gap	(b)small gap	(c)middle gap	(d)large gap
(1)airplanes to be instructed [Most important]	Color[RGB]	111,224,210	162,162,234	131,249,131	247,80,247
	Salience	2.35	2.20	3.10	4.31
(2)airplanes to be controlled [Important]	Color[RGB]	111,224,210	179,215,179	179,215,179	131,249,131
	Salience	2.35	2.00	1.51	3.05
(3)the other airplanes [Unimportant]	Color[RGB]	111,224,210	82,170,170	88,125,88	88,125,88
	Salience	2.35	1.13	0.15	0.10
Salience gap between (1)and(3)		**0.00**	**1.07**	**2.95**	**4.21**

provided from six to ten heading instructions to a part of the airplanes to be controlled.

In addition, there were airplanes going through the boundaries of sectors during execution of heading instructions. The participants had to execute hand in and hand off operations to the airplanes. One of the authors who is a retired air traffic controller and has enough experience of educating novice controllers made these scenarios and ATC plans.

2.4 Screen Design

Four patterns of screen design for the experimental simulator in terms of salience of airplane displays were prepared. They included (a) no, (b) small, (c) middle and (d) large gap of salience between important airplanes and the others. The salience of each airplane display was adjusted using the color salience model developed in our previous works [6][7][8]. The model quantifies total color salience of an object by a linier sum of the absolute color salience (feature salience) and the relative color salience against the surrounding colors (heterogeneous salience). The feature color salience of an object is defined as a linier combination of saturation, hue and gap of lightness between the object and background color. On the other hand, the heterogeneous color salience of the object is defined as a linier of saturation gaps, hue gaps, and lightness gaps between the screen average and the object.

Table 1 shows the screen design conditions for color and salience values of airplane displays in each screen pattern. The salience value shown as Table 1 was calculated by the color salience model above mentioned. In this experiment, we classified airplanes displayed on the simulator into three categories, (1) airplanes to be instructed, (2) airplanes to be controlled, and (3) the other airplanes, considering the ATC instruction tasks. The order of importance for airplanes was the order of (1), (2) and (3). We controlled salience according to the order of importance in the screen design (b), (c) and (d). The airplanes (1), airplanes to be instructed, are the airplanes to be instructed immediately in the defined ATC plans. The colors of their information displayed are changed from the colors of design (2) to design (1) at ten seconds before the moment of instruction according to the ATC plans, then, the colors are set back to design (2) after the instruction has been issued by participants. The airplanes to be controlled are those that cannot keep separation without any air traffic control. Namely, the airplanes to be instructed are a part of the airplanes to be controlled.

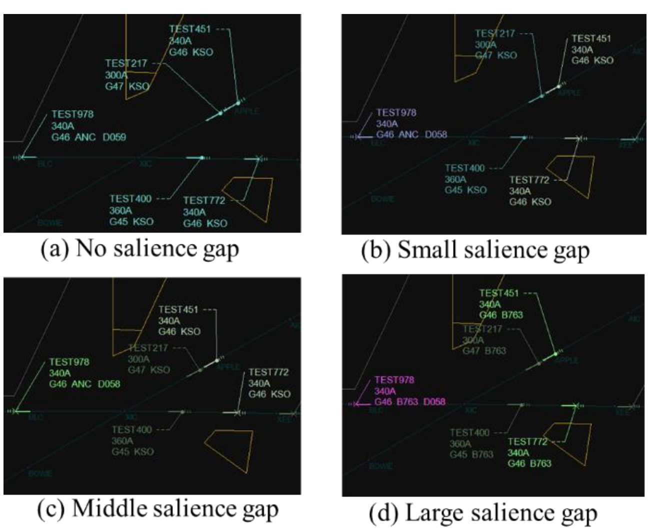

Figure 1 A part of each screen design example displaying information of airplanes (1), (2) and (3).

Fig. 1 shows an example of the each screen design (a) - (d) at the timing that the TEST 978 airplane has to be instructed. The airplane TEST 978 corresponds to the (1) airplanes to be instructed, the airplane TEST 451 and TEST 772 correspond to the (2) airplanes to be instructed, the TEST 217 and TEST 400 correspond to (3) the other airplanes.

2.5 Procedure

The following list of steps shows the task flow of participants. In this experiment, we defined giving heading instructions as the main tasks, and hand in and hand off operations as the sub tasks. Duration of experimental session was about ten minutes.

<u>Step 1</u>: Explanation of a task scenario and the corresponding ATC plan by the experimenter.
<u>Step 2</u>: Starting the ATC simulator and giving heading instruction to target airplanes according to the ATC plan.
<u>Step 3</u>: Hand in and hand off operations at an appropriate timing in parallel.
<u>Step 4</u>: Answering oral questions about situation awareness (SA) asked by the experimenter.
<u>Step 5</u>: Finishing the all planed ATC instructions.

