

2018

Design of Revising Proximity between Space and Time Cues on Flight Deck Displays to Support NextGen – The Second Phase

Chang-Geun Oh

Kent State University - Kent Campus, coh1@kent.edu

Jennie J. Gallimore

Bowling Green State University, Jgallim@bgsu.edu

Pamela S. Tsang

Wright State University - Main Campus, pamela.tsang@wright.edu

Follow this and additional works at: <https://commons.erau.edu/ijaaa>



Part of the [Graphics and Human Computer Interfaces Commons](#), [Human Factors Psychology Commons](#), and the [Systems Engineering and Multidisciplinary Design Optimization Commons](#)

Scholarly Commons Citation

Oh, C., Gallimore, J. J., & Tsang, P. S. (2018). Design of Revising Proximity between Space and Time Cues on Flight Deck Displays to Support NextGen – The Second Phase. *International Journal of Aviation, Aeronautics, and Aerospace*, 5(5). DOI: <https://doi.org/10.58940/2374-6793.1280>

This Article is brought to you for free and open access by the Journals at Scholarly Commons. It has been accepted for inclusion in International Journal of Aviation, Aeronautics, and Aerospace by an authorized administrator of Scholarly Commons. For more information, please contact commons@erau.edu.

Design of Revising Proximity between Space and Time Cues on Flight Deck Displays to Support NextGen – The Second Phase

Cover Page Footnote

This material is based upon work partially supported by the Federal Aviation Administration (FAA) under Cooperative Agreement No. DTFWA-10-A-80021. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the FAA. The authors would like to thank Mr. Andrew McCullough for providing the programming for this effort.

Introduction

To overcome potential pilot errors when required time of arrival (RTA) operation is applied in the NextGen era, the authors created three novel flight deck displays manipulating the display proximity between space and time information. The first phase of this study conducted human-in-the-loop (HITL) simulation experiments collecting objective and subjective situation awareness (SA) data. As the control condition, the traditional low display proximity setting (Low Proximity condition), that was composed of the traditional navigation display (ND) depicting spatial information only and traditional text-based control display unit (CDU), was compared with three novel displays; (1) Medium Proximity-Text composed of a novel ND that integrated RTA and estimated time of arrival (ETA) in duration format and a traditional text-based CDU, (2) Medium Proximity-Graphics that was composed of the novel ND and the novel CDU that added horizontal bar graphics to indicate the temporal conformance to the assigned RTAs, and (3) High Proximity that integrated all space and time data in a single display without the use of CDU. The objective measures showed the three novel display conditions had a similar SA levels as the traditional display condition. However, the subjective measures showed High Proximity was significantly easier to use to maintain SA than Low Proximity in all the three levels of SA defined by Endsley (1995). Although the objective measures did not show a comparative advantage of the design strategy of higher display proximity between space and time cues, the subjective measure showed a possibility of SA enhancement with more training or by developing other objective measure methods. This second phase of study measured pilots' mental workload during RTA operations as another evaluation approach. During the first phase, pilots only involved in the query tests in the autopilot flights due to the limitation of allowed time to repeat the SA measurement. In this phase, they conducted flight tasks with RTA obligations at each scheduled waypoint in the simulated environment to differentiate the operational perception while interacting with different flight deck display settings. The pilots rated their perceived mental workload after conducting simulated RTA operations with all four display conditions that had been applied to the first phase of study.

Experiment 2: Evaluation of Perceived Workload

Objective

The objective of Experiment 2 was to evaluate pilots' perceived mental workload for the different display proximity levels. In this experimental phase, pilots conducted simulated flights in the different display proximity levels.

Method

Experimental Design

This experiment was a one-factor design with four display proximity levels. The display proximity levels were the same as those applied in Experiment 1. The applied four proximity levels were shown again as Figure 1, Figure 2, and Figure 3 as references for this second part of paper. The dependent variable (DV) was subjective mental workload ratings using the Modified Cooper-Harper Scale (MCH).

