Journal of Experimental Psychology: Applied
2003, Vol. 9, No. 1, 12–22

1076-898X/03/$12.00 DOI: 10.1037/1076-898X.9.1.12

Using Auditory Streaming to Reduce Disruption to Serial Memory by Extraneous Auditory Warnings

Simon Banbury
Cardiff University

Liz Fricker
Centre for Human Sciences, QinetiQ

Sébastien Tremblay
Université Laval

Lucy Emery
Lockheed Martin U.K. Integrated Systems

Three experiments were conducted to examine the effects of extraneous speech warnings (i.e., low-priority warnings initiated during high-priority tasks) on cognitive performance and whether organizing the auditory warnings into streams can attenuate any disruption. Experiment 1 demonstrated that a variety of speech warnings can be separated into perceptually distinct streams by allocating them to discrete spatial locations. Experiment 2 showed that increasing the rate of presentation of the warnings to promote streaming decreased clarity ratings but increased perceived urgency ratings. Experiment 3 demonstrated that the disruption to serial memory for navigational information by extraneous speech warnings could be attenuated by streaming. Results are interpreted in light of previous research, and practical implications for auditory warning design are discussed.

Auditory warnings have an important role in attracting attention and conveying information in a number of work domains, including commercial and military aviation. Warnings are a simple but vital "first call" to situations in which action might be required to alleviate an accident or incident (Edworthy, 1994) and are typically used in high-workload, high-stress working environments (Edworthy & Adams, 1996). The prevalence of auditory warnings is rapidly increasing because of, in part, an increase in technological advances and capabilities (Pritchett, 2001). This trend is likely to continue given that designers have tended to operate on the basis of a "better safe than sorry" philosophy with regard to the integration of auditory warnings into new systems (Edworthy & Adams, 1996).

Designers of auditory warnings seek to improve the efficiency of reorienting the attention of the operator toward a critical event during the onset of the alarm. Indeed, the majority of research conducted on auditory warnings has focused on the factors that influence the efficiency of a warning, such as detection (e.g., Patterson, 1982), relative discrimination (differentiating between two or more warnings presented together; e.g., Deatherage, 1972), identification (e.g., Pollack & Ficks, 1954), localization (e.g., Caelli & Porter, 1980), and perceived urgency (Edworthy, Loxley, & Dennis, 1991). However, relatively little research has addressed the potential side effects of auditory warnings—specifically, the effects of extraneous auditory warnings on tasks undertaken concurrently. For example, there is relatively little research into what occurs when the operator chooses to ignore an active auditory warning in favor of concentrating on a more critical task.

Pritchett (2001) differentiated between a number of warning types; however, the simplest and most prevalent are signal detectors. Typically, the alerting system monitors a sensor input and alerts the operator when the input passes a threshold. Because the warning is initiated when a certain parameter reaches threshold, it is insensitive to what the operator is doing at the time or what situation or state the system is in. It is possible, therefore, that an auditory warning for a low-priority event may occur during a high-priority task and persist until the alarm is manually deactivated. In this case, the operator is forced to engage in the high-priority task while the low-priority warning is active. In this article, we seek to examine whether this eventuality can cause disruption to working memory (and, in particular, memory for order, or *serial* memory) and how capitalizing on the perceptual system's tendency to segregate the auditory scene into coherent and distinct streams might attenuate this disruption (see Bregman, 1990).

At first blush, it may seem that tasks involving working memory are relatively infrequent in applied settings. However, memory is viewed by a number of influential authors as pivotal to performing real-world tasks that require dynamic and complex processing (Durso & Gronlund, 1999; Endsley, 1995). Examples include air traffic control (e.g., Gronlund, Ohrt, Dougherty, Perry, & Manning, 1998), driving (e.g., Gugerty, 1997), instrument flight (e.g., Sohn & Doane, 2000), and flight situation awareness (SA; e.g., Sohn & Dattel, 2001). Indeed, SA is perhaps one of the most important determinants of successful task performance in areas such as aviation, air traffic control, large-systems operation, tactical and strategic systems, and many other everyday activities

Simon Banbury, School of Psychology, Cardiff University, Cardiff, United Kingdom; Liz Fricker, Centre for Human Sciences, QinetiQ, Farnborough, United Kingdom; Sébastien Tremblay, School of Psychology, Université Laval, Quebec City, Canada; Lucy Emery, Lockheed Martin U.K. Integrated Systems, Havant, United Kingdom.

We thank Dylan Jones and Jacky Boivin for critical reading of an earlier version of this article.

Correspondence concerning this article should be addressed to Simon Banbury, School of Psychology, Cardiff University, Cardiff CF10 3YG, United Kingdom. E-mail: banburys@cardiff.ac.uk

(Endsley, 1995). Furthermore, Endsley argued that working memory is one of the critical factors that limit operators from acquiring and interpreting information to form SA. For example, the effective monitoring of displays or system parameters over time requires that the temporal order of this information be kept intact in working memory. Doing so enables trend information to be inferred. We argue that memory for order plays a pivotal role in an operator's ability to maintain awareness of system states: specifically, the comprehension of the situation and the prediction of future states. Given the importance of SA in everyday complex activities (for a review, see Endsley, 1995), and given the pivotal role of serial memory in the acquisition and maintenance of SA, we argue that the study of the effects of extraneous sound on serial memory is an important and worthwhile endeavor.

There is a large body of evidence that suggests that the occurrence of irrelevant auditory material can cause disruption to concurrent complex mental activities. This phenomenon has become known as the irrelevant sound effect (ISE; for a review, see Banbury, Macken, Tremblay, & Jones, 2001). In the paradigm that has been used to investigate this phenomenon, participants are required to undertake a serial memory task (e.g., remembering the order of a sequence of digits) in the presence of irrelevant sound (e.g., narrative speech or a series of changing tones). The effects of irrelevant sound on serial memory are replicable and apply to most individuals (e.g., Ellermeier & Zimmer, 1997, reported that approximately 85% of participants are disrupted by irrelevant sound). Typically, the accuracy of report in the presence of background sound declines by some 30% to 50%. In addition, tasks such as mental arithmetic (Banbury & Berry, 1998), memory for prose (Banbury & Berry, 1997), reading comprehension (Martin, Wogalter, & Forlano, 1988), and transcription from video display terminals (Morris & Jones, 1991) have also been shown to be susceptible to disruption by extraneous sound. Finally, a number of researchers have shown changing-state effects, in which repeated irrelevant sounds (e.g., repeated utterances of the letter *a* or repeated presentation of a single tone) show less disruption to short-term memory than changing irrelevant sounds (e.g., repeated utterances of the letters *abcd,* or repeated presentation of tones of different frequencies; Baddeley, 2000; Jones & Macken, 1993; Jones & Tremblay, 2000; Larsen, Baddeley, & Andrade, 2000; Neath, 2000).

