Australian Government
Department of Defence
Defence Science and
Technology Organisation

Auditory Warnings in the Military Cockpit: A Preliminary Evaluation of Potential Sound Types

Sean E. Smith, Karen L. Stephan and Simon P.A. Parker

Air Operations Division
Systems Sciences Laboratory

DSTO-TR-1615

ABSTRACT

This document reports the results of two experiments assessing the viability of speech, auditory icons (environmental sounds) and abstract sounds (complex tones) for use as auditory warnings in military cockpits. Experiment One evaluated the comparative ease of learning and retention of these three sound types, with the results demonstrating that speech warnings and auditory icons were learnt and retained with equal ease, while abstract-sound warnings were found to be learnt and retained with far greater difficulty. Experiment Two examined a user's ability to respond to speech, icon and abstract auditory warnings under varying degrees of workload. The results of this experiment show that speech warnings provide an advantage in reaction time and response accuracy over auditory icon warnings, which in turn hold an advantage over abstract sounds. The results of these two experiments indicated that, under low levels of workload, speech warnings may prove to be an effective supplement to traditional visual warnings, while icons also warrant further investigation.

Published by

DSTO Systems Sciences Laboratory
PO Box 1500
Edinburgh South Australia 5111 Australia

Telephone: (08) 8259 5555
Fax: (08) 8259 6567

AR- 013-186
September, 2004

Auditory Warnings in the Military Cockpit: A Preliminary Evaluation of Potential Sound Types

Executive Summary

As military aircraft become increasingly sophisticated, operators are required to deal with an ever-increasing flow of information. Traditionally, the bulk of information presented to aircrew has been transmitted via the visual modality in the form of head-down and head-up displays. The increasing amount of information that operators are required to process, particularly in high risk situations, can potentially lead to an overload of the visual sense, which could result in crucial information being overlooked. Therefore exploration of other methods of information transfer which aim to ameliorate the current potential for visual overload are required.

The two experiments presented in this report demonstrate the potential effectiveness of auditory warnings in supplementing information currently presented on visual displays. Both experiments associated speech, auditory icons (environmental sounds) and abstract sounds (complex tones) with events typical of those represented on a radar warning receiver device. Experiment One examined the comparative ease with which these three sound types were learnt and retained. The results demonstrated that speech warnings and auditory icons were learnt and retained with equal ease, while abstract-sound warnings were found to be learnt and retained with far greater difficulty. Experiment Two examined a user's ability to respond to speech, icon and abstract auditory warnings under varying degrees of workload. The results of this experiment show that speech warnings provide an advantage in reaction time and response accuracy over auditory icon warnings, which in turn hold an advantage over abstract sounds. In combination, the results of these two experiments warrant further investigation into the viability of speech and icons for use as auditory warnings.

The work detailed in this report continues the development of advanced auditory displays at DSTO's Air Operations Division which aims to reduce the potential for visual overload that aircrew may experience in the modern military cockpit.

Authors

Sean Smith
Air Operations Division

Sean Smith is a Professional Officer working in Human Factors within the Air Operations Division of DSTO. He has a B.A.(Hons) and an M.A. in Psychology from Deakin University. During his six years at DSTO he has worked predominantly in the area of advanced auditory display design, but has also worked on analysis of noise in operational environments, the production of audio for use in RAAF simulators and cockpit digitisation.

Karen Stephan
Air Operations Division

Karen Stephan is a Professional Officer working in Human Factors within the Air Operations Division of DSTO. Karen completed a BSc(Hons) in Psychology at Monash University in the area of auditory neuroscience, after which she worked as a research assistant in the same department. Karen commenced employment with DSTO in 1996. Since then she has been involved in a variety of human factors research projects including investigations into visual and auditory display design issues. Current research efforts focus on improving the understanding of auditory cognitive processes and the application of this knowledge to advanced auditory display design.

Simon Parker
Air Operations Division

Simon Parker is a Senior Research Scientist working in the Air Operations Division of DSTO as a human factors specialist. He has a BA (Hons) and a PhD in Psychology and investigated three-dimensional sound localisation in his doctoral work. Simon has interests and expertise in human factors issues associated with electronic warfare equipment, the design of auditory warnings, the design and synthesis of 3D audio displays and the use of simulation in human factors research.

Contents

1. Introduction

Successful and safe operation of military aircraft requires the monitoring, interpretation and integration of multiple streams of information. As military aircraft have become increasingly more sophisticated, the amount of information that aircrew are required to process has also increased significantly (occasionally coinciding with a reduction in the number of aircrew). Traditionally, this information has been presented to aircrew via the visual modality. The large amount of visual information which aircrew are required to process, particularly in high risk situations, can lead to visual overload. Under such conditions, it is possible that critical information could be overlooked or ignored as a result of a lack of resources to deal with increased demand on the visual system. Consequently it is imperative to consider alternative methods of presenting information to aircrew which do not contribute to visual workload while also enhancing operational effectiveness. One potential solution under investigation in the Air Operations Division (AOD) at the Defence Science and Technology Organisation (DSTO) Melbourne is the use of auditory warnings. It is possible that demand on the visual system may be eased by re-channelling the presentation of some information via the auditory modality.

Currently, limited use is made of the auditory modality as a means of informing aircrew about dangerous or potentially dangerous conditions, such as low altitude or the presence of a military threat (Doll & Folds, 1983). Where auditory warnings are in use, in most situations they provide little information to aircrew. There are three key criteria to be addressed when designing auditory warnings (Stanton & Edworthy, 1999). First, auditory warnings should serve to alert the operator to the fact that there is a situation, or emergency that requires his/her attention. Second, auditory warnings should provide the operator with information about the nature or identity of the situation or emergency that he/she is required to attend to. Third, auditory warnings should guide the user toward the appropriate course of action to deal with the situation or emergency. Generally, the auditory warnings currently in use meet only the first of these three criteria. That is, they cue aircrew to the fact that there is an emergency, but do not provide any information about the emergency. Specific information is often only obtained after interrogation of a visual display, which leaves considerable scope for improvement in the design and implementation of auditory warnings for use in the cockpit. This report provides details of two experiments aimed at investigating issues related to the design of appropriate auditory warnings.

Auditory warnings can be divided into two distinct categories, verbal or non-verbal. Both have their relative advantages and disadvantages, which can be used as a guide for effective implementation (see Table 1). Verbal (or speech) signals have the advantage that negligible learning is required. They are also effective for conveying complex information. They can, however, take a long time to convey this information. In addition, they are not suitable for some situations, for example where loud verbal messages would be indiscreet (Edworthy & Stanton, 1995). Although all auditory warnings are subject to masking by background noise, speech warnings have the additional problem of being particularly susceptible to masking by other speech present in the cockpit, or informational masking as described by Brungart (2001).

Abstract sounds (e.g. simple tones) are commonly used as a non-verbal alternative in warning systems. They can convey information quickly, and they are not masked by speech-rich environments as verbal warnings can be. The main disadvantages of abstract sounds lie in the difficulty that humans have in learning and remembering large sets of them (Meredith & Edworthy, 1991; Montahan, Hetu, & Tanskey, 1993; Patterson & Milroy, 1980) and the limited amount of information they can convey. Due to this, it has been suggested that no more than five to eight items should make up a warning set if abstract sounds are to be used (Patterson, 1982; Sanders & McCormick, 1987).

Table 1. Summarised advantages and disadvantages of potential auditory warning types.