After an experimental task, namely the step 5, the participants formed a questionnaire of NASA-TLX [10] to be assessed subjective workload of the task. As for an experimental environment, participants executed all tasks on 22 inched displays whose resolutions were 1920*1080.

2.6 Measurement

In order to clarify relationship between the task performance and salience, the following measurements were introduced in this experiment.

Performance of the Instruction Task (Main Task)

The measurement for performance of the instruction task was gaps between participants' and the ATC plan's instruction timing. The smaller gap means the better timing of instruction, namely better task performance.

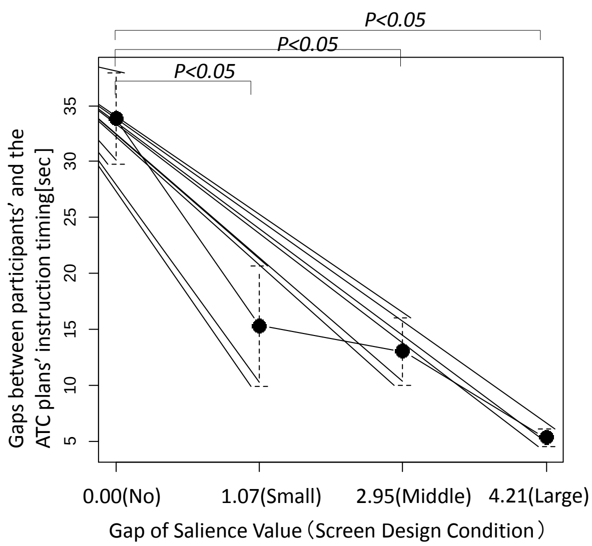

Figure 2 The gap between participants' and the ATC plans' instruction timing.

Performance of the Handing Task (Sub Task)
The measurement for performance of the handing task was the time gaps between airplanes' crossing the sector boundaries and participants' hand in and hand off operations. The time of boundary crossing is the time limit of the hand in or hand off operation. A smaller gap means a delay of the hand in or hand off operation, namely worse task performance.

Level of Situation Awareness (SA)
To assess the level of SA, the online SA probe method [9] was introduced. In this experiment, the experimenter asked the participants some oral questions about air traffic situations. The measurement for the level of SA was accuracy rate and required time for answering (reaction time). A shorter reaction time means better level of situation awareness.

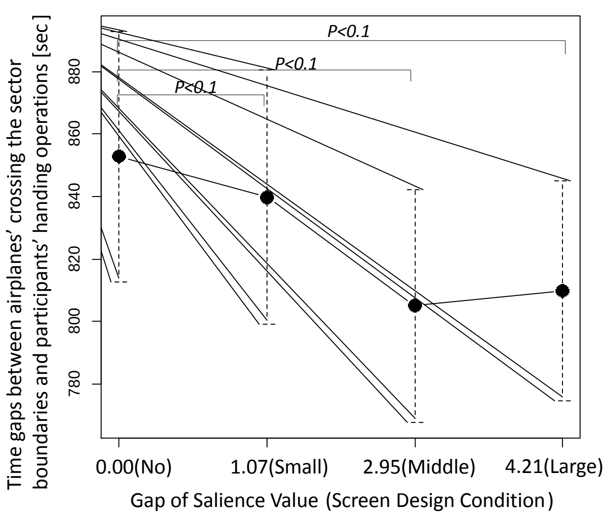

Figure 3 Time gaps between airplanes' crossing the sector boundaries and participants' handing operations.

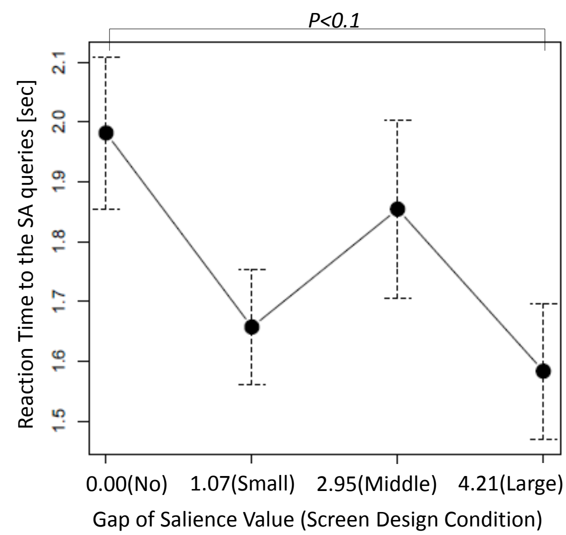

Figure 5 Reaction time to the SA queries.

