Cockpit Displays of Traffic Information and Pilot Bias in Time-to-Contact Judgments

Carl Pankok, Jr., and David B. Kaber

Pankok C Jr, Kaber DB. ***Cockpit displays of traffic information and pilot bias in time-to-contact judgments.*** **Aviat Space Environ Med 2014; 85:597–604.**

Introduction: Pilots are susceptible to over-reliance on distance when making relative time-to-contact (TTC) judgments of surrounding intruders, referred to as "the distance bias." We tested the effect of adding perceptual cues and an information feature to cockpit displays of traffic information to mitigate this bias. ***Method:*** There were 14 general aviation pilots who participated in a simulated flight scenario and were asked to make relative TTC judgments. Three levels of perceptual cue (blinking, color-change, and no-cue) were crossed with two levels of velocity data tag (present and absent) with identification of the highest risk intruder as a response. ***Results:*** Perceptual cues were associated with more accurate high-risk intruder selection (color = 95.95% correct, blinking = 95.98%, no-cue = 87.89%), decreased response time (color = 3.68 s, blinking = 3.19 s, no-cue = 6.08 s), reduced visual attention demand (color = 57% of attention, blinking = 58%, no-cue = 62%), lower workload ratings (color = 28.38/100, blinking = 29.66/100, no-cue = 48.91/100), and higher performance confidence ratings (color = 83.92/100, blinking = 82.71/100, no-cue = 58.85/100) than the no-cue displays. There was no difference between blinking and color cue displays. The data tag was associated with lower response times (present = 4.13 s, absent = 4.50 s) and higher confidence ratings (present = 78.69/100, absent = 71.63/100) than displays without. Displays including the blinking cue, color-change cue, and data tag were preferred over displays that did not include these features (color = 8 pilots, blinking = 6, no-cue = 0). ***Discussion:*** The added display features were effective in mitigating the effect of the distance bias on pilot performance measures and received favorable subjective ratings.

Keywords: cockpit automation, traffic information displays, aircraft pilot performance, human factors in display design.

AS CURRENT RADAR-based air traffic control systems transition to higher-fidelity systems based on global positioning system technology, new aviation technologies are being developed for the aircraft cockpit (7). One tool that is currently available is the cockpit display of traffic information (CDTI), which uses global positioning system technology to present pilots with accurate position information on surrounding aircraft. The use of CDTIs has the potential to positively impact general aviation (GA), in which pilots typically fly without the assistance of air traffic control. Due to demands on pilots to monitor for other aircraft, incident rates in GA are typically higher than those in commercial or military flights (14), highlighting the need for effective GA cockpit technologies. Furthermore, it has been shown that GA pilots struggle to safely use advanced displays in the cockpit (13).

GA pilots can use a CDTI presenting "intruder" aircraft position, velocity, and heading information as a guide for maneuvering their "own ship" in order to avoid conflicts. To perform this task, pilots must be able to accurately estimate the time to contact (TTC) of their own ship with each intruder and identify which one poses the highest risk (6). Research has shown that display users tend to overly rely on distance information when making relative TTC judgments (2,23,25). Furthermore, it has been suggested that display users are capable of accurately estimating relative distances and velocities of objects, but struggle with integrating this information in working memory for TTC judgments (10,20). The over-reliance on distance information when making TTC judgments has been labeled as the "distance bias" and is considered to be, in part, a product of human working memory limitations in multitarget judgment situations (22).

There are numerous theories that have been proposed to explain the occurrence of the distance bias. Wickens cited limitations in working memory capacity as the main reason (22). Of the four basic mathematical operations (addition, subtraction, multiplication, and division), division has been found to be the most difficult operation to perform in working memory (22), thus increasing the difficulty of judging TTC, which is obtained by dividing the distance of an object to a target by the velocity. There are many situations in which humans do not have the capacity to process all requisite information from a display for error-free decision-making. Sanders and McCormick stated that, "humans are generally conservative and do not extract as much information from sources as they optimally should" (18). Consequently, we tend to adopt decision-making heuristics (or "rules of thumb") which are often adequate, but not always accurate (21). Thus, in making TTC judgments, pilots tend to use heuristics primarily based on distance information, since the resources required for mental estimation, extrapolation, and arithmetic may exceed the capacity of working memory.

From the Edward P. Fitts Department of Industrial & Systems Engineering, North Carolina State University, Raleigh, NC.

This manuscript was received for review in July 2013. It was accepted for publication in February 2014.

Address correspondence and reprint requests to: Dr. David B. Kaber, 400 Daniels Hall, 111 Lampe Dr., Raleigh, NC 27695-7906; dbkaber@ncsu.edu.

DOI: 10.3357/ASEM.3810.2014

The objective of this research was to test the effect of various forms of CDTI content and formatting toward mitigating the effect of the distance bias when pilots assign conflict risk to intruder aircraft. In specific, this research tested the effect of textually presenting intruder velocity information as well as blinking or changing colors of display icons representing high-risk intruders in a prototype CDTI. The sequence of using a perceptual cue to capture pilot attention followed by the presentation of a textual display has been shown to be an effective method of conveying information (1,3,11) and the use of a blinking cue has been shown to be more effective than color change in some cases (19). In addition to assessing the overall effect of additional information and CDTI formatting, the research sought to identify which perceptual cue, blinking or color change, most effectively captured pilot attention and facilitated quick and accurate intruder risk assessments.