	Advantages	**Disadvantages**
Speech	• Convey complex information • Little learning required	• Can be lengthy • Can easily be masked by already existing speech in the cockpit
Abstract Sounds	• Less likely to be masked • Convey information rapidly	• Difficult to learn and remember • Convey a limited amount of information • Limited in the amount that can be learnt
Auditory Icons	• Little learning required • Less likely to be masked • Convey information rapidly • Can potentially convey complex information	• Not all events to be represented by an icon produce a sound • Not all sounds are appropriate to an operational environment

Auditory icons (or environmental sounds) are another type of non-verbal sound with potential for use as auditory warnings. Everyday we effortlessly extract information from our environment through the identification of naturally occurring sounds (eg traffic noise). It is considered possible that we can increase the effectiveness of aircrews' responses to auditory warnings by exploiting this pre-existing ability. Icons should share the advantages of abstract sounds with respect to informational masking and the speed of information transfer. A further advantage of icons is that they are not as limited in terms of how many can be learnt, as humans demonstrate a huge memory capacity for environmental sounds (Lawrence & Banks, 1973; Lawrence, Cobb, & Beard, 1979). Auditory icons also have the potential to convey more information than abstract sounds, and it is believed that their association with naturally occuring events means that they are easier to learn than abstract sounds (Gaver, 1986).

When considering the use of an icon as an auditory warning, some care is required in the selection of an environmental sound that adequately represents the intended target event. It must be considered that sometimes events will need to be represented which do not make a sound, or make a sound that is not appropriate for the operational environment. This leads to a situation where, if icons are to be used as an auditory warning, sounds are paired with events based upon some sort of pre-existing relationship between the sound and the event. The pairing of an event with the sound that it actually produces is referred to as a "direct" association, for example the sound of an aircraft to identify an enemy jet fighter. Whereas pairing a sound and an event based upon some other relationship is referred to as an "indirect" association, for example the cry of a bird of prey to identify an enemy jet fighter, the relationship being that both fly in the sky and have attacking capabilities. In Experiment One, a preliminary assessment of the effectiveness of auditory icons in comparison to speech and abstract sounds was conducted. In this experiment, a set of eight abstract sounds, eight speech messages and eight auditory icons (each of which shared either a direct or indirect relationship with the event they were intended to represent) were compared for ease of learning and retention over a one week period.

An auditory warning may prove to have advantages in terms of learning and retention, yet be relatively ineffective when used in an operational environment where attention is divided amongst many tasks. Experiment Two examined this issue by comparing the effectiveness of abstract sounds, speech and icon auditory warnings under varying conditions of workload. Although auditory icons can be categorised on the basis of the relationship they share with the event they are intended to represent, the effect that this relationship can have upon their performance is not investigated in either study detailed in this report. Rather, icons as a sound type, irrespective of the relationship between the icon and event, as well as speech and abstract sounds, are evaluated on their effectiveness for use as auditory warnings.

2. Experiment One: Learning & Retention

2.1 Introduction

A key ingredient of an effective auditory warning is that it is easily learnt and remembered. As outlined in Table 1, speech and auditory icon warnings should require little or no learning and should therefore be easy to remember over time. Abstract sound warnings on the other hand are considered more difficult to learn and remember over a period of time. While individual researchers have investigated each of these sound types in isolation, there are no studies that directly compare them for ease of learning and retention. This experiment compared the ease of learning and retention of sets of auditory warnings comprised of speech, abstract sounds and auditory icons. It is expected that icon and speech warnings will be more easily learnt and retained than abstract sounds. Of particular interest is how easily icon warnings are learnt and retained in comparison to speech warnings.

2.2 Method

Participants

Eighteen unpaid participants, fourteen males and four females, ranging in age from 24 to 45 years (average age = 29 years) were recruited from staff and students at Swinburne University of Technology and the Air Operations Division. All participants were naïve to the sounds being used in this experiment.

Stimuli

For each type of auditory warning a set of eight examples was generated that could be used to represent events that may occur in a military aircraft cockpit. These included seven threat warnings that may be associated with a radar warning receiver and one advisory message that may be associated with self-protection counter-measures equipment (see Table 2).

Speech sounds were created synthetically by DECTalk V4.2 (Digital Equipment Corporation), which utilises text-to-speech technology. The voice used was the adult male default voice (Paul). The choice of terms used for the speech warnings was determined by interviewing Royal Australian Air Force (RAAF) aircrew.

Auditory icons were environmental sounds that, where possible, were selected on the basis of a direct relationship between the sound and the event it was associated with. For example, for the threat event “Anti-Aircraft Artillery” a gun sound was used. During the design of the auditory icons a number of candidates were generated for each event. These candidates were then demonstrated to aircrew and the one regarded by the majority of aircrew as being the most suitable for each event was chosen. The recognisability of the icons used in this experiment was not tested empirically, however they were judged by the experimenters, on a purely subjective basis, to be easily recognisable.

The abstract sounds were pulses of synthetic complex sound that varied in both the frequency and time domains and had been designed for maximum discriminability as detailed by Patterson (1982). The pulses were 100-150 ms in duration with onsets and offsets of 20-30 ms in duration. Each sound consisted of five or more pulses in a distinctive temporal pattern. Spectral components were primarily between 0.5-5.0 kHz with four or more harmonically related components and the fundamental harmonics in the range of 150-1000 Hz. An abstract sound was randomly assigned to each of the eight events.

The sounds used in this experiment were not matched for duration. It was not possible to provide easily recognisable auditory icons and speech that could be matched for duration with abstract sounds. The duration of sounds ranged from 0.8 to 2 seconds.

Equipment

This experiment was run on a desktop PC using custom-made learning software. Sounds were presented to participants over Sennheiser HD 265 linear circumaural headphones.

Procedure

Each participant completed nine experimental sessions conducted over a period of three weeks. For each of the three auditory warning types, participants were tested on three separate days: initially on Day 1 and again two (Day 3) and seven (Day 8) days later. The initial experimental session on Day 1 consisted of several trials, each of which consisted of two phases: a learning phase and a test phase. In the learning phase, each of the eight warnings was played to the participant over a set of headphones and the name of the event paired with the sound was displayed on a computer monitor as indicated in Figure 1. The participant was asked to acknowledge each association by pressing a computer keyboard function key labelled with the name of the appropriate event as in Figure 1. Each event was always associated with the same function key across subjects and warning type.

