Running head: COCKPIT 3-D AUDIO WARNINGS

Effectiveness of 3-D Audio for Warnings in the Cockpit

A.B. Oving, J.A. Veltman and A.W. Bronkhorst

TNO Human Factors, Soesterberg, The Netherlands

Abstract

We investigated the application of three-dimensional (3-D) audio for presenting auditory warnings in the cockpit in two simulator experiments. The primary task was to follow a flight path, and additionally to respond to warnings of the Traffic alert and Collision Avoidance System (TCAS) and to system malfunction warnings. All warnings were presented both visually and aurally. The first experiment showed 12% faster response times when 3-D audio was used for the auditory warnings, compared to mono sound. This was only observed when flight path error was presented on a visual display. No effect was observed when the error was presented aurally, because this enabled pilots to pay more attention to the visual warning displays. In the second experiment, only a visual error display was employed, and the effects of 3-D audio and verbal directional information (e.g., 'up') were tested. Each type of cue resulted in a 12% reduction of response times. The combination of 3-D audio and verbal cues resulted in the best performance with response time reductions of approximately 23%. These results indicate that 3-D audio displays can improve flight safety.

7465 words

Requests for reprints should be sent to Arjen Oving, TNO Human Factors, P.O. Box 23, 3769 ZG, Soesterberg, The Netherlands; Oving@tm.tno.nl

Keywords: 3-D audio, auditory displays, warnings, flight simulator

Effectiveness of 3-D Audio for Warnings in the Cockpit

Introduction

Technological developments have made it possible to generate sounds via a set of earphones, which seem to originate from a specific location in three-dimensional (3-D) space. This 3-D audio requires the use of head-related transfer functions (HRTFs) that specify sound intensity and delay as a function of frequency for given source positions in 3-D space. Localization accuracy with such 3-D audio is almost the same as with natural sound sources (Wightman & Kistler, 1989; Wenzel, Arruda, Kistler & Wightman, 1993; Bronkhorst, 1995; Kulkarni & Colburn, 1999; Langendijk & Bronkhorst, 2000). With the use of HRTFs, 3-D audio displays can be created that can signal relevant directional information to a human operator in a natural manner.

Several flight simulator studies have investigated the use of a 3-D audio display to present the traffic advisory (TA) warning of the Traffic alert and Collision Avoidance System (TCAS) that is installed in most commercial aircraft. This system presents an auditory warning (i.e., 'Traffic.....Traffic') and visual information on a radar display when another aircraft enters a pre-defined airspace around the pilot's aircraft (see Begault, 1993, for a more detailed description). In these situations, fast localization of the intruding aircraft is important.

In a first study, Begault (1993) examined the use of 3-D audio to present the standard auditory TA warning of TCAS. With 3-D audio, the auditory warning signaled the relative direction of the intruding aircraft to the pilots. The 3-D audio display only employed azimuth cues to indicate the horizontal location of the target aircraft. The results showed that search time for the out-the-window target aircraft was reduced considerably when 3-D audio was used instead of mono sound to present the warning (i.e., 2.5 s versus 4.7 s, a reduction of 47%). However, this study did not use any visual displays to present TCAS-information to the pilots, which is common in aviation. This may have hampered search performance in the mono sound condition considerably, because pilots had no source of information regarding the location of the target.

In a second study, Begault and Pittman (1996) did use a visual head-down display (HDD) for TCAS information, but only in the mono sound condition. The HDD consisted of a 2-D map display that indicated the relative position of nearby and intruding aircraft to the pilots. The 3-D audio display provided both azimuth and elevation cues. Again, an advantage of 3-D audio over mono sound for target search time was observed (i.e., 2.13 s versus 2.63 s, a reduction of 19%), although this advantage was reduced compared to the first study. This reduction in effectiveness can be attributed largely to the inclusion of the visual TCAS display in the mono sound condition.

Begault, Wenzel and Lathrop (1997) investigated the use of 3-D audio in combination with a visual TCAS display. This study employed a comparable paradigm as the study of Begault and Pittman (1996), but now, a visual HDD with TCAS information was present in all experimental conditions. Again the results showed faster search performance with 3-D audio than with mono sound (i.e., 1.89 s versus 2.20 s, a reduction of 14%). However, compared to the previous study, the relative advantage of 3-D audio is reduced again. It is possible that differences in the experimental setup contributed to this reduction (e.g., a higher frequency of warnings). It may also be due to the presence of the visual TCAS display in the 3-D audio condition. The pilots may have used this display on some occasions to determine or confirm the location of the target, in spite of the presence of 3-D audio. The extra time needed to visually inspect the HDD may have reduced the benefits of 3-D audio.

Begault et al. (1997) do report on a subsequent analysis of the results of the previous study. In that study, it was observed that the elevation cues incorporated in the HRTFs were not effectively utilized by the pilots. This is not surprising because such cues tend to be perceived less well than azimuth cues (Bronkhorst, 1995; Wenzel et al., 1993). Therefore, Begault et al. (1997) suggest that the use of a verbal cue for elevation may have been better.

The application of 3-D audio can also be extended to other types of auditory signals in the cockpit. For instance, Haas (1998) used 3-D audio as a warning display for system malfunctions in helicopters. In this study, the spatial source of the 3-D audio warning corresponded to the location of a system malfunction of the aircraft, or to the location of a visual indicator light inside the cockpit. For instance, the warning signal for a fire in the left engine fire was presented to the left of the pilot's head. The results showed faster response times to the warnings when they were presented with 3-D audio (i.e., 3.6 s, on average), compared to the condition when only visual warning signals were present (i.e., 5.0 s). This reduction of 1.4 s (i.e., 28%) can be considered as an important reduction regarding the nature of the task. However, it should be noted that no auditory signals were presented in the visual only condition, which makes the contribution of spatialized sound to the performance improvement less clear.