Hypothesis

There will be a significant difference in the mental workload rating among display proximity levels.

Participants

Fourteen pilots participated: 1 female / 13 males; 9 commercial pilots / 5 private pilots, 13 instrument ratings / 1 no instrument rating; Mean Age = 42 years, Age Range = 20 ~ 63 years old, STD = 15.50 years. Their mean flight time was 3438 hours (Range of Flight Hour = 46.50 ~ 15000 hours, SDev = 3942 hours). Due to the difficulty recruiting licensed pilots, the same pilots from Experiment 1 were asked to participate in Experiment 2. However, five pilots from Experiment 1 were unable to participate. Four new pilots were recruited in addition to the 10 who completed Experiment 1. We believed that adding additional subjects was important for increasing statistical power. All pilots required training for this experimental phase, and the experimenter varied the scenarios so that pilots could not memorize them.

Therefore, we determined that the mental workload for this experiment should not be highly affected by adding the additional new participants. No compensation was provided for participation. Institutional Review Board (IRB) approval was granted and followed throughout this research.

Apparatus

The basic apparatus for Experiment 2 was identical to that applied in Experiment 1. However, unlike Experiment 1, the yoke and throttle were functional to fly the aircraft and pilots were instructed to fly and meet assigned RTAs at every waypoint. The simulation started with the aircraft already en route.