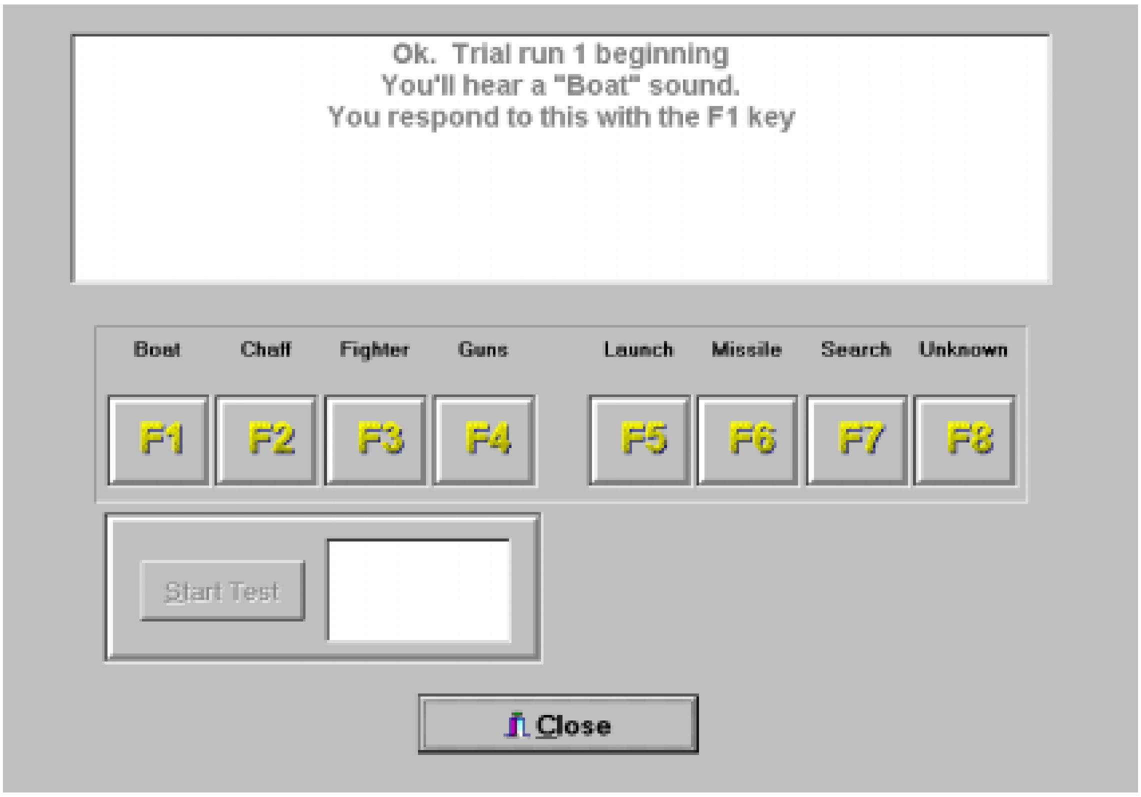

Figure 1. Screen shot of learning and retention software during the learning phase.

On completion of the learning phase, the participant proceeded to the test phase where each of the eight warnings was presented randomly over the headphones without the name of the event being indicated as displayed in Figure 2. The participant was required to identify the event associated with each warning and press the corresponding function key. The appropriate function key for each event was still displayed on the computer monitor during the test phase as demonstrated in Figure 2. Participants were instructed to respond as accurately as possible. At the end of each test phase, the percentage correct out of eight was presented to the participant on the computer monitor. Test phases were repeated until the participant achieved a criterion level of proficiency set at 100 percent correct on two consecutive test phases. The sessions on Days 3 and 8 followed the same procedure as that on Day 1.

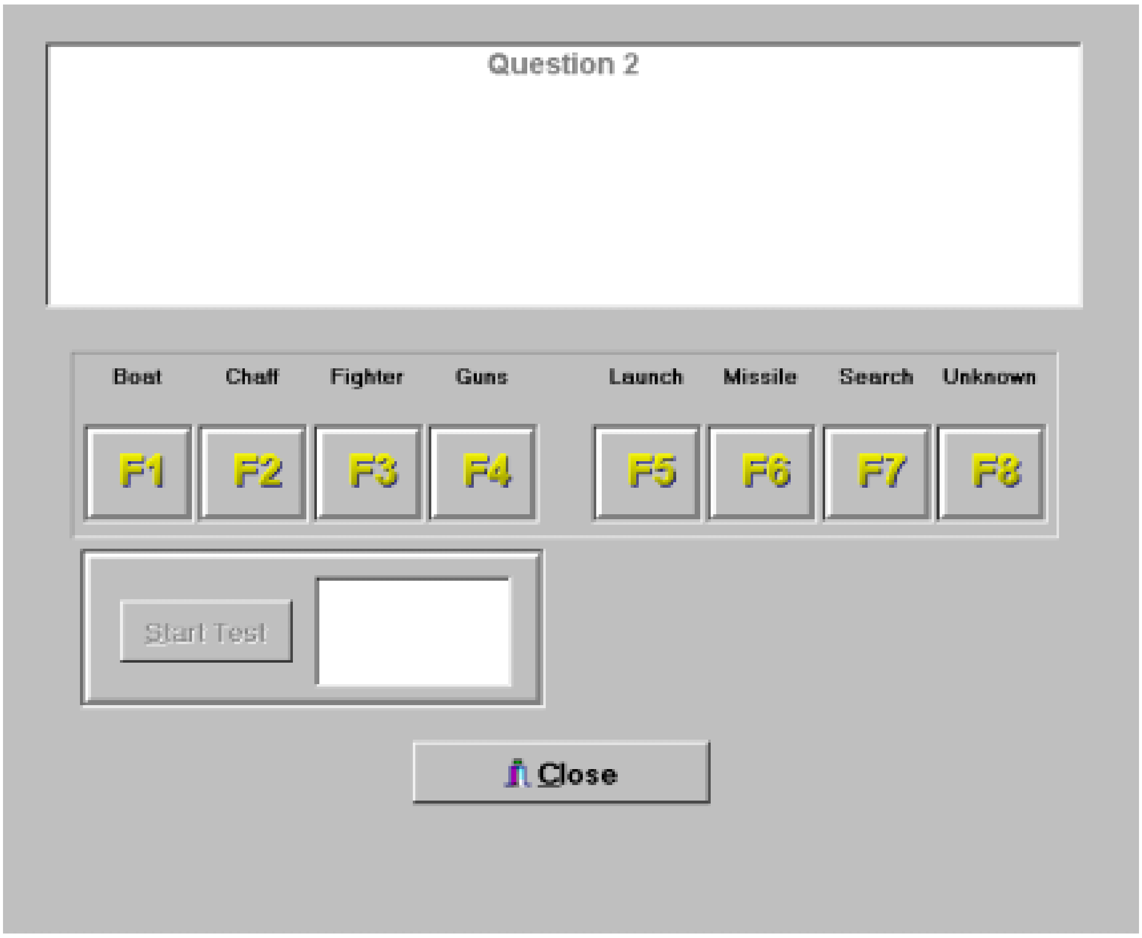

Figure 2. Screen shot of learning and retention software during the test phase.

Table 2. Speech and auditory icon sets used in this experiment

Event	Speech	Auditory Icons
Ship	"Boat"	Fog horn
Air Interceptor	"Fighter"	Bird of prey attack call
Anti-Aircraft Artillery	"Guns"	Machine gun firing
Missile Launch	"Launch"	Monkey Screech
Surface to Air Missile	"Missile"	Arrow
Search Radar	"Search"	Sonar beep
Unknown Threat	"Unknown"	Bird territorial call
Chaff Dispensed	"Chaff"	Camera flash & Windchimes

2.3 Results

For each session, two measures of performance were calculated: number of trials required to reach criterion performance and number of errors made during those trials. Reaction time was not analysed due to a lack of reliability in the software's ability to accurately capture response times. Due to the lack of variability in the number of errors and trials to criterion made in the speech condition, it was not possible to perform a parametric test of statistical significance. Therefore, a non-parametric statistical test, the Friedman Two-Way Analysis of Variance by Ranks with correction for ties, was used to test for significant differences in participant performance across the three sound types on each of Days 1, 3 and 8 for both measures of performance.

Trials to criterion

The average number of trials required to reach criterion performance for the three sets of auditory warnings on each of Days 1, 3 and 8 are shown in Figure 3. Performance on Day 1 provides an indication of the ease with which the specified associations were learnt, while that on Days 3 and 8 reveals the extent to which the associations were retained.