There are two types of explanation that have been offered for the disruption to working memory by background sound. The first type suggests that there is some form of attentional recruitment (Cowan, 1995; see also Broadbent, 1982) produced by the sound as the to-be-remembered material is registered. For example, Cowan proposed that new incoming auditory stimuli, such as irrelevant sound, attract attention away from the memory task. This attention-based framework is related to the concept of habituation to the orienting response (i.e., a conscious or unconscious shift of attention; see Sokolov, 1963), a phenomenon that designers of auditory warnings capitalize on to reorient the attention of the operator toward a critical event. The orienting response allows a neural model of the stimulus to be formed (Cowan, 1995). Each incoming stimulus is then compared with the neural model and elicits another orienting response only if the neural model is either not yet developed enough to describe the stimulus or the stimulus has changed in such a way as to make it different from the neural model. In light of this, the propensity of a warning to reorient the attention of the operator, and disrupt the operator from the task at hand, should diminish with repeated exposure. Although this view is parsimonious and compelling at first encounter, there are two key empirical objections to the habituation-like account of the ISE: (a) Contrary to the prediction that one should habituate to the presence of irrelevant sound, the disruption by irrelevant sound is enduring (Hellbrück, Kuwano, & Namba, 1996; Jones, Macken, & Mosdell, 1997; but see Banbury & Berry, 1997, for contrary findings), and as exposure is increased (by speeding up the rate of presentation or increasing the exposure duration), the disruption also increases (Bridges & Jones, 1996); and (b) capture of attention should affect all tasks to a similar degree, but the effect is more pronounced on tasks involving serial recall (for a review, see Banbury et al., 2001).

The second type of explanation suggests that interference results from the similarity of events (i.e., objects and processes) represented in memory. However, there is some disagreement between researchers as to what similarity refers to. A number of researchers have proposed that interference is based on a conflict between what is seen and what is heard, either in terms of phonological similarity (Salamé & Baddeley, 1982), shared temporal cues (LeCompte, 1996), or the overlap of modality-independent features (Neath, 2000). Alternatively, the changing-state hypothesis (Jones, Madden, & Miles, 1992) proposes that the level of interference is related to the similarity of concurrent processes of seriation (i.e., the maintenance of order): one maintaining the order of the to-be-remembered material and the other relating to the perceptual organization of the irrelevant sound (see Jones, Beaman, & Macken, 1996, for a discussion on the latter hypothesis).

The phenomena pertaining to the ISE raise some concern over the presentation and timing of auditory warnings. For example, it is entirely feasible that a low-priority warning may be initiated while the operator is performing a higher priority task. In this case, the operator may choose to refrain from deactivating the alarm (or may be unable to) and attempt to ignore the warning and continue with the task. However, the body of evidence suggests that the background warning might disrupt performance if the task is memory based (and especially if it is also seriation based) and the warning consists of a changing sound (e.g., speech or tones). At first, there seem to be a number of options open to the system designer to reduce the disruptive effects of the irrelevant sound. One option would be to reduce the sound pressure level of the warning. However, a number of researchers have shown that the effect of irrelevant speech is independent of sound pressure level, at least between 48 dB (A) and 76 dB (A) (Colle, 1980; Ellermeier & Hellbrück, 1998). Clearly, an effective reduction in disruption can be achieved only by reducing the level of the sound to that below the threshold of audibility, which is of course incompatible with the purpose of auditory warnings. Another option would be to reduce the disruption by masking the changes in energy within the irrelevant stream by degrading the signal or adding white noise (see Northwood, Warnock, & Quirt, 1979). Once again, the first strategy is incompatible with the purpose of auditory warnings, and the success of the second strategy is doubtful given that the loudness of the masking noise has to be greater than that of the speech signal in order for the disruptive effects to be reduced (Ellermeier & Hellbrück, 1998).

An alternative approach might be to capitalize on the phenomenon of *auditory streaming*—which refers to the result of perceptual processes involved in the organization of sound—as it has been shown by a number of researchers to modulate the disruption by irrelevant sound (e.g., Jones & Macken, 1995; Jones, Saint-Aubin, & Tremblay, 1999). In physical terms, sound appears to be a jumble of undifferentiated pressure changes, whereas in perceptual terms it appears to be partitioned into streams of relatively stable and distinct auditory objects (see Bregman, 1990). Crucially, there is evidence that several streams can be formed concurrently and that the objects that compose the streams can correspond to any source of sound. Finally, the organization of sound by streaming is mediated by a number of factors, such as pitch and location.

Jones and Macken (1995) investigated the role of organizational factors (i.e., streaming) in the ISE paradigm—specifically, the role of spatial location and timing. Using a changing sequence of simple utterances (i.e., *x, j,* and *w*), they compared two conditions: one in which each utterance was assigned stereophonically to the left, right, or center spatial channel and another in which the same utterance sequence was assigned to the center channel only (i.e., monophonic presentation). When the cycle of events was repeated quickly, participants' perception of the auditory scene was different for each condition. In the stereophonic condition, participants reported hearing three separate repeating streams corresponding to the three spatial locations, whereas in the monophonic condition, only one changing stream was heard. When these sequences were used as irrelevant material, they found that disruption to serial recall was significantly less in the stereophonic (three repeating streams) condition, compared with the monophonic location (one changing stream) condition.

The challenge is to reduce the disruptive effects of warnings (if they are to be ignored in favor of conducting a task of higher priority) without diminishing the important characteristics of warnings, that of conveying meaning to the operator. In line with Jones and Macken (1995), one way of achieving this is to capitalize on the perceptual system's tendency to segregate the auditory scene into coherent and distinct streams (see Bregman, 1990). It is thought that doing so may attenuate the disruption. In the present study we sought to test these assumptions using spoken auditory warnings. Although Tremblay, Nicholls, Alford, and Jones (2000) found that nonspeech sounds (e.g., tones) can be as disruptive to performance as speech when equated in terms of their acoustic variation (however, for contrasting views, see Buchner, Irmen, & Erdfelder, 1996; Salamé & Baddeley, 1982), speech warnings were used in this study as the majority of previous research on auditory streaming has used tone or utterance-based stimuli. Furthermore, little is known about what effect streaming has on the properties of speech-based warnings (e.g., intelligibility). This is an unfortunate oversight given the increase in the prevalence of speech-based warnings due to advances in speech synthesis technologies (see Spiegel & Streeter, 1997).

The attenuation of the disruptive effects of extraneous warnings through auditory streaming was investigated in Experiment 3. However, it was necessary to conduct two preliminary experiments to ascertain whether auditory warnings can be streamed (Experiment 1) and what effect this has on their properties (Experiment 2).

Experiments 1 and 2: General Introduction

The aim of the current study was to ascertain whether auditory warnings that are streamed (i.e., components of the warning allocated to separate spatial locations) are less disruptive than normally presented ones (i.e., unstreamed). To this end, two preliminary studies were conducted. Experiment 1 examined the effects of spatial location and timing on auditory warnings, which previous research had shown to modulate the phenomenon of streaming. Experiment 2 examined the effects of streaming on the properties of warnings, in terms of clarity and perceived urgency.