The following hypotheses were formulated based on the literature review:

1. The addition of blinking cues, color-change cues, and/or velocity data tags to the CDTI will reduce/mitigate the effect of distance bias, as evidenced by increased accuracy in first arrival intruder selection, in scenarios in which the bias is expected to influence pilot judgments.
2. The mean response time (RT) for pilots using the blinking cue display will be lower than the RT when pilots use the color-change display. Both blinking and color-change will produce lower RTs than the baseline (no cue) display and these differences will be greatest in scenarios in which the distance bias is expected to affect pilot RT.
3. The mean RT for displays including the data tag will also be shorter than for displays without. These differences will be greatest in scenarios in which the distance bias is expected to affect pilot RT.
4. Displays including blinking cues, color-change cues, and/or the velocity data tags will require a smaller proportion of pilot attention (i.e., a reduced gaze frequency for the CDTI) than the baseline (no cue, no tag) displays.
5. Pilots will report lower workload and higher confidence for displays that include the blinking cue, color cue, and/or velocity data tag vs. baseline displays.
6. The majority of pilots (> 50%) will prefer color-change displays, followed by blinking displays, as compared with the baseline displays. Pilots will also prefer the presence of velocity data tags to displays with no tags.

METHODS

Subjects

The study protocol was approved by the North Carolina State University Institutional Review Board. Each subject provided written informed consent. There were 14 male certified pilots with an average age of 37.1 yr (range: 21-66, SD: 16) who participated in a flight simulation experiment. All subjects had 20/20 or corrected vision. Of the pilots, nine had a private certification and five had a commercial certification. Those who had a commercial rating indicated no commercial flight hours; they simply had enough total hours to qualify for the commercial certification. The mean total flight time for the subjects was 646.5 h (range: 73-2500, SD: 717.3). Experience using a CDTI (either in an aircraft or in a simulator) was indicated by eight subjects. Among those with experience, the mean rating was 26.9 out of 100 (range: 1-70, SD: 21.1), generally indicating a low level of display use experience.

Independent Variables

Three independent variables were manipulated in the experiment, including: first-arrival start distance (FASD; as simulated in the prototype CDTI); perceptual cue type; and the presence or absence of the velocity data tag. The experiment tested two levels of the FASD: 3.70 km (2 nmi) and 5.56 km (3 nmi). In the 3.70-km FASD scenarios, the intruder aircraft designated to collide first with the own ship started the scenario at 3.70 km from the point-of-conflict while a second intruder started at 5.56 km. In the 5.56-km FASD scenarios, the intruder designated to collide first with the own ship started the scenario 5.56 km from the point-of-conflict while a second intruder started at 3.70 km. Performance in the 5.56-km FASD scenarios was expected to be influenced by the distance bias since the relative start distances of the intruders did not reflect the relative TTCs (in the case of equal velocities). The three levels for the perceptual cue were baseline (i.e., no cue), blinking, and color-change. When the blinking cue was active, the intruder icon blinked at a rate of 2.51 Hz (based on the processing speed of the computer system used to present the CDTI). When an intruder changed colors, it changed from the display baseline color of teal to yellow. These colors are commonly used in existing CDTIs for intruder icons (e.g., Garmin GMX200). The velocity tag, when present, appeared directly above the intruder and indicated the intruder's speed in knots (kt; nmi $\cdot$ h^{-1}). We chose to include the velocity rather than an explicit presentation of the TTC since the former information is currently broadcast between aircraft. Furthermore, in order to present exact TTC information in a real-life cockpit, a CDTI must have the capabilities to: 1) project future trajectories of two or more aircraft; 2) predict a time and physical location where the two will collide; and 3) continually update the TTC as the aircraft approach each other. These functions are complicated from sensing and algorithm development perspectives due to greatly varying conflict geometries that can develop. For this reason, we considered the addition of velocity to be more realistic compared with TTC.

Design

A within-subject design was utilized for the experiment. A full crossing of all levels of perceptual cue type, data tag, and FASD was presented to subjects. In each trial, two intruders appeared on the prototype CDTI and both were on a collision course with the pilot's own ship. The trials representing the two FASD levels are derived in **Table I** (denoted 3.70-km FASD Trial and 5.56-km FASD Trial). The two intruders in each of the two FASD trials were randomly assigned to four conflict angles (45, 135, 225, and 315°; see **Fig. 1**). There were no scenarios in which both intruders approached from the same conflict angle. Any scenarios

TABLE I. SCENARIO TYPES.

Intruder	Starting Distance	Velocity	TTC
3.70-km FASD Trial			
First Arrival	3.70 km	444.48 km/h	30 s
Second Arrival	5.56 km	222.24 km/h	90 s
5.56-km FASD Trial			
First Arrival	5.56 km	444.48 km/h	45 s
Second Arrival	3.70 km	222.24 km/h	60 s

TTC = time to contact; FASD = first arrival start distance.

with two trailing intruders (i.e., conflict angles of 45 and 315°) or two head-on intruders (i.e., conflict angles of 135 and 225°) were not included in order to promote the complexity of conflict judgments and reduce the total number of experimental trials. The combinations of start distances, velocities, and conflict angles resulted in 16 unique traffic scenarios; the first arrival was balanced among the various conflict geometries in order to eliminate any influence of the conflict angle or starting distance on pilot TTC judgments or subjective ratings of the CDTI configurations. Each scenario was tested with each display format, resulting in a total of 96 trials (3 cues * 2 data tag conditions * 16 scenarios) for each subject. The experiment was performed in blocks of each combination of perceptual cue and data tag to allow for pilot subjective ratings on each unique display content and format. Within each block, the order of the 16 scenarios was randomized and, for each experiment, the order of the blocks was randomized within-subject in order to prevent condition order effects.