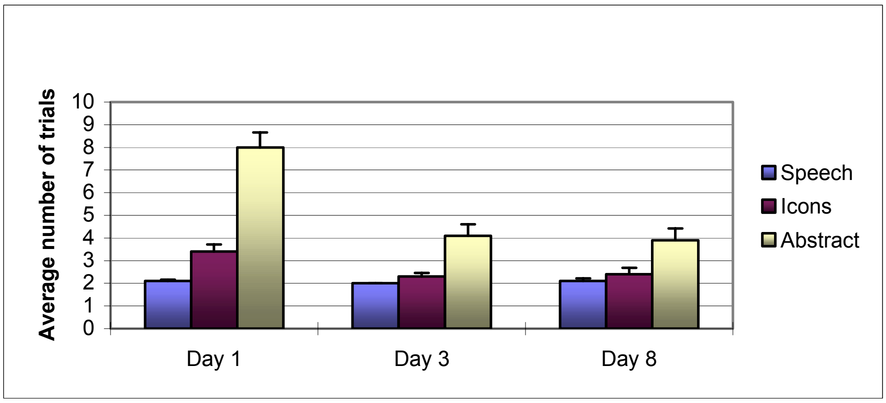

Figure 3. Mean number of trials required to reach criterion.

Day 1

It is evident that the speech auditory warnings were learnt most readily, requiring an average of 2.1 trials. As the minimum possible number of trials in which criterion could be reached was two, this represents almost perfect performance. The auditory icons were learnt slightly less readily, requiring an average of 3.4 trials, and the abstract sounds proved much more difficult to learn, requiring an average of eight trials. Analysis of the average number of trials required to reach criterion on Day 1 provided evidence of a significant difference across auditory warning type ($\chi_r^2(2)=37.26$, $p<.05$). Further analysis involved multiple comparisons between pairs of auditory warning types using the procedure outlined by Siegel and Castellan (1988). The first of these comparisons indicated that there was a significant difference between the average number of trials required to reach criterion for abstract sounds and icons (Z=19, $p<.05$). A second comparison, between abstract sounds and speech, also identified a significant difference (Z=22.5, $p<.05$). However, there was no significant difference between the average number of trials required to reach criterion for speech and icons.

Days 3 and 8

The results for Days 3 and 8 show that the learned associations between sounds and events were retained almost perfectly for speech and icons. For abstract sounds however, some re-learning was needed on both days, although the average number of trials required to reach criterion on each day was markedly less than that on Day 1. Analysis of the average number of trials required to reach criterion on Day 3 provided evidence of a significant difference across auditory warning type (χ_r^2 (2)=23.63, $p<.05$). This was also the case for Day 8 (χ_r^2 (2)=14.1, $p<.05$). Multiple comparisons between pairs of auditory warning types revealed a significant difference between the average number of trials required to reach criterion for abstract sounds and icons on each of Days 3 (D=18, $p<.05$) and 8 (D=16, $p<.05$). They also showed significant differences for abstract sounds and speech (Day 3, D=22.5, $p<.05$; Day 8, D=19.5, $p<.05$). However, there was no significant difference between the average number of trials required to reach criterion for speech and icons on either day.

Errors to criterion

The average number of errors made while reaching criterion performance for each of the three sets of auditory warnings on each of Days 1, 3 and 8 are shown in Figure 4. It can be seen that the pattern of the data in this figure is similar to that for trials to criterion.

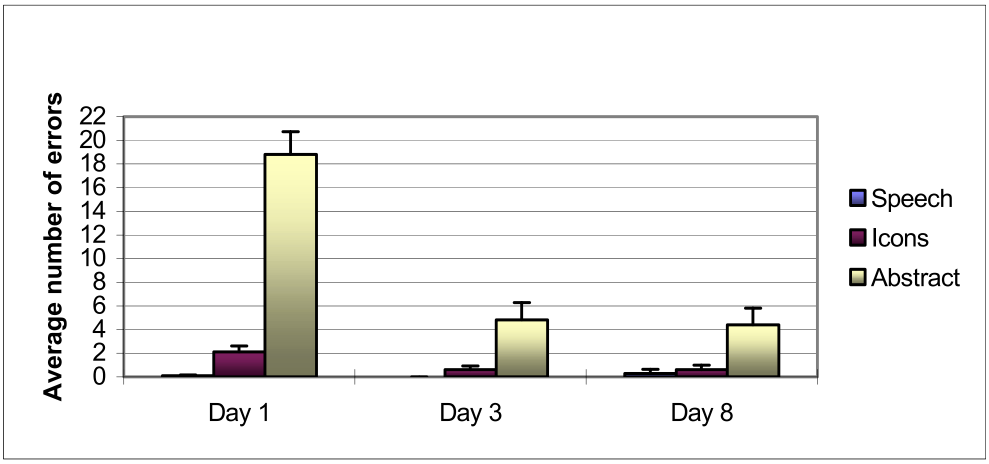

Figure 4. Mean number of errors required to reach criterion.

Day 1

Results on Day 1 again highlight the greater difficulty participants experienced while learning associations between abstract sounds and events in comparison to speech or icons. Analysis of the average number of errors made in reaching criterion on Day 1 provided evidence of a significant difference across auditory warning type (χ_r^2 (2)=32.36, p<.05). Further analyses revealed significant differences between the numbers of errors made in reaching criterion for abstract sounds and icons (D=22, p<.05) and abstract sounds and speech (D=33, p<.05). However, there was no significant difference between the numbers of errors made in reaching criterion for speech and icons.

Days 3 and 8

On Days 3 and 8 the average numbers of errors made for speech and icons were negligible, again demonstrating that associations between these sound types and events were retained almost perfectly. For abstract sounds, however, the average number of errors were markedly greater, reflecting a need for re-learning. Analysis of the average number of errors made while reaching criterion provided evidence of a significant difference across auditory warning type on each of Days 3 (χ_r^2 (2)=23.63, p<.05) and 8 (χ_r^2 (2)=24.6, p<.05). Further analyses revealed significant differences between the numbers of errors made while reaching criterion for abstract sounds and icons (Day 3, D=18, p<.05; Day 8, D=18, p<.05) and abstract sounds and speech (Day 3, D=22.5, p<.05; Day 8, D=21, p<.05). However, there was no significant difference between the numbers of errors made in reaching criterion for speech and icons on either day.

2.4 Discussion

The results presented here indicate that, as expected, associations between auditory icons and events are learnt and retained far more readily than those between abstract sounds and events. They also demonstrate that associations involving icons are learnt and retained just as readily as those involving speech. As noted above, several authors have demonstrated the difficulty of learning and remembering large sets of abstract-sound warnings (Patterson & Milroy, 1980; Montahan, et al, 1993; Meredith & Edworthy, 1991) and others have shown we have a huge capacity for remembering environmental sounds (Lawrence & Banks, 1973; Lawrence, et al, 1979). The current research replicates these findings and provides a direct comparison of the ease of learning and retention of sets of auditory warnings comprised of these sound types.

The fact that associations involving auditory icons were learnt and retained as readily as those involving speech indicates that icons do indeed have potential for use as auditory warnings. Like abstract sounds, auditory icons can convey information rapidly and should be less likely than speech warnings to be susceptible to informational masking by background speech. Unlike abstract sounds, however, they are not as limited with respect to the set size that can be learnt.