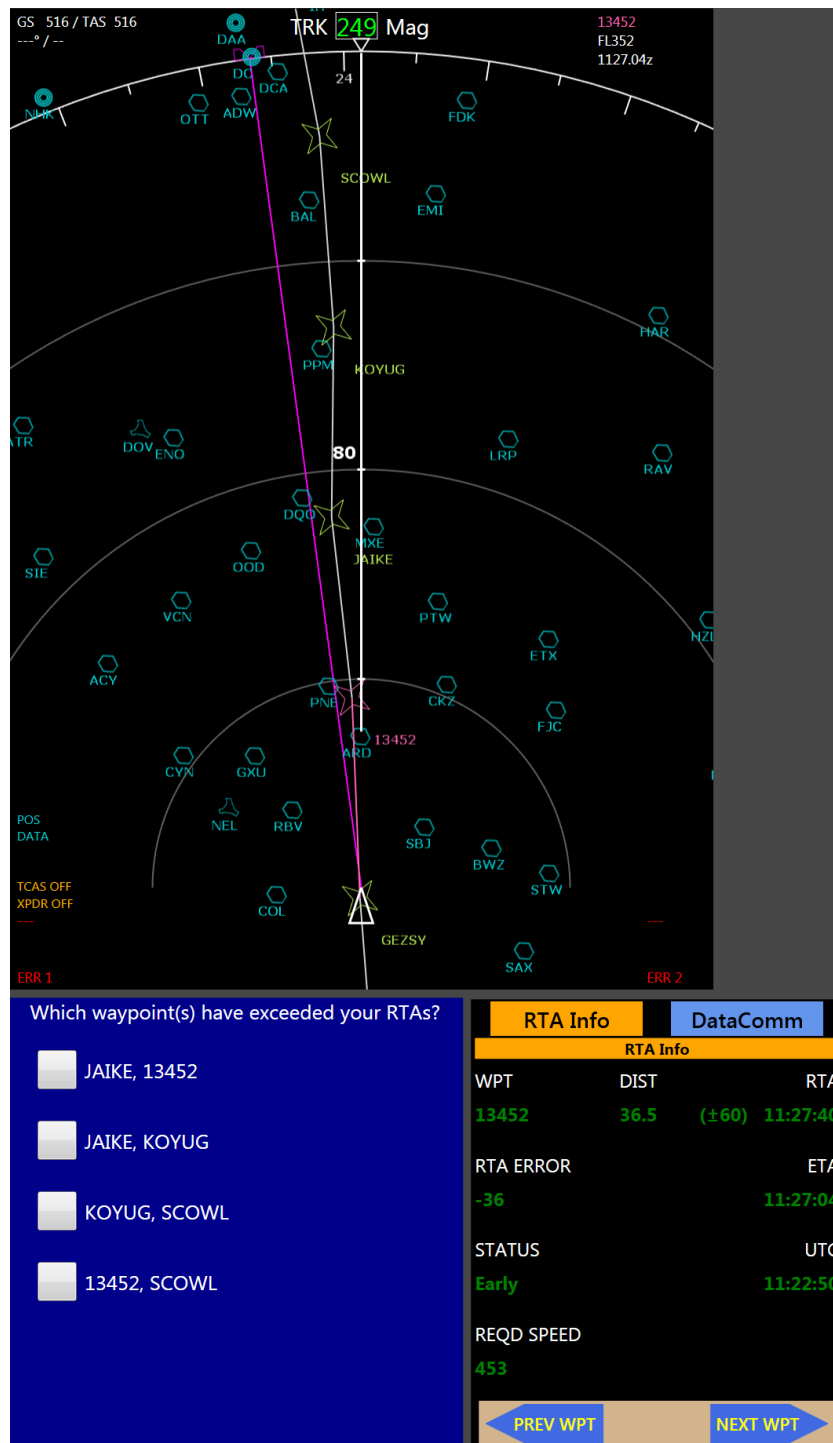


Figure 1. Low Proximity Condition. Bottom CDU is Optional Datalink Communication (DataComm) Window.



Figure 2. Medium Proximity-Text with CDU and Medium Proximity-Graphics Adopting Separated CDU with Graphics (Right); Bottom CDU is Optional DataComm Window.