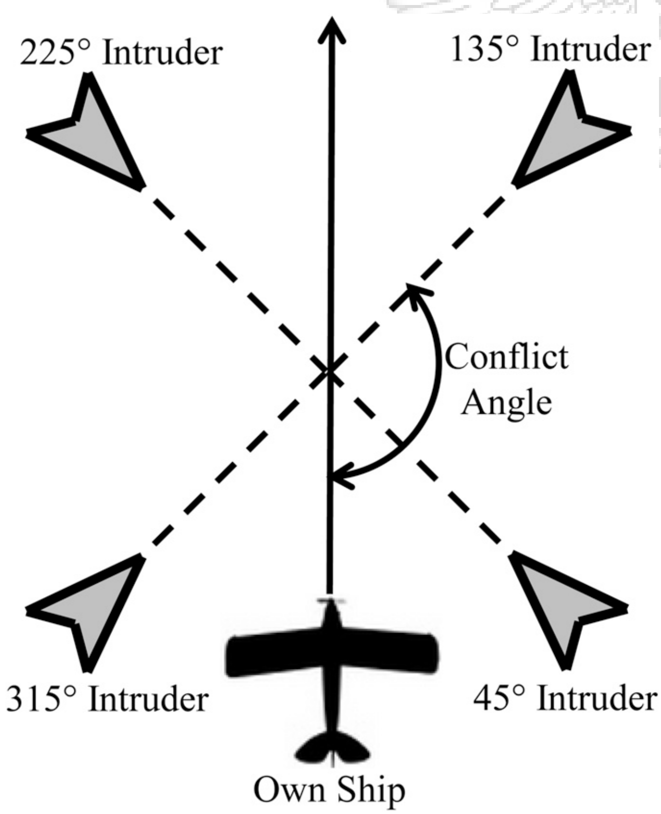

Fig. 1. Labeling of intruders by conflict angle.

Equipment

The flight simulator hardware included an X-Plane workstation, yoke control, throttle controls, and rudder pedals. A top-down view of the simulator setup is presented in **Fig. 2**. X-Plane Version 10 (Laminar Research, Columbia, SC) was used for the simulated flight and the aircraft model flown was a Cessna 172. The aircraft out-of-cockpit view and cockpit displays and controls were presented on a computer screen directly in front of the subject and the CDTI was presented on a touch-screen to the right of the out-of-cockpit view, approximately 30-35° from the subject's line of sight. During each block of trials, the subject was required to track a heading of 0° and a vertical velocity of $0\ \mathrm{m} \cdot \mathrm{min}^{-1}$ (i.e., maintain straight and level flight). The simulation was not integrated with the intruders shown in the prototype CDTI; the flying task served only as a secondary loading task while the pilot used the CDTI to make TTC judgments. To increase the realism of the simulation, speakers were set to project an in-cockpit sound level of 88.3 dBA, which is the approximate noise level experienced by pilots in a Cessna 172 when wearing a circumaural headset (9).

The prototype CDTI was developed with the C++ programming language using the Open Graphics Library. The prototype was a replicate of the Garmin GMX200 model CDTI, which is a popular display currently used in GA aircraft. Subjects were told that all intruders were at the same altitude as the own ship and the miss distance of all of intruders was 0 km, placing them on collision courses with the own ship. The subject selected the first arrival (highest risk intruder) by touching the particular intruder icon on the touch-screen. The perceptual cues (e.g., a blinking or color changing intruder icon) and velocity data tags used in the prototypes

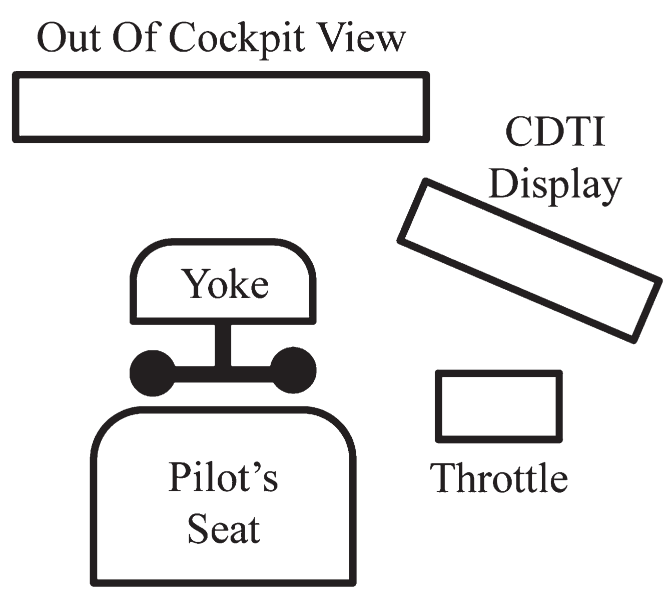

Fig. 2. Top-down view of simulator layout.

were activated according to the current Traffic Collision Avoidance System (TCAS) algorithm (4); that is, when an intruder crossed the 40-s TTC threshold, the cues were triggered.