Experiment 1

Jones and Macken (1995) used streaming by spatial location as a means of manipulating the level of disruption from background sound. In the present experiment a similar technique was used to examine whether the components of verbal auditory warnings can be streamed into three separate spatial locations. Participants were presented with verbal auditory warnings consisting of three syllables, with each syllable presented to the left, right, or center channel. For each warning, the rate of presentation was increased systematically, and participants were asked to indicate the point at which segregation into three separate streams occurred.

Method

Participants. Twelve volunteers were recruited among Cardiff University students. The age range of participants was 20–35 years old. All reported normal hearing and normal (or corrected-to-normal) vision.

Materials. Eight warnings were selected from a database of helicopter voice warnings using the following criteria: first, that the warnings consisted of exactly three syllables and, second, that the comprehension of the warning assumed no technical knowledge of helicopter operations. The warnings were: *al-ti-tude, fire war-ning, fif-ty feet, at-ten-tion, e-lec-trics, en-gine fail, cab-in hot,* and *fuel fil-ter.* All warnings were edited using digital editing software, so that the first syllable was allocated to the left channel only, the second syllable to the center channel (simultaneous presentation to both left and right channels), and the third syllable to the right channel only.[1] Following this, the rate of presentation for all warnings was increased in increments of 10%. Finally, a sound track for each of the eight warnings was compiled from these recordings so that the warning was presented twice[2] each time the rate of presentation increased by 10%. Pairs of warnings were separated by 2 s of quiet. Thus, each sound track comprised pairs of warnings whose rate of presentation systematically increased. The warnings were presented to participants using stereophonic headphones.

Design. For each warning participants were required to indicate the pair of warnings in which they perceived streaming to commence. The

[1] It has been shown that the changing-state effect is at the phoneme or utterance level and not at the word level. For example, Tremblay, Macken, and Jones (2000) found that repeating a polysyllabic word is more disruptive than repeating a monosyllabic word. We would therefore expect repeating one polysyllabic word (e.g., monophonic–unstreamed presentation of *attention*) to be more disruptive than repeating three monosyllabic words (e.g., stereophonic–streamed presentation of *at, ten,* and *tion*). This is investigated directly in Experiment 3.

[2] Bregman (1990) reported that the segregation of tones into separate streams does not occur straightaway. The warnings in the present study were therefore presented twice and in quick succession to encourage the onset of streaming.

modal presentation rate (repetitions per second) of each warning (i.e., the pair most frequently selected by participants as the threshold for streaming) was then calculated from the participant responses. The mode was used, rather than the mean, in order to select the pair of warnings that received the greatest consensus as to when streaming was perceived to commence. The order of presentation of the warnings was counterbalanced between subjects.

Procedure. Participants were tested individually in a soundproof laboratory. Before beginning the trial, participants were presented with an auditory demonstration of streaming by location. After the demonstration, the understanding of participants was verified by asking them to verbally report what the demonstration had shown. Each of the eight sound tracks was then played twice in full to the participants. On the second audition, participants were asked to indicate, by way of a button press, the point at which they perceived streaming (i.e., three separate streams corresponding to the three spatial locations) to occur. Participants were instructed to make a response only after the audition of the warning pair (i.e., before the next pair was presented).

Results and Discussion

The overall length values of each warning (in seconds) are presented in Table 1. In addition, the presentation rate (the number of repetitions per second) at which participants reported streaming to occur and the percentage of agreement between participants are presented. It is clear from the results that for some of the warnings there was some variation between participants as to the point at which streaming was perceived to have occurred. Given that fission into separate streams does not necessarily occur immediately (see Bregman, 1990), it is probable that if more repetitions of the stimuli were given, the consensus between participants could be improved and streaming could be perceived at slower rates of presentation. However, our intention in the present study was to examine whether capitalizing on the phenomenon of streaming can reduce disruption from auditory warnings in applied domains. In these contexts, it is desirable for fission into separate streams to occur as quickly as possible (even if this necessitates higher rates of presentation), so that interference with the primary task by extraneous warnings is reduced to the minimum. Nevertheless, for three of the warnings the consensus was reasonably high: attention (83%), engine fail (67%), and fire warning (67%). Furthermore, if we consider the percentage of participants who perceived streaming to occur at these points or slower, the rates of agreement are even higher: attention (91%), engine fail (100%), and fire warning (84%). These three warnings were selected for use in Experiment 3 to examine whether streamed warnings cause less disruption to a memory task undertaken at the same time, compared with warnings that were not streamed. In addition, the results indicate that for streaming to occur in voice warnings, the rate of presentation has to be relatively high. As such, the effects on the intelligibility of the warning are unclear, as are the effects on other characteristics of a warning such as perceived urgency (i.e., how pressing the warning is perceived to be) and clarity (i.e., how legible the warning sounds are). These issues were examined in Experiment 2.

Table 1
Modal Warning Length, Presentation Rate, and Percentage of Agreement Between Participants for Streamed Warnings (Experiment 1)

Warning	Overall length (s)	Presentation rate (repetitions/s)	Percentage of agreement
Attention	0.26	3.83	83
Altitude	0.38	2.61	50
Cabin hot	0.46	2.18	50
Electrics	0.45	2.22	42
Engine fail	0.45	2.22	67
Fire warning	0.48	2.10	67
Fifty feet	0.43	2.33	58
Fuel filter	0.63	1.59	58

Experiment 2

Overview

This experiment examined two important characteristics of a verbal warning: clarity and perceived urgency. The three verbal warnings from Experiment 1 that achieved the highest consensus as to when streaming occurred were used in the present experiment. These were attention, engine fail, and fire warning. Using the results of the subjective reports of when streaming occurred, we decreased the presentation rate of each warning by 10%. In addition, all versions of each warning were manipulated digitally to create stereophonic (i.e., streamed) and monophonic (i.e., unstreamed) versions. The purpose of this manipulation was to ascertain whether faster rates of presentation above the threshold of streaming (i.e., the stimuli selected from Experiment 1) produce less disruption when used as irrelevant material in Experiment 3. Conversely, slower presentation rates (below the threshold of streaming) were expected to show more disruption. This pattern of results was not expected to occur with the monophonic (i.e., unstreamed) recordings.

Perceived urgency. The perceived urgency of auditory warnings has been investigated at some length (Edworthy, 1994; Edworthy et al., 1991; Hellier, Edworthy, & Dennis, 1993). It has been shown that variations in acoustic features have strong and consistent effects on the perceived urgency of warnings. Perceived urgency is an important characteristic of warning given that studies have shown that increases in perceived urgency correlate with increased levels of response, such as a faster reaction time (Burt, Bartolome, Burdette, & Comstock, 1995; Haas & Casali, 1995). In addition, Edworthy and Stanton (1995) found that the rate of presentation was the strongest influence on the perceived urgency of a warning. They also argued that increasing the number of times a warning is presented also increases its perceived urgency. Clearly, the last finding is highly pertinent to the present study in that the rate of presentation is increased to encourage streaming to occur. A useful, albeit unexpected, side effect of this is that the perceived urgency of the warning might be increased.

