International Journal of Aviation, Aeronautics, and Aerospace

Volume 5 | Issue 5 Article 8

2018

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

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

Cover Page Footnote

This material is based upon work partially supported by the Federal Aviation Administration (FAA) under Cooperative Agreement No. DTFAWA-10-A-80021. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors 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

According to the Federal Aviation Administration (FAA), the demand of air passenger and cargo was projected to continue to grow by 2037 (FAA, 2017). This growth is expected to strain the current air transportation system. Next Generation Air Transportation System (NextGen) concepts are under development to increase U.S. airspace capabilities. One NextGen concept is to change from a three-dimensional trajectory (latitude, longitude, altitude) environment to a four-dimensional trajectory (4DT; latitude, longitude, altitude, time) based operation where specific spatial and temporal constraints are applied to better predict aircraft location at specific times. Better prediction would support the movement of traffic throughout the airspace more efficiently, increase the number of aircraft that can land, and reduce fuel waste. Using 4DT, aircraft will navigate waypoints along a flight route that will have required time of arrivals (RTAs) with specified tolerances.

Current commercial flight deck displays provide RTA and Estimated Time of Arrival (ETA) information within the Flight Management System (FMS) Control Display Unit (CDU) in a textual format. To view time information at waypoints pilots must move through multiple CDU screens requiring manual input. Information related to aircraft 3D position is displayed on the navigation display (ND) in a graphical layout. The textual ETA only for the next immediate waypoint is displayed on the top right of the ND. The cockpit developers and aviation communities have placed the ETA information for all scheduled waypoints on separate pages of the CDU before the advent of the RTA concept. When RTA operations become required, the frequency of information use and current presentation formats may not be best-suited to support operations. Pilots must integrate pertinent but separated information sets for successful situation awareness (SA) of space-time during 4DT-based navigations. Under the current configuration of separate spatial and temporal information, the ability to maintain SA related to time and space requires pilots to manipulate display screens to conduct cognitive work such as calculation of remaining time to arrive at the next waypoint, while also considering spatial locations. Text and graphics are processed through different cognitive channels (Schnotz, 2005) and spatial and temporal information is thought to be processed in different regions within the brain (Coull & Nobre, 1998).

A review of the literature on aircraft cockpit displays revealed that there was limited published research that focused on display design for enhanced SA of space-time and very few related to the RTA operations.

Krishnan, Kertesz, and Wise (1999) designed and evaluated three non-traditional time traffic displays indicating a volume of airspace around ownship. The first test condition was a 1st person perspective display using concentric circles that specify time intervals (in minutes) instead of spatial distance found on the traditional ND. Outside traffic was presented with respect to the time to contact ownship within the time circles. The second condition was a split screen with a lateral (latitude/longitude) display and a vertical (altitude) display where the distance from the ownship to traffic was indicated in minutes, not in a spatial distance. The third display was a 3rd person perspective display showing concentric circles projecting in front of ownship with traffic indicating time to pass in minutes on the circles. Experimental results evaluating the display concepts indicated that the 3rd person perspective display did not support pilot performance. However, the 1st person perspective display performed best for reaction time and error rates to determine which way to maneuver to avoid traffic. Pilots could potentially become confused between a spatial display showing miles and similar display showing time from a waypoint. This research example focused on time to contact another aircraft (closure rate) but was not designed for navigation using RTA.

NASA developed an enhanced cockpit display of traffic information (CDTI) (Battiste, Johnson, Johnson, Granada, & Dao, 2007) that presented an innovative graphic representation of ownship with surrounding air traffic in a three-dimensional (3D) perspective to facilitate detection of a collision threat. The 3D CDTI displays textual data for waypoint name, RTA, and ETA in a small table at the bottom right. Operators could manipulate the information on the table directly. The 3D CDTI employed a design for temporal spacing status indication using ‘a rectangle indicating the ownship spacing goal with a straight line drawn from the middle of the rectangle to the distant traffic attaching time intervals in seconds.’ Color coding was used to indicate early, late, or within tolerance. If the pilot was behind a required spacing goal, a white rectangle was placed in front of the ownship symbol. If the pilot was within the specified time tolerance, a green rectangle surrounded the ownship. If the pilot was ahead of the spacing goal, a yellow rectangle was placed behind the ownship. Eight pilots conducted a merging and spacing task using the 3D CDTI during en route and arrival operations. According to their subjective feedback, they would be willing to perform merging and spacing tasks with appropriate training and generally agreed that the CDTI was beneficial and usable for acquiring and tracking traffic. The 3D CDTI was different from current ND cockpit displays and evaluated for far-term use in the NextGen airspace. The concept will require significant training and does not provide a direct solution to near-term or midterm use. However, several design concepts could be integrated into existing displays or on auxiliary displays.

Lancaster et al. (2011) redesigned an ND as a Graphical Flight Planning (GFP) display, a CDU, and an electronic flight bag (EFB; the secondary electronic display employed by most airliners that is placed in left and right corners of cockpits) to support RTA operations. Throughout their design iterations they considered the use of NASA’s time symbology for the GFP. The final GFP design was composed of a north-up map display (on the left) and a table of scheduled waypoints (on the right). The map display component included a 'leaned picnic table' symbol placed immediately over the ownship to indicate its temporal status to RTAs. This symbol had been designed by NASA Langley Research Center (Barmore, Abbott, and Krishnamurthy, 2004 as cited by Lancaster et al.). When the aircraft was within the RTA tolerance, the symbol appeared as a straight line. If the aircraft was out of tolerance (early or late) the legs of the picnic table pointed toward the ownship. In addition to the symbol, a separate text tag to indicate the next waypoint and current airspeed was presented on the map display. A time scale was provided on the edge of a circle surrounding the ownship indicating the travel time from the center of the circle to the edge. The right table included scheduled waypoint lists and altitudes, required speeds, and ETA/RTAs to the waypoints. The CDU was modified to the format that included the texts "Early" and "Late" for easier understanding of temporal status. The EFB also included an information pane with a list of waypoints. Pilots could touch an interested waypoint to see a pop-up window presenting RTA, ETA, and the difference between ETA-RTA with the "Early" or "Late" indicator. The EFB provided a heading-up redundant map display with integrated ETA and RTA and the delta between RTA an ETA as text placed close the ownship symbol. Pilots could change the zoom level to see multiple waypoints and to support reading the text.

Lancaster et al. conducted human-in-the-loop (HITL) simulation tests with scenarios of RTA operation by seven pilots to evaluate the use of all three displays. According to the pilots' subjective responses, they liked the interactive capability of the GFP display and the familiarity of the CDU. The mental workload ratings were within the acceptable ranges for these two displays. The EFB display design was not well-accepted due to the interaction style and lack of familiarity. Pilots indicated they wanted to understand the reason they were not going to meet the RTA and a current time indicator in a prominent position. Pilots also perceived that the text-information display on the GFP appeared complex and somewhat cluttered. The amount of training that might be needed for this concept was not known, but pilots did mention the need for training. With this design, more information was available on the GFP. However, pilots still needed to manipulate and cognitively process RTA data on the CDU for every waypoint. Pilot feedback related to the picnic table and RTA tags varied and included possible design changes such as adding color rather than just legs on the picnic table symbol to indicate earliness or lateness. The authors noted that the feedback from pilots was useful but stated the need for objective performance measures based on the varied responses to the different designs employed.