Dependent Variables

Three performance measures were examined in the experiment, including pilot accuracy in selecting the correct target (highest risk intruder) and the RT for selection. Additionally, pilot visual attention allocation to the aircraft "dashboard" vs. the CDTI was assessed using video-based glance analysis for comparison of the different CDTI formats and content. A work sampling technique (15) was used in which an observer with no knowledge of the hypotheses of the experiment recorded pilot gaze direction at 1-s intervals during times when the CDTI was active (no eye-tracking hardware or software was used). The CDTI was considered "active" when it was displaying the two intruders moving on a collision course with the pilot's own ship (i.e., not between trials, when a blank screen was displayed). Slow-motion video recordings of pilots' faces during experimentation were synchronized with the CDTI activation times for each flight scenario. The observer painstakingly watched each video to record the gaze direction from moment to moment. The lead researcher of the study verified the observer records.

After each block of 16 trials, subjective ratings of workload and performance confidence were collected. All ratings were made on a continuous scale with anchors of "low" and "high." Ratings were measured from the "low" anchor with a resolution of 0.5 mm and transformed to values from 0 to 100 points. Upon completion of the experiment, pilot display preferences were recorded.

Procedure

Subjects were initially asked to complete an informed consent form and demographic questionnaire. Subjects were subsequently shown the various flight controls and were allowed up to 5 min to fly the simulated Cessna 172 aircraft on a straight and level course without using the prototype CDTI. Once the subject was comfortable with the flight controls, he was introduced to the prototype CDTI functionality and was required to complete a series of training trials using the software until comfortable with its operation.

At the beginning of each block of trials, the subject was informed of the display content and format to be presented in the subsequent block of scenarios (i.e., which cues would be used and whether the velocity tag would be present). When the subject was ready, the experimenter started the flight simulator and the subject was told that, when comfortable, he could start the CDTI scenarios. Before starting a scenario, the CDTI screen was blank and simply said, "Touch Anywhere to Begin." When the subject touched the screen, the air traffic scenario started immediately. The RT timer started when the blinking cue, color cue, or velocity was displayed and ended when the subject chose one of the two intruder aircraft. In the baseline trials (i.e., trials in which no cues or data tags were presented), the RT timer started when the air traffic scenario started. Since the subjects were informed of the automation they would see in the coming block of trials, there were no instances of subjects waiting for the cues or data tags to appear before choosing an intruder. This approach for calculating the RT allowed for unbiased comparison of the performance between the baseline trials and the trials that contained one of the added features.

When an intruder was selected, the CDTI returned to the blank introductory screen. Therefore, as the subject flew on a straight and level flight path, he was able to start the next scenario at his convenience. After a block of 16 trials, a blank screen appeared, indicating that the block was complete, and the experimenter paused the simulator and handed the subject a subjective rating form for assessment of workload and confidence. This procedure was repeated for the remaining five blocks of trials. The experiment lasted approximately 1 h and all subjects were compensated.

Statistical Analysis

Contingency analyses were used to assess the effect of the three levels of perceptual cue and the two levels of the velocity data tag on pilot accuracy in selection of high-risk intruders. If an intruder was selected before a display feature was presented (cue and/or data tag at the 40-s TTC threshold), the trial was classified as a "baseline-absent" trial (i.e., no perceptual cue or data tag was present at the time of the pilot's decision). This approach also applied to the RT and gaze proportion responses. For all contingency analyses, the likelihood ratio Chi-square test statistics are reported.

Analysis of variance (ANOVA) procedures were used to assess the effects of display format and content on pilot RT and the proportion of pilot visual attention allocated to the CDTI. A split-plot model was structured for both responses (12) and the analyses included only trials in which the correct intruder was selected. The model included two error terms with one based on variability attributable to trial blocks (defined by subject, cue type, and tag) and another based on individual scenarios (traffic geometries) within blocks. A log transformation was applied to the RT to satisfy all normality and homoscedasticity assumptions, which were not satisfied by the untransformed response. Regarding the gaze proportion, a nonparametric ANOVA was conducted by transforming the proportions to global ranks. Tukey's Honest Significant Difference (HSD) test was used to further investigate any significant main effects or interactions using a whole-plot error term (subject*cue*tag). The remaining post hoc tests used the subplot error (scenario within block). The letters in the Tukey's HSD Grouping column in **Tables II-IV** represent the significance groupings of means based on the post hoc tests; levels sharing at least one letter are not significantly different from each other at an $\alpha = 0.05$ level. All ANOVA and post hoc tests were based on the transformed responses.

TABLE II. POST HOC TESTS FOR THE RT ANOVA.