Clarity. Given that the purpose of an auditory warning is to convey information, it is important that the content of the warning can be clearly recognized. However, the effects of auditory streaming on warning clarity are untested. Fortunately, there are a number of studies that afford some insight into the effects of auditory streaming on warning clarity. Bregman (1990) showed that as the rate of presentation increases, the perception of streaming becomes more salient. However, one concern is that participants may find it difficult to reassemble the components of the streamed warning into one, coherent utterance. For example, Warren, Obusek, Farmer, and Warren

(1969) found that participants were unable to identify the temporal order of items within a rapidly repeated sequence of unrelated sounds. Bregman and Campbell (1971) argued that was because each of the sounds was allocated to a separate stream. As a result, the perception of order was good for items within the same stream and poor for items belonging to different streams. Thus, the research of Warren and his colleagues has an important implication for the present study. The main tenet of the present study is that interference of extraneous warnings can be reduced by allocating the components of the warning to discrete spatial locations, thereby creating three separate streams, and further increasing the propensity for streaming to occur by increasing the rate of presentation. This endeavor is at odds with Warren et al.'s findings insofar as the reconstruction of a streamed warning into a recognizable word may prove difficult because of the poor perception of order between streams.

In light of these findings we would expect participant ratings of clarity to be lower when the warnings are presented stereophonically (i.e., streamed). Moreover, in line with Edworthy and Stanton (1995) we would expect that the rate of presentation should increase ratings of perceived urgency.

Method

Participants. Twenty volunteers were recruited among Cardiff University students. The age range of participants was 20–35 years old. All reported normal hearing and normal (or corrected-to-normal) vision. None had participated in the previous experiment.

Materials. Three verbal warnings from Experiment 1—the three that had achieved the highest consensus as to when streaming occurred—were used in the present experiment. These were attention, engine fail, and fire warning. The presentation rate of each warning was also decreased by 10% (in relation to the subjective reports from Experiment 1). This procedure created two versions of each warning: above threshold and below threshold. In addition, all versions of each warning were manipulated digitally to create stereophonic (i.e., streamed) and monophonic (i.e., unstreamed) versions. All warnings were presented through stereophonic headphones at a mean sound level of 65 dB (A). Finally, a response booklet was prepared containing two 9-cm lines as indices of perceived urgency and clarity for all 12 warning presentations. The scale was labeled *low* on the left side and *high* on the right.

Design. A 3 (warning: attention, engine fail, and fire warning) × 2 (presentation: streamed and unstreamed) × 2 (speed: below threshold and above threshold) within-subjects factorial design was used. Rating scores for clarity and perceived urgency were derived from measuring from the left-hand point of the 9-cm line to the point where participants put a cross. A score out of nine was calculated whereby the higher the score, the higher the rating. The order of presentation of the warnings was counterbalanced between subjects.

Procedure. Participants were tested individually in a soundproof laboratory. All of the 12 warnings were, in turn, presented twice to participants. After the second audition, participants were asked to rate each warning's urgency and clarity by making a cross on each line corresponding to the two indices.

Results

Clarity. The group mean ratings for clarity are presented in Table 2. A 3 (warning: attention, engine fail, and fire warning) × 2 (presentation: streamed and unstreamed) × 2 (speed: below threshold and above threshold) within-subjects analysis of variance (ANOVA) showed a significant effect of warning and presentation and a significant Warning × Presentation × Speed interaction. This analysis is reported in Table 3. All significant statistical differences equate to a large effect size as defined by Cohen (1988).

Perceived urgency. The group mean ratings for perceived urgency are presented in Table 2. A 3 (warning: attention, engine fail, and fire warning) × 2 (presentation: streamed and unstreamed) × 2 (speed: below threshold and above threshold) within-subjects ANOVA showed a significant effect of warning, presentation, and speed. This analysis is reported in Table 3. No significant interactions were found. Once again, all significant statistical differences equate to a large effect size as defined by Cohen (1988).

Table 2
Clarity and Perceived Urgency Ratings for Attention, Engine Fail, and Fire Warning (Experiment 2)

	Below threshold[a]				Above threshold[a]			
	Streamed		Unstreamed		Streamed		Unstreamed	
Warning	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Clarity								
Attention	3.98	1.71	6.40	1.66	4.40	2.06	5.23	2.17
Engine fail	6.75	1.74	6.64	2.06	5.27	2.09	6.28	1.56
Fire warning	4.38	2.18	6.38	1.86	4.69	2.08	6.62	1.76
Perceived urgency								
Attention	5.68	1.94	6.05	1.56	5.81	1.80	6.76	1.08
Engine fail	5.16	2.16	5.28	2.24	6.15	2.15	6.66	1.93
Fire warning	4.32	1.59	4.40	1.88	4.79	1.95	5.97	1.48

Note. Ratings could range from 0 to 9.
[a] Refers to rate of presentation.

Table 3
Analysis of Variance for Clarity and Perceived Urgency Ratings (Experiment 2)

Source	Clarity				Perceived urgency			
	df	*MSE*	*F*	Cohen's *f*	*df*	*MSE*	*F*	Cohen's *f*
Warning	2	30.66	5.45**	0.53	2	32.08	9.47**	0.70
Error (warning)	38	5.63			38	3.40		
Presentation	1	108.81	28.22**	1.22	1	17.39	5.97*	0.56
Error (presentation)	19	3.86			19	2.91		
Speed	1	6.94	2.95	0.39	1	45.94	39.91**	1.46
Error (speed)	19	2.35			19	1.15		
Warning × Presentation	2	12.58	6.29**	0.58	2	0.75	0.62	0.18
Error (Warning × Presentation)	38	1.99			38	1.22		
Warning × Speed	2	7.19	4.21*	0.47	2	3.24	2.87	0.39
Error (Warning × Speed)	38	1.71			38	1.13		
Presentation × Speed	1	0.50	0.21	0.10	1	7.14	2.33	0.35
Error (Presentation × Speed)	19	2.45			19	3.06		
Warning × Presentation × Speed	2	9.23	6.38**	0.58	2	0.68	0.53	0.18
Error (Warning × Presentation × Speed)	38	1.45			38	1.29		

$^{*}p < .05$. $^{**}p < .01$.

