Design and Evaluation of Warning Sounds in Frigate Control Centres

Johan S. van Balken
Sander J. van Wijngaarden

TNO Defence and Security
PO Box 23, 3769 ZG Soesterberg
The Netherlands

{johan.vanbalken, sander.vanwijngaarden}@ tno.nl

ABSTRACT

In the ship control centre of the Air Defence and Command Frigates of the Royal Netherlands Navy, auditory warnings are used to indicate critical system conditions and dangerous situations. By means of subjective experiments, the suitability of the applied auditory warnings was assessed, and possible improvements were evaluated. It appeared that for one of the most urgent warning conditions (platform alarm) a less suitable auditory signal was applied, the perceived urgency of which is relatively low, while it is also easily confused with other signals. A straightforward way to modify the perceived urgency (through manipulation of the signal inter-onset interval) was derived from the experimental results.

1.0 INTRODUCTION

Auditory warning signals are intended to draw the attention of somebody (user, operator, driver, pilot,…) to a certain event. A wide range of auditory warning signals have been designed for an equally wide variety of applications, some serving their purpose better than others.

Several guidelines for designing auditory warnings have been reported previously (e.g. [1]). Some researchers have focussed on auditory aspects of warning signal design ([2]), while others have focussed on ergonomic and cognitive aspects ([4]). However, no "fool proof" procedure to design optimal warnings has been suggested yet. This has to do with the complexity of the process (in which different design aspects sometimes impose conflicting requirements), but probably also with the fact that it is essentially a *design* process, featuring traits of art as well as science.

However, by adopting careful procedures for evaluating (candidate) audio signals for warning applications, auditory warning systems can sometimes be optimised considerably. This paper describes the signal evaluation approach adopted by TNO for a single case, namely the Ship's Control Centre of the Royal Netherlands Navy's Air Defence and Command Frigates. Although a description of the entire optimisation approach for a complex audio interface (including evaluation of interactions with communications links and "live" speech communication) would be beyond the scope of this paper, a description of the most relevant evaluation tests can be given. These tests are concerned with the perceived urgency, perceptual dissimilarity and audibility of the signals.

2.0 DESCRIPTION OF WARNING CONDITIONS AND SIGNALS

The ship control centre contains working stations for several operators. Each working station produces auditory warnings which can be heard by all operators. Most of these warning signals are generated by the IMCS, the Integrated Monitoring and Control System, which is the interface between the operators and the ship's machinery. The warning signals evaluated in the current study had already been selected

van Balken, J.S.; van Wijngaarden, S.J. (2005) Design and Evaluation of Warning Sounds in Frigate Control Centres. In *New Directions for Improving Audio Effectiveness* (pp. 24-1 – 24-8). Meeting Proceedings RTO-MP-HFM-123, Paper 24. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.aps.

independently by the designers of the IMCS. Our purpose was to evaluate how each of these signals fits it specific purpose, in the context of the total set of signals and other sounds in the control centre.

2.1. Warning conditions

The ship control centre features a large range of warning and alarm conditions, varying from notifications that a new user has logged onto the IMCS to alarms signalling failure in vital parts of the machinery. The warning signals are structured into three categories, derived from the relative importance of the associated system events. This structure is given in table 1.

Table 1. The events and conditions for which auditory warnings are generated.

		Alert group	Description of condition
Alarm	Alarm	NBCD Alarm	Alarm for Nuclear, Biological or Chemical danger or damage.
		Platform Alarm	Alarm for defects or problems with machinery on board.
	Warning	NBCD Warning	Warning for Nuclear, Biological or Chemical danger or damage.
		Platform Warning	Warning for possible defects or problems with machinery on board.
Error	Error	IMCS Error	An error in the IMCS system because of drop outs in sensors or networkcomponents.
Event	Event	Operation Event	Notification of a non-critical condition, marking the execution of an operation or the transition in system states.
		Attention Event	Notification of a change in control of the IMCS.
		Propulsion Warning I	Indication that the requested propulsion condition can not be met.
		Propulsion Warning II	General alarm of the Propulsion Control System.

The conditions listed in table 1 are subdivided into three main categories or *alert groups.* An *alarm* indicates an problem in the machinery or warns for attacks. An *error* is a problem in the IMCS itself, indicating that it is possible that an alarm condition can not be raised when one should be. A change in the operation of the IMCS is marked by an *event*. The propulsion warning conditions also result in auditory warnings, but these are not generated by the IMCS; they have not formally been assigned an alert group category. Informally, their urgency is ranked just above or below “IMCS error”.

