3-D Audio Versus Head Down TCAS Displays

Durand R. Begault and Marc T. Pittman

San Jose State University Foundation
One Washington Square
San Jose, CA 95192

Table of Contents

3-D Audio Versus Head Down TCAS Displays

Durand R. Begault and Marc T. Pittman

SUMMARY

The advantage of a *head up* auditory display was evaluated in an experiment designed to measure and compare the acquisition time for capturing visual targets under two conditions: Standard head down Traffic Alert and Collision Avoidance System (TCAS) display, and three-dimensional (3-D) audio TCAS presentation. Ten commercial airline crews were tested under full mission simulation conditions at the NASA Ames Crew-Vehicle Systems Research Facility (CVSRF) Advanced Concepts Flight Simulator. Scenario software generated targets corresponding to aircraft which activated a 3-D aural advisory or a TCAS advisory. Results showed a significant difference in target acquisition time between the two conditions, favoring the 3-D audio TCAS condition for 500 ms.

INTRODUCTION

The current implementation of the Traffic Alert and Collision Avoidance System (TCAS II) uses both auditory and visual displays of information to supply flight crews with real-time information about proximate aircraft. However, the visual display is the only component designed to convey spatial information about surrounding aircraft, while the auditory component is used as a redundant warning or, in the most critical scenarios, for issuing instructions for evasive action.

Within its standard implementation, three categories of visual-aural alerts are activated by TCAS, contingent on an intruding aircraft's distance. The first category, an informational visual display, presents *proximate traffic*. In this case, TCAS functions more as a situational awareness system than as a warning system. The second category, a visual-aural cautionary alert, is a *traffic advisory*. The threshold for activating a traffic advisory is a potential conflict within 40 s; an amber filled circle is generated on a visual map display, and an auditory warning consisting of a single cycle of the spoken words TRAFFIC-TRAFFIC is given. The third category, a visual-aural warning alert, is a *resolution advisory*. The threshold for activating a resolution advisory is a potential conflict within 20-25 s; a red filled square is generated on a visual map display, and an auditory warning enunciating the necessary appropriate evasive action (e.g., CLIMB-CLIMB-CLIMB) is given.

Chappell, et al. (ref. 1) evaluated the effectiveness of TCAS during a full-mission simulation experiment. Three TCAS conditions were evaluated, each involving a different level of visual-aural information about the location of conflicting aircraft. In addition, a non-TCAS condition was evaluated where only spoken traffic advisories from air traffic controllers (ATC) were used. Their measure of performance focused on the time to make an evasive maneuver in response to a TCAS resolution advisory. The findings suggest that, although the TCAS displays are superior to ATC radio communication, no significant benefit is gained in increasing the complexity of the TCAS display itself. Specifically, no advantage was found in providing pilots with a head down planform display of traffic information.

1

Perrott, et al. (ref. 2) found that spatial auditory information can significantly reduce the acquisition time necessary to locate and identify a visual target. They used a 10 Hz click train from a speaker that was either spatially correlated or uncorrelated to a target light. The results showed that spatially correlated information from an auditory source substantially reduced visual search time (between 175-1200 ms). In an experiment by Sorkin, et al. (ref. 3), localization accuracy rather than target acquisition time was studied in a simulated cockpit environment. A magnetic head tracker was either correlated or uncorrelated with a 3-D audio display that corresponded to the locations of visual targets. Results of the study found that accuracy of azimuthal localization was improved when head movement was correlated with the 3-D audio display, but that elevation localization was no better than chance.