Discussion

Overall, the results from the clarity and perceived urgency ratings are as expected. First, the findings of Warren and his colleagues (1969) are supported insofar as clarity ratings were generally poorer when the auditory warnings were presented stereophonically (i.e., streamed). Warren et al. would argue that this is because temporal order is poorly perceived when it involves determining the order of items in different streams. In other words, the loss of order due to streaming prevented participants from reassembling the component words or phonemes. Second, the findings of Edworthy and Stanton (1995) are supported insofar as perceived urgency ratings were higher for the higher rate of presentation. However, these ratings were significantly worse when the warnings were streamed.

A limitation of the present experiment is that there is a potential confound between the warnings in terms of both their meaning (i.e., perceived urgency) and the rates of presentation needed to induce streaming (i.e., clarity). However, the impact of this potential confound on Experiment 3 is minimal given the considerable body of research that shows marginal or no effect at all of the manipulation of meaning on disruption of serial memory (for a review, see Banbury, Macken, Tremblay, & Jones, 2001).

Experiment 3

The experiment examined the effects of extraneous sound on performance on serial memory for longitude and latitude information. Specifically, the aim of the experiment was to test the assumption that stereophonically presented warnings (i.e., streamed) are less disruptive to serial recall than are monophonically presented ones (i.e., unstreamed). In addition, the rate of presentation was also manipulated to ascertain whether faster rates of presentation (i.e., above threshold) create more salient streaming effects and, as a result, produce less disruption. Conversely, the slower presentation rates below the threshold of streaming were expected to show more disruption. For the monophonic (i.e., unstreamed) recordings, it was expected that differences in the rate of presentation would not produce differences in disruption.

A total of five conditions were tested: below threshold–streamed, above threshold–streamed, below threshold–unstreamed, above threshold–unstreamed, and finally a quiet control condition. In line with Jones and Macken (1995), it was predicted that only the streamed warning (i.e., above threshold–streamed) would produce significantly less disruption to the serial memory task compared with the other warning treatments.

Participants from Experiment 2 were used in this experiment so that the same participants undertook both (a) the ratings of clarity and perceived urgency of the warnings and (b) the serial memory task in the presence of the warnings. In line with Cowan's (1995) account of the ISE, it can be argued that this procedure could lead to a potential confound in which prior exposure to the irrelevant material could lead to a reduction in the level of disruption. If this is the case, it is possible that any reduction in disruption can be explained by habituation effects. However, the effect should be consistent across all warning conditions (and not differential as we are predicting). In addition, studies of habituation within the irrelevant sound paradigm suggest that this is unlikely. For example, Banbury and Berry (1997) observed that habituation was attenuated after relatively short periods of quiet, and Tremblay and Jones (1998), using a different methodology, found no evidence of habituation.

Method

Participants. Participants from Experiment 2 took part in this study. None had participated in Experiment 1.

Materials. Thirty-two navigational messages were presented using a personal computer with sound card and external speakers. Sounds were presented at a mean sound level of 65 dB (A). Each navigation message consisted of the spoken word *longitude,* followed by 5 digits, and the spoken word *latitude,* followed again by 5 digits. The 10 digits consisted of all integers between zero and nine, with no repeats, and were presented at the rate of one each second. The background sound comprised the 12 verbal warnings from Experiment 2. All sounds were presented through stereophonic headphones at a mean sound level of 65 dB (A). Participant responses were recorded in a response booklet.

Design. A 3 (warning: attention, engine fail, and fire warning) × 2 (presentation: streamed and unstreamed) × 2 (speed: below threshold and above threshold) × 10 (serial position: 10 items in the list) within-subjects factorial design was used. Responses were scored in terms of a strict serial recall criterion; the correct item had to be in the correct position for it to be scored as correct.

Procedure. Participants were tested individually in a soundproof laboratory and seated directly in front of the computer monitor and speakers. Each trial consisted of three discrete phases. The encoding phase consisted of a visual cue on the screen to "Listen," followed by the auditory presentation of the words *longitude* and five digits and *latitude* and five digits. The rehearsal phase required participants to maintain the sequence of digits in the correct order by subvocal rehearsal. During this phase, a visual cue was presented on the computer screen instructing participants to "Rehearse" for 10 s. The auditory warnings were presented during this phase (and in this phase only) appropriate to condition. The recall phase required participants to recall the material, by way of a written response, in the order in which it was presented. They had 15 s to do this before the message "Next Trial" was displayed on the computer screen. Participants were instructed to ignore any background sound presented during the experiment.

Results

A 3 (warning: attention, engine fail, and fire warning) × 2 (presentation: streamed and unstreamed) × 2 (speed: below threshold and above threshold) × 10 (serial position: 10 items in the list) within-subjects ANOVA showed no significant main effect of warning and significant main effects of presentation, speed, and serial position. In addition, there was a significant Presentation × Speed interaction. This analysis is reported in Table 4. All significant statistical differences equate to a large effect size as defined by Cohen (1988). Finally, there was a significant Warning × Presentation × Serial Position interaction, equating to a medium effect size (Cohen, 1988). Analysis of the main effect of serial position showed that performance was significantly better on List Items 1, 6, and 10 ($p < .05$). This may reflect primacy and recency advantages within each of the longitude and latitude lists (i.e., that the 10-item list was remembered as two 5-item lists—longitude and latitude). Analysis of the Presentation × Speed interaction showed that performance was significantly better in the above threshold–streamed condition ($p < .05$). This interaction is presented in Figure 1. Analysis of the Warning × Presentation × Speed interaction showed that performance was significantly better for unstreamed attention compared with unstreamed engine fail and fire warning for Serial Position 8.

Finally, the warning data were collapsed and the quiet control data included so that a 5 (auditory condition: below threshold–streamed, above threshold–streamed, below threshold–unstreamed, above threshold–unstreamed, and quiet) × 10 (serial position) ANOVA could be carried out. Although collapsing across the different warnings runs a substantial risk of erroneous interpretation (because of the previous determination of significant effects

Table 4
Analysis of Variance for Serial Memory Performance (Experiment 3)

Source	*df*	*MSE*	*F*	Cohen's *f*
Warning	2	26.04	0.06	0.00
Error (warning)	38	451.48		
Presentation	1	106,666.67	68.99**	1.88
Error (presentation)	19	1,546.05		
Speed	1	31,537.50	29.95**	1.25
Error (speed)	19	1,052.85		
Serial position	9	71,565.28	20.26**	1.04
Error (serial position)	171	3,532.87		
Warning × Presentation	2	51.04	0.21	0.10
Error (Warning × Presentation)	38	239.64		
Warning × Speed	2	84.38	0.21	0.10
Error (Warning × Speed)	38	395.78		
Presentation × Speed	1	35,266.67	48.64**	1.60
Error (Presentation × Speed)	19	725.00		
Warning × Presentation × Speed	2	51.04	0.21	0.10
Error (Warning × Presentation × Speed)	38	239.64		
Warning × Serial Position	18	921.88	1.58	0.27
Error (Warning × Serial Position)	342	584.64		
Presentation × Serial Position	9	3,032.41	2.30*	0.35
Error (Presentation × Serial Position)	171	1,318.23		
Warning × Presentation × Serial Position	18	868.17	1.76*	0.31
Error (Warning × Presentation × Serial Position)	342	493.90		
Speed × Serial Position	9	1,093.06	0.62	0.18
Error (Speed × Serial Position)	171	1,763.38		
Warning × Speed × Serial Position	18	369.09	0.87	0.20
Error (Warning × Speed × Serial Position)	342	424.65		
Presentation × Speed × Serial Position	9	1,336.11	0.82	0.20
Error (Presentation × Speed × Serial Position)	171	1,633.63		
Warning × Presentation × Speed × Serial Position	18	1,044.09	1.63	0.30
Error (Warning × Presentation × Speed × Serial Position)	342	640.59		

$^{*}p < .05$. $^{**}p < .01$.