Subjective Workload
To assess the subjective workload of each task, the participants formed the NASA-TLX questionnaire [10] after each task.

3. RESULT

3.1 Performance of the Instruction Task (Main Task)
Fig. 2 shows the gap between participants' and the ATC plans' instruction timing for each screen design. According

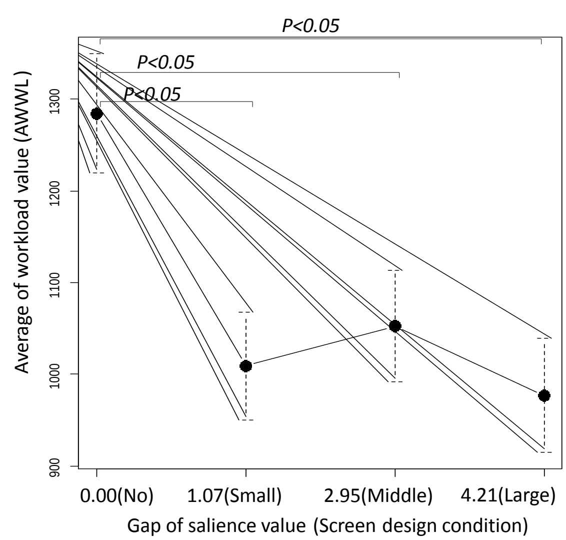

Figure 4 Average of workload value

to the graph, the gap of instruction timing are significantly smaller for the screen with small, middle and large salience gap between important and unimportant airplanes. In addition, there was no significant difference of instruction timing among the screen design with a certain gap of salience, namely a small, middle and large gap. From the result, novices can execute instructions more timely if there is a salience gap among airplanes displayed.

3.2 Performance of the Handing Task (Sub Task)

Fig. 3 shows the time gaps between airplanes' crossing the sector boundaries and participants' handing operations. In this measurement, a small value represents a delay of hand in or hand off operations, namely a low performance of sub tasks. According to the graph, hand in and hand off operations on the screen with a certain salience gap tended to be delayed than that on the screen without salience gap. From this result, we found that the performance of sub tasks tends to degrade on the screen design with a certain salience gap.

3.3 Level of Situation Awareness (SA)

Fig. 4 shows the reaction time for oral questions about SA during task execution for each screen design. According to the graph, the reaction time tended to be shorter with larger salience gap between important and unimportant airplanes. From the result, we found that more than 4.21 salience gap between information was able to enhance the SA of participants.

3.4 Subjective Workload

Fig. 5 shows the average of workload for each gap of salience value, namely screen design condition. The workload value was the AWWL [11] calculated from the result of NASA-TLX questionnaire. In this measurement, larger value represents higher workload. According to the graph, the workload value became significantly lower for the screen with a certain gap of salience than without salience gap. On the other hand, there was no significant difference among screens with small, middle and large salience gap.

3.5 Observed Error

Some errors were observed in the giving instruction task, the handing operation task and answering the SA questions through this experiment. Table 2 shows the number of observed errors for each salience gap (screen design) and each kind of tasks. The errors observed in the giving instruction task were giving incorrect instructions or giving instructions to different airplanes. Most of them were observed for the screen without or with small salience gap. The errors observed in the handing operation task were forgetting operations. The number of them were almost the same in each screen design. The errors observed in answering the SA questions were incorrect answers. The participants were displayed air traffic situations during the oral SA questions. Therefore, some assumptions by wrong memories or false recognitions of displayed information might cause the errors. Most of them were observed for the screens with middle and large salience gap.

Table 2 The number of observed errors in the experiment.

Salience gap / Task	0.00 (No)	1.07 (Small)	2.95 (Middle)	4.21 (Large)
Giving instruction (main)	7	5	2	1
Handing operation (Sub)	2	3	2	1
Answering SA questions	2	1	4	5

4. DISCUSSION

Effects of Salience on Task Performance and SA

In the task of giving instruction (the main task), the accuracy of instruction timing significantly improved for the screen with more than 1.07 salience gap (Fig. 2). The number of errors tended to reduce for the screen with more than 2.95 salience gap (Table 2). Visibility of displayed information corresponding to the gap of salience could affect improving the task accuracy and reducing the number of errors. Namely, by adjusting salience corresponding to the importance of information, visibility of displayed information and difficulty of the task were improved. These improvement resulted in better performance of the main task. From the result, as for the giving instruction task, accuracy of both timing and target can be improved.