An important consideration for RTA operations is that under the current cockpit setting, pilots have to divide their attention to separate ND and CDU to obtain space and time information and then mentally integrate for the successful

SA. One way to provide an easy access to both information sources visually is to place time and space information proximal to each other (e.g. spatially close to each other) assuming that users can reduce the eye movement time. It is also important to understand what kind of display configuration pilots are helped better, commit less error, and perceive easier for their navigation operations. Currently RTAs and ETAs are provided in the clock time standard, and pilots calculate the differences between them to determine if they are early or late and by how much. A quick indication of earliness or lateness to assigned RTAs may provide a more immediate decision related to whether steps must be taken to meet RTA.

Development of Space-Time Displays

The objective of this study was to develop and evaluate novel flight deck display formats to support RTA operations for near to midterm NextGen. The previous study findings (Krishnan et al. ; Battiste, et al., ; Lancaster et al.), along with human factors design considerations, were used as input to the design of three formats which were evaluated against the control condition; the traditional ND with spatial information and CDU for temporal information. The traditional ND and CDU were redesigned to create the new concepts. Two of the concepts provided RTA task information on both the ND and CDU. One concept integrated all RTA task information only on the ND.

Navigation Display Designs

Figure 1a is a traditional ND illustrating waypoints along the aircraft flight route. A change to the ND illustrated in Figure 1b placed RTA (**R** on the ND) and ETA (**E** on the ND) information next to each waypoint.

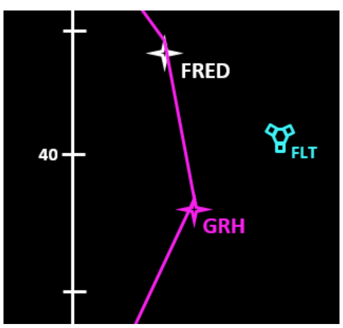

Figure 1a. Traditional ND

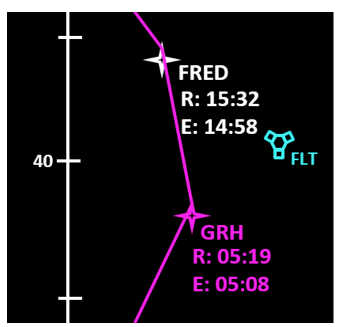

Figure 1b. Novel ND Adding RTA/ETA Texts

The time format was the 'duration' to arrive at a waypoint from the current location (i.e. the response to "how long does it take from here to there; travel time" in the [min:sec] format) rather than clock time (i.e. the response to "what time is it when I get there" in the [hr:min:sec] format). The duration until the next scheduled waypoints in the provided map scale did not exceed one hour in this study. The reason for adopting this time format was to provide a different aspect of time information while the CDU component still provided the clock time information. In the Figure 1b example, the required (“**R**”) duration to reach waypoint “FRED” is 15 minutes and 32 seconds, and the estimated (“**E**”) duration to reach the waypoint is 14 minutes and 58 seconds; this means that the aircraft should arrive at “FRED” in 15 minutes and 32 seconds, but it is actually estimated to take 14 minutes and 58 seconds to arrive at the waypoint.

Figure 2 illustrates one design that provides another close display proximity by integrating all required information onto the ND. The temporal information in Figure 2 shared the common scenario as Figure 1b (same UTC: Coordinated Universal Time, RTA and ETA at each waypoint).

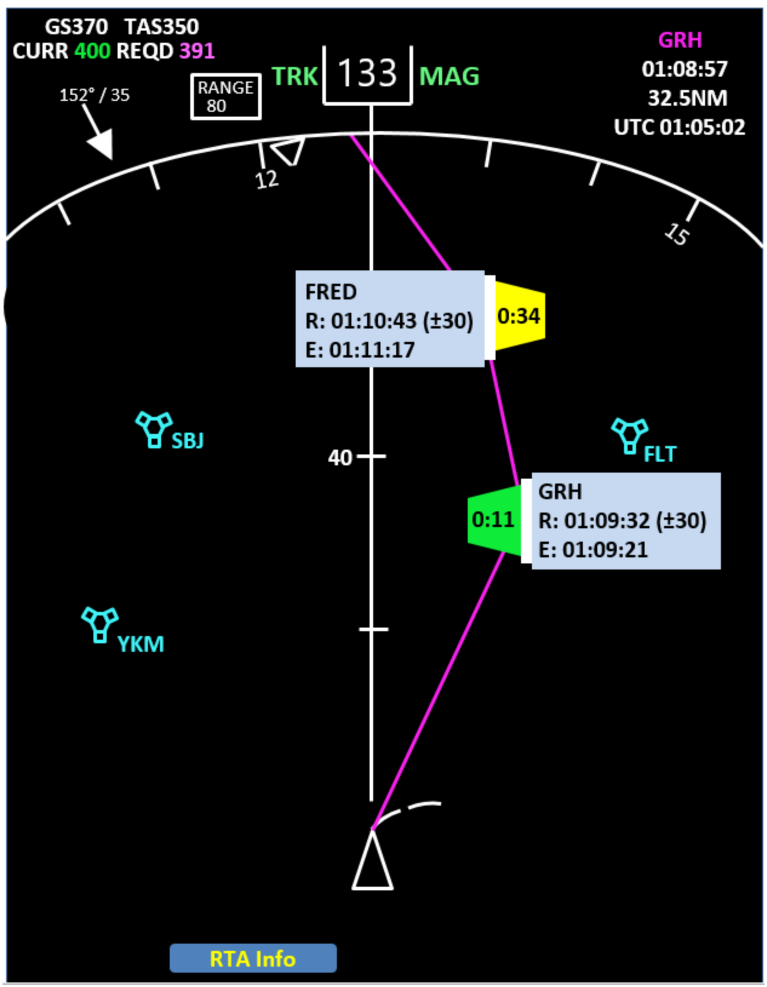

Figure 2. Novel ND Adding Trapezoid Graphics

The display provided an indication of being early or late using shape and color coding. The display used a trapezoid shape to indicate earliness and lateness. If aircraft was early, the shorter side of the trapezoid pointed to the left. If late, it pointed to the right. The number inside the trapezoid specified the difference between RTA and ETA ([min:sec] format). The trapezoid color showed differently as the temporal conformance to assigned RTAs changed: green (on schedule), yellow (slightly early or late), and red (significantly early or late). The display added a time data box to present RTA and ETA by each trapezoid. The pilot could turn off the box when they needed display decluttering. This display kept the RTA and EAT in the clock time format because it had no CDU component. It did not have the duration information support. Figure 2 illustrates a situation where the pilot will be on time at waypoint "GRH", but slightly early at "FRED" (i.e. the difference between RTA and ETA is 34 seconds as indicated inside the trapezoid, which is longer time than ±30 seconds, the RTA tolerance in the parenthesis next to the RTA).

