EVALUATION OF AUDIO AND VISUAL ALERTS DURING A DIVIDED ATTENTION TASK IN NOISE

Ann Nakashima[1] and Jacquelyn M. Crébolder[2]
[1]Defence Research and Development Canada – Toronto, 1133 Sheppard Avenue West,
PO Box 2000, Toronto, ON, Canada M3M 3B9, ann.nakashima@drdc-rddc.gc.ca
[2]Defence Research and Development Canada – Atlantic, Dartmouth, NS, Canada B2Y 3Z7

ABSTRACT

The Halifax class frigate operations room is a demanding environment in which operators are required to monitor multiple visual displays and auditory communication channels. The current alerting system is ineffective, as the visual alerts tend to be ignored or dismissed without being read, and the auditory alerts are turned off completely. Visual alerting strategies have already been investigated. The current study compared the response times (RT) to visual, auditory and combined (audiovisual) alerts as subjects performed a visual divided attention task using three displays (secondary task). Another objective was to investigate the effects of alert type on the performance of the secondary task. The experiment was performed in quiet and in recorded frigate control room noise (69 dBA). There were no significant differences in RT between the visual and audiovisual alerts in quiet or noise. The RT for the auditory alert was significantly higher than the audiovisual alert in quiet, and than both the visual and audiovisual alerts in noise. There was no main effect of alert type on the performance of the secondary task. The audiovisual alert could be beneficial for detection in the operations room because 1) the RT was not significantly different from the visual alert, indicating that the auditory component was not distracting, and 2) it is more likely to be detected over the visual alert when the operators are looking away from the displays. Future studies should investigate the psychoacoustic properties of the auditory component of the alert for perceived urgency, in the interest of prioritizing the alerts.

RÉSUMÉ

La salle des opérations d'une frégate de classe Halifax est un environnement exigeant dans lequel les opérateurs doivent surveiller de nombreux affichages visuels et canaux de communication sonore. Le système d'alerte actuel n'est pas efficace, car les alertes visuelles sont souvent ignorées ou rejetées sans être lues, et les alertes sonores sont coupées complètement. Les stratégies d'alerte visuelle ont déjà été étudiées. La présente étude visait à comparer le temps de réaction (TR) aux alertes visuelles, sonores et combinées (audiovisuelles) de sujets qui effectuaient une tâche visuelle en situation d'attention partagée au moyen de trois afficheurs (tâche secondaire). Un autre objectif consistait à étudier les effets du type d'alerte sur le rendement pour la tâche secondaire. L'expérience a été réalisée dans un milieu silencieux et un milieu bruyant où jouait le bruit enregistré d'une salle de contrôle de frégate (69 dBA). On n'a remarqué aucune différence significative dans le TR entre les alertes visuelles et audiovisuelles dans un milieu silencieux ou bruyant. Le TR pour l'alerte sonore était beaucoup plus élevé que celui pour l'alerte audiovisuelle dans un milieu silencieux et ceux pour les alertes visuelles et audiovisuelles dans un milieu bruyant. Le type d'alerte n'a eu aucun effet principal sur le rendement de la tâche secondaire. L'alerte audiovisuelle aiderait à la détection dans la salle des opérations pour les raisons suivantes : 1) le TR était peu différent de celui pour l'alerte visuelle, ce qui indique que les éléments sonores ne détournaient pas l'attention; 2) l'alerte audiovisuelle a plus de chances d'être détectée que l'alerte visuelle lorsque les opérateurs ne regardent pas les afficheurs. De nouvelles études devraient porter sur les propriétés psychoacoustiques de l'élément sonore de l'alerte pour les urgences perçues, de façon à établir l'ordre de priorité des alertes.