The Present Experiments

In the present two experiments we examined the use of 3-D audio for presenting both TCAS warnings and system malfunction warnings. There are a number of differences with the Begault studies (Begault, 1993; Begault & Pittman, 1996; Begault et al., 1997) and the study of Haas (1998). We always tested 3-D audio as a supplement to visual warning signals presented on a HDD, because it is expected that such visual aids will never be absent. In addition, we always compared the 3-D audio displays to mono sound displays, to determine the effects of using spatialized sound for auditory warnings. Furthermore, we used a head tracking system to make the external position of the sound source with 3-D audio invariant under head movements. Changes in the differences between the two ears with respect to stimulus characteristics that resulted from head movements (e.g., shifts in arrival time and frequency spectrum of the signal), could then be used as a cue for the localization of the sound sources. In combination with the use of individualized HRTFs, this can improve localization accuracy to approximately the level of accuracy with real sound sources (Bronkhorst, 1995). The 3D audio system also provided a relatively high resolution of about 5°. The direction of the sound source thus closely corresponded to the direction of the target location (within the limits of the system resolution). Another difference is that a part-task flight simulator with only one seat was used in the present experiments, limiting the number of tasks that the single pilot had to perform. This provides greater experimental control, but has the downside that subjects can put all their effort in the limited number of tasks and subsequently reduce potential effects of the manipulations. To keep the pilots busy, we made the primary flight task a continuous tracking task, by having the pilots follow a specific flight path.

It was hypothesized that pilots will respond faster to the auditory warnings when they are presented with 3-D audio, because of the natural ability of humans to perceive the spatial information included in such an auditory signal. Even in the presence of visual displays that provided spatial information regarding the warnings, benefits of 3-D audio were expected. Visual displays have several drawbacks for providing aircraft-referenced spatial information compared to 3-D audio (although exceptions such as head-up and helmet-mounted displays exist). For instance, the spatial information in a planview radar display is compressed and requires integration of spatial dimensions in order to determine the relative location of other

aircraft in real space. In addition, the pilot has to switch visual attention between the in-cockpit display and the out-the-window view. This is not the case with 3-D audio, which can provide uncompressed and integrated spatial information in an aircraft-referenced display, while the pilot flies head-up.

In the first experiment, we studied the effects of 3-D audio for presenting cockpit warnings. This was done in two different conditions regarding the primary flight task. In the first condition, information on flight path deviations was presented on a visual HDD. In the second condition, this information was presented by means of an auditory display. Using mono sound, verbal directions were presented when deviations became too large, enabling pilots to fly more head-up. When 3-D audio is used to present the TCAS or system failure warnings in this latter situation, the pilots do not need to transition from the out-the-window scene to the in-cockpit visual display in order to retrieve spatial information for these warnings, thereby minimizing the time needed to respond to the warnings. Therefore, it was expected that the effects of 3-D audio would be more pronounced in the second condition.

In the second experiment, we examined the effects of both 3-D audio and verbal directional information for the auditory warnings (e.g., verbal indicators such as 'up' and 'left'). This was done to see if more cognitive spatial information could also assist the pilot, as suggested by Begault et al. (1997), compared to the perceptual spatial information with 3-D audio. In addition, such verbal cues may support the perception of source elevation when they are combined with 3-D audio.

Experiment 1

Method

Subjects

Twelve male student pilots participated in the study. They were all students at the Royal Dutch Airlines Flight Academy (KLS) and were in the final stage of their training or had recently completed this. They were aged between 19 and 29 years.

Tasks

Flight task. The primary task for the subjects was to follow a predetermined flight path that consisted of a holding phase and an approach phase. The holding phase involved occasional horizontal turns and changes in altitude, and the approach phase involved a straight and gradual descent towards a runway. The flight path was presented on a HDD as the center of a tunnel-in-the-sky (see Figure 1). The width and height of this tunnel represented 300 m, and it showed the flight path up to a distance of 4000 m. A small cross in the display indicated the current position of the aircraft. The task was performed in two modes: a visual and an auditory mode. In the visual mode, a second and smaller tunnel was presented within the tunnel-in-the-sky on the HDD. The height and width of this tunnel represented 50 m and provided the pilots with more precise information on deviations from the flight path. In the auditory mode, this second tunnel was not present. Now, verbal directional information was presented whenever the deviations from the flight path became larger than 25 m (either horizontally or vertically). This verbal information indicated the direction in which the pilots had to steer the aircraft in order to minimize the error (e.g., "up left" or "down"). These verbal cues were always presented in mono sound. One flight run lasted approximately 27 min.

Warning tasks. For the TCAS task, subjects had to respond to a TA warning of the system. This warning consisted of the standard verbal warning (i.e., 'Traffic.....Traffic') via the earphones, and a visual symbol on a map display in the cockpit, indicating the relative azimuth and elevation of the intruding aircraft (see Figure 1). No information about the target aircraft was visible prior to the warning. The target was always kept at a fixed position relative to the own aircraft, so that any movements of the own aircraft did not affect the relative direction of the target location. The subjects were required to press a button on the cockpit console when they visually detected the intruding aircraft. To avoid guessing, the intruding aircraft was not visible in the out-the-window scene in 2 of the 8 warnings during a trial. In these cases, the subjects had to press another button, signaling that the aircraft was not present in the outside view. The target could appear in one of eight different locations: -55°, -22°, 22° and 55° azimuth combined with 0° or 12° elevation (note: the available field of view from the mockup allowed the target on the far left side to be presented at an elevation of -10°). All targets were directly visible from the position of the pilot. There were eight TCAS warnings during a flight run, each in a different location. The warnings were presented in random order and at random moments during a flight run.