Control Display Unit Designs

Figure 3 shows a traditional CDU illustrating the temporal status which scenario is compatible with Figure 1b and Figure 2 for the sample waypoint "FRED."

Figure 3. Traditional CDU

A redesign of the traditional CDU format by adding graphics as a temporal conformance indicator created a novel CDU as illustrated in Figure 4. The UTC and RTA within the left example of Figure 4 were the same as Figure 1b, Figure 2, and Figure 3. The right example revised the ETA to represent a significantly early status.

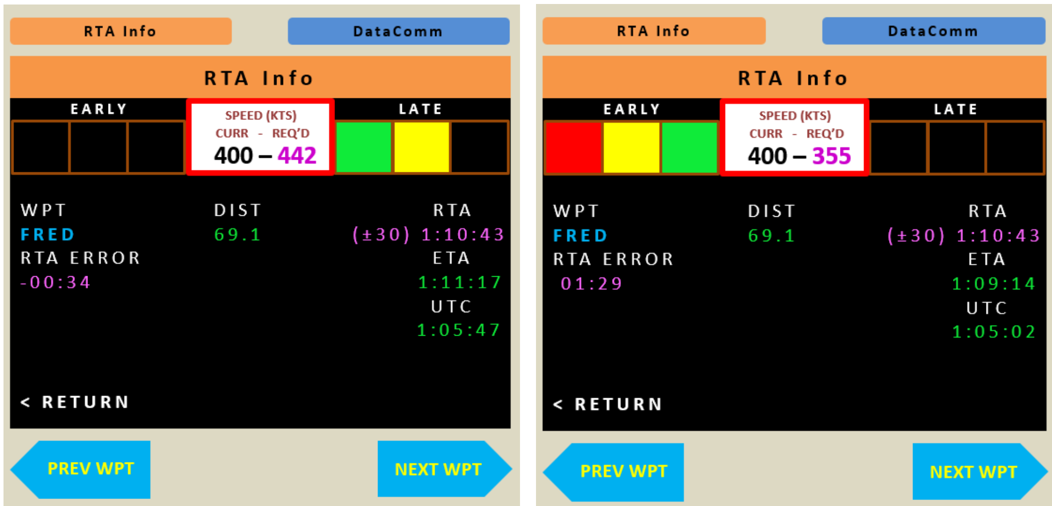

Figure 4. Novel CDU Prototype with Graphical Temporal Conformance Indicator (Left: Slightly Late, Right: Significantly Early)

In Figure 4, the center box by a series of color block indicates current and required speed so that pilots can maintain the "on time" status. The three blocks filling color toward the left or right of the center box form a bar similar to sound level bars, to provide a quick indication of the aircraft temporal conformance status. The bar was presented to the left side if aircraft was early, and to the right if aircraft was late. If aircraft was 'on time', the graphic showed the first block from the center to the right or left in green. If aircraft was 'slightly early or late', the second block turned yellow while the green block remained. If aircraft was 'significantly early or late', the third block turned red while the yellow box and green box remained. This novel CDU still provided textual time data for the waypoint except the data of status and required speed under the graphics. The reason that the graphics used a horizontal bar rather than a vertical bar was based on the research by Ishihara, Keller, Rossetti, and Prinz (2008). They discovered that most people perceived the representation of earliness and lateness using horizontal direction rather than vertical direction, and interpreted the left as early and the right as late. For both the traditional CDU and the novel CDU, pilots had to move between screens to view future or past waypoints by choosing "PREV WPT" (previous waypoint) and "NEXT WPT" (next waypoint) buttons at the bottom. A Textual Datalink Communication (DataComm) page was also available on a separate tab. Four combinations of these different ND and CDU designs created four display conditions for this study.

This study applied these designs to the evaluation of pilots' SA of space-time and perceived mental workload during HITL simulation tests with RTA navigation scenarios through two phases of experiment.

Experiment 1: Evaluation of Situation Awareness

Objective

The objective of Experiment 1 was to evaluate pilots' SA of space-time using the novel displays with higher display proximity between spatial and temporal cues and additional graphical representations of temporal conformance, compared to a traditional format with low display proximity.

Display Proximity Level

Based on the manipulation of display proximity level between space and time information stated above, a traditional display and three novel display conditions were created resulting in four display conditions. This study defined the four display conditions as Low Proximity, Medium Proximity-Text, Medium Proximity-Graphics, and High Proximity. All conditions had a question window under the ND and placed a CDU to the right of the question window.

Low Proximity referred to the traditional display condition with the ND (Figure 1a) and CDU (Figure 3) where space and time information was separated. To view RTA and ETA information, pilots had to select the waypoint of interest by choosing the previous or next waypoint on the CDU. The ND presented only space information (geographic location).

Medium Proximity-Text (see Figure 5 with the left CDU component) was composed of the ND shown in Figure 1b that included RTA and ETA text information for duration next to each waypoint. The traditional CDU shown in Figure 3 was used in this condition as the Low Proximity. When touching the DataComm tab on the CDU, the CDU component showed the DataComm page.

Medium Proximity-Graphics (see Figure 5 with the right CDU component) included the ND shown in Figure 1b and the novel CDU design in Figure 4. The ND and the DataComm page of the CDU component were the same as Medium Proximity-Text.

High Proximity (see Figure 6) eliminated the use of the CDU as the time information source. The ND included the novel designs illustrated in Figure 2. Pilots could touch the "RTA Info" Boolean button at the bottom of the ND to make the time information box next to the trapezoid visible or invisible. The trapezoids were always visible as the waypoint symbols that had RTA obligations. The current time (UTC) was added at the top-right. The current and required speeds to meet the RTA requirement until the next immediate waypoint were added textually at the top-left of the ND. With all information integrated on a single display, the display proximity was defined to be "high." Pilots could monitor all space and time information via the ND. The CDU component in this condition only provided a DataComm information page.