The subdivision of alert groups into categories implies differences in *urgency* between the various warning conditions. These differences in urgency should be reflected by the choice of warning signals; each signal should suggest an appropriate amount of urgency, while urgent signals will be allowed to be more intrusive than those of lesser urgency. The appropriateness of the warning signals associated with table I was evaluated in a subjective comparison procedure. Also, it was evaluated whether the chosen signals were sufficiently dissimilar, and whether these signals were audible in the acoustic control centre environment.

2.2.Warning sounds

Different sounds were proposed to be used for each of the alert groups. These sounds are described in table 2. Most sounds are played in a loop and stop when the alert condition is acknowledged by the operator in a computer program. Events do not require such a reaction and are only repeated 2 or 5 times.

Table 2. Description of the sounds to each of the alert groups. A short verbal description is given, as well as the fundamental frequency and duration of the sound and the inter-onset interval (IOI) between two successive sounds and the number of repetitions.

Alert group	Description of associated sound	f_0 (Hz)	dur (s)	IOI (s)	n
NBCD Alarm	buzzer with linear frequency sweep up	400-670	0.5	0.5	∞
Platform Alarm	beep (dominating component: 5.5 kHz)	1000	0.1	0.2	∞
NBDC Warning	6-step-fading synthetic tone	370	0.75	0.75	∞
Platform Warning	sharp buzzer	300	0.25	0.55	∞
IMCS Error	four gradually fading beeps (1 beep:) (complete signal:)	1000	0.15 1.2	0.3 1.2	∞
Operation Event	chime	800	0.52	0.52	5
Attention Event	chime	800	0.52	0.52	2
Propulsion warning I	triangular wave beep	2000	0.5	1.0	∞
Propulsion warning II	triangular wave beep	2000	0.15	1.0	∞

3.0 METHODS

By means of three different subjective experiments, the suitability of the applied auditory warnings was assessed, and possible improvements were evaluated. These experiments were aimed at evaluating the perceived urgency of the signals, their perceptual dissimilarity and their audibility.

3.1.Participants

Sixteen university students, 18-25 years of age, participated in the experiments. All participants reported normal hearing. They were not familiar with operations on board of frigates, nor were they in any way introduced to the IMCS; prior knowledge of the sounds or systems would potentially have biased the outcome of the results, since participants would be inclined to adjust their responses towards the current situation. The participants were paid for their attendance.

3.2.Stimulus Material

Three sets of stimuli were synthesized, based on the original signals:

1. The signals of table 2, occurring only once without repetitions
2. These same signals, but with repetitions added (looped) according to the timings of the original set
3. A set of signals based on the IMCS error signal of table 2, but with different repetition rates and temporal structures

Signal sets 1 and 2 are smaller than the original set, since some were based on the same (chime) sound, but a different number of repetitions.

3.3.Pairwise comparison (urgency)

A pairwise comparison test was carried out to determine the relative urgency of the sounds. Participants were presented with a pair of sounds, which they could play back by means of two buttons. Only one sound at a time could be played, but they could switch back and forth between sounds at will. Once they felt that they had heard enough to form an opinion ("which alarm signal sounds more urgent?"), they indicated their choice by pressing one of two "decision" buttons.

All possible pairs of sounds were presented once, in random order. From the participants' responses, a so-called preference matrix was constructed. This matrix indicates how often each of the signals was chosen over its competitor, the competitor being each other signal in the set.

Relative urgencies were calculated by taking z-scores for each of the cells, and summing row-wise ([5]). This yields a metric that indicates how often each of the signals was perceived to be "the more urgent one", using the assumption that the underlying statistical distribution is normal. Variance was calculated by means of a jackknifing procedure: for each person the urgencies per sound were determined by substracting the urgencies calculated from the matrix with the responses of all participants from those with one person left out. The significance of differences in urgency scores was determined by a binomial test. The pairwise comparison experiment was carried out with signal sets 1, 2 and 3.

3.4.Triadic comparison (dissimilarity)

A triadic comparison test was performed to determine the relative dissimilarity between the sounds. The setup was similar to the pairwise comparison, only groups of three sounds were presented. All sounds could be re-played by means of different buttons. For each triad participants had to successively point out the most similar pair and the most dissimilar pair by means of pressing buttons representing pairs of sounds.