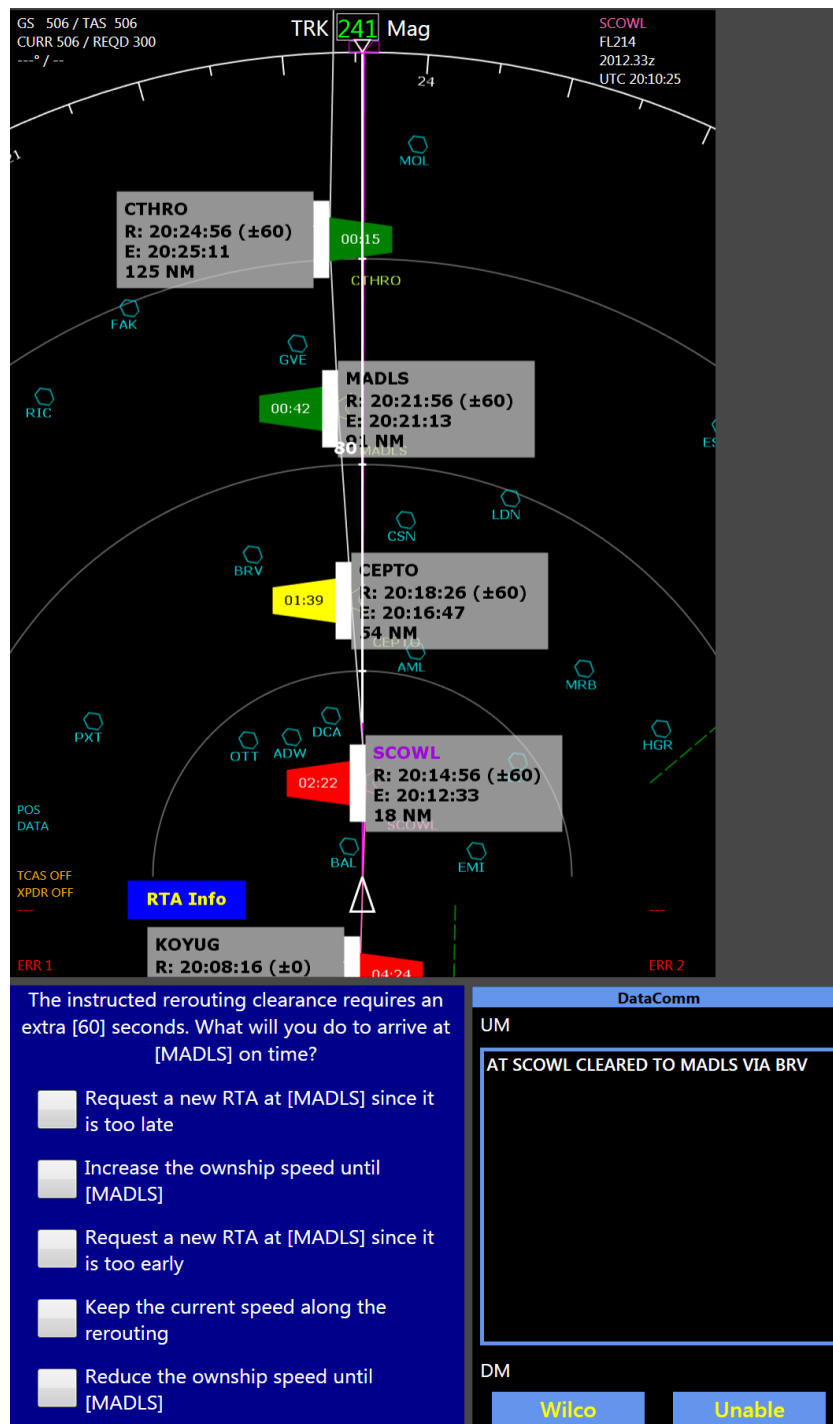


Figure 3. High Proximity Condition Showing Added Design Elements with No CDU as a Time Information Source.

Scenarios

The flight plan sets applied in Experiment 1 were reused for this phase. However, each scenario had four successive “next” waypoints that were used as measurement points within the scenario. The starting point for each scenario began immediately after passing a “previous” waypoint. Using this scenario, all four flight route segments per display proximity level session were defined as “the closest distance from waypoint A to waypoint B.” Each display proximity level used different waypoints. Each pilot participated in four simulated flights, one for each display proximity level. The order of display proximity level was randomized and an order repetition was carefully avoided. The RTA tolerance could vary between two levels (± 8 seconds or ± 10 seconds) so that the pilots would not know the specific tolerance in advance, which required them to closely monitor the RTA information. The tolerance values were based on the very short simulation scenarios.

No comparison of performance was conducted between the two levels. The initial temporal status (RTAs and ETAs) varied for each scenario. For example, a scenario may have started with either late or early condition to the next waypoint. The conditions of early or late would vary between waypoints. Each scenario included a segment between the 3rd and 4th waypoint that required the pilot to respond to an ATC textual datalink communication (DataComm) message asking the pilot to reroute, as an added task. The experimenter prepared two simple spatial rerouting message sets and selected one for each display proximity level. The clearance messages were taken from the Special Committee 214 (SC214) standard, which was the official FAA committee for DataComm message creation and evaluation for National Airspace System (NAS). One message was “AT [position] OFFSET [specified distance] [direction] OF ROUTE.” The second message was “AT [position] CLEARED TO [position] VIA [route clearance enhanced].” The information specific to the route was placed in the brackets for the experiment. For example, the second message can create: “AT [SCOWL] CLEARED TO [MADLS] VIA [BRV]” (the situation illustrated in Figure 2 and 3). The pilot was required to evaluate the clearance based on the flight plan and answer WILCO (Will Comply) or UNABLE on the CDU. The purpose of adding this task within the flight was to provide a more realistic situation where the pilot was performing additional tasks, and the ATC clearance would directly influence the RTA tasks.