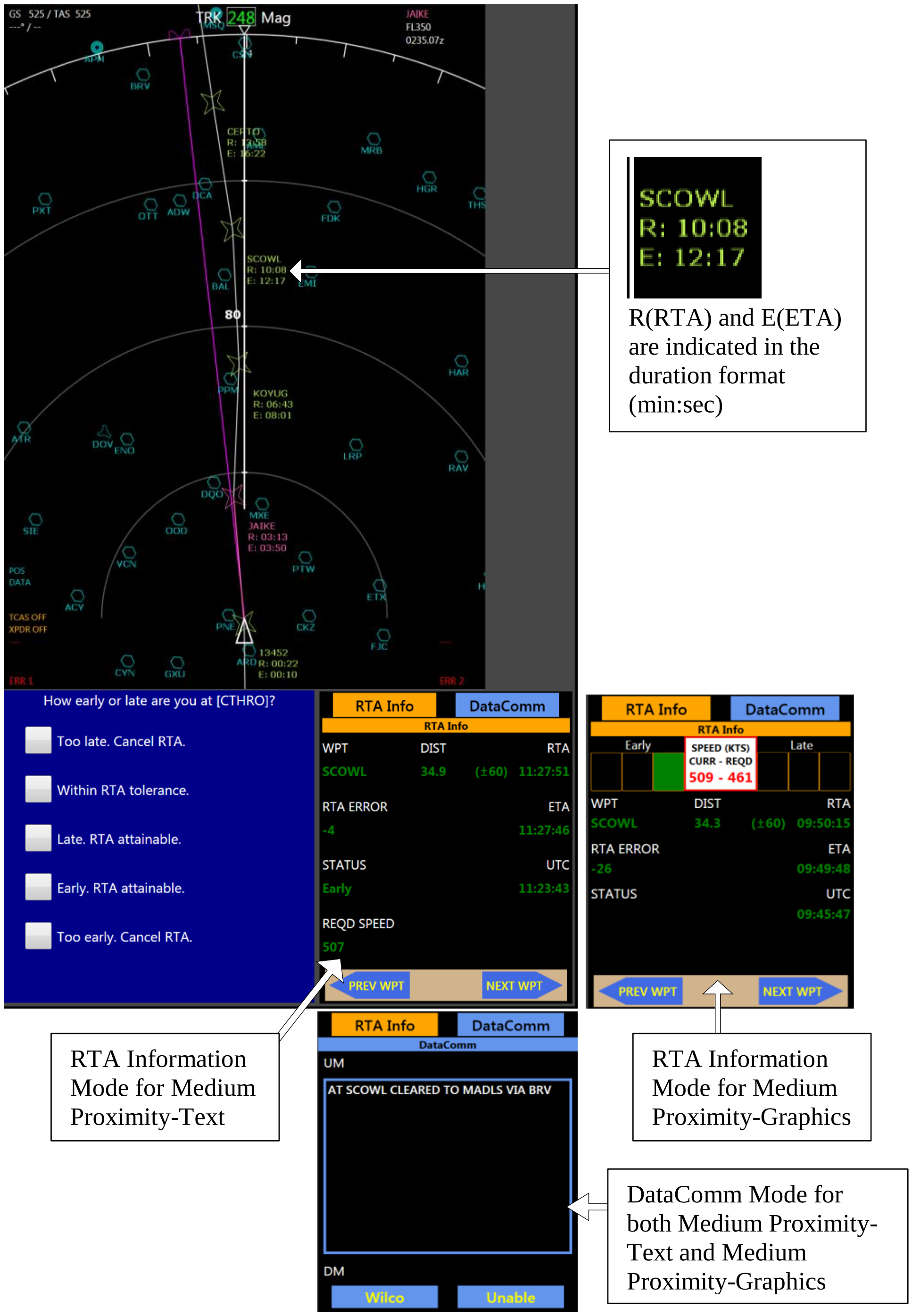

Figure 5. Medium Proximity-Text with CDU and Medium Proximity-Graphics Adopting Separated CDU with Graphics (Right); Bottom CDU is Optional DataComm Window.

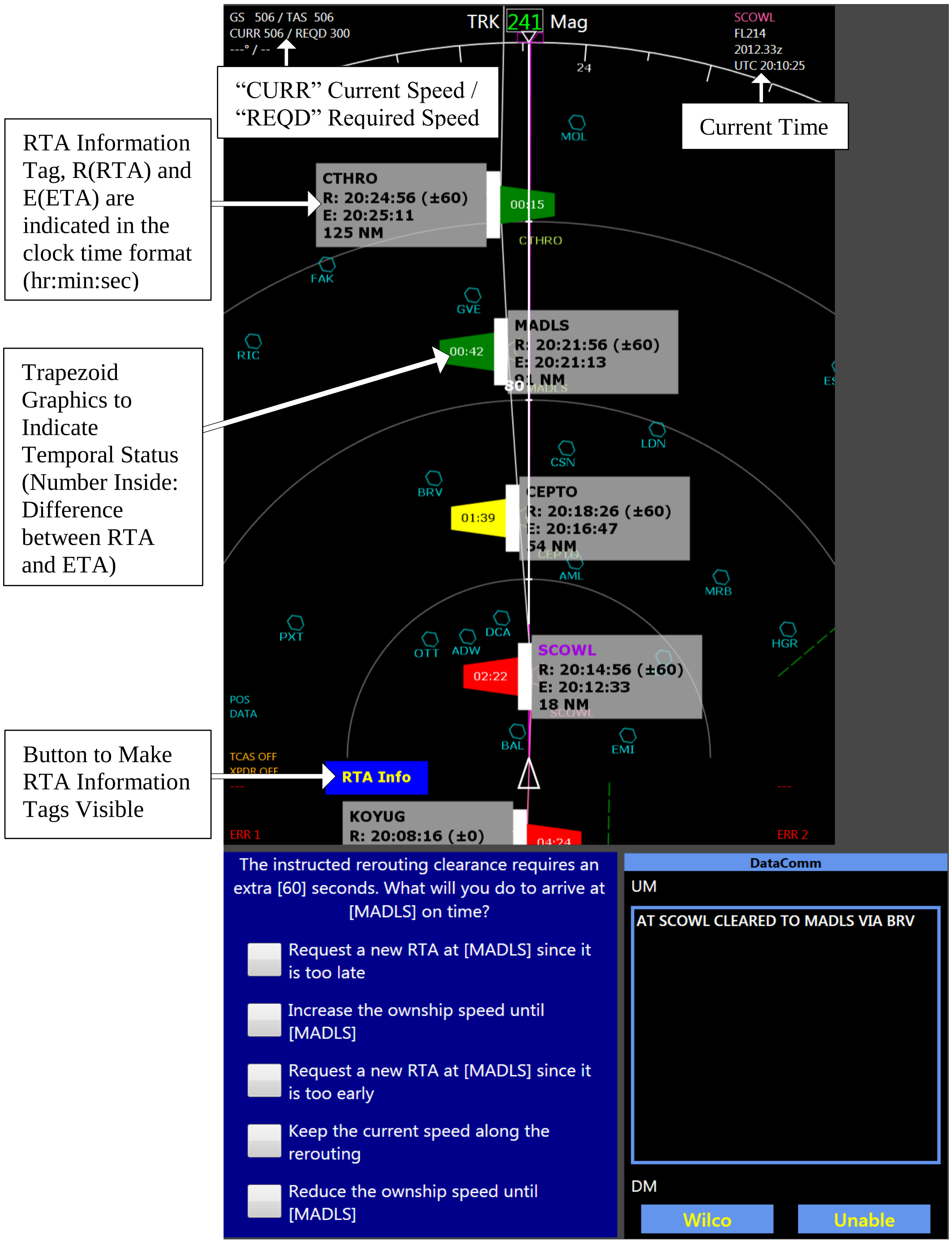

Figure 6. High Proximity Condition Showing Added Design Elements

Method

experimental design. This experiment was a 4 (display proximity level) x 3 (SA level question ask) within-subjects design. The display proximity level was already explained. The experimenter created SA questions to evaluate the SA considering pilots' expected perspectives while focusing on space and time information for RTA tasks. The SA level question ask was manipulated by the researchers based on the three levels of SA defined by Endsley (1995). The SA levels were differentiated to evaluate how participants performed the given tasks differently per SA level. The questions are shown in Table 1.