It is presumed that learning and retention of icon-event associations is enhanced by a direct relationship between the icon and the event. The strongest relationship should be between sounds and the events that cause them and, as noted by Gaver (1986), icon-event pairings of this nature should be the easiest to learn. The one example of such a pairing in this study was between the sound of a machine gun being fired and the event "anti-aircraft artillery". Providing some support for this suggestion is a post-hoc observation that the anti-aircraft artillery icon was one of only two that were identified correctly on all occasions.

Events will sometimes need to be represented for which icons with a direct relationship are either inappropriate or not available, therefore requiring the use of an indirect icon-event relationship. The fact that most of the icon-event relationships used in this experiment were indirect, and that overall performance was statistically equivalent to speech, suggests that indirect icon-event relationships can be as effective as direct icon-event relationships. It is important to note that it was not the aim of this experiment to explicitly evaluate performance differences between direct and indirect icon-event relationships. At this stage, any evidence which suggests that learning and retention performance is equivocal between direct and indirect icon-event relationships is purely anecdotal. Consequently a more systematic investigation which explicitly compares the learning and retention of icon-event pairings with direct or indirect relationships is required to explore this issue fully.

3. Experiment Two: Effect of Workload on Auditory Warning Response

3.1 Introduction

Auditory warnings must be of benefit in an environment where attention is divided and workload varies. It is possible that an auditory warning that is easily learnt and retained, and effective in normal circumstances may lose its effectiveness when workload is high, and the operator's attentional capacity is decreased. To investigate this issue, an experiment was conducted to assess participants' performance using speech, icon and abstract warnings under varying workload conditions. In this experiment participants were required to perform a visual tracking task, which provided a very low-fidelity simulation of flying an aircraft, as well as various combinations of other visual and auditory tasks, while also responding to auditory warnings. Of particular interest was the accuracy with which participants responded to the three auditory warning types, and the speed with which they responded. It is expected that participants' accuracy and reaction time in responding to abstract sounds may be negatively affected by the inclusion of competing tasks, which should serve to increase their workload.

3.2 Method

Participants

Twelve unpaid participants, ten males and two females, ranging in age from 22 to 41 years (average age = 28 years) were recruited from staff within AOD at DSTO Melbourne.

Stimuli

The sound stimuli consisted of the eight events used in Experiment One (see Table 2), represented as speech, abstract sounds and icons in addition to Air Traffic Control (ATC) messages. The speech warnings were generated by a speech synthesis software package (DECTalk 4.2, Digital Equipment Corporation). The auditory icons were digital recordings of environmental sounds. Patterson's (1982) 'attensons' were used as the abstract sound stimuli. The ATC messages were recordings made by an adult male recruited from DSTO Melbourne and contained instructions to change direction or heading (eg, "Condor ascend 1000 feet.").

The visual stimuli consisted of a computer-generated tracking task, where the participant had to align a moving cross-hair, using a joystick, against a stationary cross in the centre of a computer monitor. Participants were also required to monitor the movement of two gauges placed on either side of the tracking display, and respond appropriately when a gauge went into an out-of-range condition (see Figure 5 for an example of the tracking and gauges tasks).

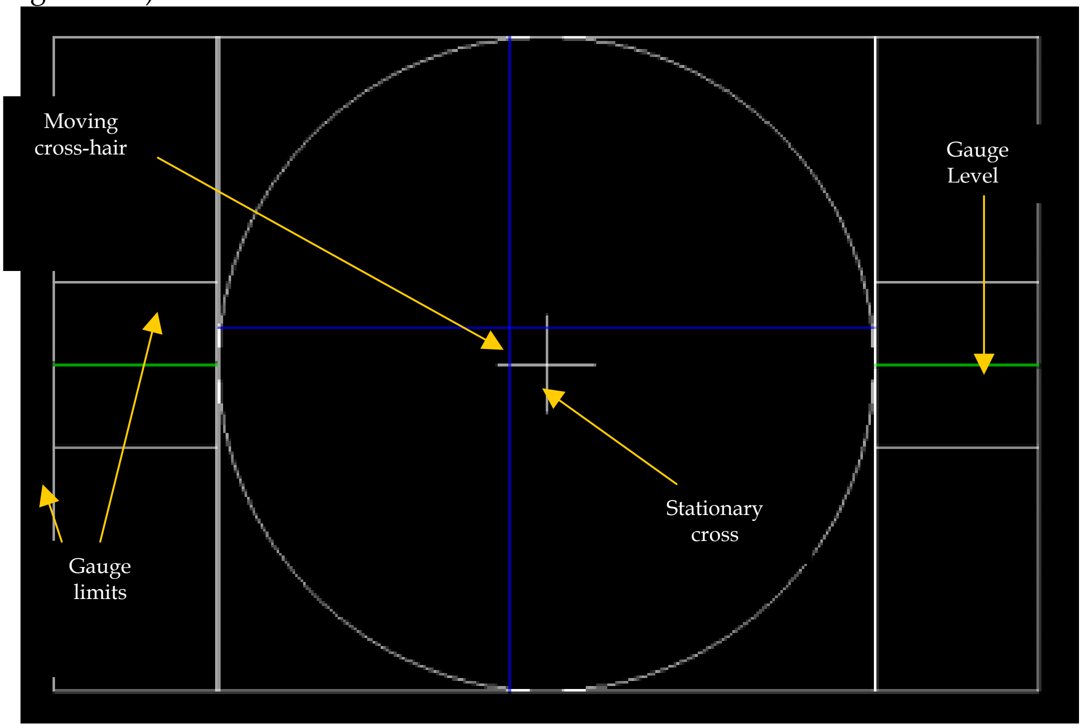

Figure 5. Screen shot of workload software depicting the tracking and gauges tasks.

Procedure

This experiment consisted of four workload levels that are summarised in Table 3. In the baseline level, participants were required to perform the tracking task, which involved keeping a moving cross-hair aligned with a smaller stationary cross centred on the computer screen (see Figure 5), and respond to auditory warnings. A total of eight auditory warnings were presented throughout a trial at various pre-determined intervals. Participants responded to auditory warnings by pressing an appropriate key on the computer keyboard.

Workload was manipulated by adding tasks to the baseline level. In level two, participants performed a visuomotor task in addition to the baseline tasks. The participants were required to monitor moving-line gauges and respond when the gauges moved out of range by pressing a button on the joystick used for the tracking task. A gauge was deemed out of range when the green lines represented in Figure 5 moved either above or below the white boundary lines and turned red.

The third workload level involved an auditory attention task as well as the baseline tasks. ATC messages were played to the participant throughout the trial in the presence of continuous, non-related, background ATC chatter. Participants were required to attend to the messages and to remember any message which began with the call-sign "Condor" (eg. Condor, ascend 1000 feet).

The fourth workload level encompassed all components of the other three levels, i.e. the baseline tasks (tracking and auditory warning responses), the visuomotor gauge task and the ATC messages.

Table 3. Tasks per workload condition.

Workload Level	Tasks
1:	Tracking & auditory warning responses
2:	Tracking, auditory warning responses & gauge monitoring
3:	Tracking, auditory warning responses & ATC messages
4:	Tracking, auditory warning responses, gauge monitoring & ATC messages

Participants were required to attend one familiarisation and six test sessions (two for each of the three sound types). The familiarisation session consisted of training on each of the four workload levels (see Table 3). The test sessions commenced with participants learning the association between the auditory warnings to be used in that session and the eight events, utilising the same software as in Experiment One and the same procedure as described for Day 1 of that experiment. Each test session also consisted of eight experimental trials of 180 seconds duration. The eight experimental trials consisted of two presentations of each workload level, with presentation order of the workload levels counterbalanced. Order of presentation of warning types was counterbalanced across test sessions.