Begault (ref. 4) evaluated the effectiveness of a 3-D head up auditory TCAS display during a full-mission simulation by measuring target acquisition time. All crews used visual out-the-window search in response to a TCAS advisory, since no planform display was used. Half the crews heard the standard loudspeaker audio alert, and half heard an alert that was spatialized over headphones using 3-D sound techniques. The direction of the spatialization was linked to the target location's azimuth, but not its elevation. In addition, the spatialized audio stimuli were exaggerated by a factor of three in relationship to the visual angle to facilitate head movement in the aurally guided visual search (e.g., visual targets at 10° azimuth would correspond to spatialized stimuli at 30° azimuth). Results of the study found a significant reduction in acquisition time when using spatialized sound (4.7 vs. 2.5 s).

The current study evaluated the feasibility of using either a head down visual display (standard TCAS), or a head up audio display (3-D TCAS). 3-D sound was used for aurally guided visual search as in the study by Begault (ref. 4), but without inclusion of the exaggeration factor mentioned above. In addition, the 3-D audio display in the current experiment included three categories of elevation cues. Two groups consisting of 5 crews were evaluated during a full-mission simulation. It was hypothesized that a significant difference in both acquisition time and the number of targets acquired might occur between the two conditions.

EXPERIMENT METHOD

Subjects

Ten two-person flight crews served as subjects for this study. Crews were composed of airplane pilots employed by a major U.S. air carrier and were rated in a glass cockpit aircraft (e.g., Boeing 757, 767, 737-300/400 or 747-400). Each crew member was paid a nominal amount for participating. Since all crew members had current medical certificates, they had been previously evaluated for normal hearing within the last year (First Officers) or six months (Captains) by company and FAA medical examiners.

2

Experimental Design

Two groups, each comprised of five two-person crews, were evaluated in a between-subjects design. The standard TCAS group used a audio-visual system approximating the TCAS system currently implemented in U.S. commercial air carriers. This consisted of an audio traffic advisory presented via an overhead speaker and a standard TCAS head down map display.

The 3-D TCAS group used stereo headsets and were presented a binaurally-processed version of the audio portion of the traffic advisory, but were not supplied with any visual system information. The perceived direction of the 3-D auditory advisory was adjusted to correspond to the azimuth of the target out-the-window.

The 10 crews were assigned randomly to either the standard TCAS or the 3-D TCAS group. The dependent variables were: (1) the time interval between the appearance of a visual target in conjunction with an aural advisory and the verbal response from a crew member indicating acquisition of the target; and (2) the number of targets acquired.

The crew members were instructed to call out verbally when they had visually acquired the aircraft outside the window (a consistent utterance, such as "got it!"). Acquisition time (the difference between the time the visual target was generated and the beginning of the verbal utterance) was observed on video tapes by an unbiased researcher. Each verbal acquisition increased the count for the number of targets acquired. The acquisition time was determined by the time code generated on the video tapes. The accuracy of determining the beginning of the verbal utterance was within 2 video frames (0.066 s). Target acquisition times and the number of targets acquired were also categorized according to whether the target was visible to both or to only one crew member.

Stimuli

A total of 24 targets were presented to the crews for evaluation during the cruise phase of the flight. Six additional targets were included as "dummy" targets to provide a realistic context for the TCAS system in the vicinity of the airports (3 during takeoff and 3 during landing phases of flight) and were not included as part of the experiment's data set. This was because of the relatively high amount of variability which can occur between crews during takeoff and landing phases of flight (e.g., workload, ATC communications). Also, in the vicinity of the airports (but not during the cruise phase of flight), simulated city lights are visible, making the out-the-window scene difficult to control across crews. The relative luminosity and contrast ratios between modeled airport data and the targets would otherwise be an uncontrolled variable in acquisition time, since crews approach airports in a slightly different manner and time.

In order to maintain a consistent visual image size, all targets were fixed at a 3 mi. distance from the aircraft. This made the target appear as a flashing dot of light similar to that seen out the cockpit window of a real aircraft. However, the target did not change size from the perspective of the subjects since its position was always linked to the position of the simulator aircraft; in other words, it visually appeared to remain at a fixed distance and identical speed to the simulator. This was done to eliminate

3

movement of the target as a variable, and to eliminate differences between crews as a function of movement of the aircraft.