Level	Mean (seconds)	SD (seconds)	Tukey's HSD Grouping
Cue*Tag Interaction			
Baseline, Absent	6.54	6.43	A
Baseline, Present	5.61	7.28	B
Color, Present	3.82	6.14	BC
Color, Absent	3.54	4.52	BC
Blinking, Absent	3.43	4.87	BC
Blinking, Present	2.95	4.34	C
Cue*FASD Interaction			
5.56 km, Baseline	7.21	8.49	A
3.70 km, Baseline	5.01	4.26	A
5.56 km, Color	4.08	6.82	B
5.56 km, Blinking	3.72	6.36	B
3.70 km, Color	3.35	3.82	C
3.70 km, Blinking	2.76	2.32	C
Tag*FASD Interaction			
5.56 km, Absent	5.51	7.11	AB
5.56 km, Present	4.78	8.03	C
3.70 km, Absent	3.81	3.43	A
3.70 km, Present	3.60	3.92	B

RT = reaction time; FASD = first arrival start distance.

Descriptive statistics (i.e., means and SDs) are presented in original units for easy interpretation.

An ANOVA was also used to analyze pilot subjective ratings of workload and intruder selection confidence. The statistical model included subject as a blocking variable, main effects of perceptual cue type and data tag presence, and the two-way interaction. All statistical assumptions of the ANOVA were met by the subjective rating data. Finally, counts are presented for pilot display feature preferences.

RESULTS

The contingency analysis on intruder selection accuracy used FASD as a grouping variable in order to determine whether perceptual cues mitigated the effect of distance bias in scenarios for which the FASD was longer. The analysis revealed a significant effect of cue type when the FASD was longer ($\chi^2 = 13.47$, $P = 0.001$, blinking = 95.98% correct, color = 95.92%, baseline = 87.89%), but not when the FASD was shorter ($\chi^2 = 2.20$, $P = 0.333$, baseline = 100.00% correct, color = 100.00%, blinking = 99.55%). The percentage of correct intruder selections was significantly higher when blinking ($\chi^2 = 10.30$, $P = 0.001$) or color-change ($\chi^2 = 8.82$, $P = 0.003$) cues were presented, as compared with the baseline display when the FASD was longer (i.e., the condition in which pilots were expected to be susceptible to distance bias). However, there was no significant difference in target selection accuracy between the blinking and color change displays for the longer FASD ($\chi^2 = 0.06$, $P = 0.807$).

The contingency analyses examining the effect of the velocity tag were also separated based on FASD setting to determine whether the tag mitigated the effect of the distance bias. Results revealed no significant effect of tag on intruder selection accuracy in scenarios in which the FASD was shorter ($\chi^2 = 1.39$, $P = 0.239$, tag present = 100.00% correct, tag absent = 99.70%) or in scenarios in which the FASD was longer ($\chi^2 = 2.42$, $P = 0.120$, tag present = 94.64% correct, tag absent = 91.62%).

Analysis of the log-transformed RT revealed significant main effects of cue type [$F(2,65) = 32.98$, $P < 0.001$], data tag [$F(1,65) = 10.89$, $P = 0.002$], and FASD [$F(1,1092) = 13.82$, $P < 0.001$], as well as interactions between cue type and data tag [$F(2,65) = 4.99$, $P = 0.010$], cue type and FASD [$F(2,1092) = 10.19$, $P < 0.001$], and data tag and FASD [$F(1,1092) = 15.70$, $P < 0.001$]. The three-way interaction was not significant [$F(2,1092) = 2.57$, $P = 0.077$]. Tukey's test for each two-way interaction is presented in Table II.

As shown in Table II, RT was slowest for the "baseline-absent" display type, followed by the "baseline-present" type (i.e., including data tags). The displays integrating a blinking or color-change cue generally resulted in faster RTs than the baseline displays. Additionally, with the exception of the baseline displays, the presence of the velocity data tag had no significant effect on RT.

Concerning the interaction between cue type and FASD, mean RTs were slower for the baseline displays than for the blinking or color-change displays, regardless of whether the FASD was shorter or longer. Additionally, the mean RTs and their SDs were generally greater when the FASD was longer. Tukey's HSD groupings provided evidence that the mean RT for the baseline display was significantly slower than for either the blinking or color-change displays within each FASD setting ($P < 0.05$). Between FASDs, the mean RTs for the blinking and color-change displays were significantly faster when the FASD was shorter ($P < 0.05$). However, there was no significant difference between the mean RTs for the FASDs when the baseline displays were used.

The effect of the interaction between the data tag and FASD on RT revealed that displays without the tag yielded slower RTs than when the tag was present. Additionally, as with the cue by FASD interaction, the mean RTs and corresponding SDs were generally greater when the FASD was longer. Tukey's HSD test results supported these trends.

The split-plot ANOVA on the rank-transformed proportion of gazes to the CDTI revealed a significant main effect of the FASD [$F(1,1092) = 133.78$, $P < 0.001$], but no effect of cue type [$F(2,65) = 1.65$, $P = 0.199$] or data tag [$F(1,65) = 0.75$, $P = 0.389$]. The analysis also revealed a significant interaction between cue type and FASD

TABLE III. POST HOC TESTS FOR VISUAL ATTENTION ALLOCATION ANOVA.

Level	Mean Proportion	SD Proportion	Tukey's HSD Grouping
Blinking, 3.70 km	0.69	0.27	A
Baseline, 3.70 km	0.66	0.26	A
Color, 3.70 km	0.65	0.31	A
Baseline, 5.56 km	0.58	0.25	B
Color, 5.56 km	0.48	0.24	C
Blinking, 5.56 km	0.47	0.17	C

TABLE IV. POST HOC TESTS FOR SUBJECTIVE RATINGS ANOVAs.