Table 1
List of SA Question for Experiment 1

SA Level	[Question ID] Question
SA1: Level 1 SA (Perception)	**[SA1 Space]** At which waypoint must you arrive between [Specific Clock Time] and [Specific Clock Time]? **[SA1 Time ET]** What is the ETA at [Specific Waypoint]? **[SA1 Space Status]** Which waypoints have exceeded your RTAs?
SA2: Level 2 SA (Comprehension)	**[SA2 Time TL]** How long will it take to arrive at [Specific Waypoint]? **[SA2 Time Conform]** How early or late are you at [Specific Waypoint]?
SA3: Level 3 SA (Projection)	**[SA3]** The instructed rerouting clearance requires an extra [Specific Seconds] seconds. What will you do to arrive at [Specific Waypoint] on time?

The *SA1* question group (three question types) asked the pilots' 'perception of elements' in the given situation as the definition of level 1 SA (Endsely, 1995). The *SA1 Space* question asked pilots to indicate the waypoint (space information) pertinent to a certain ETA. The *SA1 Time ET* question asked pilots to find the ETA of a specific waypoint. The *SA1 Space Status* question asked to determine which waypoint exceeded an RTA (i.e. going to be early). Participants selected one answer from four choices presented on a question window.

The *SA2* question group(two question types) referred to 'the comprehension of the given situation' as the definition of level 2 SA (Endsley). The *SA2 Time TL* question asked about the estimated duration from the current location to a specific waypoint. This question required calculation using current time and ETA with the Low Proximity display. It had four answer choices. The *SA2 Time Conform* question asked about the temporal conformance related to an RTA at a specific waypoint. There were five possible selections to answer this question: "Too Early, Cancel RTA", "Early, RTA Attainable", "On Time", "Late, RTA Attainable", and "Too Late, Cancel RTA."

The *SA3* question (single type) referred to 'the projection to future states' as the definition of level 3 SA (Endsely, 1995). For the *SA3* questions, the CDU presented a DataComm text clearance message to request for exploring alternative flight routes. The participants were required to find alternative flight routes by themselves and then determine if they could comply with the rerouting within the given RTA. This question had five answer selections: "Keep the current speed until the waypoint", "Reduce the speed until the waypoint", "Increase the speed until the waypoint", "Cancel the RTA since it is too early", and "Cancel the RTA since it is too late."

All these six question types had only one correct answer. The experiment asked one question type three times varying scenarios during each display proximity level session, resulting in 18 questions per display proximity level, for a total of 72 trials per participant. The dependent variables (DVs) were (a) the response time (RT) to answer the SA question, (b) the percent of correct answers (accuracy), and (c) the perceived difficulty for each question type within each display proximity level; this was to evaluate the consistency between objective measures and subjective measures. The evaluation of separated cues vs. integrated cues within a display by measuring RT and accuracy of answers to SA questions had been conducted by Koch et al. (2013) for a healthcare application.

hypotheses. Table 2 specifies the hypotheses with respect to three DVs for Experiment 1.

Table 2
Hypotheses of Experiment 1

Number	Hypothesis Specific to [DV]
Hypothesis 1	[RT] There will be a significant difference among display proximity levels.
Hypothesis 2	[Accuracy] There will be a significant difference among display proximity levels.
Hypothesis 3	[Subjective Difficulty Ratings] There will be a significant difference among display proximity levels.
Hypothesis 4	[RT] There will be a significant difference among SA levels.
Hypothesis 5	[Accuracy] There will be a significant difference among SA levels.
Hypothesis 6	[Subjective Difficulty Ratings] There will be a significant difference among SA levels.

Hypothesis 7	[RT] There will be an interaction between SA level and display proximity level.
Hypothesis 8	[Accuracy] There will be an interaction between SA level and display proximity level.
Hypothesis 9	[Subjective Difficulty Ratings] There will be an interaction between SA level and display proximity level.

participants. Fifteen licensed pilots participated in the experiment (2 females / 13 males; 11 commercial / 4 private pilots, 14 instrument / 1 no instrument rating; Age Range = 21 - 63 years old, Mean Age = 45 years, STD = 13.8 years). Their mean flight time was 3697 hours (Range of Flight Hours = 76 - 15000 hours, STD = 3974 hours). Due to the difficulty in finding licensed pilots, it was not possible to create a homogenous group with respect to flight hours. The levels of flight experience varied among the participants. They were not compensated for their participation. Institutional Review Board (IRB) approval was granted and followed throughout this research.

apparatus. This experiment used a medium-fidelity PC-based flight simulator. It employed X-Plane® flight simulation program (Version 10 Professional). The simulator had two touch screen monitors (50 cm wide x 30 cm tall for each). The left screen was presented in landscape mode and provided the Boeing 777 instrument panels from the X-Plane program. The right screen, in portrait mode, displayed the ND and CDU that were developed for this research, along with the question window. The test ND was approximately 22 cm x 31 cm, the test CDU was 12.6 cm x 14.4 cm, and the question window was 13.7 cm x 14.4 cm. The outside visual world was presented on a large projector screen in front of the pilot. The outside visual world did not provide any specific reference for this experiment.

test scenarios. Authors developed flight plans using Goodway flight planning software (Ver. 4.0) for this study. The aircraft in these simulation flights was flown on autopilot mode through randomized flight scenarios. Each display proximity level included scenarios of flight situations with eighteen different route segments between waypoints. During the flight situation in each segment, the test program provided a different SA question on the question window. The three repetitions of a question type within each display proximity level used different waypoints and time data, and the same question type was not presented consecutively. The question order was randomized. For the SA3 question, the CDU displayed a DataComm clearance indicating a rerouting from the original path. Only the en route phase of flight was included for the scenarios.

procedures. The simulation depicted an en route flight for 20 seconds as the acclimation period, and then the test program presented a new question on the question window. The simulator made a sound when each question was presented on the screen. The ND and CDU were visible during the question presentation. The pilots' task was to monitor the flight and answer the SA questions as the flight progressed along the given route. Pilots touched the screen for one answer out of multiple selections. The question then disappeared and the test program automatically moved to the next segment. The pilot continued responding to the 18 queries and then moved to the next display proximity level until all 72 query trials (in all the four display proximity levels) were completed. The order of display proximity levels per pilot was randomized and an order repetition was carefully avoided. Flight scenario sets and question sets were randomized for each individual pilot. Upon completion of all four sessions, pilots entered another session of rating their perceived difficulty in an Excel table composed of display proximity level (row) and question type (column), using a scale of 1 (very easy) to 10 (very difficult). The experimenter provided them with a re-explanation of each specific display proximity level and question type so that they recalled their specific perceptions effectively. The pilots did not have any difficulty recalling their perceptions according to their feedback. The purpose of this subjective evaluation was to compare their perceptions with the objective measures. Pilots also provided oral feedback to the experimenter indicating what they liked and did not like about the displays.