For the system malfunction task, subjects had to respond as quickly as possible to a system malfunction warning. There were four systems for which a malfunction could occur: engines, flaps, ailerons and spoilers. Each malfunction could occur on either the left side or the right side of the aircraft, resulting in eight different warnings. The subjects responded by pressing one of eight buttons that corresponded with a particular warning. The visual warning display always indicated the type and location of the malfunction (see Figure 1), whereas the auditory display only indicated the specific system (e.g.. "Flaps.....Flaps"). All eight malfunction warnings were presented during a flight run, in a random order and at random moments. However, the warnings for the flaps and the spoilers were only presented when these systems were used. When a malfunction was scheduled, but the respective system was not active, the warning was postponed until the pilot did use that system.

In total, there were 16 warnings during a single flight run. The subjects did not receive any feedback on their response to a warning. The warnings were presented continuously until the subject had pressed one of the response buttons. All auditory signals were presented to both ears, and consisted of digitally recorded broadband speech.

Apparatus

The experiment was conducted in a fixed-base simulator with a spherical dome projection system. The projected scene with 156° horizontal and 42° vertical field of view was generated with a three-channel high performance graphics system (Evans & Sutherland ESIG-2000) at an update rate of 30 Hz. A mockup of a partially instrumented cockpit of a commercial passenger aircraft was placed in the center of the dome, at an observation distance of 3 m. The pilot sat on the left side of the mockup. Controls were installed for direction and speed control, as well as switches for flaps and spoiler settings.

A monitor located in front of the pilot functioned as the HDD (see Figure 1). The upper left section of this HDD showed system status information, and displayed the visual warning signals for the system malfunction task. The upper right section showed, among other things, information about aircraft heading, roll and pitch, airspeed, and altitude. The flight path display with the tunnel-in-the-sky was located in the lower right section. The smaller tunnel consisted of lined-up squares at a distance of 100 m from each other. For the larger tunnel, the lines of the four corners were drawn, with cross-sections located at every 100 m on these lines. The TCAS radar display was presented in the lower left section. This display is configured heading-up, providing aircraft-referenced positional information about other

aircraft to the pilot. A dot indicated the relative horizontal orientation of the intruding aircraft, while numbers placed above the dot indicated the relative elevation of the target. Only one target was presented during a warning.

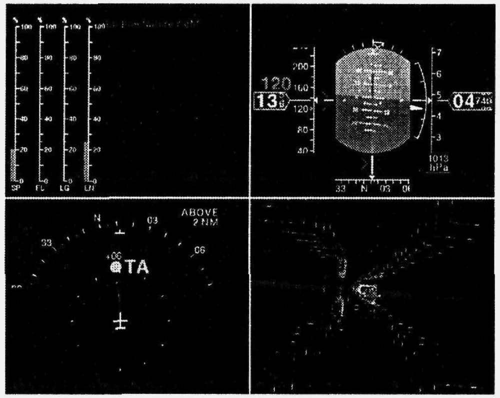

Figure 1. The main visual head-down display that was implemented in the cockpit console in front of the pilot.

3-D audio system

Three-dimensional audio was generated with a set-up consisting of an AKAI S3200 XL sampler, a PC equipped with two Loughborough DSP32C boards, a mixer, a headphone amplifier driving Sennheiser HD 530 headphones and a Polhemus Fastrak head tracker. The PC was synchronized to and controlled by the simulator PC through a parallel port connection. The two DSP boards in the sound PC performed the convolution with HRTFs. The outputs of the boards were then mixed, amplified and fed into the headphones. Head movements of the participant were registered by the head tracker and transmitted through an RS232 interface to the sound PC.

The HRTFs that were used to create 3D audio were measured for each subject individually on a day prior to the experiment. A blocked ears measurement procedure was applied, using microphones in foam earplugs, that were inserted into the ear canals (Møller, Sørensen, Hammershøi & Jensen, 1995). The measurements involved 976 different source positions, covering all directions with elevations greater than -60°, with a resolution of 5-6° (Bronkhorst, 1995). The raw HRTFs, measured with a 10-ms test sound, were transformed to minimum phase impulse responses with lengths of 3.4 ms. The actual filter used for 3D sound generation had a length of 13.7 ms and contained both the direct sound and the six first-order reflections. Reflections were calculated for a rectangular room with a volume of 25 m^3 and surfaces with an absorption factor increasing from 0.41 at 250 Hz to 0.93 at 8 kHz. The reflections were added in order to increase the perceived distance of the virtual sound sources (Bronkhorst & Houtgast, 1999). HRTF filtering was performed in the frequency domain, using the overlap-add algorithm (see e.g. Oppenheim & Schafer, 1989). This introduced a delay of about 34 ms between an update of the virtual sound source position by the sound PC and the actual modification of the 3D sound. In creating the HRTFs, no interpolation between measured directions was performed.

Experimental design

The flight task was performed in either a visual mode or in an auditory mode. Within each mode, the two warning tasks were performed with either mono sound or with 3-D audio. This resulted in four experimental conditions. The subjects only performed one run in each condition. In the mono sound conditions, the unprocessed auditory warnings were presented to both ears. With 3-D audio, the direction of the sound source corresponded to the direction of the target aircraft location (i.e., the azimuth and elevation of the target location) or to the direction of the malfunction location (i.e., 90° or -90° azimuth at 0° of elevation).

A within-subjects design was employed. Half of the subjects started in the visual mode for the flight task, while the other half started in the auditory mode. Within each of these groups, half of the subjects started with mono sound and the other half with 3D audio. The order of the sound conditions remained the same for all subjects when they switched to the other flying mode.