Level	Mean Rating	SD Rating	Tukey's HSD Grouping
Workload			
Baseline	48.91	26.11	A
Blinking	29.66	22.76	B
Color	28.38	23.38	B
Confidence			
Color, Present	85.05	8.36	A
Color, Absent	82.79	8.04	A
Blinking, Absent	82.72	11.86	A
Blinking, Present	82.69	7.84	A
Baseline, Present	68.33	21.34	B
Baseline, Absent	49.37	22.62	C

[F(2,1092) = 5.55, $P = 0.004$]. The three-way interaction between cue type, data tag, and FASD was also significant [F(2,1092) = 4.63, $P = 0.010$]. Due to the lack of relevant trends associated with the three-way interaction, the analysis focuses on the significant two-way interaction between cue type and FASD, reported in Table III. Within each level of cue type, the proportion of gazes to the CDTI was significantly greater for the shorter FASD vs. longer ($P < 0.05$). Additionally, in the longer FASD scenarios, the proportion of visual attention required by the CDTI was significantly less when the display included the blinking or color change features compared to the baseline displays.

An ANOVA performed on the subjective workload ratings revealed a significant effect of cue type [F(2,65) = 23.55, $P < 0.001$], but no effect of the velocity data tag [F(1,65) = 0.70, $P = 0.406$] nor a significant interaction effect [F(2,65) = 1.05, $P = 0.355$]. Table IV shows the baseline displays produced the highest subjective workload rating, followed by displays integrating the blinking cue and then those with the color-change cue. Tukey's post hoc groupings revealed workload ratings for the blinking and color-change cues to be significantly lower than the ratings for the baseline display ($P < 0.05$); however, there was no significant difference between the mean workload ratings for these displays.

An ANOVA performed on pilot confidence in high-risk intruder selection revealed significant effects of cue type [F(2,65) = 38.18, $P < 0.001$] and data tag [F(1,65) = 7.14, $P = 0.010$] as well as a significant interaction of cue type and data tag [F(2,65) = 5.14, $P = 0.009$]. Due to the significant interaction, simple effects were examined, as shown in Table IV. All displays including a blinking or color-change cue received a higher confidence rating than either of the no-cue displays. Additionally, no-cue displays that included a velocity data tag were associated with higher confidence ratings than the baseline display without data tag.

Upon completion of the experiment, pilots were asked to identify preferences for display features and settings. Eight pilots preferred the color cue, six preferred the blinking cue, and none preferred the baseline display. In addition, 10 out of the 14 pilots preferred the displays in which the velocity data tag was present as opposed when it was not.

DISCUSSION

The addition of blinking and color-change cues as well as velocity data tags to the prototype CDTI was expected to mitigate the effect of the distance bias (Hypothesis 1), as indicated by greater pilot accuracy in intruder selection. The significance of the cue type when the FASD was longer (and lack thereof when the FASD was shorter) suggests that the distance bias affected pilot intruder selection and that cues were effective for reducing the bias. The lack of significance of the data tag indicated that any improvement in pilot information processing associated with the addition of display features was attributable to the perceptual cues. Even when pilots were provided with explicit velocity information, they did not make the effort to perform long division in their working memory to estimate the TTC for each intruder and to make accurate high-risk intruder selections. It is likely that time pressure of the traffic scenarios presented in the experiment led to reliance on the distance bias in the presence of data tags, premature intruder selection, and inaccurate responses. This finding might be contingent upon the starting distances of intruders to the pilot's own ship; i.e., less time pressure on pilots might lead to greater use of the data tags. These results support Hypothesis 1 for the addition of display cues, but refute the hypothesis when considering velocity data tags.

The mean RT was expected to be fastest when pilots used displays with blinking cues, followed by those with color-change cues, followed by the baseline displays (Hypothesis 2). It was also expected that mean RT would be faster for displays that included the velocity tag as compared to those without (Hypothesis 3). These trends were expected to be most pronounced in scenarios in which the distance bias was expected to affect pilot decision making (i.e., when the FASD was longer). The interaction between the cue type and FASD indicated that the difference in RT among the cues was more pronounced for the longer FASD scenarios than for the shorter scenarios. The cues were effective in mitigating the effect of the distance bias by reducing the RT as compared to the baseline displays. The cues and data tags allowed pilots to more quickly make a TTC judgment and select the highest risk intruder aircraft. However, there was no evidence of blinking being superior to color-change for information acquisition and processing when pilots used the CDTI. The interaction between the cue type and the data tag showed that the presence of the data tags resulted in significantly faster RTs than the baseline displays, but not in displays including either the blinking or the color-change cues. This finding indicates that when there was no perceptual cue present in the CDTI, pilots might have used the added velocity data tags to reduce effort in estimating intruder TTC through calculations in the working memory for high-risk identification, as compared to the baseline display with no tag. However, the results demonstrate that cues attracted greater pilot attention to the CDTI and had a more profound effect on RT than the tag. These results support Hypotheses 2 and 3. It is likely that pilots developed a strategy in which they exploited the perceptual cues when presented simultaneously with the data tags and used the data tags only when there was no perceptual cue.