1. INTRODUCTION

The Halifax class frigate operations room is a demanding, high-intensity environment, manned by about twenty Navy personnel. Many of the personnel are sensor operators, who are required to monitor several visual displays showing tactical and administrative data. The operators also use headsets to monitor two channels of communication while keeping track of face-to-face communication within the room. Automated systems, both auditory and visual, are in place to warn operators of impending system errors and tactical threats (e.g., detection of submarines, mines, torpedoes). However, recent discussions with Navy personnel indicate that the alerting system is not effective. The alerts are not prioritized, and all alerts are sent to all operators, regardless of relevance. The auditory alert is a single tone that is activated constantly and presents no information about the urgency of the alert. As a result, the operators tend to turn the auditory alerts off due to annoyance. Visual alerts appear as flashing text at the bottom of one of the screens, and can be easily ignored or dismissed without being read. The Halifax Class Frigate is undergoing a complete modernization upgrade and it is therefore of interest to explore methods of enhancing the way operators are alerted.

For high-priority alerts, it is important to capture the attention of the operator as quickly as possible. Although simple reaction time (RT) has been reported as faster for auditory than for visual stimuli (130 vs 170 ms), it has been argued that this is only the case when the subjective intensity differences between the modalities was not controlled (Wickens, 1992). When a visual stimulus is presented simultaneously with an auditory stimulus (bimodal), it has been shown that subjects tend to respond to the visual stimulus, and the auditory component is often ignored (Colavita, 1974). This phenomenon is known as visual dominance (Colavita and Weisberg, 1979; Sinnett et al, 2007). It is unknown if the results for experiments of simple or serial RT can be generalized to real-world environments in which a complex task must be accomplished in the presence of environmental stressors. Sinnett et al (2007) were able to replicate the visual dominance results of Colavita (1974) and extended the experiments by increasing the perceptual load in both the visual and auditory domains. This was done by adding a number of distracting (i.e., non-target) images and sounds to the presented stimuli. There were no significant differences in RT for auditory and visual alerts with the increased perceptual load (Sinnett et al, 2007). Anaesthetists have been shown to respond more quickly to auditory alarms from the monitoring equipment than visual (Morris and Montano, 1996); however, another study showed no differences in RT between the two modalities (Loeb and Tecumseh, 2002). Conflicting results have also been shown for the use of audiovisual alerts. Some studies have shown audiovisual alerts to elicit faster RT and better accuracy (fewer missed alarms) than both unimodal auditory and visual alarms (Chan and Chan, 2006), while other studies have found better accuracy, but slower RT for the audiovisual alerts (Sinnett et al, 2007).

Another factor to consider in real-world environments is the presence of background noise. Previous studies have shown that the presence of high-level background noise did not affect serial RT for visual alerts, but accuracy was adversely affected (Abel, 1990). Auditory alerts must be presented at levels that are sufficient to be heard above the background noise, but not so high that they become startling or distracting. It has been suggested that auditory warnings should be between 15 and 25 dB above the masked threshold to ensure detection without being aversive or disruptive to verbal communication (Patterson, 1982). Thus, the optimal level of the auditory alarm depends on the level and spectrum of the background noise. In addition to the level, there are a number of psychoacoustic properties known to affect the perceived urgency of an auditory alarm (Patterson, 1982; Edworthy and Hellier, 2000; Arrabito et al, 2004).

In the frigate operations room, the high workload in both the visual and auditory modalities (multiple visual displays and communication channels) and the presence of background noise will impact how the operators react to different types of alarms. Different types and locations of visual alerts using three displays have been investigated (Crébolder and Beardsall, 2009; Crébolder & Beardsall, 2009b; Roberts, 2008, Roberts and Foster-Hunt, 2008). It was found that a vertical red bar appearing on one or all the displays was detected more quickly than a red border around the displays. This was true whether the bar appeared at the top left side, or centered across the top or bottom of the display. In addition, static alerts (red bar or border displayed for four seconds) were detected more quickly than flashing alerts, and detection was faster when the alerts were shown simultaneously on all three displays than when shown on one. The purpose of the current study was to compare the RT for the previously tested visual alerts to simple auditory and audiovisual alerts. A secondary objective was to investigate the effects of alert type on the performance of a simulated operations room task.