The response time for the warning tasks, defined as the time between the onset of the warning and the first button press, was used to analyze the performance on these tasks. In addition, the trials in which the target aircraft was not present in the outside view were excluded from the analysis. For each subject, the mean response time per condition was calculated and subsequently used in the statistical analyses. The percentage of errors (i.e., incorrect responses) for each warning task was also determined. The results were analyzed with an analysis of variance (ANOVA) with flying mode (auditory or visual) and sound (mono sound or 3-D audio) as within-subjects variables. Where appropriate, significant effects were further analyzed by post-hoc Tukey tests.

Procedure

Two subjects were present on an experimental day. During the morning, they were trained in the experimental tasks and conditions for approximately two hours each. Regarding the experimental tasks, they were instructed to follow the flight path as accurately as possible, and to respond as quickly as possible to any warning without making any errors. The experimental runs were performed in the afternoon, with the subjects taking turns after each completed experimental run. They were seated in a separate room between experimental runs, to prevent any interaction between them. No feedback was given regarding task performance.

Results and discussion

Due to technical problems, parts of the data were lost for three subjects. Therefore, their data were excluded from the statistical analysis. This resulted in an unbalanced design, in which six of the remaining nine subjects had started in a 3-D audio condition. Regarding the malfunction warning task, only the response times were registered properly, making it impossible to determine the percentage of errors.

TCAS task. The ANOVA on the response times to the TCAS warnings revealed a significant interaction between flying mode and sound [$F(1, 8) = 11.58$, $p < .01$]. This interaction is shown in Figure 2. The response times in the auditory mode increased somewhat when 3-D audio was used (i.e., from 3.11 s to 3.33 s), but this difference did not prove significant in a post-hoc test. However, the response times decreased considerably in the visual mode when 3-D audio was present (i.e., from 3.64 s to 3.16 s). The post-hoc test revealed that this difference between the two sound conditions was significant ($p < .05$). No effects were observed for the percentage of errors in detection of the target (i.e., 8.7% overall).

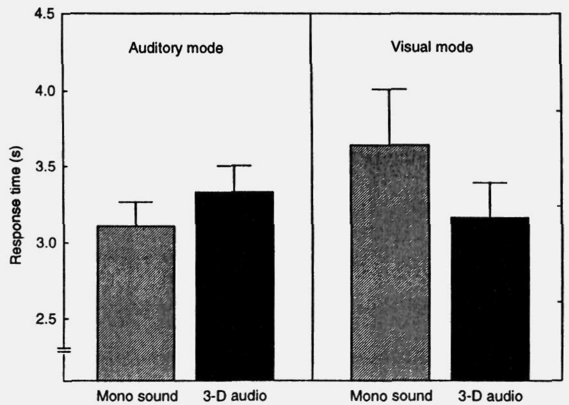

Figure 2. Mean response time to a TCAS warning, as a function of flying mode and type of sound (standard errors included).

In the visual mode, 3-D audio thus resulted in faster response times to the warning. The reduction in response time was approximately 0.48 s or 13%, which is comparable to the results of Begault et al. (1997). This indicates that the pilots did use the directional cues available with 3-D audio to search for the intruding aircraft. However, such an effect of 3-D audio was not observed in the auditory mode, even though a more pronounced advantage of 3-D audio was expected in this condition. Pilots were hypothesized to fly more head-up in this condition and subsequently take more advantage of the 3-D audio display, because they were able to pay less attention to the visual flight path display. Apparently, the subjects did not fly more head-up in the auditory mode, but instead paid more attention to the in-cockpit visual warning displays. This hypothesis is supported by the observation that the response time with mono sound in the auditory mode was faster than the response time with mono sound in the visual mode ($p < .05$). This difference can only be explained if the subjects were already focused more on the warning display, and not on the out-the-window scene. Because the pilots had to perform a limited number of tasks and knew that warnings would occur, this strategy enabled them to be more prepared for any oncoming warnings, thereby negating the benefits of 3-D audio.

System malfunction task. The analysis of the response times for the system malfunction task also showed a significant interaction between flying mode and sound [$F(1, 8) = 6.85$, $p < .05$]. A comparable pattern as for the TCAS warnings was observed (see Figure 3). The response times in the two sound conditions in the auditory mode did not differ statistically from each other. This can also be attributed to the reduced need for the subjects to visually scan the flight path display in the auditory mode, and subsequently pay more attention to the visual warning display. This enabled them to pick up the visual system malfunction warning quicker, reducing the need for other directional cues as provided with 3-D audio.

The post-hoc test on the interaction result did show a trend towards significance for the difference between the two sound conditions in the visual mode ($p < .07$). The responses were faster with 3-D audio than with mono sound: 2.23 s and 2.52 s, respectively, a reduction of approximately 11%. This demonstrates that, in conditions when visual attention is not already directed at the visual warning displays, the spatial coupling of malfunction location and warning sound by means of 3-D audio, can be beneficial for the speed with which pilots react to in-cockpit warnings. Thus 3-D audio also proves to be advantageous for tasks that do not require visual search, as is the case in the TCAS task. This finding is comparable to the results of Haas (1998), although this time, the effect can be attributed directly to the application of spatialized sound.

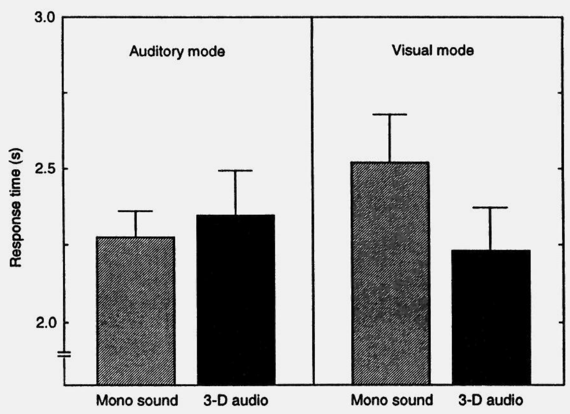

Figure 3. Mean response time to a system malfunction warning, as a function of flying mode and type of sound (standard errors included).