3.3 Results

Participants' responses to the auditory warnings were measured and analysed in terms of reaction time (time between sound onset and participant response) and accuracy of response.

3.3.1 Reaction Time Data

Participants' reaction times to auditory warnings were analysed across the variables of sound type (speech, auditory icons, abstract sounds), gauges (present/absent) and ATC messages (present/absent).

A 3 X 2 X 2 Analysis of Variance conducted on the data revealed no significant interactions or main effect for the gauges variable ($p > .05$). However, a significant main effect was detected for sound type ($F(2,22) = 54.15$, $p < .01$). Planned comparisons indicated that participants responded significantly faster to speech warnings than to icons ($F(1,11) = 16.67$, $p < .01$) and significantly faster to icons than to abstract sounds ($F(1,11) = 88.78$, $p < .001$).

There was a significant main effect for the ATC messages variable ($F(1,11) = 36.07$, $p < .01$), with reaction time to auditory warnings increasing in the presence of the ATC message monitoring task. A slight trend was detected in the reaction time data for both the ATC messages and gauges variables. While the addition of the the gauge monitoring task (see Figure 6) and the background ATC messages (see Figure 7) led to an increase in reaction time for both speech and abstract sounds, the reaction time to icons reversed this trend (see figure 7). It is worth noting that the statistical power of this experiment was reasonably low (ranging from .358 to .621 for the non-significant tests) indicating that a replication of this experiment with an increased number of participants may produce different results.

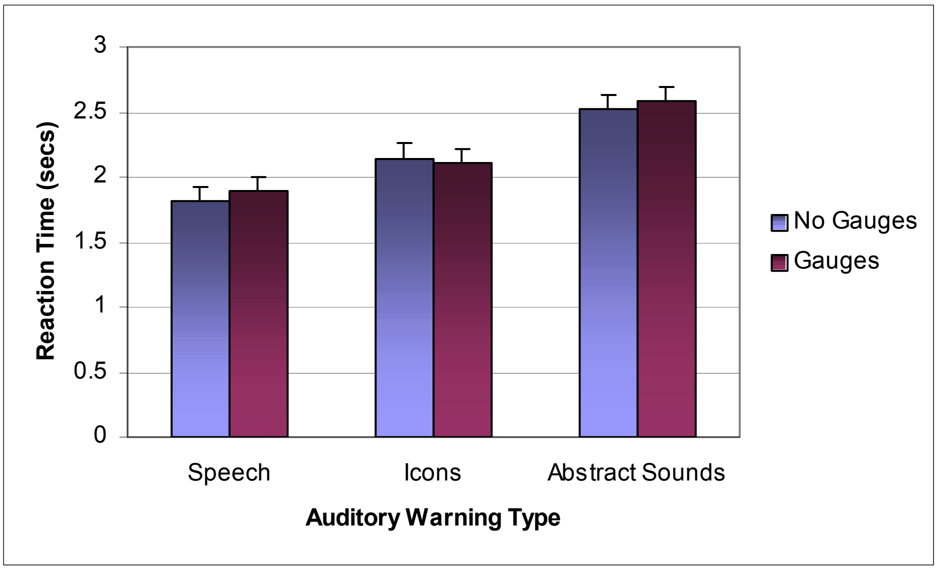

Figure 6. Mean reaction time to auditory warnings across the gauges condition

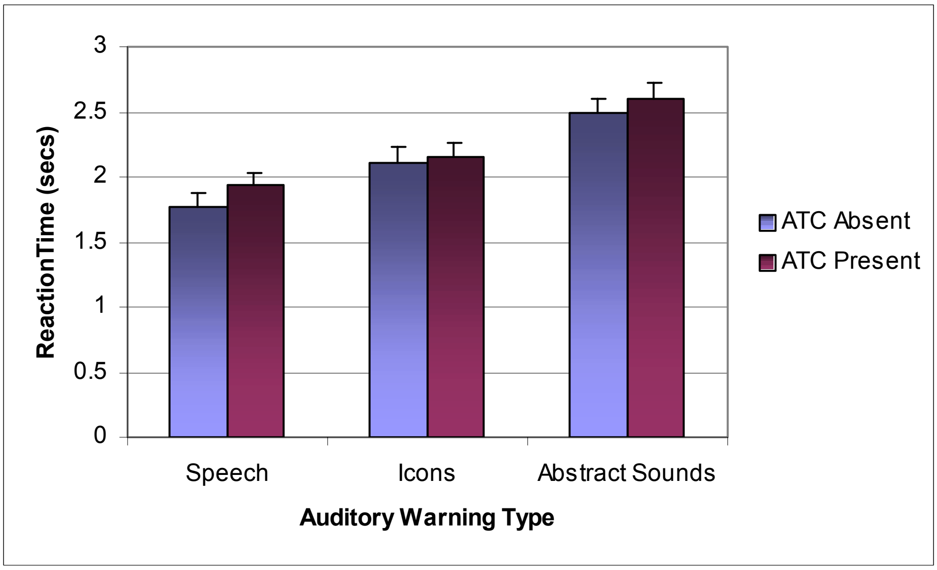

Figure 7. Mean reaction time to auditory warnings across the ATC messages condition

3.3.2 Accuracy Data

Participants' responses to auditory warnings were also analysed in terms of response accuracy. Due to the lack of variability in the accuracy of responses to speech warnings, a non-parametric Analysis of Variance (Friedman 2-way ANOVA) was applied to the accuracy data collapsed across the ATC messages and gauge variables. The data had to be collapsed as the non-parametric test used does not permit evaluation of interactions, therefore it was not possible to evaluate the impact, if any, of workload on participants' accuracy in responding to the different warning types. This analysis indicated a significant effect for sound type (χ_r^2 (2) = 18.17, p<.05). Further analyses revealed significant differences between the numbers of errors made while reaching criterion for abstract sounds and icons (Day 3, D=18, p<.05; Day 8, D=18, p<.05) and abstract sounds and speech (Day 3, D=22.5, p<.05; Day 8, D=21, p<.05). Further analyses of mean response accuracy revealed that speech warnings were identified significantly more accurately than both icons (D = 14.5, p<.05) and abstract sounds (D = 20, p<.05). There was no significant difference between mean accuracy scores for icons and abstract sounds (D = 5.5, p<.05) (refer to figure 8).