All triads were presented once, in random order. From the participants' responses a so-called dissimilarity matrix was formed. For every pair of sounds a dissimilarity score was calculated by adding 1 unit for each "most dissimilar" rating and distracting 1 unit for every "most similar" rating. The most similar pairs were determined by finding the pairs with the lowest scores. For the triadic experiment signal set 1 was used.

Dissimilarity matrices can be subjected to multidimensional scaling techniques to explore the dimensions underlying the participants' similarity judgments. For brevity the results of such an analysis (which is routinely performed as part of this type of experiments) is not given here. Instead, the analysis presented here (in section 4) is based on the raw dissimilarity scores, converted into equivalent z-scores.

3.5.Audibility

Audibility was measured by determining the threshold of audibility when presenting the warning sounds in background noise. Participants could control the overall sound level of the looped sounds with a slider and had to adjust the loudness to make the sounds just inaudible.

All sounds were presented twice. The background noise was recorded in the control room of the frigate. The audibility results are reported here as the average signal-to-noise ratio corresponding to the threshold of audibility, where signal and noise levels are expressed as A-weighted sound pressure levels. The audibility threshold was determined for signal set 2.

4.0 RESULTS

4.1.Perceived urgency

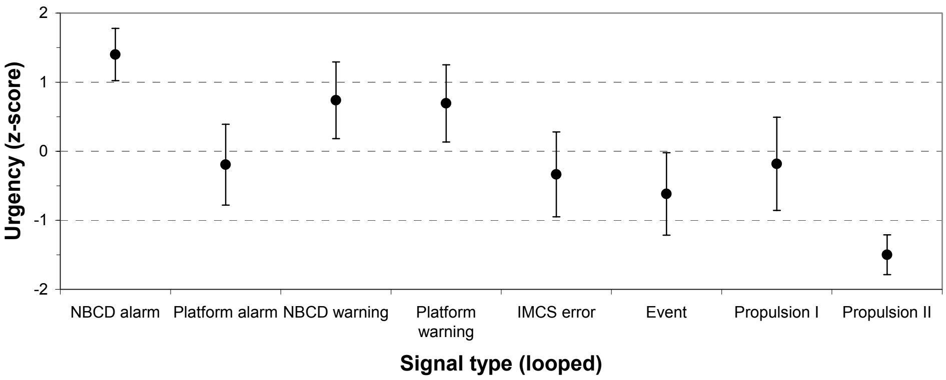

Figure 1. Perceived urgency of the looped warning sounds (signal set 2). The error bars represent one standard deviation.

Figure 1 shows that the perceived urgency rated by the participants matches the urgency of most of the alarm conditions, except for platform alarm. The platform alarm sound is judged to be less urgent than the NBCD warning ($p<0.05$). The Propulsion II sound is least urgent of all the sounds. Propulsion II is less urgent than Event ($p<0.05$).

4.2.Perceived urgency - tempo and structure

The urgency decreases with a longer inter-onset interval (Fig. 2). The three signals with an average inter-onset interval of 400 ms were judged to be more urgent than the signals with an average interval of 700 ms ($p<0.01$).

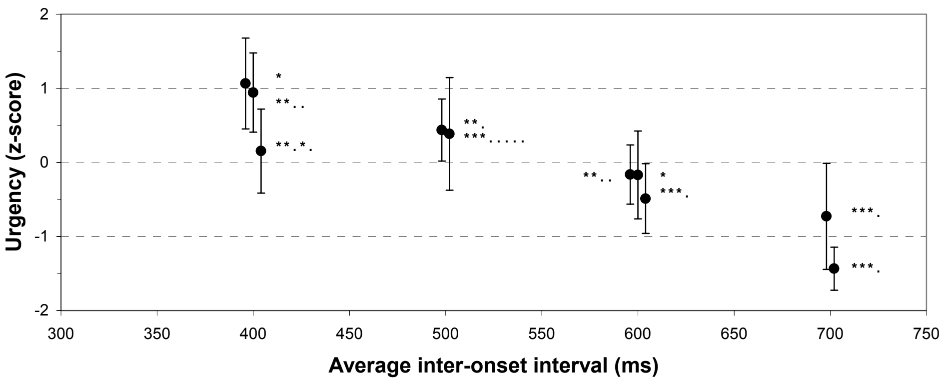

Figure 2. Urgency of IMCS beep-patterns (signal set 3) versus their average inter-onset intervals. The structure of the repetition patterns is indicated by the codes next to the points: each character denotes a period with the same duration, a '*' marks a 150 ms IMCS error beep, a '.' marks a silence. Error bars represent one standard deviation.