Taken together, the results of Experiment 1 showed that the effects of 3-D audio were only noticeable when the primary task information was presented only visually (i.e., when pilots paid more attention to the visual flight path display). Apparently, the pilots did not have to switch attention to the respective visual warning displays, and subsequently could respond faster to the warning when 3-D audio was used in the visual mode. This finding is of importance since the increasing level of automation in modern aircraft tends to increase the amount of time that pilots spent looking at displays inside the cockpit, especially during take-off and landing when the traffic density is relatively high (Damos, John & Lyall, 1999). When subjects could pay more attention to the visual warning displays (i.e., prepare themselves for the next warning), as was possible in the auditory mode for the flight task, 3-D audio did not enhance performance. However, this latter observation may not be representative for normal cockpit operations, because such anticipation is hardly possible in real world situations. So it can be expected that 3-D audio will be even more effective for the unexpected warnings during real flight.

In addition to the use of verbal cues in the second experiment, we modified the flight task and the visual flight display to see if the effect of 3-D audio would increase when pilots would need to pay more attention to this display. This was done by making the flight path display less informative (i.e., removing the tunnel-in-the-sky that provided preview) and by making the flight task more demanding (i.e., introducing simulated crosswinds that acted on the aircraft). It was hypothesized that pilots would need to focus more on the flight path display then, and subsequently could pay less attention to the visual warning displays and the outside view. In addition, the visual display for the malfunction warnings was relocated to make it more difficult to check this display frequently. This may also enhance the effect of 3-D audio further.

Experiment 2

Method

Subjects

Sixteen male student pilots were recruited from the Royal Dutch Airlines Flight Academy (KLS). Of these sixteen subjects, four had participated in the previous experiment. All subjects were in the final stage of their pilot training or had recently completed this. Their age ranged between 20 and 28 years.

Tasks

Flight task. The primary task for the subjects was to approach a runway and subsequently land the aircraft. For this task they had to follow a specific flight path towards the runway (i.e., comparable to the approach phase in Experiment 1). Information about deviations from the flight path was displayed to the subjects on the main HDD by means of two cross-hairs (see Figure 4). The subjects had to align these cross-hairs to nullify the error to the flight path. No auditory cueing was employed in this experiment. The subjects were instructed to follow the flight path as closely as possible and to maintain a target speed. To make the task more demanding, a continuous and changing crosswind was added that affected the horizontal position of the aircraft. The direction of the crosswind was perpendicular to the flight path, and was either coming from the left side or from the right side. An arrow on the HDD indicated the direction of the crosswind.

Warning tasks. The TCAS task was modified somewhat compared to Experiment 1. This time, the subjects were required to identify the specific orientation of the target aircraft,

instead of only detecting a target. Two distinct orientations were used; the target aircraft was either flying level (i.e., looking at the side of the fuselage) or climbing (i.e., looking at the underside of the wings and fuselage). These orientations subtended approximately the same visual angle. The subjects responded by pressing one of two buttons that indicated the specific target orientation. Also, the number of target locations was increased to 32, with the relative azimuth of these locations ranging from -60° (left side) to 60° (right side) azimuth in steps of 8°, and an elevation of either 14° (above eye level) or -6° (below eye level). All targets were directly visible from the pilot's position.

The system malfunction task remained unchanged, except that warnings for the flaps and spoilers were presented whether or not these systems were active. Also, the visual display for the warnings was now positioned to the right of the main visual display, in the center of the cockpit console.

Apparatus

The same experimental equipment as in Experiment 1 was used. There were some changes to the main HDD (see Figure 4). In the upper left section, the visual malfunction warnings were removed and now an arrow was presented that indicated the wind direction. In the lower right section, two cross-hairs that indicated flight path tracking performance replaced the tunnel-in-the-sky display. The inner cross-hair (i.e., off-center in Figure 4) indicated the momentary horizontal and vertical position error with respect to the flight path. The outer cross-hair resembled the aircraft and remained fixed in the center of the display.

Also, the same 3-D audio system was used. Again, individualized HRTFs were used and they were measured on a day prior to the experiment.

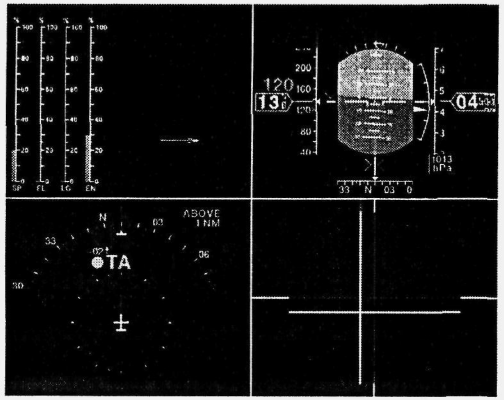

Figure 4. The main visual display in Experiment 2.

Experimental design

The auditory warnings were presented with mono sound or with 3-D audio, in combination with verbal location information (i.e., the verbal conditions) or without such information (i.e., the no-verbal conditions). This resulted in four experimental conditions. The conditions were presented within-subjects, with the conditions balanced across subjects. Each subject performed four blocks of four trials, resulting in 16 trials in total. Each block was performed in a different condition. For each trial, a different scenario was used that consisted of four warnings, with both TCAS and malfunction warnings present. This resulted in a total of 32 TCAS warnings (i.e., all 32 positions of the target) and 32 malfunction warnings for each subject. The 16 scenarios were balanced across subjects, and the order and timing of the warnings in each scenario were determined randomly prior to the experiment. Each trial lasted approximately 8.5 min.