2. EXPERIMENTAL PROTOCOL

2.1 Participants

Protocol approval was obtained from the Human Research Ethics Committee (HREC) of Defence Research and Development Canada (DRDC), and informed consent was obtained from all participants. Sixteen subjects (eight males, eight females) aged 21 to 51 years (30.6 ± 8.8 years) participated in the study. All subjects had normal or corrected-to-normal vision. A trained technician screened the subjects for a history of ear disease and hearing thresholds no greater than 25 dB HL (hearing level), bilaterally, at pure tone frequencies of 0.5, 1, 2 and 4 kHz. The hearing thresholds were measured with an audiometer (Interacoustics AC40, Eden Prairie, MN) using Békésy tracking (Brunt, 1985), in a double-walled, soundproof booth (Series 1200, Industrial Acoustics Company, Bronx, NY). All subjects were right-handed.

2.2 Experimental platform

The experiment was conducted using a personal computer with three 22” liquid crystal display (LCD) monitors. The software was programmed using E-Prime 2.0 (Psychology software tools, Pittsburgh, PA), running on a Windows XP operating system. The primary task required the subjects to respond to alerts by pressing the space bar (the alerts are described in the next section). The secondary task required the subjects to monitor a tactical display (center monitor) and classify contacts (unknown ships) as being neutral or hostile. A schematic of the experimental platform is shown in Figure 1. The task was designed to be a high-intensity task, involving multiple displays, similar to a task that a sensor operator might be required to perform. Contacts (triangles on the screen) originated in the periphery of the tactical display and moved toward the ownship (the ship to be protected), located at the centre of the display. Upon selecting a contact on the tactical display, information regarding the ship size, ship speed and whether or not weapons were on board, was displayed on the status display (left monitor). The subject used the information to determine if the contact was hostile or neutral, and then entered the corresponding three character code on the reporting display (right monitor). The codes were “qwe” for a neutral contact and “asd” for a hostile contact. The contact information is listed in Table 1. A contact was categorized as hostile if any two of its attributes were hostile, or neutral if any two attributes were neutral. When a contact was classified correctly, it disappeared from the tactical display; otherwise, it kept moving toward the ownship. The subject received feedback about the correctness of the entered response by a message displayed under the response box (correct or incorrect). If any contact, neutral or hostile, reached the ownship, it was destroyed and the displays were reloaded.

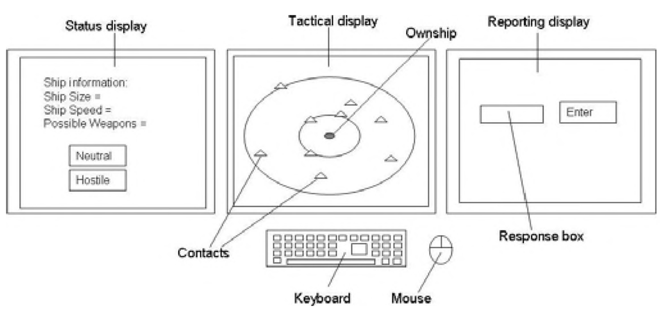

Figure 1: The experimental platform.

Table 1. Hostile and neutral attributes for contact classification.

	Categorized as Hostile	Categorized as Neutral
Ship Size	Small	Large
Ship Speed	Fast	Slow
Possible Weapons	Yes	No