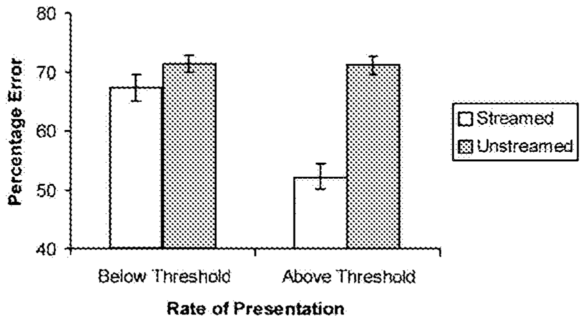

Figure 1. Overall percentage of error in serial recall when irrelevant streamed and unstreamed warnings are presented at two rates of presentation (Experiment 3). Error bars represent standard error.

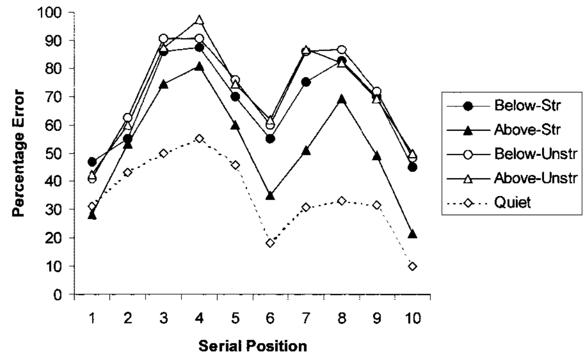

Figure 2. Serial recall performance under conditions of quiet and of below threshold–streamed (Below-Str), above threshold–streamed (Above-Str), below threshold–unstreamed (Below-Unstr), and above threshold–unstreamed (Above-Unstr) warnings (Experiment 3).

associated with the different warnings), we believe that such an analysis may be illuminating. There were significant main effects of both auditory condition and serial position, equating to a large effect size (Cohen, 1988). The interaction of the two factors was also significant, equating to a medium effect size (Cohen, 1988). This analysis is reported in Table 5. The serial position curves for each auditory condition are presented in Figure 2. Post hoc Newman–Keuls analysis of the main effect of serial position showed that recall performance was significantly better on List Items 1, 6, and 10 ($p < .05$). Analysis of the main effect of auditory condition indicated that performance on all noise treatments was significantly worse than the quiet control ($p < .05$). However, performance on the above threshold–streamed condition was also significantly better than those of the other noise treatments ($p < .05$). Post hoc Newman–Keuls analysis of the interaction showed that for Serial Position 1 there were no significant differences between the auditory conditions ($p > .1$), for Serial Positions 2 and 5 performance in all auditory conditions was significantly worse than in the quiet condition ($p < .05$), and for Serial Positions 3, 4, 6, 7, 8, 9, and 10 performance in both quiet and above threshold–streamed conditions was significantly better than the other auditory conditions ($p < .05$). Furthermore, for these positions performance in the quiet condition was significantly better than the above threshold–streamed condition ($p < .05$). For the above threshold–streamed condition, then, the disruptive effects of extraneous warnings were attenuated for 7 of the 10 serial positions.

Discussion

The results of this experiment show that extraneous speech warnings can disrupt performance on a simple serial memory task. The results also support the findings of Jones and Macken (1995) insofar as auditory streaming can attenuate the disruptive effect of the extraneous speech, but not to the extent of performance in quiet. Specifically, above threshold–streamed presentations of the warnings produced significantly less disruption than the other presentation types for the majority of serial positions, particularly the latter half of the list.

General Discussion

This series of experiments has shown that extraneous speech warnings can disrupt performance on a relatively simple serial memory task. Furthermore, the experiments have shown that capitalizing on the perceptual system's tendency to segregate the auditory scene into streams can attenuate this disruption (see Bregman, 1990). In addition, an increase in the rate of presentation of the warning increased subjective ratings of perceived urgency. However, these ratings were significantly worse when the warnings were streamed. Finally, the results have also shown that the streaming of warnings had a significant cost (Cohen's f of 0.47) insofar as reduced clarity.

Theoretical Implications

In the light of these findings, the established explanations regarding the ISE and auditory streaming can be scrutinized. The reduction of interference by streaming observed in Experiment 3 provides more evidence for the assumption that perceptual organization of incoming background sound can modulate its disrup-

Table 5
Analysis of Variance for Serial Memory Performance (Experiment 3)

Source	*df*	*MSE*	*F*	Cohen's *f*
Auditory condition	4	49,544.58	88.72**	2.13
Error (Auditory condition)	76	558.45		
Serial position	9	25,767.41	32.57**	1.31
Error (Serial position)	171	791.19		
Auditory Condition × Serial Position	36	899.19	2.67**	0.37
Error (Auditory Condition × Serial Position)	684	336.30		

** $p < .01$.

tion to short-term memory. The changing-state hypothesis (see Jones, Madden, & Miles, 1992) proposes a process-oriented account of the ISE. Disruption of serial recall is assumed to be due to a concurrence of two processes of seriation: from the deliberate rehearsal of the to-be-remembered list and from obligatory preattentive organization of the irrelevant sound sequences (Jones et al., 1996). Crucially, evidence is now emerging that seriation is a characteristic of preattentive processing of auditory sequences also, insofar as the manner in which the sound is sequentially organized by perceptual processes is analogous whether the sound is attended to or unattended (Jones, Alford, Bridges, Tremblay, & Macken, 1999). The magnitude of the ISE seems to be related to the extent of change between successive sounds, although this relation is far from straightforward. This account can therefore explain the finding that the degree of disruption can be modified considerably by changing the perceptual organization of the irrelevant sound (generically referred to as *auditory streaming*), largely independent of changes in identity.