For the TCAS warnings, the verbal cues indicated the quadrant of the out-the-window scene in which the target was located. The quadrants were defined by the vertical midline

(i.e., 0° azimuth) and the horizonline (i.e., 0° elevation), resulting in four sections (i.e., up left, up right, down left and down right). The verbal cues were included in the standard TA warning (e.g., 'Traffic up left.....Traffic up left'). For the system malfunction warnings, the verbal cues only indicated the lateral side of the aircraft where the malfunction was located (e.g., 'Flaps left.....Flaps left').

The same performance measures as in Experiment 1 were used. The statistical analyses involved ANOVAs with sound (mono sound or 3-D audio) and verbal information (absent or present) as within-subjects variables.

Procedure

Subjects came in for half a day to participate in the experiment. First, they received instructions about the tasks and conditions. They were instructed to follow the specified flight path as close as possible and to maintain the target speed. For the warning tasks, they were instructed to respond as quickly as possible, without making any errors. The subjects were then familiarized with the flight simulator and with the tasks in the different conditions. During the training, they received feedback on their warning task performance; a beep sounded when they pressed an incorrect button. No feedback was given in the experimental trials. The training lasted approximately one hour.

The 16 experimental trials were divided over two periods of eight trials each (i.e., two blocks of four trials in each period), with a short break in between. Duration of the total experiment was approximately four hours per subject.

Results and discussion

Due to scenario problems in the first few experimental trials, results for three of the 512 TCAS warnings had to be discarded from the analyses.

TCAS task. The ANOVA on the TCAS response times showed significant main effects of sound [$F(1, 15) = 15.62$, $p < .01$], and of verbal information [$F(1, 15) = 6.98$, $p < .05$] only. Both effects are depicted in Figure 5, as a function of the four conditions. The subjects responded faster when 3-D audio was present. On average, the response time was shortened by 0.60 s to 4.24 s in the 3-D audio conditions (i.e., a reduction of 12%). The presence of verbal cues also resulted in faster responses compared to the conditions without this information (i.e., 4.28 s and 4.80 respectively; a reduction of 11%). The ANOVA on the percentage of errors showed no significant effects (i.e., 7.5% overall). This indicates that the effects on response time did not result from a speed-accuracy trade-off.

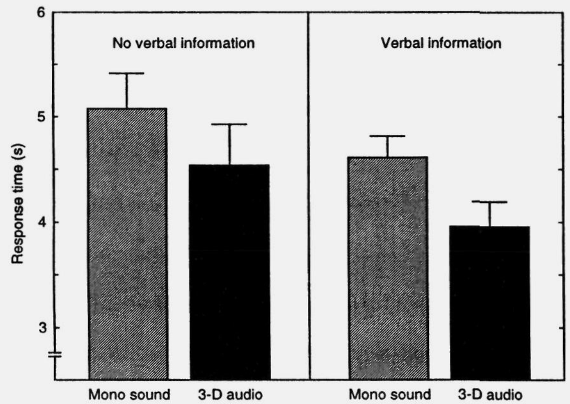

Figure 5. Mean response time to a TCAS warning in each condition (standard errors included).

The finding of main effects for both 3-D audio and verbal information indicates that the combination of these two types of spatial cues resulted in the best performance (see also Figure 5). In addition, the absence of a significant interaction between the two effects shows

that these effects are additive, resulting in a reduction in response time of approximately 23% when the cues are combined. Thus the addition of cognitive cues to the 3-D audio display effectively enhanced search performance further. This effect may be ascribed, to some degree, to the use of the unambiguous verbal cues for elevation to resolve the potential up-down confusions with 3-D audio. Apparently, the use of multiple cues for spatial location in one display can make up for the shortcomings of each individual cue (e.g., the poorer elevation perception with 3-D audio and the lower resolution of the verbal cues). In addition, it is known that the use of redundant cues can lead to faster response times through the process of probability summation (Wickens, 1992). The two types of cues can be considered redundant to some degree in that they both presented spatial information in an auditory display, even though that spatial information is presented in a different format and with a different resolution.

The observed effect of 3-D audio is comparable to the effects found in Experiment 1 and in the study of Begault et al. (1997), and again shows that 3-D audio improves response speed to TCAS alerts. However, the proposed increase in visual load did not result in a stronger effect of 3-D audio in comparison with Experiment 1. It is possible that the close proximity of the visual TCAS display to the flight path display, in combination with the absence of distractors in that display, still enabled the subjects to scan the TCAS display with a minimum of costs to search performance. It is also possible that the visual task load was not increased at all compared to Experiment 1. When the subjects already paid a lot of attention to the visual flight display in the first experiment, the changes to the flight task in the second experiment could not increase the visual load much further. The limited number of tasks that the subjects had to perform, in combination with the high frequency of TCAS warnings, may also have contributed to it. The subjects could prepare themselves for oncoming warnings, while they were watching the visual flight display. This may also explain the similar results found in the studies where 3-D audio was also tested as a supplement to a visual TCAS display. Therefore, the effect of 3-D audio is hypothesized to be larger in real flight, which generally involves unexpected warnings.

Additional analyses. The results showed no difference in performance between the 3-D audio and verbal information, even though the format (i.e., perceptual versus cognitive) and resolution (i.e., high versus low) of these spatial cues was different. However, these differences may have affected the way in which the subjects searched for the target. We therefore looked at search behavior with each type of spatial information more thoroughly. To investigate consistency in search performance, we analyzed the individual standard deviation (SD) of the response times in each condition. In addition, we examined whether the effects of 3-D audio and verbal information were dependent on the spatial location of the target, by differentiating between central and eccentric targets. Based on the absolute difference in azimuth between the target location and the head orientation of the pilot at the onset of the warning, two groups were defined: central target locations (difference ranging from 0° to 30° azimuth), and eccentric locations (difference of more than 30° azimuth). The subsequent ANOVA on the TCAS response times included sound (2 levels), verbal information (2 levels), and target location (central or eccentric) as within-subjects variables.