Experiment 1 Results

RTs, accuracies (percent correct answers), and subjective difficulty responses were statistically analyzed. The RT and accuracy data had indications how participants objectively reacted based on their situation awareness in which they collected essential information from the given simulation screen. A post-hoc simple effects F-test was used to determine statistically significant interactions and Tukey's Honest Significant Difference (HSD) test was used for paired comparisons with α set at 0.05. Results of the ANOVA and simple-effects F-tests are presented in Table 3 and illustrated in Figures 7, 8, and 9. In the figures the letters above each graph indicate Tukey test groupings. Data point marks with the same letter indicate no statistically significant difference from one another.

Table 3
Results of ANOVA and Post-Hoc Simple Effects F-Tests for Dependent Variables

Description of Result	Result
Response Time (RT)	
Significant interaction between display proximity level and SA level of question	F (6, 79) = 2.27, p = 0.045
Significant difference (S.D.) among display proximity levels for SA1 questions (simple-effect).	F (3, 57) = 5.63, p = 0.002
Statistically significant main effect of RT for SA level of question.	F (2, 26) = 111.88, $p \leq$ 0.0001.
Accuracy (% Correct Responses)	
Significant interaction between display proximity level and SA level of question	F (6, 80) = 4.67, p = 0.0004
S.D. across display proximity levels for SA3 (simple-effect).	F (3, 345) = 38.67, $p \leq$ 0.0001
S.D. among different SA levels for three display proximity levels (simple-effect).	Medium Proximity-Text: F (2, 102) = 7.43, $p \leq$ 0.0001 Medium Proximity-Graphics: F (2, 99) = 14.14, $p \leq$ 0.0001 High Proximity: F (2, 104) = 10.39, $p \leq$ 0.0001
Significant main effect of display proximity level.	F (3, 57) = 3.81, p = 0.015
Significant main effect of SA question level.	F (2, 30) = 59.52, $p \leq$ 0.0001
Subjective Difficulty	
Significant interaction between display proximity level and SA question level.	F (6, 176) = 9.87, $p \leq$ 0.0001
Significant main effect of display proximity level.	F (3, 234) = 58.56, $p \leq$ 0.0001
Significant main effect of SA question level.	F (2, 168) = 88.06, $p \leq$ 0.0001

response time (**Figure 7**). There was a statistically significant interaction between display proximity level and SA level of question. For SA1 questions, the High Proximity ($\bar{x}$ = 15.62 sec) had a significantly shorter RT than Medium Proximity-Graphics ($\bar{x}$ = 24.99 sec). For SA2 questions and for SA3 questions, there were no statistically significant differences for RT across the display proximity levels. For all display proximity levels, SA3 questions resulted in significantly longer RT than SA1 and SA2 questions given the requirement of extra time to consider the reroute.

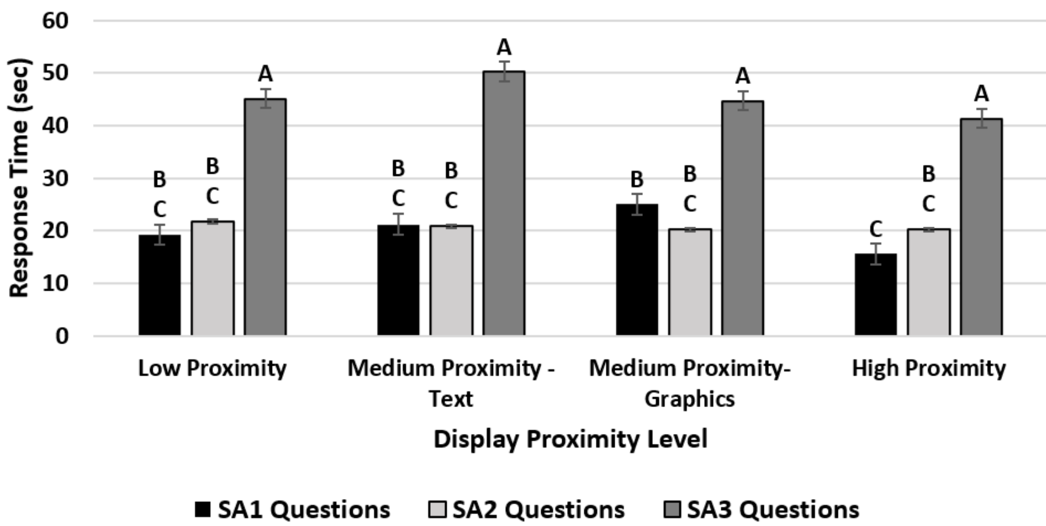

Figure 7. Mean RT with Tukey Grouping Identifiers above Each Bar. Error bars are added.

The main effect of SA level of question for RT was statistically significant; the RT for SA1 ($\bar{x}$ = 20.27 sec) and SA2 questions ($\bar{x}$ = 20.73 sec) were significantly shorter than the RT for SA3 questions ($\bar{x}$ = 45.34 sec). The main effect of display proximity level was not statistically significant. The main effects should be considered with caution given the significant interactions.

percent correct answers - accuracy (Figure 8). There was a statistically significant interaction between display proximity level and SA level of question for accuracy. For SA3 questions, the Low Proximity resulted in higher accuracy ($\bar{x}$ = 86.10 %) than the Medium Proximity-Text ($\bar{x}$ = 57.78 %). For SA1 and SA2 questions among all display proximity levels, there were no statistically significant differences in accuracy.

For the Low Proximity, there were no statistically significant differences in accuracy for all SA levels of question. However, within Medium Proximity-Text, a statistically significant difference was found between all three SA levels. Within the Medium Proximity-Graphics and High Proximity conditions, pilots answered SA1 and SA2 questions more accurately than SA3 questions.