2.3 Design

We employed a 3x2 repeated measures design (three alert types x two background noise conditions). The three alert types were visual, auditory and audiovisual, and will be described in greater detail below. The two background noise conditions were quiet and recorded Halifax class frigate operations room noise (69 dBA). The noise was a combination of speech and machinery noise, thus potentially providing both informational and energetic masking of the auditory alert (Brungart, 2001); the one-third octave band spectrum, as well as the total A-weighted and linear levels, are shown in Figure 2. Since most of the energy in the noise spectrum was below 1 kHz, an auditory alert composed of frequencies in this range would have to be presented at a relatively higher level to reduce the effects of masking. The auditory alert was a 1 kHz tone, presented at 75 dB SPL, which was well above the background noise level in that frequency band (54 dB SPL). We chose this simple auditory alert because it was similar to what is currently being used in the operations room, and therefore operationally relevant. Both the background noise and the auditory alerts were presented over headphones (Sennheiser HD 280), which were worn during both of the background conditions. The visual alert was a static red bar shown at the bottom of all three displays. We chose to show the visual alert on all three displays to match the omnidirectional nature of the auditory alert as closely as possible. Therefore, the three types of alerts were: audio only (1 kHz tone), visual only (a static red bar shown on all three of the displays) and audiovisual (1 kHz tone synchronized with the red bar). The alerts were displayed for four seconds, or until acknowledged by the subject by pressing the space bar. The duration of the alerts on the screen was chosen to maintain consistency with the previous studies (Crébolder and Beardsall, 2009; Crébolder & Beardsall, 2009b; Roberts, 2008, Roberts and Foster-Hunt, 2008). Alerts that were not acknowledged were recorded as being missed.

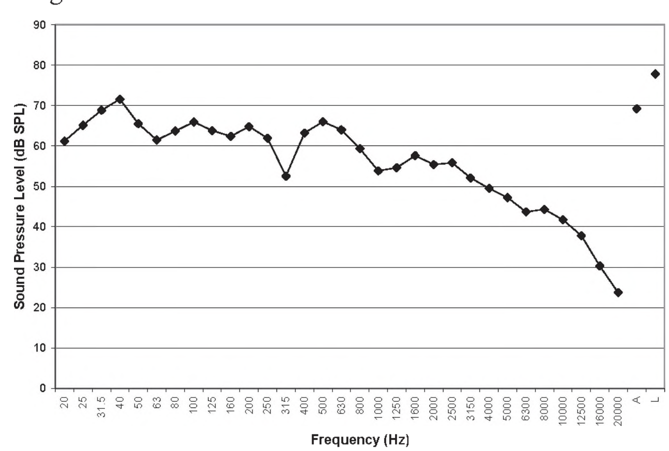

Figure 2: Background noise recorded in the Halifax Class Frigate operations room.

The experiment was divided into blocks. Between 14 and 16 alerts were presented within a block and each block lasted approximately three minutes. Only one type of alert was presented within a block. The alert types were randomized between blocks, such that nine blocks were presented in total with three of each alert type. The subjects performed the experiment twice: once in quiet, and once in noise. The order of the quiet and noise sessions was counterbalanced between subjects.

2.4 Procedure

All subjects participated in a training session to familiarize themselves with the displays, functions, and task. Using a training experiment consisting of three blocks (one of each type of alert), an experimenter talked them through the first run, and then allowed them to run through a second time without help. During the second run, the subject wore the headphones with the background noise turned on. Most subjects were comfortable with the task after completing the training experiment twice (once in quiet, once in noise).

During the experimental session, subjects first performed a warm-up set of three blocks. They then completed two sets of nine blocks: one in quiet and one in noise, with a 10-minute break in between.

2.5 Statistical Analysis

The main outcome measures of this study were RT to alert type, and efficiency of contact classification (accuracy and speed). The numbers of missed alerts, false alarms and destroyed ownship were also analyzed. Within-subjects analyses of variance (ANOVA) were applied to the data to evaluate the significance of variation in alert type and noise condition. Non-parametric analyses (Friedman) were used to analyze the contact classification data. The effects of age on the main outcome measures were calculated using correlations. All analyses were calculated using SPSS 17.0 (Statistical package for social sciences, SPSS Inc., 2008) and p values < 0.05 were considered to be significant.