The out-the-window positions of the targets patterned a 3x8 matrix (see Figure 1). The positions of the targets were randomly assigned to 1 of 8 azimuths (-50°, -37°, -22°, -10°, 10°, 22°, 37°, and 50°) within 3 elevations. Five targets were assigned to azimuths 3,000 feet above own ship; 14 targets at azimuths at the same elevation as own ship; and 5 targets at azimuths 3,000 feet below own ship.

For the Standard TCAS condition, a computer generated four to six moving symbols depicting aircraft. The symbols appeared at pseudo-random positions, and were presented on the TCAS map display. One of the symbols would be elevated to advisory status for target acquisition evaluation, while the remaining symbols would eventually vector off the display.
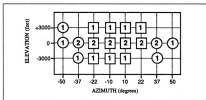

Figure 1. The relative elevations and azimuths of the 24 targets used in the experiment. The number indicates the frequency of occurrence; the squares indicate targets visible to both crew members.

Available Field-of-View

A substantial limitation inherent in all flight simulators is the available out-the-window field-of-view for each pilot. The simulator used in this experiment was a modified Lockheed-Georgia cockpit equipped with a Singer-Link Advanced Simulator Technology visual system. This system had a 3-channel, 4-screen display. Each channel contained a discrete display of visual information relevant to the scenario; the 2 center screens in the simulator displayed an identical visual scene from 1 channel, with 1 screen visible by each pilot. The center channel screen enables each pilot a field-of-view extending to approximately ±25° azimuth. In addition, 2 side screens fed by the other 2 channels gave each pilot a unique side field-of-view that extended the total field-of-view to approximately ±52° azimuth.

Figure 2 shows the available field-of-view for the Captain; the First Officer's would be the mirror image of this figure. Note that the field-of-view from 25° to 52° is available only to 1 crew member while the area between ±25° is available to both crew members.
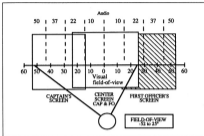
4

Figure 2. The horizontal field-of-view in the simulator, from the perspective of the left seat (Captain's position). The numbers within the dashed lines show the mapping between visual azimuths and the specific azimuth position of the 3-D sound cue that was used for the alert.

Figure 3 shows the vertical field-of-view. The immediate range is from approximately -13° to +16°, but can extend from -18° to +20° with head and body adjustments. For reference, the visible range of a target at 3 miles is shown in terms of relative elevation (in feet) above and below the simulator.
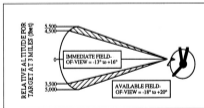

Figure 3. The vertical field-of-view in the simulator, based on the relative altitude of an aircraft at 3 miles distance.

Audio Environment and 3-D Sound Processing

A special TCAS advisory sound was formed for this experiment. Specifically, in addition to the usual TRAFFIC-TRAFFIC enunciation, a pre-advisory tone was used. The pre-advisory consisted of two brief (66 ms) complex tones (labeled BIP) separated by 39 ms of silence. These were synthesized by adding multiple square waves with different fundamental frequencies, and then giving the overall composite a rapid amplitude envelope rise time to favor the conveyance of spatial information. Because of the rich harmonic structure, it could be played at a level approximately 10 dB below the speech alert and still be noticeable. The TCAS speech alert TRAFFIC-TRAFFIC was digitally recorded by a male speaker in a soundproof booth using an electrostatic microphone, preamplifier, and a digital audio tape (DAT) recorder.

The total duration of the alert was 1.36 s: 171 ms for the pre-advisory; a 85 ms silent interval; 462 ms for the word TRAFFIC; a 180 ms silent interval; and another 462 ms TRAFFIC (see Figure 4). This recording was transferred to a desktop computer using audio recording software and hardware at a sampling rate of 50 kHz. Next, the aural alert was convolved with HRTF measurements for the spatial auditory positions: Left $10°$, $22°$, $37°$, and $50°$ azimuth. Positions for right $10°$, $22°$, $37°$, and $50°$ azimuth were obtained by reversing the output channels at playback, resulting in a total of eight available spatialized positions.