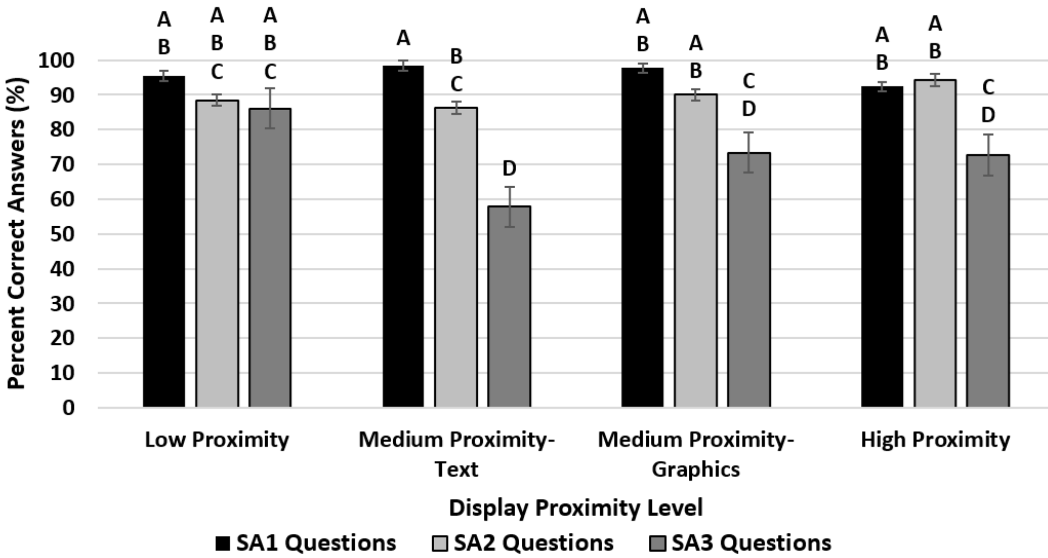

Figure 8. Mean Percent Correct Answers with Tukey Grouping Identifiers above Each Bar. Error bars are added.

Main effects for display proximity level and SA level of question were both statistically significant for accuracy. For display proximity level, the Low Proximity ($\bar{x}$ = 90.00 %) had significantly higher percent of correct answers than the Medium Proximity-Text ($\bar{x}$ = 80.82 %). All three SA levels of question were significantly different; SA1 resulted in the highest percent correct ($\bar{x}$ = 96.00 %) followed by SA2 ($\bar{x}$ = 89.75 %), then SA3 ($\bar{x}$ = 72.48 %). The main effects should be considered with caution given the significant interactions.

subjective difficulty (Figure 9). Results indicated a statistically significant interaction between display proximity level and SA level of question for the subjective difficulty response. The simple effects by display proximity level post-hoc F-test was statistically significant.

For the Low Proximity, pilots rated all three SA levels of question to be significantly different. As can be seen by Figure 9, for the Low Proximity, SA1 was rated as being the easiest ($\bar{x}$ = 3.44), followed by SA2 ($\bar{x}$ = 4.90) and SA3 ($\bar{x}$ = 6.40). For the Medium Proximity-Text, SA3 was rated to be significantly more difficult ($\bar{x}$ = 5.67) than SA1 ($\bar{x}$ = 3.20) and SA2 ($\bar{x}$ = 2.77). Also for the Medium Proximity-Graphics, SA3 ($\bar{x}$ = 5.13) was rated significantly more difficult than SA1 ($\bar{x}$ = 2.89) and SA2 ($\bar{x}$ = 2.20). There was no significant difference between SA1 and SA2 for the two Medium Proximity levels. For the High Proximity, SA1 ($\bar{x}$ = 1.24) was rated as significantly easier than SA2 ($\bar{x}$ = 2.60) and SA3 ($\bar{x}$ = 3.13), and no significant difference was found between the SA2 and SA3.

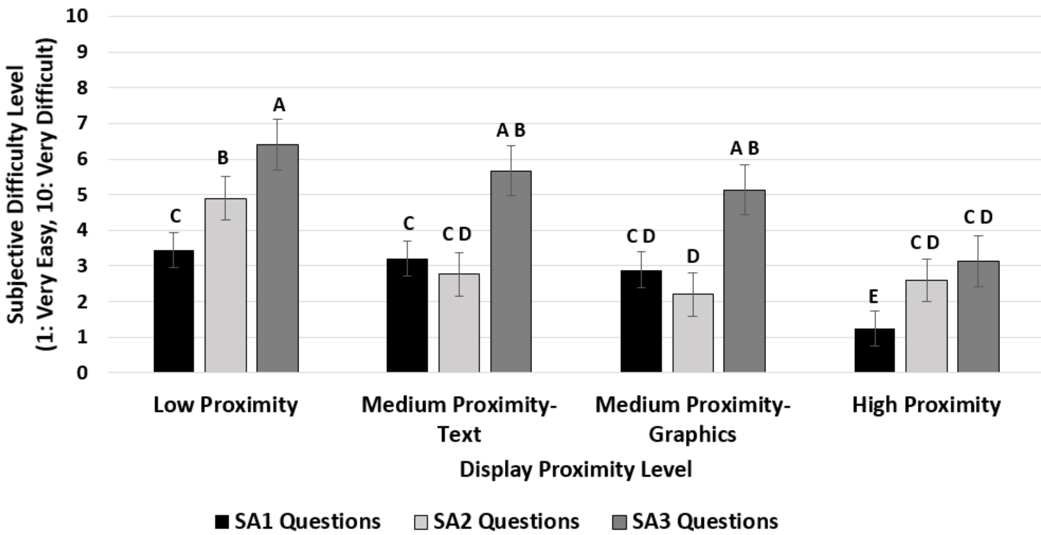

Figure 9. Mean Subjective Difficulty with Tukey Grouping Identifiers above Each Bar. Error bars are added.

For the subjective difficulty, the main effects of display proximity level and SA level question ask must be considered with respect to the interaction. For display proximity level, the High Proximity ($\bar{x}$ = 2.33) was rated significantly easier than other three levels, and the Medium Proximity-Text ($\bar{x}$ = 3.88) and the Medium Proximity-Graphics ($\bar{x}$ = 3.41) were rated significantly easier than the Low Proximity ($\bar{x}$ = 4.92). No significant difference was found between the two Medium Proximity levels. All three SA levels of question were significantly different. SA1 question ($\bar{x}$ = 2.69) was rated as the easiest, followed by SA2 question ($\bar{x}$ = 3.12), with SA3 rated as the most difficult ($\bar{x}$ = 5.08).

pilots' preference of display proximity level with respect to flight hours (Table 4). Table 4 summarizes the subjectively easiest display proximity levels according to pilots' feedback, compared with the levels of the shortest RT and the highest accuracy according to their test results, based on their self-reported flight hours. Regardless of flight hours, 11 of 15 indicated that the High Proximity was the easiest to use. Two pilots rated the Low Proximity as the easiest and indicated that they were already comfortable using the traditional display setting with time information separated. One of the two (who had 2000 flight hours) actually had the shortest RT and the highest accuracy with the Low Proximity. The other (who had 6000 flight hours) had the shortest RT with the Medium Proximity-Text and the highest accuracy with the Medium Proximity-Graphics. Since the two Medium Proximity levels used both the ND and the CDU for separate space and time information sources, we can infer that this pilot was not comfortable with the absence of the CDU when he searched the time information during the experiment.