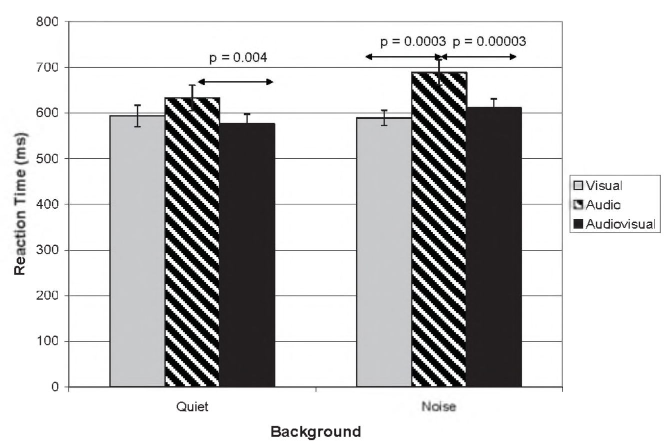

Figure 3: Reaction time in quiet and noise by alert type. The error bars indicate standard error.

3. RESULTS

3.1 Primary Task – Alert Detection

The RTs for the three types of alerts, grouped by background noise condition, are shown in Figure 3. A repeated-measures ANOVA with two factors (alert type and background) was applied to the RT data. Because sphericity was violated for alert type, a Greenhouse-Geisser correction was applied to the degrees of freedom ($\varepsilon = 0.688$). There were main effects for alert type, $F(1.38, 20.65) = 18.65$, $p = 0.0001$, partial $\eta^2 = 0.554$ and background, $F(1, 15) = 12.16$, $p = 0.003$, partial $\eta^2 = 0.448$. There was also a significant interaction between type and background, $F(2, 30) = 9.321$, $p = 0.001$, partial $\eta^2 = 0.383$. Post-hoc analyses using a Bonferroni correction showed significantly longer RT for the audio alert (633 ±113 ms) than the audiovisual alert (575 ± 84 ms) in quiet ($p = 0.004$). In noise, the mean RT for the audio alert (689 ±110 ms) were significantly higher than both the visual (589 ± 67 ms, $p = 0.0003$) and audiovisual (611± 78 ms, $p = 0.00003$) alerts.

The mean RTs, grouped by alert type, are shown in Figure 4. T-tests were calculated to look at the effect of background noise on RT for each of the alert types. There were significant differences between the RT for the auditory alert in quiet (633 ± 113 ms) and noise (689 ± 110 ms), $t(15) = 0.001$ (two-sided) and the audiovisual alert in quiet (576 ± 84 ms) and noise (611 ± 78 ms), $t(15) = 0.004$ (two-sided).

Pearson correlation coefficients were calculated between age and RT for each alert type and background condition (Table 2). Significant positive correlations were found for the audiovisual RT in quiet ($r = 0.514$, $p = 0.04$), and all alert types in noise (visual: $r = 0.558$, $p = 0.025$; audio: $r = 0609$, $p = 0.012$; audiovisual: $r = 0.663$, $p = 0.005$). The correlations indicated an increase in reaction time with age, especially in noise.

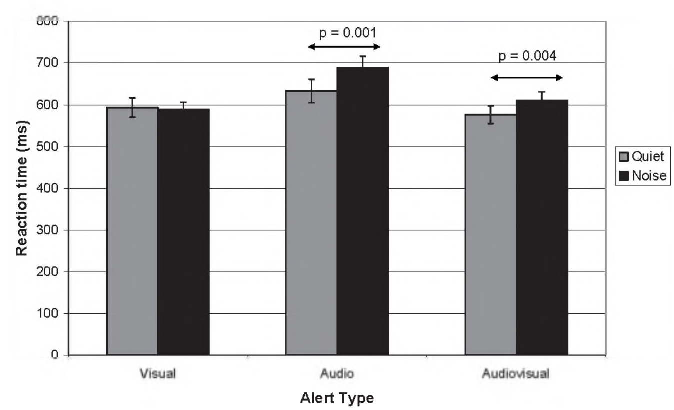

Figure 4: Reaction time by alert type in quiet and noise. The error bars indicate standard error.

There was only one occurrence of a missed alert by one subject (a visual alert in the quiet condition), so no analysis was performed on the miss data. The number of false alarms (hitting the space bar when no alert was presented) was also small. The average number of false alarms in quiet was 1.3± 1.5 (collapsed across alert type) and 2.4 ± 1.7 in noise, out of approximately 135 presentations (nine blocks of 14 to 16 presentations).