The convolution was performed in non-real time on the desktop computer by supplying formatted versions of the measurements to a standard signal processing package. The resulting signals were then converted to a 33.3 kHz sample rate in 12-bit signed integer form and subsequently stored in a stereo audio sampler. The stimuli were played back in coordination with the scenario software via *note on/off* commands inherent to the Musical Instrument Digital Interface (MIDI) specification.

Each pilot wore a stereo headset (a modified Sennheiser HME 1410-KA) that was selected for comfort and fidelity. The headphone frequency response ranged between 20 Hz-18 kHz and weighed 250 g. The headset had a supra-aural design (the drivers rested on the outside of the ears), allowing outside conversation to be monitored more easily than with a circumaural design. Playback of the speech portion of the alert was at approximately 74 dB SPL at the transducer; the simulator's ambient background noise was approximately 70 dB (C weighting) measured in the center of the cockpit with an omnidirectional microphone during the cruise phase of flight. The spectrum of the ambient sound was approximately that of white noise (for wind simulation) combined with engine sound simulations.
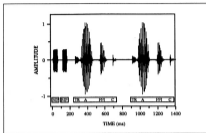
6

Figure 4. Arrangement of the pre-alert and traffic enunciation used for TCAS advisories in the experiment.

PROCEDURE

Training

Each crew spent 2 days at the simulator, with the first day-and-a-half devoted to familiarization and training. The training period focused on the particular handling capabilities of the aircraft, the touch screen displays, electronic checklist, and procedures to be used. It also included a brief demonstration of the 3-D audio system for the 5 crews using that system. This consisted of a 2 minute demonstration of several targets accompanied by the 3-D audio traffic alert. No other information was given to the pilots about the nature of the experiment.

Scenario

The crews flew the experimental flights on the afternoon of the second day. The experiment was conducted during the cruise phase of the fourth and final leg (SFO – LAX) flown. The first three legs of the scenario were considered practice and therefore were excluded from the analysis. The 24 targets were designed to occur at an approximate rate of 1 every 3 minutes during the cruise phase of flight (more than 15 miles from departure or destination). Each individual target was activated according to the distance in miles from the destination. During the experiment, all normal operations were realistically simulated, including conventional VOR navigation and communications with ATC (ground, tower, approach, departure, and center). Complete darkness was simulated with approximately 50 mi. visibility throughout the flights. Crews were instructed to follow their normal company standard operating procedures as closely as possible.

7

RESULTS

A target was considered to have been "acquired" if the crew obtained it within a 10 s time window, which is the limit before the traffic could potentially be elevated to *traffic resolution* status in a real situation. Only 2 targets were acquired outside this time window and were treated as outliers.

Based on the examination of acquisition time, a total of 20 outliers (acquisition times > 3 SD) were found. The standard TCAS group had 7 outliers, 2 being extreme outliers (±5 SD), while the 3-D TCAS group had 13 outliers, 1 extreme. All outliers greater than 3 SD were excluded from the analysis. These outliers appeared in a random manner among crews and condition, and did not correlate to specific targets.

A 2-way analysis of variance (ANOVA) with acquisition time as the dependent variable was conducted. This analysis (*Condition x View*) was conducted to determine if significant differences existed between targets in the field-of-view available to both crew members versus individual field-of-view. The mean acquisition time for the standard TCAS group was 2.63 (SD, 1.19), while the mean for the 3-D TCAS group was 2.13 (SD, 0.78). The ANOVA revealed a significant main effect for condition, $F(1, 187) = 15.09$, $p < 0.0001$, as well as a significant main effect for view, $F(1, 187) = 50.37$, $p < 0.0001$, although there was no interaction present $F(1, 187) = 1.76$, $p > 0.05$.