Procedures

The experimenter instructed participants to maintain their altitudes at 35,000 feet. Pilots flew the aircraft through four waypoints with an obligation of meeting RTAs in each waypoint. They were required to increase or decrease speed to meet the RTAs. Between the third and fourth waypoint, a DataComm clearance was sent to the aircraft using the DataComm mode of CDU. After answering WILCO or UNABLE, pilots continued to fly until passing the fourth waypoint and the trial ended. Upon completion of the flight session within each display proximity

level, pilots were asked to provide a mental workload rating from 1 (lowest) to 10 (highest) for the display proximity level using the MCH, based upon their retrospective perceptions. Based on the definition of MCH, pilots could easily follow the given MCH flow-chart to select one level among 10 according to their perception without any significant training. After providing a rating they moved to the next randomly assigned display proximity level trial. They were allowed to rearrange their MCH ratings after finishing the second or later display proximity level sessions and comparing the results with prior ones. The experimenter collected the four rating numbers per individual pilot.

Experiment 2 Results

The MCH ratings were analyzed using a one-way ANOVA for the display proximity level. The main effect of display proximity level was statistically significant; $F(3, 24) = 8.10, p = 0.02$. Tukey's test results revealed that the MCH rating for the High Proximity ($\bar{x} = 1.93$) was significantly lower (lower perceived workload) than the Low Proximity ($\bar{x} = 3.50$). Figure 4 indicates the workload rating results. The letters above the graph indicate the Tukey grouping letters assigned to indicate the different workload levels. The MCH rating technique defines acceptable workload to be a rating of 3 or below (Gawron, 2000).

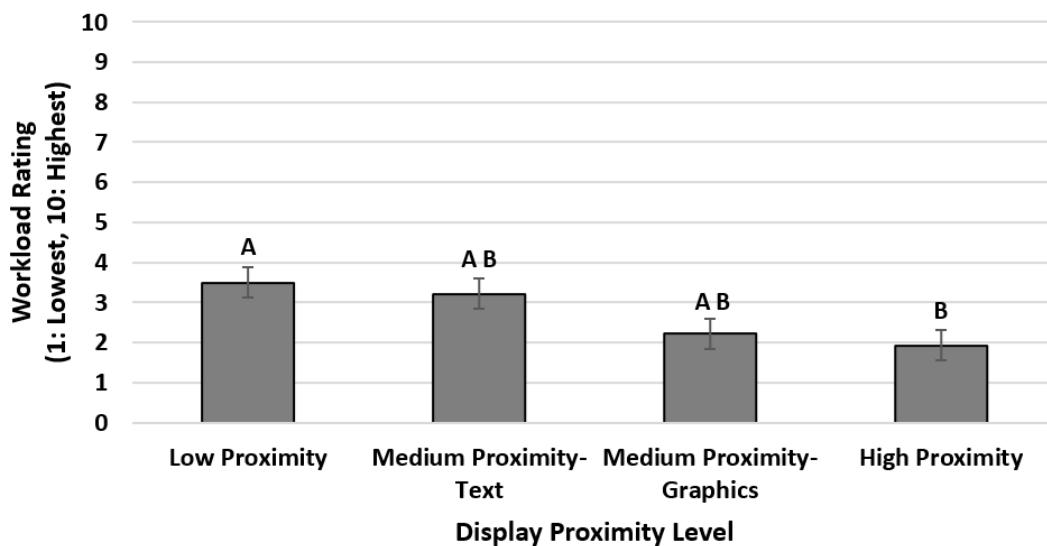


Figure 4. Mental Workload Rating Results for Experiment 2 Using the MCH with Tukey Grouping Identifiers above Each Bar. Error bars are added.

Ten pilots who participated in both Experiment 1 and Experiment 2 indicated they monitored the information somewhat differently during Experiment

2. They stated that the required speed information shown on the novel CDU of the Medium Proximity-Graphics and on the ND of the High Proximity was helpful to conduct the simulated RTA tasks during Experiment 2. They had not needed to focus on the speed information during Experiment 1 because the aircraft was on autopilot mode. Pilots also indicated that they paid more attention to distance information to next waypoints during Experiment 2. Their comments included that High Proximity was beneficial because they could view full future space-time situations at multiple waypoints using a single screen.