The ANOVA on the response time SD showed a main effect for sound [$F(1, 15) = 11.58$, $p < .01$]. The SD was smallest with 3-D audio (i.e., 1.34 s versus 1.98 s in the mono sound conditions). No other effects were observed. This indicates that the subjects searched for the target in a more consistent manner in the conditions with 3-D audio, while this was less so with verbal information. This may be due to the use of head tracking in the 3-D audio display, which resulted in an invariant sound source location when head movements were

made. This may have enabled the subjects to track down the target by using the changes in the azimuth cues that resulted from head movements, and subsequently provided them with continuous feedback on their search performance. This was not possible with the verbal cues, since these cues remained unchanged during a warning, regardless of any head movements of the pilot.

The ANOVA on response times regarding target location revealed, in addition to the known main effects of sound and verbal information, a main effect of target location [$F(1, 15) = 5.05$, $p < .05$], with faster response times to more central targets. More interestingly however, was the significant interaction between verbal information and target location [$F(1, 15) = 15.68$, $p < .01$]. As shown in Figure 6, the verbal cues only aided the subjects when the targets were located eccentric ($p < .01$), not when they were positioned central. This suggests that the subjects were biased to the eccentric part of the out-the-window scene to search for the target when the warning was supplemented with verbal cues, which may be due to the low resolution of the verbal cues. This is in contrast with the effect of 3-D audio, since no interaction of sound and target location was observed. Thus 3-D audio resulted in faster response times regardless of target location.

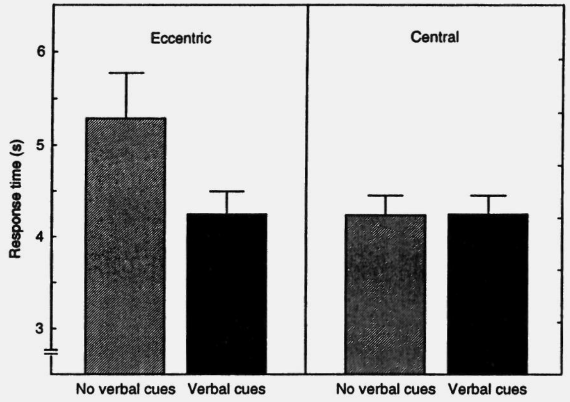

Figure 6. Mean response time to a TCAS warning, as a function of verbal location information and target location (standard errors included). The absolute azimuth for central targets was equal to or smaller than 30°, and for eccentric targets more than 30°.

The results of these additional analyses indicate that the subjects used the two types of spatial information differently. The verbal cues were only beneficial when the target was located more eccentric and also resulted in more variable search behavior. With 3-D audio, the subjects searched for the target in a consistent and efficient manner. These findings suggest that the absence of performance differences between the two types of spatial information may be an artifact of the experimental set-up (e.g., high frequency of warnings, known target locations, and good visibility of the targets).

System malfunction task. The analysis of the response times to the malfunction warnings also showed significant main effects of sound [$F(1, 15) = 22.24$, $p < .001$], and of verbal information [$F(1, 15) = 19.70$, $p < .001$] only. The two effects are shown in Figure 7, as a function of the four conditions. On average, the response time decreased from 2.64 s to

2.33 s (i.e., a reduction of about 12%) when 3-D audio was present. Similar results were found for the presence of verbal cues (i.e., a reduction from 2.64 s to 2.33 s). The analysis of the percentage of errors did not show any significant effects (i.e., 1.1% overall), indicating that the effects on response time were not the result of a speed-accuracy trade-off by the subjects.

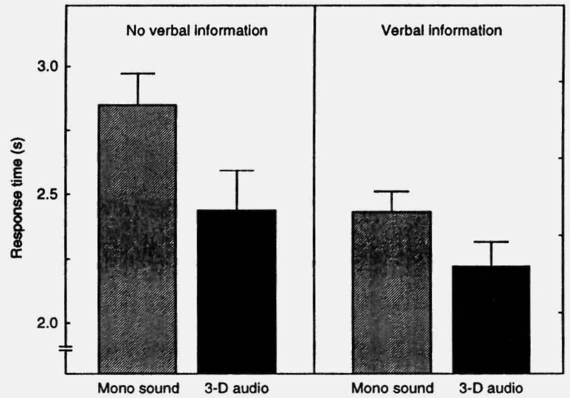

Figure 7. Mean response time to a system malfunction warning in each condition (standard errors included).

The observed effect of 3-D audio is in line with the findings of the first experiment. Compared to the study of Haas (1998), the effect was considerably smaller (i.e., 28% versus 12% reduction in response time), but this difference in results can be attributed largely to the absence of auditory warning signals in the comparison condition of that study. As was observed in the TCAS task, the effects of 3-D audio and of verbal information are additive, as indicated by the absence of an interaction. The combination of 3-D audio and verbal cues thus results in the best performance, as can be seen in Figure 7 (i.e., a reduction in response time of approximately 23%). This can not be attributed to a resolution of up-down confusions due to the presence of verbal cues, since the malfunction task did not require elevation cues. *It is more likely the result of using fully redundant cues in the auditory display. The two types* of spatial information for the malfunction task can be considered fully redundant, since the resolution of each cue was comparable and matched the resolution required for the malfunction task.