An additional ANOVA (*Condition x Elevation*) was conducted to determine if there were significant differences in target acquisition time for targets at the aircraft's elevation versus targets from above and below (i.e., those that fell into the upper or lower horizontal sections of the grid). This analysis also showed a significant main effect for condition, $F(1, 187) = 11.19$, $p < 0.001$, but no significant main effect for elevation, $F(1, 187) = 1.01$, $p > 0.05$, or interaction present, $F(1, 187) = 0.11$, $p > 0.05$. Figure 5 displays the mean target acquisition time and standard deviations for the 24 targets.

An additional set of analyses were conducted using the number of targets acquired as the dependent variable. The mean number of targets acquired for the standard TCAS group was 19.4 (SD, 1.95), while the mean for the 3-D TCAS group was 18.2 (SD, 2.95). There were no significant main effects or interactions for these analyses, although there was approaching significance for the number of targets acquired at a particular elevation, $F(1, 219) = 3.65$, $p < 0.06$.
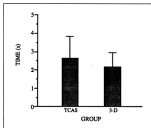
8

Figure 5. Mean target acquisition times (2.63 vs. 2.13) and standard deviations for the 24 targets used in the experiment.

CONCLUSIONS

The results of this experiment imply that the presence of a spatial auditory cue can significantly reduce the time necessary for visual search in an aeronautical safety environment. This result is in line with the studies of Perrott, et al. (ref. 5) and Perrott, et al. (ref. 2) that found advantages for aurally guided visual search using analgous conditions in the laboratory. Although 500 ms may seem to be a modest improvement, it does suggest that, in an operational setting, an aural 3-D TCAS display may be desirable in addition to a standard TCAS display. This is because pilots can keep their head "out the window" looking for traffic without needing to move the head downwards to the platform map display and then back up. In other words, by accessing an alternative perceptual modality—sound— the visual perceptual modality is freed to concentrate on other tasks, if necessary. In an actual cockpit with 3-D sound added to the current TCAS system, the pilot flying could use the auditory information for immediate head up search while the pilot not flying could gain numerical altitude information and verify the direction for the other pilot. Future experiments will focus on evaluating the combination of the two systems.

Begault (ref. 4) evaluated 3-D and monaural traffic alerts in a similar experiment, but without use of a head down map display. In that experiment, the spatialized positions were "exaggerated" in relation to the visual display by a factor of three. Spatialization of the aural alert resulted in a decrease in the mean target acquisition time from 4.7 s to 2.5 s. This result, along with the current data, suggests that spatial processing of an auditory alert is useful for guided visual search; in other words, aural alerts have greater potential in human-machine interfaces than to function merely as "attention getting" mechanisms. Begault (ref. 4) suggested the use of the exaggerated auditory azimuths may have contributed to the faster acquisition times for the 3-D display. The mean target acquisition time of 2.5 s (SD, 0.8) in that study

was greater than that found in the present experiment (M, 2.13 s; SD, 0.78), suggesting that exaggerated auditory stimuli are not necessary for effective aurally guided visual search.

Overall, the results presented here must be evaluated provisionally, particularly for the reason that a simulator's field-of-view is not at all equivalent to that in an actual aircraft, in spite of the substantial efforts to insure realism. Unlike actual cockpits, the field-of-view in the simulator is such that the person sitting on the left side cannot see beyond 25° to the right, and the person on the right side cannot see beyond 25° to the left. So it may have been that the spatial auditory cues were used for crude estimates of visual target positions, in order to transcend the limitations of the simulator environment (i.e., if it sounds to the right, the First Officer searches, and if it sounds to the left, the Captain searches). This is indicative of a task delegation procedure. However, in an actual operations context, visual search is usually conducted most actively by the pilot not flying, depending on the context of the phase of flight and the relative urgency of the TCAS alert. Even if this trade-off feature were not an element, the spatial auditory cue could still have been utilized as a crude way for determining where to begin visual search. If it is true that the spatial sound cue provides a general direction for search that is subsequently refined by visual search, then the additional azimuthal accuracy provided by a head-coupled 3-D auditory display (Sorkin, et al., ref. 3) is probably unnecessary.