Table 2. Pearson correlation coefficients between subject age and reaction time.

Alert type (Background)	Correlation coefficient	P
Visual (quiet)	0.458	0.074
Audio (quiet)	0.465	0.069
Audiovisual (quiet)	0.514	0.04*
Visual (noise)	0.558	0.025*
Audio (noise)	0.609	0.012*
Audiovisual (noise)	0.663	0.005*

3.2 Secondary Task – Contact Classification

Performance on the secondary task was analyzed based on the accuracy and the total number of contacts classified (neutral and hostile combined). The accuracy data (percentage correct) for contact classification were non-normally distributed. Non-parametric Friedman tests applied to the data in quiet and noise showed no main effect of alert type for either background condition. In addition, there was no main effect of background on accuracy. The average accuracies (collapsed across all alert types) were 99.0 ± 0.8% and 98.6 ± 1.3% in quiet and noise, respectively. Similarly, Friedman tests applied to the data for number of contacts classified showed no main effect of alert type or background. The average numbers of contacts classified (collapsed across all alert types) were 335 ± 17 and 335 ± 18 in quiet and noise, respectively.

Ownship explosions occurred if a contact reached the ownship without being classified. There were only two occurrences of exploded ownship in quiet (one subject, once each during the visual and audio alert conditions) and four in noise (four different subjects, twice each during the visual and audiovisual alert conditions).

4. DISCUSSION

It has been demonstrated that the auditory modality is superior to the visual for alerting (Wickens, 1992); therefore, one might expect that a unimodal auditory alert would be responded to more quickly than a unimodal visual alert. However, factors such as intensity differences between the stimuli, workload in the two modalities, and environmental noise, seem to complicate these general conclusions. In the current study, there was no significant difference in RT between the audio and visual alerts in quiet, but the RT was significantly higher for the audio alert than the visual in noise. This finding might be explained by the fact that the perceptual load in the auditory modality is low in the quiet condition, allowing more modality-specific resources to process the target (Sinnett et al, 2007). In the noise condition, since the perceptual load increased in the auditory modality but not the visual, audio RT increased.

The choice of modality for an alert in part depends on the attentional resources being consumed by the task at hand. The auditory modality is omnidirectional, meaning that an auditory alert can be heard regardless of the source location. By contrast, in order for a visual alert to be effective, the subject must process it by actively attending to the spatial location of the alert (Posner, Nissen, & Klein, 1976). For example, it has been suggested that anaesthetists spend less than one-third of their time looking at the monitors in the operating room, which places limitations on the use of unimodal visual alarms (Loeb and Fitch, 2002). For this type of environment, it may be safer to rely on auditory alarms when a time-sensitive (i.e., high priority) response is required (Morris and Montano, 1996). In the current study, the subjects were engaged in a visually intensive task, requiring them to attend to the displays at all times. This was likely why the RT for visual and audio alerts were not significantly different in the quiet condition.

The use of a bimodal (e.g., audiovisual) alert can help to enhance detection in complex, multimodal environments. It is often necessary to use auditory alerts in conjunction with visual alerts because operators need to read or otherwise examine the context of the warning message (Chan and Chan, 2006). The literature has suggested that audiovisual alerts lead to fewer detection errors, but may or may not be detected more quickly than their unimodal counterparts (Chan and Chan, 2006; Sinnett et al, 2007). Our results showed that the RT for the audiovisual alert in both the quiet and noise conditions was significantly faster than the RT for the unimodal audio alert, but not significantly different from the unimodal visual alert. It is possible that the subjects displayed visual dominance in their response to the audiovisual alerts (responded to the visual component while ignoring the audio); however, since the RT for audiovisual alerts was not significantly different from the visual alerts, it appeared that the auditory component did not hinder detection. None of the alert types affected performance on the secondary task. A combined alert would likely be advantageous because operators will shift their attention from the displays for various reasons, such as when talking to someone in the room.