The relocation of the visual display for the malfunction warnings did influence the effect of 3-D audio somewhat. Looking at the conditions that were similar in the two experiments (i.e., the visual mode conditions of Experiment 1 and the conditions without verbal cues of Experiment 2), the response times were somewhat longer in the present experiment (i.e., 0.27 s on average). In addition, the use of 3-D audio resulted in a reduction of the response time with 0.29 s (i.e., 11.5%) in these conditions of Experiment 1, whereas the reduction in the present experiment was 0.41 s (i.e., 14.7%). This indicates that the effect of 3-D audio became somewhat more pronounced when the visual display was placed more peripherally to the pilot.

General discussion

The results of the two experiments showed advantages of 3-D audio for auditory warnings in the cockpit. For both the TCAS warnings, that required the pilot to visually search for an intruding aircraft, and the system malfunction warnings, that required the pilot to localize a malfunction, the use of 3-D audio resulted in faster reaction times compared to mono sound. These findings are in line with the results of previous studies on the use of 3-D audio for auditory warnings in the cockpit (Begault, 1993; Begault & Pittman, 1996; Begault et al., 1997; Haas, 1998).

In both experiments, 3-D audio was tested as a supplement to visual head-down displays and was always compared to a monaural display. This was also the case in the study of Begault et al. (1997). The results of these three experiments are also comparable. In the two experiments reported here, the 3-D audio displays reduced the response times to both TCAS and malfunction warnings with approximately 12%, while Begault et al. (1997) observed a reduction of 14% for TCAS warnings. This suggests that an advantage of 3-D audio of approximately 12% to 14% on response time to warnings can be expected in real world situations. This can be considered a conservative estimation, since the results were obtained under controlled conditions in (part-task) flight simulators (e.g., a high frequency of warning incidents, expected by the pilots). Also, the TCAS targets in the different studies were all visible in a restricted forward field of view. In real flight, most intruding aircraft will be in that region. However, it can be expected that 3-D audio may be even more beneficial when targets are outside the direct field of view of the pilots. Other simulator studies have shown that 3-D audio can be used effectively to signal target positions all around the own aircraft (Bronkhorst, Veltman & Van Breda, 1996; Veltman, Oving & Bronkhorst, this issue).

The second experiment showed that the application of verbal location information in the warning resulted in comparable performance improvements as observed with the application of 3-D audio, and thus may serve as an alternative to 3-D audio for warning presentation. However, it was also observed that search performance in the TCAS task with 3-D audio was independent of target location and more consistent than with verbal information. The combination of the two types of spatial information always resulted in the best performance. Thus the application of multiple cues for spatial location in an auditory warning display may enhance flight safety even further.

References

Begault, D.R. (1993). Head-up auditory displays for traffic collision avoidance systems advisories: a preliminary investigation. Human Factors, 35, 707-717.

Begault, D.R. & Pittman, M.T. (1996). Three dimensional audio versus head-down traffic alert and collision avoidance system displays. International Journal of Aviation Psychology, 6, 79-93.

Begault, D.R., Wenzel, E.M. & Lathrop, W.B. (1997). Augmented TCAS advisories using a 3-D audio guidance system. Proceedings of the Ninth International Symposium on Aviation Psychology (pp. 353-357). Columbus, OH: Ohio State University.

Bronkhorst, A.W. (1995). Localization of real and virtual sound sources. Journal of the Acoustical Society of America, 98, 2542-2553.

Bronkhorst, A.W. & Houtgast, T. (1999). Auditory distance perception in rooms. Nature, 397, 517-520.

Bronkhorst, A.W., Veltman, J.A. & Van Breda, L. (1996). Application of a three-dimensional auditory display in a flight task. Human Factors, 38, 23-33.

Damos, D.L., John, R.S. & Lyall, E.A. (1999). The effect of level of automation on time spent looking out of the cockpit. International Journal of Aviation Psychology, 9, 303-314.

Haas, E.C. (1998). Can 3-D auditory warnings enhance helicopter cockpit safety? Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting (pp. 1117-1121). Santa Monica, CA: The Human Factors and Ergonomics Society.

Kulkarni, A. & Colburn, H.S. (1999). Role of spectral detail in sound-source localization. Nature, 396, 747-749.

Langendijk, E.H.A. & Bronkhorst, A.W. (2000). Fidelity of 3D-sound reproduction using a virtual auditory display. Journal of the Acoustical Society of America, 107, 528-537.

Møller, H., Sørensen, M.F., Hammershøi, D. & Jensen, C.B. (1995). Head-related transfer functions of human subjects. Journal of the Audio Engineering Society, 43, 300-321.

Oppenheim, A.V. & Schafer, R.W. (1989). Discrete-time signal processing. London: Prentice Hall.

Veltman, J.A., Oving, A.B. & Bronkhorst, A.W. (this issue). 3-D audio in the fighter cockpit improves task performance. International Journal of Aviation Psychology.

Wenzel, E.M., Arruda, M., Kistler, D.J. & Wightman, F.L. (1993). Localization using nonindividualized head-related transfer functions. Journal of the Acoustical Society of America, 94, 111-123.

Wickens, C.D. (1992). Engineering psychology and human performance. New York: HarperCollins.

Wightman, F.L. & Kistler, D.J. (1989). Headphone simulation of free-field listening. II: Psychophysical validation. Journal of the Acoustical Society of America, 85, 868-878.

VOORNAAMSTE BEVINDINGEN

Geef in niet meer dan 30 woorden weer wat de voornaamste bevindingen uit dit onderzoek waren. Uw tekst wordt onder de rubriek "uit het onderzoek" in het INFO gepubliceerd.

Een tweetal vliegsimulator experimenten lieten zien dat piloten sneller reageerden op de auditieve waarschuwingen van het TCAS systeem in de civiele cockpit, wanneer deze waarschuwingen werden gepresenteerd met 3D-audio in vergelijking tot mono geluid.

......

....

......

....

......

Trefwoorden publikaties:

3D audio

Simulators

Performance

Flight interfaces

Warnings

....