However, theories of interference that invoke mechanisms relying on similarity of content between the irrelevant sound and the to-be-remembered sequence have difficulty in accounting for any of these effects. For example, Salamé and Baddeley's (1982) phonological similarity account argues that the disruption by irrelevant sound is related to the conflict of phonological code from the to-be-remembered material and the irrelevant sound. Moreover, the level of disruption is related to the degree of similarity between these two codes. However, the pattern of results in Experiment 3 is not consistent with this account. Specifically, the relation between the phonology of the to-be-remembered material and the irrelevant warnings was identical in both the above threshold–streamed and –unstreamed conditions, yet the degree of disruption was greater in the streamed condition. Similarly, Cowan's (1995) attentional recruitment account, which proposed that disruption by irrelevant sound is modulated by the habituation to the orienting response, is also not supported for the same reasons; given the similarity of content between the streamed and unstreamed material, the rate of habituation should be identical, leading to similar levels of disruption.

The presence of streaming effects, on the one hand, and reasonably good recognition and clarity ratings, on the other, needs some reconciliation. Recent studies by Macken, Tremblay, Culling, and Jones (2002) found a relationship between the level of disruption by irrelevant sound and the availability of order information. Specifically, when order was difficult to ascertain (because of rapid presentations or lack of transitional information), disruption was attenuated. When transitional information was provided (by connecting sounds with continuous glides), the order report was significantly better and disruption by these sounds significantly worse. Indeed, the findings of the current study support these findings insofar as the above threshold–streamed warnings showed slightly reduced levels of disruption at a cost of reduced levels of clarity. However, the fact that the warnings were not incomprehensible merits some discussion. We suggest that transitional information implicit in the warning itself assists the fusion of the three streams into a single, coherent, and comprehensible one. In other words, the utterances that make up the warning can be recombined in only one unique way: a comprehensible phrase. Indeed, there is good evidence for the role of top-down processing in the reconstruction of streamed tokens (see Bregman, 1990).

Our intention in the present study was to demonstrate that principles of auditory streaming could be used to reduce the deleterious effects of background sound where low-priority warnings are to be ignored in favor of attending to a high-priority task, without the risk of disruption. In this case, the general assumption is that the onset of streaming is independent of the locus of attention. However, the notion that auditory streaming does not require deliberate attention is controversial. Carlyon, Cusack, Foxton, and Robertson (2001) claimed that attention is required for the auditory streaming process to take place, whereas Macken, Tremblay, Houghton, and Jones (2003) argued that preattentive perceptual organization can also occur. In other words, it seems that there is a natural tendency for the brain to stream auditory information. If this is indeed the case, auditory streaming can be used to attenuate disruption by extraneous auditory warnings whether they are attended to, partly attended to, attended to at first, or ignored completely (i.e., unattended).

Practical Implications

As far as the practical implications of this study are concerned, the results demonstrate both disruption to serial memory by extraneous background noise and the potential for using the principles of auditory streaming to attenuate this disruption.

The size of the effects found in Experiment 3 is of interest to the study of noise abatement in the workplace. The disruption from extraneous speech warnings equates to a large effect size (Cohen's d of 3.14). Considering that the task of correctly remembering and recalling navigation information is safety critical, the results from the present study showed that virtually none of the participants were able to recall the navigation messages perfectly in the presence of extraneous auditory warnings. The attenuation of the disruption using auditory streaming is therefore very notable; the reduction in disruption by extraneous warnings equates to a large effect size (Cohen's d of 2.40). However, the streaming of warnings also caused a significant reduction in warning clarity. This finding is a major concern to designers of auditory warnings as the perceived clarity of speech warnings is crucial to their effectiveness. Although the ratings for the streamed warnings were significantly lower than for the unstreamed ones, these ratings were at the midpoint of scale. In other words, participants reported that the clarity of the streamed warnings was neither especially good nor especially poor. In addition, the warnings were novel to participants. Given sufficient exposure to streamed warnings, ratings of perceived clarity might improve. On balance, then, the cost of reduced clarity is outweighed by the benefit of reduced disruption, especially if only secondary warnings are streamed.

The main premise of the present study is the context in which the operator may be undertaking a high-priority task at the same time as the presentation of a lower priority warning. However, rather than ignore the warning and risk disruption by the extraneous sound, the operator could choose to interrupt the primary task and manually cancel the warning. Unfortunately, the cost to the primary task from interruption can also be severe (e.g., Gillie & Broadbent, 1989). The findings of the present study are also timely given the growing use of the auditory modality, as an alternative to the often overloaded visual modality, to present warning and status information to the operator (for an overview, see Wickens & Hollands, 2000).

From a naturalistic point of view, it is a pity that the warnings used in the present study were attention, engine fail, and fire warning. Clearly, it would be nonsensical to suggest that there are more urgent tasks that would require the operator to ignore these kinds of warnings. However, the choice of warnings was an artifact of the constraints imposed on the selection of suitable warning stimuli (i.e., the requirement to have exactly three syllables and be generic enough for nonspecialist participants to be able to understand them). Further research could make use of more realistic scenarios and warnings through the use of three-dimensional sound presentation equipment to increase the spatial locations available for streaming to above three. Finally, the novelty of the warnings selected for evaluation is also a limitation of the present study, in that the unfamiliarity of the warnings may have reduced the ratings of clarity. Further studies could either recruit participants familiar with the warning stimuli or incorporate a short training phase for nonspecialist participants.

In summary, the results of the present study have shown that extraneous auditory warnings can severely disrupt performance on a simple short-term memory task, that these warnings can be streamed by manipulating spatial location and timing, and that the use of these principles of auditory streaming can attenuate the level of disruption. However, operators and system designers may find the reduction in clarity when warnings are streamed unacceptable. Other means of reducing this disruption will have to be found. One option is through the use of context-sensitive management of speech and tone warnings so that noncritical warnings or radio communications can be suppressed or postponed during safety-critical phases of the task.

References

Baddeley, A. D. (2000). The phonological loop and the irrelevant speech effect: Some comments on Neath. *Psychonomic Bulletin & Review, 7,* 544–549.

Banbury, S., & Berry, D. C. (1997). Habituation and dishabituation to speech and office noise. *Journal of Experimental Psychology: Applied, 3,* 181–195.

Banbury, S., & Berry, D. C. (1998). The disruption of office-related tasks by speech and office noise. *British Journal of Psychology, 89,* 499–517.

Banbury, S., Macken, W. J., Tremblay, S., & Jones, D. M. (2001). Auditory distraction: Phenomena and practical implications. *Human Factors, 43,* 12–29.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences in tones. *Journal of Experimental Psychology, 89,* 244–249.

Bridges, A. M., & Jones, D. M. (1996). Word-dose in the disruption of serial recall by irrelevant speech: Phonological confusions or changing state? *Quarterly Journal of Experimental Psychology, 49,* 919–939.

Broadbent, D. E. (1982). Task combination and the selective intake of information. *Acta Psychologica, 50,* 253–290.

Buchner, A., Irmen, L., & Erdfelder, E. (1996). On the irrelevance of semantic information for the "Irrelevant Speech" effect. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49A,* 765–779.

Burt, J. L., Bartolome, D. S., Burdette, D. W., & Comstock, J. R. (1995). A psychophysiological evaluation of the perceived urgency of auditory warning signals. *Ergonomics, 38,* 2327–2340.