Displays integrating blinking cues, color change cues, and/or velocity data tags were expected to require less pilot visual attention resources, as compared with the baseline displays (Hypothesis 4). The results of the split-plot ANOVA on the rank-transformed proportion of gazes to the CDTI demonstrated this to be true when the FASD was longer, but not when the FASD was shorter. The lower proportion of attention required by the CDTI when the blinking or color-change cue was used suggests that the added features reduced the amount of time the pilot needed to comprehend a traffic situation and decreased the visual workload in judgments on intruder TTC. However, these differences may be a result of the experiment's design rather than a true representation of the effect of the FASD. When the FASD was shorter, the perceptual cue was active when the scenario started, but this was not the case for scenarios in which the FASD was longer. In the shorter FASD scenarios, pilots could immediately identify the high-risk intruder, limiting the total number of glances to the display (i.e., the denominator in the proportion calculation) and increasing the proportion of glances to the CDTI. In the longer FASD scenarios, intruders traveled for 5 s before the display features were activated. Pilots had to focus on the flight simulation display during this time in order to maintain the aircraft in straight and level flight. They infrequently sampled the CDTI, watching for the perceptual cue and/or data tags to appear. However, as a result of concentration on the flight simulation, the proportion of gazes to the CDTI remained low, as in the shorter FASD scenarios, but for this alternate reason. These findings support Hypothesis 4 in the long FASD scenarios, but refute Hypothesis 4 in the short FASD scenarios.

Pilots were expected to rate workload as being lower for displays that included blinking cues, color-change cues, and/or the velocity data tags than those in which no cues or tags were present (Hypothesis 5). The main effect of cue type on perceived workload demonstrated this hypothesis to be partially true. It is not surprising that the blinking and color-change cue displays resulted in lower workload ratings than the baseline displays since they facilitated more efficient information processing (i.e., precluding the need for pilots to mentally calculate intruder TTC). However, there was no significant effect of the data tag, indicating no workload benefit or cost compared to displays in which the velocity information was not presented. This is not surprising given that the data tags provided no benefit for intruder selection accuracy or pilot attention allocation.

Pilots were also expected to indicate higher confidence in their intruder selections with displays that included blinking cues, color-change cues, or velocity data tags compared to displays that did not use these features (Hypothesis 5). Post hoc tests on the cue type by data tag interaction supported this hypothesis. Although the tags had a significant effect for the baseline (no cue) display, results indicated cues had a stronger effect on confidence ratings than tags alone. It is likely that the perceptual cues drew pilot attention to the displays more effectively than the data tags and the cues supported more efficient pilot information acquisition and analysis. Furthermore, the results support the contention that pilots were using the data tag only in scenarios in which no cue was presented.

Finally, the majority of pilots were expected to prefer the color-change displays, followed by blinking displays, as well as both of these displays to baseline displays (Hypothesis 6). Additionally, it was expected that pilots would prefer the velocity data tags to displays without tags. The fact that one perceptual cue was not selected significantly more than the other suggests the "best" cue for pilot information processing is likely a matter of preference. Color is an effective cue since it creates luminance changes, which can be detected in peripheral vision. The color yellow is also often used to indicate system states requiring caution. Blinking can be even more effective than color-change in terms of drawing pilot attention, especially compared to low (target-to-background) contrast displays. On the other hand, some pilots may find blinking cues to be annoying and, thus, prefer color-change. A majority of pilots preferred the presence of data tags because they simply provided more information than the baseline display. Although results indicated that tags did not enhance performance, it is possible they inspired the greater pilot confidence observed for TTC judgments when using displays with tags. These findings and inferences are all in line with Hypothesis 6.

Limitations of the present study include the use of a personal computer-based flight simulator. Future research should make use of a more realistic cockpit simulation to facilitate greater pilot immersion in flight task performance and display use. Another limitation is the small sample (16) of air traffic geometries used to test the prototype CDTIs. In specific, only two FASDs and two intruder velocities were used, resulting in four possible TTCs. Testing additional levels of these variables would provide more definitive results on the effects of display format and content on pilot performance. One final limitation is that this research focused on the effect of blinking and color-change cues and velocity data tags in CDTIs, when there are many other types of cues that could be implemented in such displays, which might be more effective. Other CDTIs have integrated aural alerts (8), increasing/decreasing size of features (8,16), different colors (17,24), and velocity trend vectors (5). Future research should examine whether there are other, more effective ways to convey intruder distance, velocity, or TTC information to pilots through CDTIs.

The objective of this research was to assess the effectiveness of adding features to a prototype CDTI, including perceptual cues and velocity data tags for intruder aircraft characterization in order to mitigate the effect of the distance bias exhibited by pilots when making relative TTC judgments. Results of the experiment indicate that the addition of color-change or blinking cues and velocity data tags provide a benefit to pilots in terms of assessing the relative risk of multiple intruders manifested in RT and selection accuracy. There was also evidence that the proportion of pilot visual attention allocated to the CDTI was reduced when cues were featured in the display, specifically when the FASD was longer. Furthermore, subjective ratings revealed that pilots preferred the added display features and felt they provided a significant benefit to performance.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Nancy Currie, senior engineer at NASA Johnson Space Center, for her assistance and expertise in conducting the literature review and providing input for the flight simulation, prototype display design, and air traffic scenario design. The authors would also like to thank Dr. David Dickey for his valuable statistics input. This research was supported by a grant from the U.S. National Institute for Occupational Safety and Health (NIOSH: No.2 T42 OH008673-08). The opinions expressed in this report are those of the authors and do not necessarily reflect the views of NIOSH.