Discussion

General Overview

This section discusses a comprehensive implication derived both from the first (Experiment 1) and second (Experiment 2) phases of study. Table 1 specifies the results of Experiment 2 hypothesis testing. Only the display condition integrating all information necessary for RTA operation onto the ND (High Proximity) provided with lower mental workload than the traditional display. However, the two Medium Proximity conditions did not show any advantage with respect to the mental workload. Even the MCH ratings of the Medium Proximity-Text was above the acceptable workload range along with the Low Proximity.

Table 1
Hypothesis Testing Results of Experiment 2

Number	Hypothesis	Reject / Not Reject
Hypothesis 10	[Mental Workload Ratings] There will be a significant difference among display proximity levels.	Not Reject

The following sections describe what were implied from the entire study.

Display Proximity Level

Table 2 provides a comparison of design differences among the four display proximity levels.

Table 2
Comparison of Display Proximity Level

Criteria	Low Proximity	Medium Proximity-Text	Medium Proximity-Graphics	High Proximity
Level of Displayed Space-Time Information	Low proximity of space & time info. Pilots must compute the predicted status.	Time data added to ND (duration), Traditional CDU.	Time data added to ND (duration). Graphic indication of temporal conformance on CDU for quick indication of early or late status.	All space-time information on presented ND and temporal conformance graphic.
Visual Information	Text, baseline	Same Text as Low, added graphic	Same Text as Low, added graphic	Text & graphics combined (but text can be removed)
Amount of Required Manipulation	Baseline, move between screens	Locate some info on ND or move between screens if not on ND	Locate some info on ND or move between screens if not on ND	No need to move across screens to see info on waypoints
Attention to Monitor Space-Time Information	ND and CDU	ND and CDU	ND and CDU, quick glance to see time conformance on CDU	ND only

Through Experiment 1 and Experiment 2, pilots perceived the High Proximity to be the easiest condition to use. Pilots commented that they liked the ability to view the temporal status of multiple waypoints on a single ND in this condition. Although it was limited to be shown in the objective measure, the effectiveness of design strategy for High Proximity (minimizing the distance between space and time information, and graphical indication of temporal conformance) was successfully shown in the subjective difficulty and mental workload rating in this study. As analyzed in Table 2, all space and time information were presented in a single display, pilots may have saved their time to search and mentally integrate the two pertinent information elements. Texts and graphics that were processed in the heterogeneous cognitive channels (Schnotz, 2005) could be well-supported for the space-time SA when they were spatially close to each other. It might also eliminate the cumbersome activity of manipulating the CDU screen to search for the time information at any waypoint and did not require eye moving between ND and CDU: the study of eye-gaze data in this area

needs to be conducted for this perspective. However, as the amount of information on a display increases to include the necessary text information with graphics, the display may be perceived as more cluttered and increase search time. In this study, the High Proximity had a decluttering technique, which may have enabled pilots to avoid increasing the search time based on the response time (RT) results.

The Medium Proximity-Text and the Medium Proximity-Graphics also displayed time information for multiple waypoints on ND. However, in this study, the ND component of the two Medium Proximity displays presented duration information only; pilots still needed to view the CDU when they were required to check the clock time information. In addition, the ND component itself here did not have any temporal conformance indicator. The temporal conformance graphics in the novel CDU was appreciated by some pilots according to their feedback. However, the advantage of this design was limited: only the lower subjective difficulty ratings for SA2 questions.

SA Level of Question

For RT, participants took significantly longer to answer SA3 questions than SA1 and SA2 questions for all display proximity levels. This was expected because the concept of projection to the future requires more cognitive work. This difference was the same for subjective difficulty; the SA1 and SA2 questions were perceived easier than SA3 questions in all display proximity levels. The accuracy of SA3 question became lower in the three novel displays while the accuracies of all three SA levels were similarly high in Low Proximity display. This may be a negative aspect of novel display designs in this study. Enough training could improve the accuracy in the higher display proximity levels assuming no participants in this study had seen such novel displays before. No consistency was found between the SA1 and SA2 questions for the RT, accuracy, and subjective difficulty of all display proximity levels.