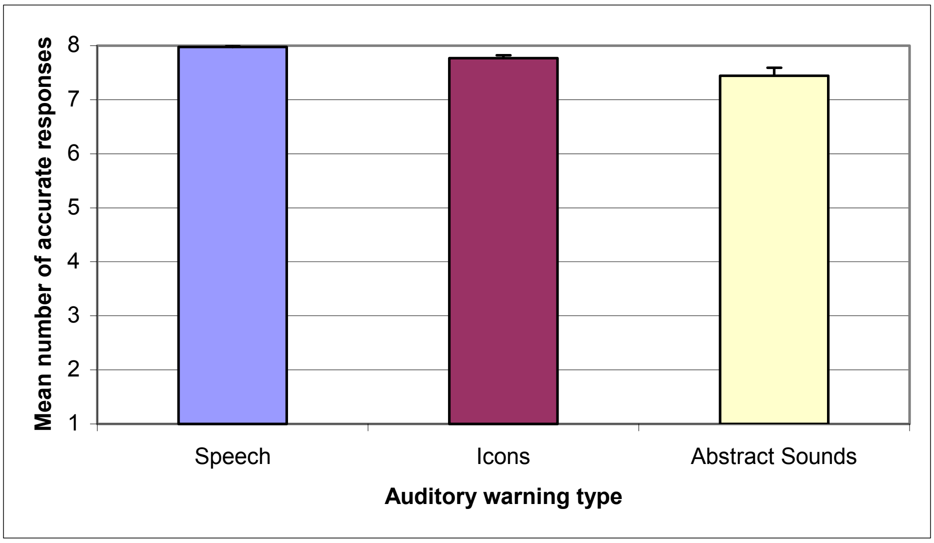

Figure 8. Mean number of accurate responses per trial to auditory warnings

3.4 Discussion

Participants' reaction time performance using abstract sounds under conditions of varying workload was significantly worse (that is slower) than for the other two sound types. Whilst the statistical analysis of the accuracy data identified a significant difference between abstract sounds and speech, participants correctly identified abstract sounds 7.4 times out of 8 indicating that once learnt, abstract sounds do demonstrate some potential for use as warnings. However, this level of accuracy came at a cost of increased response time. Consequently, it is reasonable to infer that abstract sounds are not the most effective sound type to use as an auditory warning in situations where there are eight warnings in a set and participants are required to use the warnings under conditions of high workload.

The results from Experiment One indicated that speech and icons were equivalent in terms of learning and retention. In Experiment Two, however, where participants were required to perform additional tasks, speech warnings were found to provide an advantage, as demonstrated by the faster and more accurate responses to speech warnings in all of the workload conditions. There was a slight trend in the reaction time data for the addition of extra tasks to the baseline workload condition to affect performance using speech warnings more than icons.

Overall reaction times were reasonably slow in this experiment, ranging from 1.85 seconds for speech to 2.55 seconds for abstract sounds. These lengthy reaction times can be attributed in part to the response mechanism that required participants to reach for a key on a keyboard and identify the appropriate response button from a group of eight.

It is important to note that while the overall level of workload in this experiment was designed to be demanding, participants had little difficulty with the baseline tracking condition. In fact, it was not until both the visuomotor task and the auditory attention task were added that the participants reported any difficulties. The workload resulting from this laboratory-based experiment did not approach the workload pilot's encounter in the real cockpit. Considering that the results suggest that is possible that the advantage of speech warnings over icons may be reduced with an increase in workload, further investigation is required with greater workload demands placed upon the participants.

4. Conclusions

Speech, abstract sounds and auditory icons have been identified as potential auditory warning types. Each of these sound types have advantages and disadvantages associated with their use as auditory warnings. For example, speech warnings have the advantage of requiring little learning and being capable of presenting complex information, yet they suffer from the potential of being masked by speech already present in the cockpit. Abstract sounds should be less likely to be masked by cockpit speech but are more difficult to learn. Auditory icons share the advantage of abstract sounds in that they should be less likely to be masked by cockpit speech, and in addition are less difficult to learn. However, at times it could be difficult to match an auditory icon with a warning event, as not all warning events produce a sound. These advanatges and disadvantages directly impact upon the effectiveness of these sound types' use as auditory warnings. Some of these issues were addressed by the experiments described in this report.

The results of Experiment One indicate that auditory icon warnings are learnt and retained as readily as speech warnings, indicating their potential for use as auditory warnings. The results of Experiment Two further demonstrated the potential of icons, as there was a trend for response time to icon warnings to be slightly less disrupted by competing tasks than speech and abstract sounds. It was observed that the auditory icons used in these two studies varied with respect to the type of relationship they had with the event they were chosen to represent (i.e. direct vs indirect). Whether the relationship between an auditory icon and the event it is intended to represent impacts upon the effectiveness of the icon requires further investigation. Therefore, before auditory icons can be recommended for use as auditory warnings, a systematic investigation that explicitly compares the learning and retention of icon-event pairings with direct or indirect relationships is required. Further investigation is also required into the impact of increased workload on response times to icon and speech warnings.

The results of Experiments One and Two suggest that there is an advantage for speech and icons over abstract sounds, which can perhaps be attributed to their naturalistic nature. However, abstract sounds should not be disregarded altogether. It is worth noting that identification of abstract sounds in Experiment Two was almost perfect under various workload conditions (see Figure 8), although reaction time was significantly increased relative to the other two sound types. It is possible that there is a place for abstract sound warnings in the cockpit, e.g., in operational environments where there are few auditory warnings to distinguish between or the informational content of auditory warnings is low. It is also worth considering that abstract sounds are easier to design than effective icon warnings and are not as susceptible as speech to the filtering effects of the communication systems.

It can be surmised from the results of these two experiments that speech and auditory icons hold demonstrative potential for use as auditory warnings. However, it is evident that further research is required before strong recommendations can be made as to whether one sound type is preferable to the other, and if there are any circumstances under which one sound type would be preferrable to the other. Further research has focused on the effect that relationship type between an auditory icon and the event it represents has upon learning and retention (Stephan, Smith, Parker, McAnally & Martin, 2000). This research has demonstrated that sound-event pairs with an indirect relationship are learnt as easily those with a direct relationship. Another study is currently underway to evalaute the effect that workload levels, higher than those induced in Experiment Two, have upon response to speech and icon auditory warnings, with icon warnings broken into the categories of direct, indirect and arbitrary relationships.

5. References:

Brungart, D.S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the American Acoustical Society, 109 (3)*, 1101-1109.

Doll, T.J., Folds, D.J., &, Leiker, L.A. (1983). *Auditory information systems in military aircraft: Current configurations versus state of the art* (Report USAFSAM-TR-84-15). USAF School of Aerospace Medicine, Brooks Air Force Base, Texas.

Edworthy, J., & Stanton, N. (1995). A user centred approach to the design and evaluation of auditory warning signals: 1. Methodology. *Ergonomics 38 (11),* 2262-2280.

Gaver, W.W. (1986). Auditory icons: Using sound in Computer Interfaces. *Human-Computer Interaction*, *2*, 167-177.

Lawrence, D.M. & Banks, W.P. (1973). Accuracy of recognition memory for common sounds. *Bulletin of the Psychonomic Society, 1*, 298-300.

Lawrence, D.M., Cobb, N.J., & Beard, J.I. (1979). Comparison of accuracy in auditory and tactile recognition memory for environmental stimuli. *Perceptual and Motor Skills, 48*, 63-66.

Meredith, C.S., & Edworthy, J. (1991). Sources of confusion in intensive therapy unit alarms. In N. Stanton (Ed.) *Human Factors in Alarm Design* (pp. 207-219). Taylor and Francis: London.

Momtahan, K, Hetu, R., & Tansley, B. (1993). Audibility and identification of auditory alarms in the operating room and intensive care unit. *Ergomonics, 36*, 1159-1176.

Patterson, R.D. (1982). *Guidelines for auditory warning systems on civil aircraft (CAA Paper 82017).* London: Civil Aviation Authority.

Patterson, R.D. and Milroy, R. (1980). *Auditory warnings on civil aircraft: the learning and retention of warnings. (CAA Paper 7D/S/0142).* London: Civil Aviation Authority.