4.3. Perceptual dissimilarity

Dissimilarity scores were calculated for all pairs of signals out of signal set 1. Table 3 lists the five most similar pairs.

Table 3. The five most similar pairs. Similarity is expressed the z-score of the similarity scores, and the associated cumulative normal distribution for that value.

Pair		z-score
IMCS error	Platform Alarm	2.52
IMCS error	Platform Warning	1.42
Event	Propulsion I	1.34
Platform Alarm	Propulsion I	0.96
Platform Alarm	Platform Warning	0.92

The signals for “IMCS error” and “platform alarm” in set 1 are most similar. A remarkable result is that “platform alarm,” one of the more urgent alert groups, is listed three times in the table with the five most similar pairs.

4.4.Audibility

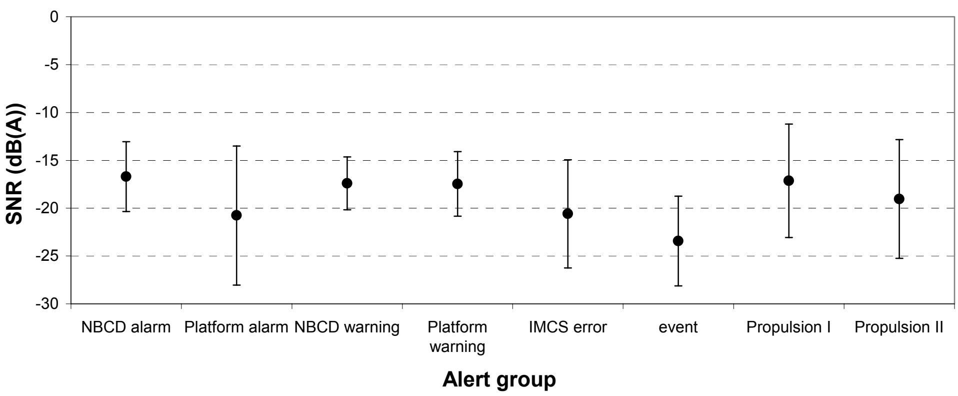

Figure 1: The audibility threshold for the warning sounds (signal set 1), expressed as the A-weighted signal-to-noise ratio (SNR). Error bars represent one standard deviation .

The average audibility threshold over all sounds lies at -19.0 dB(A) SNR. Per sound, every threshold lies within 5 dB from this average.

5. DISCUSSION AND CONCLUSIONS

Most of the signals used in the control centre were designed adequately. However, it appears that for one of the most urgent warning conditions (*platform alarm*) a less suitable auditory signal was chosen. The perceived urgency of the platform alarm sound is relatively low. Moreover, it is easily confused with a the IMCS error signal. Both sounds consist of 1000 Hz beeps, although the IMCS error signal consists of a series of 4 fading beeps, while the platform alarm uses the beeps continues. These two signals show up as the most similar pair in the triadic comparison experiment. We recommended to change either of the signals to avoid recognition problems.

Figure 2 shows that a mismatch in perceived urgency can be repaired relatively easily by adjusting the inter-onset interval. In case of the platform alarm, this means that we recommended increasing the tempo of the signal.

The audibility thresholds are satisfactory for all signals. It is commonly recommended that warning signal levels are adjusted to 15 – 25 dB above the detection threshold ([2],[3]). In this case, this is feasible with warning signal levels that do not exceed the level of the background noise.

REFERENCES

[1] Edworthy, J. Loxley, S. & Dennis, I. (1991). "Improving auditory warning design: relationship between warning sound parameters and perceived urgency." *Human Factors* 33, 205-231

[2] ISO/DIS 7731 (2002). *Ergonomics – Danger signals for public work areas – Auditory danger signals.* ISO, Geneva, Switzerland.

[3] Patterson, R.D. (1982). *Guidelines for Auditory Warning Systems on Civil Aircraft.*(CAA Paper 82017). London, UK: Civil Aviation Authority

[4] Stanton, N.E. & Edworthy, J. (1999). *Human Factors in Auditory Warnings*. (Ashgate, Aldershot, UK)

[5] Torgerson, W.S. (1958) *Theory and methods of scaling*. (Wiley and Sons, New York, USA).