Authors and affiliation: Carl Pankok, Jr., B.S., M.S., and David B. Kaber, M.S., Ph.D., Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC.

REFERENCES

1. Beringer DB, Allen RC, Kozak KA, Young GE. Responses of pilots and nonpilots to color-coded altitude information in a cockpit display of traffic information. In: Proceedings of the Human Factors Ergonomic Society's 37th Annual Meeting; 1993 October 11-15; Seattle, WA. Santa Monica, CA: HFES; 1993:84-7.
2. Bolton ML, Bass EJ. Using relative position and temporal judgments to identify biases in spatial awareness for synthetic vision systems. Int J Aviat Psychol 2008; 18:183–206.
3. Christ RE. Review and analysis of color coding research for visual displays. Hum Factors 1975; 17:542–70.
4. Federal Aviation Administration. Introduction to TCAS II, Version 7.1. Washington, DC: Federal Aviation Administration; 2011.
5. Funabiki K, Iijima T, Nojima T. CDTI design in a trajectory-based operations concept for small aircraft. In: Proceedings of the 23rd Digital Avionics Systems Conference, DASC 2004; 24-28 October 2004; Salt Lake City, UT. New York: IEEE; 2004.
6. Hickey DT. Individual differences in strategic processing in a dynamic spatial reasoning task. In: Proceedings of the Human Factors Ergonomic Society's 34th Annual Meeting; 1990 October 8-12; Orlando, FL. Santa Monica, CA: HFES; 1990:934-8.
7. Joint Planning and Development Office. Concept of operations for the next generation air transportation system, Version 3.2. Washington, DC: Joint Planning and Development Office; 2010.
8. Jones DR, Prinzel III LJ, Shelton KJ, Bailey RE, Otero SD, Barker GD. Collision avoidance for airport traffic simulation evaluation. In: Proceedings of the 29th Digital Avionics Systems Conference, DASC 2010; 3-7 October 2010; Salt Lake City, UT. New York: IEEE; 2010.
9. Lamm E, Lawrence N. Interior sound levels in general aviation aircraft 2010; Retrieved 15 November 2012 from http://ohsonline.com/articles/2010/07/12/interior-sound-levels-in-general-aviation-aircraft.aspx.
10. Law DJ, Pellegrino JW, Mitchell SR, Fischer SC, McDonald TP, Hunt EB. Perceptual and cognitive factors governing performance in comparative arrival-time judgments. J Exp Psychol Hum Percept Perform 1993; 19:1183–9.
11. Lorenz B, Biella M. Evaluation of onboard taxi guidance support on pilot performance in airport surface navigation. In: Proceedings of the Human Factors Ergonomic Society's 50th Annual Meeting; 2006 October 16-20; San Francisco. Santa Monica, CA: HFES; 2006:111-5.
12. Montgomery DC. Design and analysis of experiments, 3rd ed. New York: John Wiley & Sons; 1991.
13. National Transportation Safety Board. Safety study: introduction of glass cockpit avionics into light aircraft. Alexandria, VA: National Technical Information Service; 2010.
14. National Transportation Safety Board. Review of U.S. civil aviation accidents 2007-2009. Alexandria, VA: National Technical Information Service; 2011.
15. Niebel BW, Freivalds A. Methods, standards, & work design, 12th ed. Boston: WCB/McGraw-Hill; 2003.
16. Palmer EM, Clausner TC, Kellman PJ. Enhancing air traffic displays via perceptual cues. ACM Trans Appl Percept 2008; 5:4/1-22.
17. Riley V, Sierra E, Mogford R, Johnson W, Kopardekar P, et al. Pilot perceptions of airspace complexity. In: Proceedings of the 22nd Digital Avionics Systems Conference, DASC 2003; 12-16 October 2003; Indianapolis, IN. New York: IEEE; 2003.
18. Sanders MS, McCormick EJ. Human factors in engineering and design. New York: McGraw-Hill, Inc.; 1993.
19. Thackray RI, Touchstone MR. Effects of monitoring under high and low taskload on detection of flashing and coloured radar targets. Ergonomics 1991; 34:1065–81.
20. Tresilian JR. Perceptual and cognitive processes in time-to-contact estimation: analysis of prediction-motion and relative judgment tasks. Percept Psychophys 1995; 57:231–45.
21. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science 1974; 185:1124–31.
22. Wickens CD. Spatial awareness biases. Moffett Field, CA: National Aeronautics and Space Administration; 2002. Report No: AHFD-02-6/NASA-02-4.
23. Xu X, Rantanen EM. Effects of air traffic geometry on pilots' conflict detection with cockpit display of traffic information 2007. Hum Factors 2007; 49:358–75.
24. Xu X, Wickens CD, Rantanen EM. Effects of conflict alerting system reliability and task difficulty on pilots' conflict detection with cockpit display of traffic information. Ergonomics 2007; 50:112–30.
25. Zuschlag M, Chandra D, Grayhem R. The use and understanding of the proximate status indication in traffic displays. In: Proceedings of the 30th Digital Avionics Systems Conference, DASC 2011; 16-20 October 2011; Seattle, WA. New York: IEEE; 2011.