Table 4
Pilots Perceived Easiest Condition, RT, and Accuracy Results Based on Self-Reported Flight Hours

Pilot Flight Hours	Subjective Perception as the Easiest Condition	His/Her Condition with the Shortest RT	His/Her Condition with the Highest Accuracy
76	High Proximity	Medium Proximity-Graphics	Medium Proximity-Text
120	High Proximity	High Proximity	Low Proximity
240	High Proximity	Medium Proximity-Graphics	Low Proximity
260	High Proximity	Low Proximity	Low Proximity, Medium Proximity-Text, Medium Proximity-Graphics, High Proximity
500	Medium Proximity-Text	High Proximity	Medium Proximity-Graphics
2000	Low Proximity	Low Proximity	Low Proximity
2480	High Proximity	Medium Proximity-Graphics	High Proximity
3000	High Proximity	Medium Proximity-Text	High Proximity
3220	High Proximity	High Proximity	Low Proximity
4500	High Proximity	Medium Proximity-Graphics	Low Proximity
5000	High Proximity	High Proximity	Low Proximity, Medium Proximity-Graphics
5000	Medium Proximity-Graphics	High Proximity	Low Proximity, Medium Proximity-Text, Medium Proximity-Graphics
6000	Low Proximity	Medium Proximity-Text	Medium Proximity-Graphics
8060	High Proximity	Medium Proximity-Text	Medium Proximity-Graphics
15000	High Proximity	High Proximity	Low Proximity, Medium Proximity-Text

There was inconsistency in pilots' subjective perceptions and their actual RTs and accuracies in this table. RTs were mostly shorter when pilots conducted tasks in higher display proximity levels than Low Proximity (13 of 15 pilots). However, the tendency was different for accuracy; only 6 out of 15 pilots had higher accuracies in higher display proximity levels. As shown in Table 5, 9 pilots' accuracies were not lower than any other conditions or even highest with Low Proximity.

Discussion

General Overview

This paper provides discussion on the first phase of study. The paper about the second phase of study will provide the specific discussion with respect to display proximity level, SA level of question, expertise level, and study limitations for the whole study. The novel design strategies applied in this study did not show any better SA performance than the traditional setting objectively. There could be limitations in this study such as limited training time for new flight deck displays. Experiment 1 did not find any difference in RT and accuracy with novel display conditions compared to the traditional format (Low Proximity) for all three SA levels of questions, except for the case of decreased accuracy with Medium Proximity-Text for SA3 questions. As seen in the accuracy results of Medium Proximity-Text condition, adding RTA and ETA information on the ND in the duration format combined with the traditional CDU sometimes degraded the pilots' SA performance. However, integrating all information necessary for RTA operation onto the ND (High Proximity) resulted in equal RT and accuracy compared to the traditional method, and it provided better subjective ratings in all SA levels and mental workload. This indicate there is a possibility of showing higher SA performance using the High Proximity design if any other objective measure methods are developed. However, the two Medium Proximity conditions showed limited improvements even subjectively: easier to answer SA2 questions only.

Hypothesis Testing

Table 5 specifies the results of Experiment 1 hypothesis testing based on Table 2.

Table 5
Hypothesis Testing Results of Experiment 1

Number	Hypothesis	Reject / Not Reject
Hypothesis 1	[RT] There will be a significant difference among display proximity levels.	Reject
Hypothesis 2	[Accuracy] There will be a significant difference among display proximity levels.	Reject
Hypothesis 3	[Subjective Difficulty Ratings] There will be a significant difference among display proximity levels.	Not Reject
Hypothesis 4	[RT] There will be a significant difference among SA levels.	Not Reject
Hypothesis 5	[Accuracy] There will be a significant difference among SA levels.	Not Reject
Hypothesis 6	[Subjective Difficulty Ratings] There will be a significant difference among SA levels.	Not Reject
Hypothesis 7	[RT] There will be an interaction between SA level and display proximity level.	Not Reject
Hypothesis 8	[Accuracy] There will be an interaction between SA level and display proximity level.	Not Reject
Hypothesis 9	[Subjective Difficulty Ratings] There will be an interaction between SA level and display proximity level.	Not Reject

Table 5 represent the results shown in Figure 7, 8, and 9, not the statistical significance of fixed effects between SA levels or between display proximity levels. There were interactions between SA level and display proximity level for all RT, accuracy, and subjective difficulty ratings. However, we could not find any difference for the RT and accuracy among display proximity levels. More discussion about this study will be stated in the paper about the second phase of the study.

Conclusions

The design strategy of providing close spatial proximity between space and time information and graphic indications of temporal conformance information themselves did not show any objective evidence for improved RTA navigation, but they showed subjective evidence for it from this first phase of 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 and subjectively perceived to be easier to use than the traditional display. The overall conclusion will be discussed again in the paper of the second phase of study.

References

Barmore, B., Abbott, T., and Krishnamurthy, K. (2004). Airborne-managed spacing in multiple arrival streams. In *Proceedings of the 24th International Congress of the Aeronautical Sciences (ICAS 2004).* Yokohama, Japan, August 29-September 3, 2004.

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). Springer, Berlin, Heidelberg.

Coull, J. T., and Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *The Journal of Neuroscience*, *18*(18), 7426-7435.

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.

FAA (2017). Press release - *FAA aerospace forecast fiscal years 2017 - 2037*. Retrieved Dec. 20, 2017, from https://www.faa.gov/data_research/aviation/aerospace_forecasts/media/FY2017-37_FAA_Aerospace_Forecast.pdf

Ishihara, M., Keller, P. E., Rossetti, Y., and Prinz, W. (2008). Horizontal spatial representations of time: Evidence for the STEARC effect. *Cortex*, *44*(4), 454-461.

Koch, S. H., Weir, C., Westenskow, D., Gondan, M., Agutter, J., Haar, M., ... and Staggers, N. (2013). Evaluation of the effect of information integration in displays for ICU nurses on situation awareness and task completion time: a prospective randomized controlled study. *International journal of medical informatics*, *82*(8), 665-675.

Krishnan, K., Kertesz Jr, S., & Wise, J. A. (1999, November). Space-time display: a human-centered concept for cockpit-centered air traffic systems. In *Digital Avionics Systems Conference, 1999. Proceedings. 18th* (Vol. 1, pp. 4-A). IEEE.

Lancaster, J., Riddle, K., Feyereisen, T., Olufinboba, O., Rogers, B., Gannon, A., Suddreth, J., He, G. (2011). Trajectory based operations and the legacy flight deck: envisioning design enhancements for the flight crew. Golden Valley, MN: Honeywell Advanced Technology.

Schnotz, W. (2005). An integrated model of text and picture comprehension. *The Cambridge Handbook of Multimedia Learning*, 49-69.