Sanders, M.S. and McCormick, E.J. (1987). *Human Factors in Engineering and Design, Sixth Edition.* New York: McGraw-Hill.

Siegel, S. and Castellan, N.J. (1988). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Stephan, K., Smith, S., Parker, S., McAnally, K., & Martin, R. (2000). Auditory Warnings in the cockpit: An Evaluation of Potential Sound Types. In G. Edkins & P. Pfister (Eds.), *Innovation and Consolidation in Aviation* (pp. 231-241). Burlington, VT USA: Ashgate Publishing.

Stanton, N.A., & Edworthy, J. (1999). Auditory warning and displays: An overview. In N. Stanton & J. Edworthy (Eds.) *Human Factors in Auditory Warnings* (pp.3-30). Aldershot: Ashgate.

DISTRIBUTION LIST

Auditory Warnings in the Military Cockpit:
A Preliminary Evaluation of Potential Sound Types

Sean E. Smith, Karen L. Stephan and Simon P.A. Parker

AUSTRALIA

DEFENCE ORGANISATION

	No. of copies
S&T Program	
Chief Defence Scientist FAS Science Policy AS Science Corporate Management Director General Science Policy Development	shared copy
Counsellor Defence Science, London	Doc Data Sheet
Counsellor Defence Science, Washington	Doc Data Sheet
Scientific Adviser to MRDC, Thailand	Doc Data Sheet
Scientific Adviser Joint	1
Navy Scientific Adviser	Doc Data Sht & Dist List
Air Force Scientific Adviser	1
Army Scientific Adviser	Doc Data Sht & Dist List
Scientific Adviser to the DMO M&A	Doc Data Sht & Dist List
Scientific Adviser to the DMO ELL	Doc Data Sht & Dist List
Systems Sciences Laboratory	
Chief of Air Operations Division	Doc Data Sht & Dist List
Research Leader	Doc Data Sht & Dist List
Task Manager	1
Sean E. Smit	1
Karen L. Stephan	1
Simon P.A. Parker	1
DSTO Library and Archives	
Library Fishermans Bend	Doc Data Sheet
Library Edinburgh	1
Defence Archives	1
Capability Systems Division	
Director General Maritime Development	Doc Data Sheet
Director General Capability and Plans	Doc Data Sheet
Assistant Secretary Investment Analysis	Doc Data Sheet
Director Capability Plans and Programming	Doc Data Sheet
Director Trials	Doc Data Sheet

Office of the Chief Information Officer

Deputy CIO	Doc Data Sheet
Director General Information Policy and Plans	Doc Data Sheet
AS Information Strategy and Futures	Doc Data Sheet
AS Information Architecture and Management	Doc Data Sheet
Director General Australian Defence Simulation Office	Doc Data Sheet
Director General Information Services	Doc Data Sheet

Strategy Group

Director General Military Strategy	Doc Data Sheet
Director General Preparedness	Doc Data Sheet

Navy

SO (SCIENCE), COMAUSNAVSURFGRP, NSW	Doc Data Sht & Dist List
Director General Navy Capability, Performance and Plans, Navy Headquarters	Doc Data Sheet
Director General Navy Strategic Policy and Futures, Navy Headquarters	Doc Data Sheet

Air Force

SO (Science) - Headquarters Air Combat Group, RAAF Base, Williamtown NSW 2314	Doc Data Sht & Exec Summ

Army

ABCA National Stabdardisation Officer, Land Warfare Development Sector, Puckypunyal	e-mailed Doc Data Sht
SO (Science), Deployable Joint Force Headquarters (DJFHQ) (L), Enoggera QLD	Doc Data Sheet
SO (Science) - Land Headquarters (LHQ), Victoria Barracks NSW	Doc Data & Exec Summ

Intelligence Program

DGSTA Defence Intelligence Organisation	1
Manager, Information Centre, Defence Intelligence Organisation	1 (PDF)
Assistant Secretary Corporate, Defence Imagery and Geospatial Organisation	Doc Data Sheet

Defence Materiel Organisation

Deputy CEO	Doc Data Sheet
Head Aerospace Systems Division	Doc Data Sheet
Head Maritime Systems Division	Doc Data Sheet
Chief Joint Logistics Command	Doc Data Sheet
Head Materiel Finance	Doc Data Sheet

Defence Libraries

Library Manager, DLS-Canberra	Doc Data Sheet
Library Manager, DLS - Sydney West	Doc Data Sheet

OTHER ORGANISATIONS

National Library of Australia	1
NASA (Canberra)	1

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy	
Library	1
Head of Aerospace and Mechanical Engineering	1
Serials Section (M list), Deakin University Library, Geelong, VIC	1
Hargrave Library, Monash University	Doc Data Sheet
Librarian, Flinders University	1

OUTSIDE AUSTRALIA

INTERNATIONAL DEFENCE INFORMATION CENTRES

US Defense Technical Information Center	2
UK Dstl Knowledge Services	2
Canada Defence Research Directorate R&D Knowledge & Information Management (DRDKIM)	1
NZ Defence Information Centre	1

ABSTRACTING AND INFORMATION ORGANISATIONS

Engineering Societies Library, US	1
Materials Information, Cambridge Scientific Abstracts, US	1
Documents Librarian, The Center for Research Libraries, US	1

SPARES	5

Total number of copies: Printed 31 PDF 1 = 32

Page classification: UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA							
				1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)			
2. TITLE Auditory Warnings in the Military Cockpit: A Preliminary Evaluation of Potential Sound Types				3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION) Document (U) Title (U) Abstract (U)			
4. AUTHOR(S) Sean E. Smith, Karen L. Stephan and Simon P.A. Parker				5. CORPORATE AUTHOR Systems Sciences Laboratory PO Box 1500 Edinburgh South Australia 5111 Australia			
6a. DSTO NUMBER DSTO-TR-1615	6b. AR NUMBER AR 013-186			6c. TYPE OF REPORT Technical Report		7. DOCUMENT DATE September 2004	
8. FILE NUMBER M1/9/1092	9. TASK NUMBER 99/157		10. TASK SPONSOR		11. NO. OF PAGES 20		12. NO. OF REFERENCES 13
13. URL on the World Wide Web http://www.dsto.defence.gov.au/corporate/reports/ DSTO-TR-1615.pdf				14. RELEASE AUTHORITY Chief, Air Operations Division			
15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT *Approved for public release* OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, SALISBURY, SA 5108							
16. DELIBERATE ANNOUNCEMENT No Limitations							
17. CASUAL ANNOUNCEMENT Yes							
18. DEFTEST DESCRIPTORS Auditory warning systems, Cockpits, Workloads, Situation awareness, Performance evaluation							

19. ABSTRACT

This document reports the results of two experiments assessing the viability of speech, auditory icons (environmental sounds) and abstract sounds (complex tones) for use as auditory warnings in military cockpits. Experiment One evaluated the comparative ease of learning and retention of these three sound types, with the results demonstrating that speech warnings and auditory icons were learnt and retained with equal ease, while abstract-sound warnings were found to be learnt and retained with far greater difficulty. Experiment Two examined a user's ability to respond to speech, icon and abstract auditory warnings under varying degrees of workload. The results of this experiment show that speech warnings provide an advantage in reaction time and response accuracy over auditory icon warnings, which in turn hold an advantage over abstract sounds. The results of these two experiments indicated that, under low levels of workload, speech warnings may prove to be an effective supplement to traditional visual warnings, while icons also warrant further investigation.

Page classification: UNCLASSIFIED