Expertise Level

Some expert pilots did not like the higher display proximity levels because they were already comfortable with the traditional display setting to conduct their flight operations. According to the interviews with expert participants in this study, experienced pilots had developed information monitoring patterns that may have been difficult to break. If novel displays were utilized, a high level of training to break old habits may have been necessary. However, new pilots may have no habit to break to interact with these novel designs.

Study Limitations

The time available to test pilots in HITL simulation flight tasks was limited. It may be very difficult to obtain pilots who can share enough time to be fully familiar with novel flight deck display designs, conduct a long series of query sessions, and conduct a long-time simulated flight tasks with RTA obligations. Making homogenous pilot groups (e.g. experience level, pilot certification class) was also very difficult because the access to the pilot group without any compensation was limited. In Experiment 2, the time data at waypoints varied under the short segments of time. A full RTA flight or longer simulations may have induced better pilot immersion into the tasks providing different results.

There was also limited time for training with the novel displays. The need of training new cockpit display concepts had been shown with Battiste, Johnson, Johnson, Granada, and Dao (2007) and Lancaster et al. (2011). Training of the novel display use for the experiments may have not been enough to show actual accurate SA as they perceived. However, some pilots performed equally well with the novel cases compared to the traditional condition. This implies that the novel display concepts were not very difficult to learn. Further investigation into the amount of training that would be needed to use these displays and an in-depth evaluation of possible errors is warranted.

The incompatibility in RT and accuracy may be due to an arbitrary tradeoff between speed and accuracy. This research assumed the indication of ‘earliness’ should be on the left and the ‘lateness’ should be on the right by interpreting Ishihara, Keller, Rossetti, and Prinz (2008). This design could be controversial; other researchers could interpret the position of ‘behind’ and ‘ahead’ indication in the opposite way. The question set per level of SA have been created based on the definition by Endsley. However, it is possible that some questions are not 100% accurate to the definition in certain situations.

Overall Conclusions

The design strategy of providing close spatial proximity between space and time information and graphic indications of temporal conformance themselves did not show any objective evidence for improved RTA navigation, but they showed subjective evidence for it from this study. While the results do not show an unequivocal advantage over the traditional display condition, the novel display that integrated all space-time information on the ND (High Proximity display) resulted in similar SA levels. The novel display that integrated the space-time information fully in a single display were subjectively perceived to be easier to use than the traditional display. The conflict between the objective and subjective outcomes imply another phase of study needs to be conducted to evaluate the findings from the subjective outcomes of this study. It is recommended that concepts be evaluated

under enhanced test environments that implement additional external factors including pilot experience, training, workload, and more complex flight environments for extended flight times. Also, the test outcomes will be more realistic if a group of pilots who are qualified in full FMC/CDU task environment such as airline pilots or corporate jet pilots can participate in the study. Furthermore, the temporal indications of aircraft are also critical for ATC displays, so similar display settings with this study can be adapted for ATC displays to evaluate air traffic controllers' space-time situation awareness.

References

- Battiste, V., Johnson, W. W., Johnson, N. H., Granada, S., & Dao, A. Q. (2007, July). Flight crew perspective on the display of 4D information for en route and arrival merging and spacing. In *International Conference on Human-Computer Interaction* (pp. 541-550). Berlin, Germany: Springer.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 32-64.
- Gawron, V. J. (2000). *Human performance measures handbook*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Ishihara, M., Keller, P. E., Rossetti, Y., & Prinz, W. (2008). Horizontal spatial representations of time: evidence for the STEARC effect. *Cortex*, 44(4), 454-461.
- Lancaster, J., Riddle, K., Feyereisen, T., Olufinboba, O., Rogers, B., Gannon, A., ... & He, G. (2011). Trajectory based operations and the legacy flight deck: envisioning design enhancements for the flight crew.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. *The Cambridge Handbook of Multimedia Learning*, 49-69.