Caelli, T., & Porter, D. (1980). On difficulties in localizing ambulance sirens. *Human Factors, 22,* 709–724.

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 115–127.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. *Journal of Verbal Learning and Verbal Behavior, 19,* 722–735.

Cowan, N. (1995). *Attention and memory. An integrated framework.* Oxford, England: Oxford University Press.

Deatherage, B. H. (1972). Auditory and other sensory forms of information presentation. In H. P. Van Cott & R. G. Konkade (Eds.), *Human engineering guide to equipment design* (p. 124). Washington, DC: U.S. Government Printing Office.

Durso, F. T., & Gronlund, S. D. (1999). Situation awareness. In F. T. Durso, R. Nickerson, R. Schvaneveldt, S. Dumais, S. Lindsay, & M. Chi (Eds.), *The handbook of applied cognition* (pp. 283–314). New York: Wiley.

Edworthy, J. (1994). The design and implementation of non-verbal auditory warnings. *Applied Ergonomics, 25,* 202–210.

Edworthy, J., & Adams, A. (1996). *Warning design. A research perspective.* London: Taylor & Francis.

Edworthy, J., Loxley, S., & Dennis, D. (1991). Improving auditory warning design: Relationship between sound parameters and perceived urgency. *Human Factors, 33,* 205–231.

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

Ellermeier, W., & Hellbrück, J. (1998). Is level irrelevant in "irrelevant speech"? Effects of loudness, signal-to-noise ratio, and binaural masking. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1406–1414.

Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the "irrelevant speech effect." *Journal of the Acoustical Society of America, 102,* 2191–2199.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37,* 32–64.

Gillie, T., & Broadbent, D. E. (1989). What makes interruptions disruptive? A study of length, similarity and complexity. *Psychological Research, 50,* 243–250.

Gronlund, S. D., Ohrt, D. D., Dougherty, M. R. P., Perry, J. L., & Manning, A. C. (1998). Role of memory in air traffic control. *Journal of Experimental Psychology: Applied, 4,* 263–280.

Gugerty, L. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied, 3,* 42–66.

Haas, E. C., & Casali, J. G. (1995). Perceived urgency of and response time to multi-tone and frequency-modulated warning signals in broadband noise. *Ergonomics, 38,* 2313–2326.

Hellbrück, J., Kuwano, S., & Namba, S. (1996). Irrelevant background speech and human performance: Is there long-term habituation? *Journal of the Acoustical Society of Japan, 17,* 239–247.

Hellier, E. J., Edworthy, J., & Dennis, I. D. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors, 35,* 693–706.

Jones, D. M., Alford, D., Bridges, A., Tremblay, S., & Macken, W. J. (1999). Organizational factors in selective attention: The interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 464–473.

Jones, D. M., Beaman, C. P., & Macken, W. J. (1996). The object-oriented episodic record model. In S. Gathercole (Ed.), *Models of short-term memory* (pp. 209–238). London: Erlbaum.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 369–381.

Jones, D. M., & Macken, W. J. (1995). Organizational factors in the effect of irrelevant speech: The role of spatial location and timing. *Memory & Cognition, 23,* 192–200.

Jones, D. M., Macken, W. J., & Mosdell, N. (1997). The role of habituation in the disruption of recall performance by irrelevant sound. *British Journal of Psychology, 88,* 549–564.

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *The Quarterly Journal of Experimental Psychology, 44A,* 645–669.

Jones, D. M., Saint-Aubin, J., & Tremblay, S. (1999). Streaming by spatial location of irrelevant sound: Further evidence. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 52A,* 545–554.

Jones, D. M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review, 7,* 550–558.

Larsen, J. D., Baddeley, A. D., & Andrade, J. (2000). Phonological similarity and irrelevant speech. *Memory, 8,* 145–157.

LeCompte, D. C. (1996). Irrelevant speech, serial rehearsal, and temporal distinctiveness: A new approach to the irrelevant speech effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1396–1408.

Macken, W. J., Tremblay, S., Culling, J. F., & Jones, D. M. (2002). *Selective attention in memory: Processing order of irrelevant events determines distraction.* Manuscript submitted for publication.

Macken, W. J., Tremblay, S., Houghton, R. J., & Jones, D. M. (2003). Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 43–51.

Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language, 27,* 382–398.

Morris, N., & Jones, D. M. (1991). Impaired transcriptions from VDUs in noisy environments. In E. J. Lovesay (Ed.), *Contemporary ergonomics 1991* (pp. 184–189). London: Taylor & Francis.

Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychological Bulletin & Review, 7,* 403–424.

Northwood, T. D., Warnock, A. C. C., & Quirt, J. D. (1979). Noise control in buildings. In C. M. Harris (Ed.), *Handbook of noise control* (pp. 24.1–24.16). London: McGraw-Hill.

Patterson, R. (1982). *Guidelines for auditory warning systems on civil aircraft* (CAA Paper No. 82-17). London: Civil Aviation Authority.

Pollack, I., & Ficks, L. (1954). Information of elementary multidimensional auditory displays. *Journal of the Acoustical Society of America, 26,* 155–158.

Pritchett, A. R. (2001). Reviewing the role of cockpit alerting systems. *Human Factors and Aerospace Safety, 1,* 5–38.

Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21,* 150–164.

Sohn, Y. W., & Dattel, A. R. (2001). Expertise effects in situation memory and awareness. *Proceedings of the 45th Human Factors and Ergonomics Society meeting* (pp. 281–285). Minneapolis, MN: Human Factors and Ergonomics Society.

Sohn, Y. W., & Doane, S. M. (2000). Predicting individual differences in situation awareness: The role of working memory capacity and memory skill. *Proceedings of the First Human Performance, Situation Awareness and Automation Conference* (pp. 293–298). Savannah, GA: SA Technologies Inc. and Mississippi State University.

Sokolov, E. N. (1963). *Perception and the conditioned reflex.* New York: Pergamon Press.

Spiegel, M. F., & Streeter, L. (1997). Applying speech synthesis to user interfaces. In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of human–computer interaction* (2nd ed., pp. 1061–1084). Amsterdam: Elsevier Science.

Tremblay, S., & Jones, D. M. (1998). Role of habituation in the irrelevant sound effect: Evidence from the effects of token set size and rate of transition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 659–671.

Tremblay, S., Macken, W. J., & Jones, D. M. (2000). Elimination of the word length effect by irrelevant sound revisited. *Memory & Cognition, 28,* 841–846.

Tremblay, S., Nicholls, A. P., Alford, D., & Jones, D. M. (2000). The irrelevant sound effect: Does speech play a special role? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1750–1754.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. (1969). Auditory sequence: Confusion of patterns other than speech or music. *Science, 164,* 586–587.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Received May 30, 2002
Revision received November 15, 2002
Accepted November 19, 2002 ■