Our results for the unimodal auditory alert (slower RT than visual and audiovisual) seem to support the behaviour of the sensor operators, who turn the auditory alerts off due to annoyance. However, since the audiovisual alert did not cause any performance decrements over the visual alert in terms of RT or secondary task performance, the design and implementation of audiovisual alerts in the operations room should be further investigated. Specifically, the addition of a divided attention auditory task to simulate the monitoring of communication channels should be added to the experimental platform. The additional load on the auditory channel will make it more difficult to alert the operator. In addition, different types of auditory alerts could be introduced to demonstrate urgency or priority.

5. CONCLUSION

An investigation of visual, audio and audiovisual alerts in the Halifax class frigate control room environment showed that while performance of the secondary visual task was not different across alert types, reaction time was slowest for the auditory alert in both quiet and noise. Reaction time for the audiovisual alert was not significantly different than for the visual alert, and the audiovisual alert may be more easily detected when the operators are not looking directly at the displays. The design and implementation of audiovisual alerts in the modernized frigate control room should therefore be further investigated.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Sharon M. Abel for her scientific input, Mr. G. Robert Arrabito for reviewing an earlier draft of this paper and Mrs. Ingrid Smith for conducting the study. This research was supported by Defence Research and Development Canada, Land Partner Group, Command Thrust: Auditory Overload in Land Operations (12oi), Internal Operating Grant.

REFERENCES

Abel SM. (1990). The extra-auditory effects of noise and annoyance: An overview of research. J Otolaryngology, 19 Suppl 1:1-13.

Arrabito GR, Mondor TA and Kent KJ. (2004). Judging the urgency of non-verbal auditory alarms: A case study. Ergonomics, 47(8):821-840.

Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. J Acoust Soc Am 2001; 109(3):1101-1109.

Brunt MA: Békésy Audiometry and Loudness Balance Testing. Baltimore, MD. Williams and Wilkins, 1985.

Chan AHS and Chan KWL. (2006). Synchronous and asynchronous presentations of auditory and visual signals: Implications for control console design. Applied Ergonomics 37:131-140.

Colavita FB. (1974). Human sensory dominance. Perception & Psychophysics, 16:409-412.

Colavita FB and Weisberg D. (1979). A further investigation of visual dominance. Perception & Psychophysics, 25(4):345-347.

Crébolder JM and Beardsall J. (2009a). Visual alerting in complex command and control environments. Poster presented at the Human Factors and Ergonomics Society 53rd Annual Meeting, San Antonio, Texas, October 2009.

Crébolder JM and Beardsall J. (2009b). Investigating visual alerting in maritime command and control. Defence Research and Development Canada – Atlantic Technical Memorandum, TM 2008-281.

Edworthy J and Hellier E. (2000). Auditory warnings in noisy environments. Noise Health, 6:27-39.

Loeb RG and Fitch WT. (2002). A laboratory evaluation of an auditory display designed to enhance intraoperative monitoring. Anesth Analg, 94:362-8.

Morris RW and Montano SR. (1996). Response times to visual and auditory alarms during anaesthesia. Anaesth Intens Care, 24:682-584.

Patterson RD. (1982). Guidelines for auditory warning systems on civil aircraft. Report 82017, London: Civil Aviation Authority.

Posner MI, Nissen MJ and Klein RM. Visual dominance: An information-processing account of its origins and significance. Psychological Review, 83:157-171.

Roberts S. (2008). Evaluation of Alerts in the maritime domain: Study 2 program modifications. Defence Research and Development Canada – Atlantic Contract Report CR2008-268.

Roberts S and Foster-Hunt T. (2008). Evaluation of visual alerts in the maritime domain: Behavioural research study. Defence Research and Development Canada - Atlantic Contract Report CR 2008-057.

Sinnett S, Spence C and Soto-Faraco SS. (2007). Visual dominance and attention: The Colavita effect revisited. Perception & Psychophysics, 69(5):673-686.

Wickens CD. (1992). Engineering psychology and human performance, Second edition. HarperCollins, New York.