In the task of handing operations (the sub task), the operation was significant delayed, namely the performance degrade for the screens with more than 1.07 salience gap (Fig. 3). The number of errors was the almost same in each screen design (Table 2). The reason is considered that participants would allocate attention to the airplanes with a high salience and could not allocate attention to the airplanes with a low salience. From this result, controlled salience according to the importance in terms of the main task would give negative effects to the performance of the sub task.

As for SA, the reaction time tended to be faster for the screens with larger salience gap and it was significantly faster for the screen with 4.21 salience gap (Fig. 4). However, the significant level was low, 10 %. Thus, the salience gap was not so effective for improving situation awareness of air traffics.

Screen Design Criteria

From the results of the experiment, we propose screen design criteria for multiple tasks considering color salience. On the other hand, appropriate design criteria for multiple tasks would be deferent since the weight of sub task is different for each business requirement in multiple tasks. Thus we considered screen design criteria in each weight of sub tasks as follows. In the criteria, the salience value assume the value calculated by the salience model.

- If the main task performance is considered to be much important, it is preferable to adjust salience corresponding to the importance of information and to set salience gap more than 2.95 between information of different importance.
- If the improvement of the SA is considered in addition to the main task performance, it is preferable to adjust salience corresponding to the importance of information and to set salience gap less than 4.21 between information of different importance.
- If the sub task performance is considered in addition to the main task performance, it is preferable to adjust salience corresponding to the importance of information and to set salience gap less than 1.07 between information of different importance. On the other hand, too small salience gap can cause degrade accuracy of main tasks and increase possibilities of error occurring.

5. CONCLUSION

The objective of this research is to develop screen design criteria considering human perceptions to improve users' task performance and situation awareness in multiple tasks such as the ATC tasks. Among several kinds of perceptions, we focused on salience here since it is one of the user interface elements causing a high-impact perception in accurate and efficient ATC tasks. This paper reported an experiment for clarifying relationship between salience and the ATC task performance including SA. Ten people who were novices for the ATC tasks participated in the experiment. The experimental task was a multiple task including giving instruction task as a main task and handing airplane task as a sub task. The experimental screens provided four conditions with no, small, middle and large salience gap between important and unimportant airplanes. The salience of the screen was adjusted using the color salience model which we have developed in our previous works. In addition, this paper proposed screen design criteria for multiple tasks based on the results of the experiment. In the future, we would like to carry out experiments with experts of the ATC tasks and clarify the effects of salience to their task performance. We also would like to update the screen design criteria based on the results of the experiments.

6. REFERENCES

[1] D. Karikawa et al., Study on Performance Visualization for Educational Purpose in Air Traffic Control Area, The Journal of Human Interface Society, vol.11, No.3, (2009) (In Japanese).

[2] M. Endsley: Situation Awareness. In: Aviation Systems, *Handbook of Aviation Human Factors*. Mahwah, NJ: Lawrence Erlbaum Associates. 257. 11, (1999).

[3] L. Itti et al., A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, vol. 40, pp. 1489—pp. 1506 (2000).

[4] Tanaka, et. al.,. An Attractiveness Evaluation Model Based on the Physical Features of Image Regions. *IEICE A JPN edition*, J83-A(5): 576-588 (2000) (In Japanese).

[5] Bosse T. et al.: Automated Visual Attention Manipulation. In: Paletta L., Tsotsos J.K. eds. Attention in Cognitive Systems. WAPCV 2008. Lecture Notes in Computer Science, vol. 5395. Springer, Berlin, Heidelberg, (2009).

[6] H. Yoshida et al., Quantifying Visual Attention - Effect of Object size -, *Advances in Ergonomics Modeling, Usability & Special Populations*, Volume 486 of the series Advances in Intelligent Systems and Computing, pp. 173– 187 (2016).

[7] H. Yoshida et al., Quantification of Visual Attention (2) - Modeling Effect of Size, Layout, and Integrated Elements - , *Human Interface Symposium 2013* (2013) (In Japanese).

[8] A. Yokomizo et al., Quantification of Visual Attention (1) - Modeling Effect of Color -, *Human Interface Symposium 2013* (2013) (In Japanese).

[9] M. Endsley, et al.: Situation awareness in air traffic control: Enhanced displays for advanced operations, (No. DOT/FAA/CT-TN00/01). Federal Aviation

Administration Technical Center Atlantic City NJ, (2000).

[10] S. Haga et al, Japanese version of NASA Task Load Index: Sensitivity of its workload score to difficulty of three different laboratory tasks, The Japanese Journal of Ergonomics, Vol. 32, No. 2, pp. 71—79, (1996) (In Japanese).

[11] S. Miyake et al, Compatibility of AWWL with WWL in NASA-TLX, The Japanese Journal of Ergonomics, Vol. 31, pp. 404—405, (1995) (In Japanese).