Parallel explorations of aurally guided visual search should continue to be evaluated under controlled laboratory conditions, and then compared to research under actual flight operations. An important factor, not evaluated here, is that out-the-window targets can move quickly across the field-of-view; the work by Strybel, et al. (ref. 6) on evaluating the Minimum Audible Movement Angle (MAMA) is particularly relevant in this regard. Future experiments at NASA Ames will evaluate standard TCAS systems augmented with 3-D audio displays of directional information. These are more likely to be implemented in actual operations than a purely 3-D audio TCAS system, since increased safety via a redundant system is more desirable than replacing one system with an equivalent one.

ACKNOWLEDGMENTS

10

REFERENCES

1. Chappell, S. L., Billings, C. E., Scott, B. C., Tuttell, R. J., Olsen, M. C., and Kozon, T. E. (1989). *Pilots' Use of a Traffic Alert and Collision-Avoidance System (TCAS II) in Simulated Air Carrier Operations. Volume I: Methodology, Summary and Conclusions.* (NASA Technical Memorandum 100094). Moffett Field, CA: NASA Ames Research Center.

2. Perrott, D. R., Sadralodabai, T., Saberi, K., and Strybel, T. Z. (1991). Aurally aided visual search in the central visual field: Effects of visual load and visual enhancement of the target. *Human Factors,* 33 (4), 389-400.

3. Sorkin, R. D., Wightman, F. L., Kistler, D. S., and Elvers, G. C. (1989). An exploratory study of the use of movement-correlated cues in an auditory head-up display. *Human Factors,* 31 (2), 161-166.

4. Begault, D. R., (1993). A head-up auditory display for TCAS advisories: *Human Factors,* 35 (4), 707-717.

5. Perrott, D. R., Saberi, K., Brown, K., and Strybel, T. Z. (1990). Auditory psychomotor coordination and visual search behavior. *Perception and Psychophysics,* 48, 214-226.

6. Strybel, T. Z., Manligas, C. L., and Perrott, D. R. (1992). Minimum audible movement angle as a function of the azimuth and elevation of the source. *Human Factors,* 34 (3), 267-275.

11

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE March 1994	3. REPORT TYPE AND DATES COVERED Contractor Report	

4. TITLE AND SUBTITLE

3-D Audio Versus Head Down TCAS Displays

5. FUNDING NUMBERS

NCC-2-327

6. AUTHOR(S)

Durand R. Begault and Marc T. Pittman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

San Jose State University Foundation
One Washington Square
San Jose, CA 95192

8. PERFORMING ORGANIZATION REPORT NUMBER

A-94066

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

NASA CR-177636

12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Unclassified-Unlimited Subject Category – 03	

13. ABSTRACT (Maximum 200 words)

The advantage of a head up auditory display was evaluated in an experiment designed to measure and compare the acquisition time for capturing visual targets under two conditions: Standard head down traffic collision avoidance system (TCAS) display, and three-dimensional (3-D) audio TCAS presentation. Ten commercial airline crews were tested under full mission simulation conditions at the NASA Ames Crew-Vehicle Systems Research Facility Advanced Concepts Flight Simulator. Scenario software generated targets corresponding to aircraft which activated a 3-D aural advisory or a TCAS advisory. Results showed a significant difference in target acquisition time between the two conditions, favoring the 3-D audio TCAS condition by 500 ms.

14. SUBJECT TERMS			15. NUMBER OF PAGES 17
Auditory localization, Human machine interfaces, Speech communication, Auditory displays			16. PRICE CODE A03

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified		