Journal of Experimental Psychology: Applied
2005, Vol. 11, No. 3, 157–174

1076-898X/05/$12.00 DOI: 10.1037/1076-898X.11.3.157

Assessing the Effectiveness of Various Auditory Cues in Capturing a Driver's Visual Attention

Cristy Ho and Charles Spence
University of Oxford

This study was designed to assess the potential benefits of using spatial auditory warning signals in a simulated driving task. In particular, the authors assessed the possible facilitation of responses (braking or accelerating) to potential emergency driving situations (the rapid approach of a car from the front or from behind) seen through the windshield or the rearview mirror. Across 5 experiments, the authors assessed the efficacy of nonspatial–nonpredictive (neutral), spatially nonpredictive (50% valid), and spatially predictive (80% valid) car horn sounds, as well as symbolic predictive and spatially presented symbolic predictive verbal cues (the words "front" or "back") in directing the participant's visual attention to the relevant direction. The results suggest that spatially predictive semantically meaningful auditory warning signals may provide a particularly effective means of capturing attention.

Keywords: auditory, driving, spatial attention, visual, warning signal

Recent laboratory-based research has provided extensive evidence for the existence of cross-modal links in spatial attention between different sensory modalities, such as audition and vision (e.g., see the chapters in Spence & Driver, 2004). In particular, research suggests that the efficiency of human multisensory information processing can be enhanced by presenting relevant information to different senses from approximately the same spatial location (see Driver & Spence, 2004, for a recent review; though see also Soto-Faraco, Morein-Zamir, & Kingstone, 2005). Cross-modal links in spatial attention have now been found between all possible combinations of auditory, visual, and tactile stimuli (Spence & Driver, 2004). These links in spatial attention have been demonstrated both for *endogenous* (i.e., voluntary) orienting and for *exogenous* (or stimulus-driven) orienting (see Klein & Shore, 2000, for a review). Research suggests that independent mechanisms may control these two kinds of attentional orienting (e.g., Klein & Shore, 2000; Spence & Driver, 2004).

The existence of such cross-modal links in spatial attention may have a number of consequences for interface design, an area that has seen a rapid growth of research interest over recent years (e.g., see Oviatt, 1999, 2002; Oviatt, DeAngeli, & Kuhn, 1997; Spence & Driver, 1997b). For instance, Oviatt (1999) suggested that users may be more inclined to interact multimodally with spatial applications, such as dynamic map systems, when given the option to choose between, for example, speech input, pen input, or both, rather than a strictly unimodal input. Traditional modality-specific accounts of information processing (e.g., Wickens, 1980, 1984, 1991; see also Kieras & Meyer, 1997) have typically proved inadequate in accounting for such user preferences for integrating inputs and outputs from multiple sensory modalities. According to the modality-specific view, separate attentional resources are assumed to be available for the processing of auditory, visual, and tactile information (see also Duncan, Martens, & Ward, 1997; though see Sinnett, Costa, & Soto-Faraco, in press). When the input and output are handled by different sensory modalities, coordination (switching) between the different modalities may therefore be costly (e.g., Spence, Nicholls, & Driver, 2001) and thus less efficient than in a case in which a common multisensory attentional resource account is assumed. Indeed, recent evidence has suggested that auditory and visual spatial attention do not represent separate resources but are instead linked (see Driver & Spence, 2004, for a recent review). Investigation into the design of multisensory interface systems may therefore offer the potential to enhance the efficiency of human–computer interaction in spatial application domains (e.g., Oviatt, 1999; Spence & Driver, 1997b).

Spence and Read (2003) recently investigated the consequences of cross-modal links in endogenous spatial attention between audition and vision under complex conditions. Participants in their study had to shadow (i.e., repeat aloud) triplets of two-syllable words presented from one of two loudspeakers placed either in front, or from the side, while trying to ignore an irrelevant speech stream presented from an intermediate location. At the same time, the participants had to drive around suburban and inner-city roads on a fixed-base driving simulator. The participants found it easier to shadow the words presented from the front rather than those presented from the side, even in trials without the driving task (illustrating the so-called *frontal speech advantage*, e.g., Hublet, Morais, & Bertelson, 1976, 1977). Under certain conditions, this spatial effect was found to be more pronounced (Cohen's d increased from 2.75 in the single-task condition to 4.58 in the dual-task condition) when participants were simultaneously engaged in a demanding driving task (i.e., when the perceptual load

---

Cristy Ho and Charles Spence, Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom.

This research was supported in part by scholarships from the Clarendon Fund, Oxford, United Kingdom, and from the Overseas Research Student Awards Scheme to Cristy Ho. We thank Juliet Slater for running the participants in Experiment 5.

Correspondence concerning this article should be addressed to Cristy Ho, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, United Kingdom. E-mail: cristy.ho@psy.ox.ac.uk

of the participant's task was high; Lavie, 2005; Rees & Lavie, 2001). Spence and Read's results therefore suggest that when people have to attend to multiple sources of information simultaneously, it can be advantageous to present the various sources of input to the different sensory modalities from the same, rather than from different, spatial locations (see also Spence & Driver, 1997b; though see also Soto-Faraco et al., 2005).

In their recent review of cross-modal links in exogenous spatial attention, Spence, McDonald, and Driver (2004) highlighted the extensive evidence documenting the influence of spatially nonpredictive auditory cues on visual target discrimination performance, and vice versa, from both behavioral and electrophysiological studies. Taken together, the available evidence now supports the view that performance of a variety of visual tasks can be facilitated by the presentation of an auditory cue from the same location shortly beforehand. Given that auditory cues can be used to exogenously capture a participant's visual attention in the laboratory, it seems plausible to apply such insights in multisensory interface design, for instance, in the domain of warning signals for car drivers (e.g., Spence & Driver, 1999).

Begault (1993) reported one of the few studies to have investigated the potential utility of using spatially localized auditory warning signals to attract visual attention in an applied setting. The spoken advisory warning alert "traffic, traffic" was presented over a 3-D head-up auditory warning display (typically used by flight crews) to pilots. These spatial auditory cues were presented from one of seven different possible visual locations that predicted (with 100% validity) the direction from which targets would appear in a visual search task presented through a cockpit window. Visual search times were significantly faster by a magnitude of 2,200 ms (and no less accurate) when spatial auditory cues were used, as compared with performance when nonspatial auditory cues (presented to one ear to a control group of participants via a headset) were used. It is interesting to note that the facilitation of visual search times for spatial relative to nonspatial auditory cues was much larger (2,200 ms) than that typically seen in laboratory studies of cross-modal cuing (e.g., see Spence et al., 2004) and could potentially be critical for an aircraft pilot attempting to avoid an approaching object, such as another aircraft or a flock of birds. Begault's results therefore demonstrate that the cross-modal orienting of visual spatial attention elicited by the presentation of an auditory cue can also influence visual performance in an applied setting.

Consistent with Begault's (1993) findings, Perrott, Sadralodabai, Saberi, and Strybel (1991) also demonstrated a performance enhancement for participants in a visual search task when aided by a spatially correlated auditory cue (a 10 Hz click train). This advantage was more pronounced (a facilitation of 200 ms in response latency) when the visual targets were presented in the periphery, although it was still present for targets presented from closer to fixation (where a 10–25 ms advantage was seen). Perrott et al. suggested that acoustic events might serve to orient the gaze of a person engaged in a different task. That is, they might lead to a cross-modal shift of overt (and covert) spatial attention (see also Perrott, Saberi, Brown, & Strybel, 1990). In another similar visual search study, the relative effectiveness of 2-D and 3-D auditory spatial cues and 2-D visual cues were compared with a baseline condition of unaided (i.e., uncued) visual search (Perrott, Cisneros, McKinley, & D'Angelo, 1996). Once again, auditory spatial cues were found to be particularly effective in aiding visual search performance (even more effective, in fact, than visual cues).

The studies described so far clearly demonstrate the potential benefits for interface design that may come from exploiting exogenous and/or endogenous cross-modal links in spatial attention between audition and vision (and presumably other modalities, such as touch; see Ho, Tan, & Spence, 2005). However, it is unclear whether the facilitatory effects from the use of spatially predictive auditory warning cues reported in these studies (e.g., in the studies of Begault, 1993; Perrott et al., 1990, 1991, 1996) reflect the consequences of cross-modal links in exogenous orienting, cross-modal links in endogenous orienting, or perhaps more likely, some unknown combination of the two effects (see Spence, 2001). One of the primary aims of the present study was therefore to try and quantify the benefits of the use of exogenous, endogenous, and combined exogenous plus endogenous auditory spatial cuing on visual task performance in a complex visual environment involving a simulated driving task. In particular, the experiments reported here were designed to investigate whether the location and/or informational content of auditory warning signals could be used to facilitate the detection of potential emergency situations seen in the rearview mirror and/or from the front in a driving situation.

Given that the auditory cues used in most auditory spatial cuing studies published to date (e.g., pure tones or white noise bursts; see Spence et al., 2004) are not associated with any particular semantic meaning, they may be inappropriate for use in more realistic situations in which various different (often semantically meaningful) sounds are frequently designed to convey independent messages and may simultaneously compete for an interface operator's attention (cf. Shinn-Cunningham & Ihlefeld, 2004). Therefore, in the present study, we decided to start by using the sound of a car horn in order to attract driver attention (cf. Oyer & Hardick, 1963). We used a car horn sound instead of any other arbitrary auditory cue sound because, as reported by Graham (1999), the use of auditory icons that are analogous to everyday events can be comprehended more quickly and easily than can abstract sounds (such as white noise bursts or pure tones). That is, the sound of a car horn carries an inherent meaning (understood by some participants in our study to be mirror checking) as well as a perceived urgency (see also Edworthy, Loxley, & Dennis, 1991). In this regard, one might wonder whether car horn sounds are in some sense inherently symbolic cues, a point to which we will return later.

In an attempt to simulate an attention-demanding situation, a rapid serial visual presentation (RSVP) task was chosen, consisting of a continuous stream of distractor letters with target digits embedded periodically within it. The RSVP task has been used extensively in laboratory-based dual-task attention research to maintain cognitive load on participants (e.g., Soto-Faraco & Spence, 2002) and to measure the temporal distribution of attention under conditions of task switching (e.g., Allport & Hsieh, 2001; Klein & Dick, 2002). Given that the main goal of the present study was to assess the effectiveness of various types of auditory cues to alert a driver of an event occurring at an unattended location, the location at which the RSVP stimuli was presented served as the attended location at which our participants' attention was primarily engaged. The RSVP task was used to model a continuously and uniformly highly attention-demanding situation (see Shapiro, 2001) such as when the visual load of a driver's attention is concentrated toward the front (Lansdown, 2002).

According to Graham (1999), the sound of a car horn does not necessarily have any inherent association with dangerous driving situations, which might, for example, be implicit in the sound of skidding tires. Instead, the car horn sound is typically understood by drivers to indicate the presence of another vehicle in the vicinity (cf. Edworthy & Hellier, in press, for a recent discussion of the classification of auditory signals). Our first experiment was designed to assess whether the very presence of a car horn sound might lead to a general alerting effect. The effectiveness of this nonspatial–nonpredictive auditory cue (the cue was always presented from the same location directly under the participant's seat) in redirecting visual attention served as a baseline measurement for comparisons with our subsequent experiments that used spatial auditory cues (spatial in the sense that the cues coincided with, or predicted, the direction of the target visual driving events on a certain percentage of trials). We hypothesized that the spatially neutral cue might act as a nonspatial alerting signal (see Posner, 1978; Zeigler, Graham, & Hackley, 2001) that would result in a general facilitation of response latencies in the visual driving task.

In order to examine any cross-modal facilitation effects of auditory car horn sounds on visual attention, we conducted Experiment 2 in which spatially nonpredictive car horn sounds were presented (i.e., the spatial auditory cue came from the relevant direction on 50% of trials and from the inappropriate direction on the remaining 50% of trials) to orient participants' attention to subsequent visual driving events presented from in front or behind (seen via a rearview mirror). In Experiment 3, we investigated whether the effectiveness of such auditory spatial cues would be enhanced if their occurrence was made predictive of the likely location of any subsequent critical visual events. Specifically, the direction of the car horn sound predicted the correct location of the critical visual event on 80% of the trials. We hypothesized that the magnitude of any interaction between the relative locations of the auditory and visual stimuli would be more pronounced in Experiment 3 than in Experiment 2, as the presentation of the auditory cues should elicit both an exogenous and an endogenous shift of spatial attention in the cued direction (see Spence & Driver, 1994, 1997a, 2004). Experiment 4 was designed to investigate the relative efficacy of the semantic component of symbolic spatial verbal instructions ("front" and "back") for orienting spatial attention when they were predictive of the likely critical driving event. Finally, in Experiment 5, we examined the effectiveness of redundant (or integrated) spatial cues by presenting the symbolic spatial verbal instructions used in Experiment 4 from the appropriate spatial locations (i.e., the word "front" from the front, and the word "back" from the rear).

Experiment 1

Method

Participants. Twelve participants (age: $M = 19$, $SD = 1$ year; age range: 18–20 years; 4 men and 8 women) took part in this experiment. All reported normal hearing and normal or corrected-to-normal vision. All of the participants had a valid U.K. driver's license and, on average, had been driving for 1.7 years (ranging from 4 months to 3 years). All the participants were right-handed by self-report. The experiment lasted for approximately 60 min. The participants received course credit for their participation. The experiments were conducted in accordance with the guidelines laid down by the Department of Experimental Psychology, University of Oxford.

Apparatus and materials. The experiment was conducted in a 220 cm × 142 cm experimental booth. The participants were seated in the center of the room and were instructed to hold a Logitech MOMO Racing Force Feedback Wheel (Logitech Inc., Fremont, CA) mounted on a desk situated directly in front of them. The foot pedals linked to the racing wheel were placed in a comfortable position on the floor in front of the participants. A mirror positioned directly in front of the participants (at a distance of 50 cm) was used to display the stimuli in the RSVP task (see Figure 1). A monitor showing a video image taken through a car windshield was also positioned 70 cm in front of the participants. A car rearview mirror (6 cm × 15 cm; Model RV-32, Summit Automotive, England) was attached to the upper left corner of the monitor. The participants could see the rear video shown on a second monitor placed 120 cm away via the reflection in the rearview mirror. A loudspeaker cone was positioned directly below the participant's chair, thus eliminating any exogenous cuing effect associated with the spatial location from which the cue sound itself was presented (Spence & Driver, 1994).

The RSVP distractor set consisted of 17 letters (B, C, D, E, F, J, K, L, M, N, P, R, S, T, X, Y, and Z) whereas the target set consisted of six digits (2, 3, 4, 5, 6, and 9; cf. Soto-Faraco & Spence, 2002). A computer monitor, occluded from the direct view of the participants, was used to present the stimuli in the RSVP stream. The monitor display was reflected by means of two mirrors so that participants could see the letters and digits on the mirror directly ahead of them (see Figure 1). The RSVP characters were 8.3 mm × 8.5 mm in size on the mirror as seen by the participants.

The video clips were recorded in a countryside area in rural Oxfordshire, England. The video of the windshield was filmed from behind the driver's seat, and showed a car in front being followed at a roughly constant distance at a speed of approximately 50 km per hr. The rear video was also filmed from behind the driver's seat and showed the same car following at approximately the same distance and speed. The critical clips in the video included the sudden fast approach (by the car in which the video camera was mounted) toward the car in front at approximately 100 km per hr or the sudden rapid approach from the car behind. The noncritical clips included the car in front moving away at the normal speed (i.e., 50 km/hr) or the car behind retreating (see Figure 2 for samples of stills taken from the video clips). Two critical and two noncritical clips were created. The sound of a real car horn (600 ms duration; 8,000 Hz; 66 dB[A]) downloaded from the Internet (retrieved from http://www.a1freesoundeffects.com/carhornshort.wav; downloaded on 28-11-2003) was used as the auditory cue in Experiment 1. A red light-emitting diode (LED) was placed on the mirror directly in front of the participants to provide them with feedback whenever they made an incorrect response in the driving task (see Figure 1B). An amber LED placed directly below the red LED was illuminated whenever a participant failed to keep the accelerator depressed appropriately during the experiment.

Design. The experimental session consisted of eight blocks that lasted for 6 min each. The RSVP stimuli consisted of a continuous stream of distractor letters with target digits periodically embedded within it. Sixty-six targets were presented in each block in the RSVP task with each of the six target digits being presented 11 times in total. Each item in the RSVP stream was presented for 40 ms, with a blank gap of 80 ms before the onset of the next stimulus. The presentation of the stimuli was synchronized to the next refresh of the monitor (the screen refresh rate was 75 Hz, i.e., 13.3 ms per frame). The temporal gap between successive target digits in the RSVP stream was in the range of 2,040–6,360 ms.

For the driving task, 24 randomized scenarios were presented in each experimental block. In a given experimental block, each of the two critical driving scenarios was presented 10 times, with each of the two noncritical driving scenarios being presented twice (i.e., the ratio of critical to noncritical trials was 83:17). Each scenario consisted of a 15,000 ms video clip

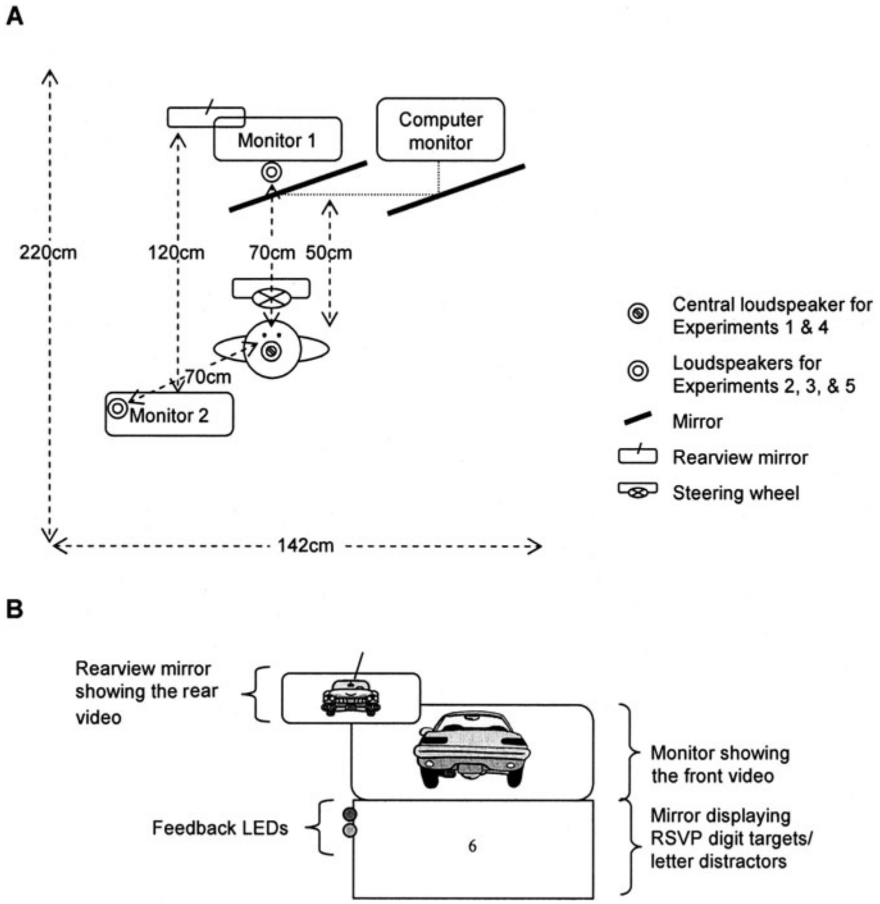

Figure 1. (A) Schematic bird's-eye view and (B) participant's-eye view of the experimental set-up used in Experiments 1–5.

together with an auditory cue. The onset of the auditory cue coincided with either the start of the critical driving event (defined as the initiation of the reduction of the intercar distance that lasted for 1,800 ms—at which point the two cars would have collided) or, in the case of a noncritical driving event, the start of an increase in intercar distance (see Figure 3). The temporal gap between successive auditory cues was 8,500–21,500 ms. Each scenario transited to the next one with fade-out and fade-in effects (using Adobe Premiere 6.0; Adobe Systems Inc., San Jose, CA), and all transitions began and ended with the same clip of the car in front and the car from behind at a constant distance. Editing was performed in order to ensure that the transitions were not noticeable and that the participants could not anticipate the critical events in the driving clips. In total over a participant's experimental session, there were 528 targets in the RSVP task and 192 driving trials (i.e., 160 critical and 32 noncritical trials).

The participants were given two practice blocks in which to familiarize themselves with the experimental set up and the operation of the response pedals. In the first practice block, the participants had only to perform the RSVP task. Initially, each visual stimulus was presented for 98 ms with a blank interval of 109 ms between successive stimuli. The rate of stimulus presentation increased gradually during the block, with the duration of visual stimulus presentation being reduced by 2 ms and the blank following the presentation of each target being reduced by 1 ms throughout the practice block, until the experimental rate of stimulus presentation was attained (though note that the actual presentation duration was limited by the refresh rate of the monitor). The stimulus timings used in the second practice block were the same as those used in the subsequent experimental blocks.

Procedure. The participants responded to targets in the RSVP task by pressing the right paddle shifter on the steering wheel. For the driving task, they were instructed to keep the accelerator depressed slightly throughout the experiment in order to model realistic driving conditions. The amber LED placed directly in front of the participants was illuminated whenever the accelerator was depressed inappropriately. When the participants detected a car approaching rapidly from behind (following the presentation of an auditory cue), they were instructed to accelerate by pressing the accelerator all the way down to the floor with their right foot. By contrast, when the participants detected that they were rapidly approaching the car in front, they were instructed to brake by pressing the brake pedal with their left foot instead. (This design was chosen to avoid the possible noise in the data that would have been introduced by moving the right foot from the accelerator to depress the brake pedal, though see van Winsum & Brouwer, 1997; van Winsum & Heino, 1996) For noncritical trials, in which the video showed the intercar distance to be increasing, the participants simply had to carry on driving normally without making any specific response.

Results

Table 1 shows the mean response time (RT) data and the percentages of correct responses in the driving task in Experiment

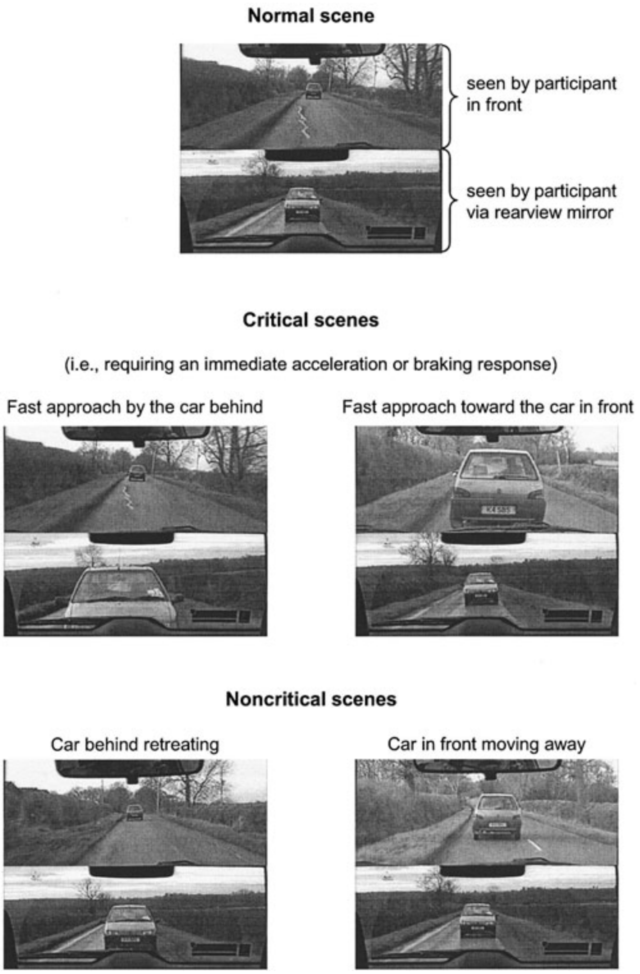

Figure 2. Sample video stills taken from the video clips used in Experiments 1–5. The upper half of each still shows the view of the windshield seen directly in front of participants, whereas the lower half shows the rear view seen by the participants in the rearview mirror. Note that in the experiment itself, the participants only saw the upper half of the video from the front and the lower half of the video from behind via the rearview mirror.

1 (together with the data from the four subsequent experiments). Trials with an incorrect response were discarded from the RT analysis; 3.8% of the trials were removed due to no response being made within 1,800 ms of the onset of the critical visual driving event. For a given critical trial in the driving task, only the first response made by a participant after the onset of the visual driving event was considered in the subsequent data analysis. Thus, an error was defined as an incorrect first response made after the onset of the critical visual driving event. Paired comparison *t* tests were performed on the RT and error data from Experiment 1 to compare responses with the front and rear critical driving events. The results revealed that participants responded significantly more rapidly to critical visual driving events occurring at the front than from the rear, $t(22) = 4.91$, $p < .01$, Cohen's $d = 1.19$. The error data, however, showed no significant difference between the front and rear, $t(22) = 0.19$, *ns*, Cohen's $d = 0.03$, suggesting that participants did not make more mistakes in response to driving events occurring in a particular direction.

Performance in the concurrent RSVP task was also measured to ensure that participants performed the central attention-demanding visual task as instructed. Responses occurring 1,500 ms or more after the presentation of the target digit were considered invalid (i.e., they were treated as false alarms). The mean percentage of correct detection responses in the RSVP task was 78.2% (*SE* = 2.6%), with a mean RT of 577 ms (*SE* = 10 ms).

Discussion

The results of Experiment 1 provide a baseline measure of the effectiveness of the presentation of the spatially uninformative car horn sound in alerting participants to orient their attention to the visual driving events. The faster responses (with a large effect size) to events occurring from the front of the participant than from the rear were as expected given the smaller distance between the location of the stimuli in the RSVP task to the front monitor than to the rearview mirror (situated further to the left of the participants) and that participants' visual attention was focused primarily toward the front.

In real-world driving, the visual attention of a driver is, as in Experiment 1, primarily focused toward the front, with events at the rear typically only monitored by means of rearview and side mirrors. Even so, there are blind spots around a vehicle that fall out of a driver's sight at any given time. The attention paid to the rear therefore depends, to a great extent, on the frequency with which a driver checks his or her mirrors. In one of the few studies to have been conducted in this area, Brookhuis, de Vries, and de Waard (1991) evaluated the effects of introducing dual-task demands (mobile telephoning) on driving performance. They used mirror checking as a performance measure to estimate how much attention participants were paying to other traffic. They found that the frequency of mirror checking depended primarily on the road situation (e.g., on a busy ring road, less attention was paid to the mirrors as compared with the situation of driving on a quiet motorway). It is interesting to note that the dual-task demands of telephoning and driving simultaneously did not seem to affect the frequency of mirror checking (i.e., of attention being directed to the rear).

Experiment 2

Given that most drivers have extensive experience of the fact that the rearview mirror represents the space behind a car, we thought that an auditory signal from the rear might facilitate the detection of a critical situation from behind albeit seen from the front via the rearview mirror. This hypothesis was based on the assumptions that (a) cross-modal links in exogenous spatial attention between audition and vision can facilitate the appropriate deployment of spatial attention, and (b) there appears to be an automatic and effortless translation between visual events seen in front (but via the rearview mirror) and the awareness of driving events to the rear. No previous study of exogenous cross-modal spatial attention has looked at the representation of the space seen indirectly in a mirror, although there is a growing interest in cognitive neuroscience in how (and whether) attention can be directed to the space seen indirectly via mirror reflection (e.g.,

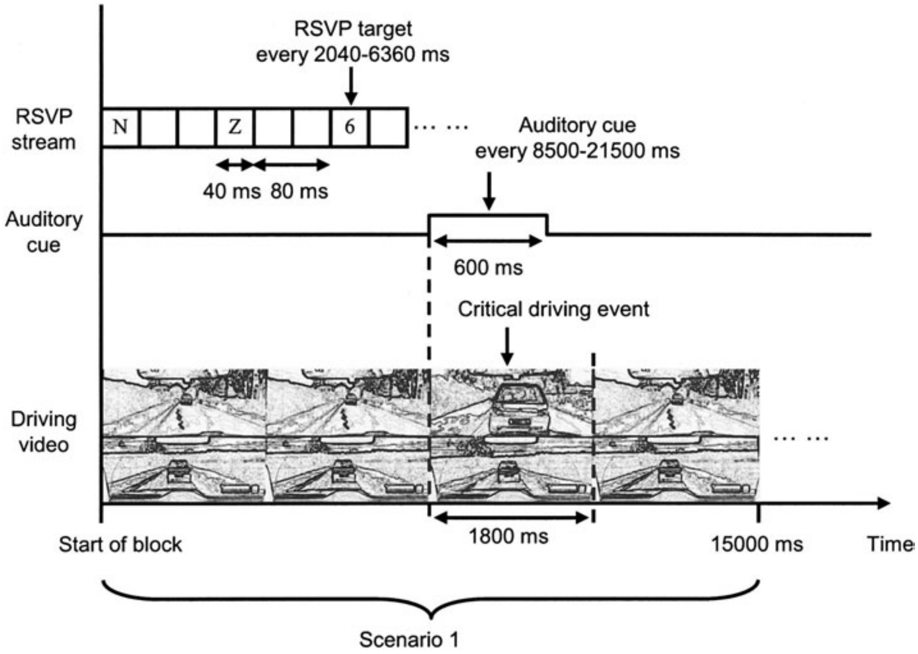

Figure 3. A schematic timeline showing the temporal sequence of events in Experiments 1–5.

Binkofski, Buccino, Dohle, Seitz, & Freund, 1999; Gregory, 1998; Maravita, Spence, Sergent, & Driver, 2002; von Fieandt, 1966). As well as being of applied interest, the results of the present study should therefore also inform our understanding of this aspect of spatial representation and cognition.

Experiment 2 was therefore designed to assess the effectiveness of spatially nonpredictive auditory cues (i.e., indicating the directions of the critical driving events on 50% of trials). We expected to find faster responses to target driving events occurring in the cued direction rather than in the uncued direction (due to the exogenous shift of cross-modal attention that we hypothesized would be elicited by the presentation of the cue) even though the cues were uninformative with regard to the likely location of the target visual events (just as was the case for the alerting nonspatial–nonpredictive cues used in Experiment 1).

Method

Twelve participants (age: *M* = 25, *SD* = 4 years; age range: 21–33 years; 6 men and 6 women) took part in this experiment. All reported normal hearing and normal or corrected-to-normal vision. All of the participants had a valid U.K. driver's license, and had, on average, been driving for at least 6 years (range: 2–15 years). All of the participants were right-handed by self-report. None of the participants had taken part in Experiment 1. They were recruited by word of mouth and did not receive any recompense for their participation.

The apparatus, materials, design, and procedure of the experiment were exactly the same as in Experiment 1, with the exception that the location from which the auditory cues were presented now varied between two possible loudspeaker cones. The two loudspeakers were placed on a virtual circle (70 cm in diameter) centered on the participants' head, one to the front and the other to their rear left-hand side (see Figure 1A). Auditory cues from the rear were presented from the left (rather than from directly behind the participant) in order to avoid any potential front-back confusions (e.g., Geissler, 1915; Stevens & Newman, 1936; cf. Kitagawa, Zampini, & Spence, 2005). The presentation of the auditory cues from the rear-left was spatially compatible in terms of the side to which participants had to turn their heads in order to fixate the rearview mirror. The auditory cues were now presented from the same direction as the critical visual driving events on half of the trials (and from the opposite direction on the

Table 1
Mean Response Times (RTs; in Milliseconds), Percentages of Correct Responses (CR), and Standard Errors for the Driving Task as a Function of the Location of the Visual Driving Event in Experiments 1–5

				Front (windshield)				Back (rearview mirror)			
Experiment	Cue sound	Cue type	Cue validity	RT	*SE*	CR	*SE*	RT	*SE*	CR	*SE*
1	Car horn	Nonspatial (alerting)	N/A	911	46	95.1	2.2	1,083	41	94.9	1.3
2	Car horn	Spatial	50%	1,115	56	94.2	1.7	1,133	24	94.1	1.5
3	Car horn	Spatial	80%	970	32	96.0	1.3	1,046	26	96.3	0.7
4	Verbal directional cues	Symbolic	80%	805	62	90.4	3.7	985	43	92.4	3.0
5	Verbal directional cues	Spatial symbolic	80%	780	32	89.3	2.8	957	29	91.5	3.3

remaining 50% of trials). In other words, the auditory car horn sound had a spatial cue validity of 50%.

Results

Table 2 shows the mean RT data and the percentages of correct responses in the driving task in Experiments 2–5. Table 3 shows the statistics for the RT and error data analyses of the driving task in Experiment 2. As in Experiment 1, trials with an incorrect response were discarded from the RT analysis; 6.5% of the trials on average were removed across participants as a result of no response being made within 1,800 ms of the onset of the critical visual driving event. An analysis of variance (ANOVA) was performed on the RT data to assess whether the spatial location from which the auditory cues were presented had any effect on our participants' ability to respond to critical visual scenes seen in either the rearview mirror or from directly ahead (corresponding to events normally seen through the windshield). The two variables in our within-participants design were auditory cue direction (front vs. back) and visual stimulus location (front vs. back).

The analysis of the RT data revealed no significant main effect of either auditory cue direction or visual stimulus location (see Table 3). The interaction between these two variables was, however, significant and had a large effect size, with participants responding more rapidly to critical visual driving events cued by an auditory cue from the appropriate (or valid) spatial direction than from the inappropriate (or invalid) direction (see Figure 4). The mean RT on valid trials (i.e., when the auditory cue and the visual stimulus were both presented from the same direction) was 1,102 ms (*SE* = 37 ms), whereas that for invalid trials (i.e., when the auditory cue was presented from the opposite direction to the visual stimulus) was 1,146 ms (*SE* = 33 ms). This spatial cuing effect had a medium effect size, Cohen's $d = 0.76$, for targets presented in the rearview mirror (i.e., for visual stimuli at the back, mean cuing effect of 70 ms), but a negligible effect size, Cohen's $d = 0.10$, for target events seen through the windshield (i.e., for visual stimuli at the front, $M = 18$ ms). Given previous claims that a driver's attention is normally directed toward the front (cf. Lansdown, 2002), spatially nonpredictive auditory cuing might be expected to produce a more pronounced benefit (i.e., a larger effect size) when used to direct attention toward the rear (i.e., toward the normally unattended, or less attended, direction) than when directing attention toward the front (where attention is likely to have been focused anyway; cf. Spence, Shore, & Klein, 2001), just as we found.

A similar analysis of the error data from Experiment 2 revealed no significant main effects of auditory cue direction or visual stimulus location, nor any significant interaction between these two variables. Note, however, that the trend in the error data was for participants to make more errors in the invalidly cued trials than in the validly cued trials, thus ruling out a speed–accuracy trade-off account of the cuing effect reported in the RT data (see Duncan, 1980; Müller & Findlay, 1987).

Performance in the concurrent RSVP task was analyzed using the same criteria for invalid responses as in Experiment 1. The mean percentage of correct detection responses in the RSVP task was 71.9% (*SE* = 3.8), with a mean RT of 594 ms (*SE* = 13 ms), which was similar to the results obtained in Experiment 1.

Discussion

The results of Experiment 2 show that participants reacted more rapidly to critical visual driving events seen in the rearview mirror when they were preceded by an auditory warning cue from the same direction (i.e., from the rear) than from the opposite direction (i.e., from the front). Our results therefore suggest that drivers tend to associate what they see in the rearview mirror with the space behind them (rather than from in front where the visual information is actually presented; cf. Higashiyama & Shimono, 2004; Higashiyama, Yokoyama, & Shimono, 2001). By contrast, our results show only marginally faster responses for critical visual driving events seen in front through the windshield when the auditory warning cue came from the front rather than from behind. It is possible that auditory warning cues might have had less effect on responses to visual driving events seen from the front (rather than from the rear), given that the visual attentional focus of drivers is typically directed to the front under normal driving conditions (Lansdown, 2002; cf. Farnè & Làdavas, 2002). This focusing of visual attention toward the front was simulated in our study by our use of the RSVP task. The results of Experiment 2 complement the previous findings of Perrott et al. (1990, 1996) in demonstrating the usefulness of spatial auditory cues in aiding visual search performance in complex visual scenes (see also

Table 2

Mean Response Times (RTs; in Milliseconds), Percentages of Correct Responses (CR), and Standard Errors for the Driving Task as a Function of the Location of the Visual Driving Event and the Direction of the Auditory Cue in Experiments 2–5

				Front (windshield)								Back (rearview mirror)							
				Front auditory cue				Rear auditory cue				Front auditory cue				Rear auditory cue			
Experiment	Cue sound	Cue type	Cue validity	RT	*SE*	CR	*SE*	RT	*SE*	CR	*SE*	RT	*SE*	CR	*SE*	RT	*SE*	CR	*SE*
2	Car horn	Spatial	50%	1,106	56	93.7	1.7	1,124	57	94.8	1.9	1,168	23	94.5	1.6	1,098	32	93.8	1.9
3	Car horn	Spatial	80%	907	30	98.2	0.7	1,033	36	93.7	1.9	1,107	30	96.3	1.2	985	25	96.3	0.5
4	Verbal directional cues	Symbolic	80%	770	60	97.3	0.8	840	67	83.5	6.9	1,025	50	89.1	4.5	945	41	95.7	1.7
5	Verbal directional cues	Spatial symbolic	80%	695	26	98.0	0.6	865	41	80.5	5.2	1,052	38	87.0	5.8	862	24	95.9	1.0

Table 3
ANOVA Results for the Response Times (RT) and Error Data Analyses of the Driving Task in Experiment 2

	RT			Error		
Source	*MSE*	*F*(1, 11)	*f*	*MSE*	*F*(1, 11)	*f*
Auditory cue direction (A)	2,519	3.12	0.51	19.7	0.02	0.04
Visual stimulus location (V)	34,076	0.11	0.10	7.8	0.00	0.02
A × V	3,977	5.93*	0.70	11.5	0.91	0.28

Note. ANOVA = analysis of variance.
* $p < .05$.

Doyle & Snowden, 2001). We were able to extend this finding to demonstrate similar effects when visual attention was directed to the rear in a simulated driving task.

In Experiment 2, we observed a spatial cuing effect of a large effect size with the use of a spatially nonpredictive car horn sound cue. At one level, the spatially nonpredictive car horn sounds used in Experiment 2 can be considered as being very similar to the nonspatial–nonpredictive sounds used in Experiment 1 because they were both uninformative with regard to the likely location of the target visual events. However, the redundant information elicited by the spatial variability in the location of the car horn sound might have required extra processing time, thus slowing down the participants' reactions (see Wallace & Fisher, 1998). This suggests that the invalid or incongruent spatial directional cue can impair performance while the correct spatial directional cue may facilitate it.

It should be noted that the driving task used in the present study was a relatively simple one. The participants only had to judge whether there would be a potential collision and react to this information by either accelerating, braking, or making no response. This may have limited the magnitude of the facilitatory effect of the spatial cues because the participants only had to make the distinction between four possible road scenarios (i.e., front-critical, rear-critical, front-noncritical, and rear-noncritical), which may have been quite easy to learn. In a real-world driving situation, by contrast, drivers may need more time to perceive and analyze the situation in order to give an appropriate response even if a spatial warning signal can direct their visual attention in a particular direction. Doyle and Snowden (2001) have suggested that spatial auditory information may be more useful in situations where the information presented visually is more demanding. Indeed, it has been argued elsewhere that cross-modal links in spatial attention may become more pronounced in general as task load increases (e.g., see Spence, Ranson, & Driver, 2000). Note also that when drivers are on the road, they typically have to monitor the information available to several different senses simultaneously (e.g., not just vision and audition, as studied here, but also proprioception and vestibular inputs; see Kemeny & Panerai, 2003), and hence the attentional load of the situation may be higher (e.g., see Lavie, 2005; Rees & Lavie, 2001).

Experiment 3

The use of spatially nonpredictive auditory cues (i.e., coming from the relevant direction on only 50% of the trials) in Experiment 2 allowed us to isolate any cross-modal attentional facilitation effects associated with purely exogenous (or stimulus-driven) orienting (see Spence, 2001; Spence & Driver, 1997a). If this spontaneous reaction to the rearview mirror can be readily acquired by drivers as the present results suggest, it might provide some ideas regarding the future design of spatial auditory warning signals in cars. Given that unreliable cues in real-life situations have been shown to adversely affect performance (e.g., see Bliss & Acton, 2003; Sorkin, 1988), our third experiment was designed to investigate whether the effectiveness of such warning signals would be enhanced by making the location of their occurrence predictive of the likely location of any critical visual driving event that might occur subsequently. The location of the auditory cue now predicted the location of the critical visual event correctly on 80% of trials (the cue being invalid on the remaining 20% of trials). Note that this experimental manipulation of cue validity is also ecologically valid given that most warning cues in realistic situations are designed to be informative (cf. Sorkin, 1988).

Method

Twelve new participants (age: $M = 22$, $SD = 5$ years; age range: 18–34 years; 4 men and 8 women) took part in this experiment. All of the participants reported normal hearing and normal or corrected-to-normal vision. The average driving experience of the participants was 3.5 years (ranging from 5 months to 15 years). Nine participants were right-handed, 2 were left-handed, and 1 was ambidextrous by self-report. None of the participants had taken part in either of the preceding experiments. Seven participants received course credit for their participation, the rest received a chocolate bar.

The apparatus, materials, design, and procedure of the experiment were exactly the same as those used in Experiment 2, with the sole exception that the auditory cues were now presented from the relevant location of the visual events on 80% of the trials. The location of the cue was invalid on the remaining 20% of trials. The instructions given to the participants emphasized the fact that the cues were spatially predictive (i.e., 80% valid) to ensure that participants formed an appropriate expectation regarding the reliability and potential informativeness of the auditory cues. The participants were instructed prior to the start of the experiment to try and direct their visual attention in the direction indicated by the auditory spatial cues.

Results

Table 4 shows the statistics for the RT and error data analyses of the driving task in Experiment 3. The results of Experiment 3 were analyzed in a similar manner to those of Experiment 2; 1.3% of the trials were removed as a result of no response being made within 1,800 ms of the onset of the critical visual driving event. A two-way within-participants ANOVA was performed on the RT data with the variables of auditory cue direction and visual stimulus location (see

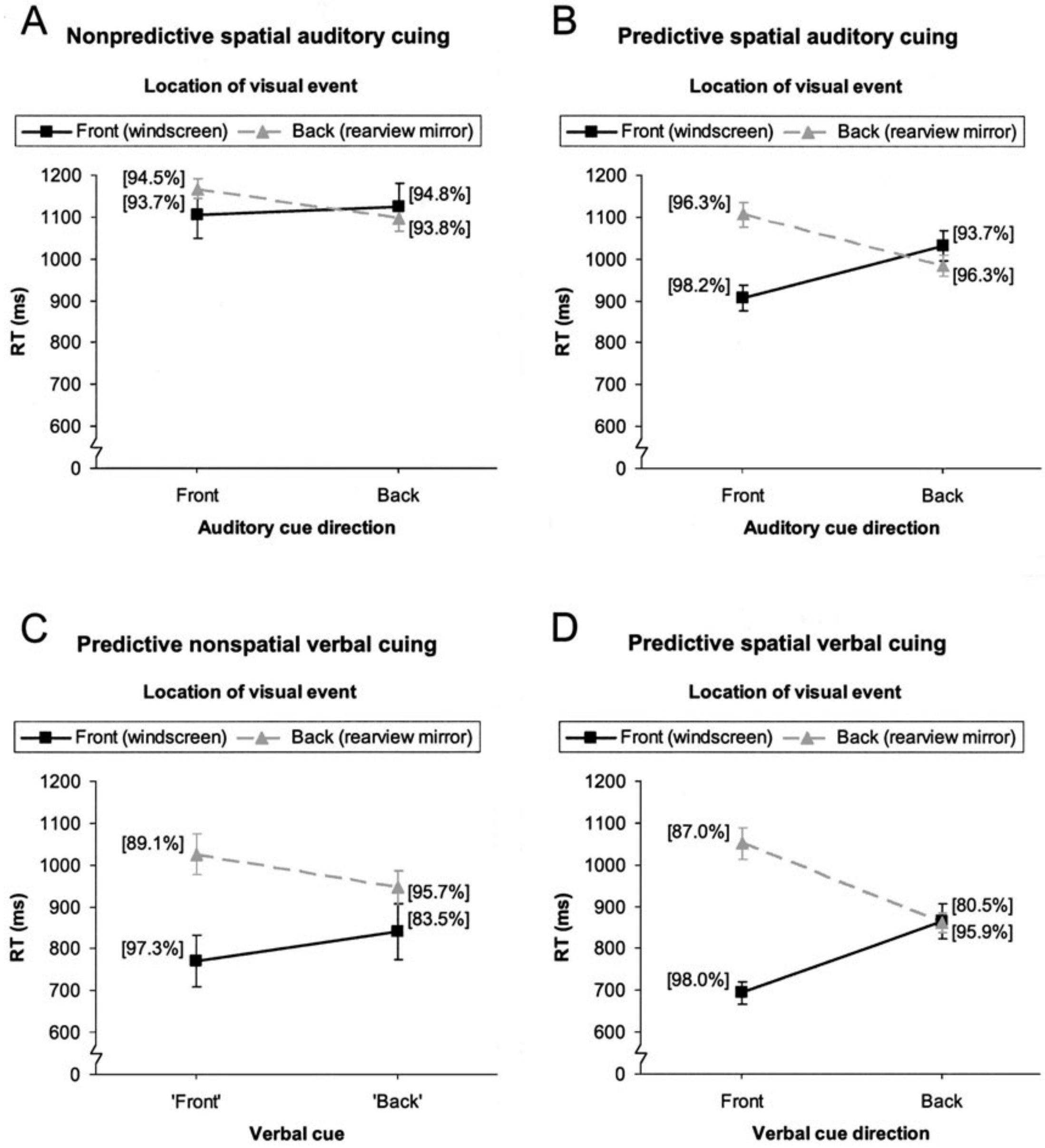

Figure 4. Summary of the interactions between the direction of the auditory cue and the location of the critical visual event in the driving task in Experiments 2 (A), 3 (B), 4 (C), and 5 (D). Percentages of correct responses are indicated by numerical values in parentheses. Error bars indicate standard errors of the RTs. The slope of the lines indicates the magnitude of the spatial cuing effects observed.

Tables 2 and 4). This analysis revealed no significant main effect of auditory cue direction. There was, however, a significant main effect of visual stimulus location with a large effect size, with participants responding more rapidly to visual driving events presented from the front than from the rear. The interaction between auditory cue direction and visual stimulus location was also significant, again revealing a large effect size, showing that participants responded more rapidly to critical visual driving events when they were validly cued by an auditory cue from the same direction than when they were invalidly cued by an auditory cue from the opposite direction to the visual driving events (see Figure 4B). The spatial cuing effects for targets presented both from the front and from the rearview mirror had a large effect size, Cohen's $d = 1.15$ and 1.33, respectively. This means that responses to visual stimuli seen either through the windshield or via the rearview mirror were significantly faster when cued by an auditory signal from a valid (as opposed to an invalid) direction. (Note, by contrast, that the spatial cuing effect only reached a medium effect size for driving events occurring at the rear in Experiment 2 but not for events occurring from the front.)

A similar analysis of the error data revealed a significant main effect of auditory cue direction with a large effect size, with participants making fewer mistakes when the car horn sounds

Table 4
ANOVA and T Test Results for the Response Times (RT) and Error Data Analyses of the Driving Task in Experiment 3

Source	RT			Error		
	MSE	*F*(1, 11)	*f*	*MSE*	*F*(1, 11)	*f*
Auditory cue direction (A)	951	0.03	0.05	7.9	7.47*	0.79
Visual stimulus location (V)	9,407	7.30*	0.78	22.2	0.07	0.08
A × V	2,532	72.25**	2.45	13.2	4.55	0.62

Note. ANOVA = analysis of variance.
* $p < .05$. ** $p < .01$.

came from the front than when they were presented from the rear. There was, however, no significant main effect of visual stimulus location, nor any significant interaction between these two variables (see Table 4).

The mean percentage of correct detection responses in the concurrent RSVP task was 78.0% (*SE* = 3.9%), with a mean RT of 595 ms (*SE* = 15 ms), which was similar to the results obtained in the two previous experiments.

Discussion

The results of Experiment 3 confirm the potential usefulness of spatially informative auditory cues as a means of capturing a driver's visual attention. They show that responses to critical driving events presented both from the front and from the rear were facilitated by valid (as opposed to invalid) spatial precuing by a car horn sound. Once again, this facilitation of response latencies cannot be accounted for in terms of a speed–accuracy trade-off in the data because participants also responded significantly more accurately on validly cued, as opposed to invalidly cued, trials overall.

In order to assess the consequences of increasing the reliability of the warning signal on driver reactions to the auditory cues, we compared the results from Experiments 2 and 3 (50% vs. 80% cue validity, respectively). As predicted, the increase in the reliability of the warning signals resulted in a larger spatial cuing effect in terms of the effect size (see Tables 3 and 4). Table 5 shows the statistics for the between-experiments analyses of the RT and error data in the driving task of Experiments 2 and 3. A three-way mixed ANOVA was performed on the RT data from the two experiments to assess the consequences of increasing the predictive validity of the auditory cue on task performance. The between-participants variable was cue validity (50% in Experiment 2 vs. 80% in Experiment 3), and the within-participants variables were auditory cue direction and visual stimulus location. This analysis revealed a significant main effect of cue validity, with a large effect size, indicating faster overall performance when the cues were spatially predictive (i.e., 80% valid with regard to the likely location of the target visual event; response latency: *M* = 1,008 ms, *SE* = 26 ms) than when the cues were spatially nonpredictive (i.e., 50% valid; *M* = 1,124 ms, *SE* = 34 ms) with respect to the location of the critical driving event (see Table 2). The interaction between the three variables was also significant, again with a large effect size, presumably reflecting the fact that the spatial cuing effect for driving events occurring at the front was significant in Experiment 3 but not in Experiment 2. There was no significant main effect of auditory cue direction or visual stimulus location. However, it should be noted that we acknowledge that there might be some possible confounds in terms of the demographic differences of our participants when drawing comparisons across experiments.

A similar analysis of the error data revealed no significant main effect of cue validity. There was, however, a significant interaction between auditory cue direction, visual stimulus location, and cue validity. Participants made more errors overall in Experiment 2 than in Experiment 3. The results seem to suggest an improvement in performance when the participants directed their endogenous spatial attention in the direction indicated by the cue (i.e., when it was informative; Experiment 3), in addition to any exogenous shifts of attention elicited by the simple presentation of the auditory cue itself from a particular spatial location (Experiments 2 and 3). These results are similar to those reported previously by Spence and Driver (1994) in a series of experiments on auditory spatial attentional orienting, in which the use of an informative auditory

Table 5
ANOVA Results for the Between-Experiments Comparisons of the Response Times (RT) and Error Data in the Driving Task in Experiments 2–3

Source	RT			Error		
	MSE	*F*(1, 22)	*f*	*MSE*	*F*(1, 22)	*f*
Auditory cue direction (A)	1,735	2.00	0.29	13.8	1.84	0.28
Visual stimulus location (V)	21,760	2.40	0.32	15.0	0.04	0.04
Cue validity (C)	43,678	7.42*	0.56	70.6	1.28	0.23
A × V × C	3,253	11.54**	0.69	12.4	4.89*	0.45

Note. ANOVA = analysis of variance.
* $p < .05$. ** $p < .01$.

precue was found to facilitate auditory target discrimination responses more than the presentation of the same cuing stimuli when they were spatially nonpredictive (see also Mondor & Amirault, 1998). Note though that both Spence and Driver and Mondor and Amirault studied covert orienting mechanisms (i.e., in the absence of eye movements), whereas the present study is concerned primarily with the overt cross-modal orienting of visual attention using auditory cues.

Experiment 4

Very few studies have attempted to vary the semantic meaning of auditory warning signals in order to see what effect, if any, this factor has on the exogenous capture of spatial attention (though see Selcon, Taylor, & McKenna, 1995; Oyer & Hardick, 1963). Given this fact, it would seem reasonable that an investigation of the effectiveness of different types of auditory warnings may enhance our understanding of the ability of operators (e.g., car drivers) to use everyday knowledge when responding to warning signals (cf. Edworthy et al., 1991; Horowitz & Dingus, 1992). In an elaborate early study of the relative effectiveness of various different auditory warning signals, Oyer and Hardick (1963) investigated the optimum acoustical dimensions of auditory alerting signals. Sounds including animal and human noises (such as elephants stampeding and baby cries), horns, sirens, bells, buzzers, pure tones, whistles, and noises of various types were assessed for their potential alerting effect by different segments of the population of the United States. The participants rated the sounds along the dimensions of pleasant–unpleasant, startling–nonstartling, and bizarre–ordinary. From their results, Oyer and Hardick were able to specify the frequency, intensity and time ranges within which the optimal warning signals should appear. Among all the alerting sounds evaluated, one of the car horn sounds was rated as the fifth most effective alerting signal, preceded by missile alarm, yelper siren, British air raid siren, and falcon horn sounds. Note though that the current approach to ergonomic alarm design puts a stronger emphasis on the informative nature of alarms (as opposed to their capacity simply to alert as was the case in the 1960s; cf. Edworthy & Hellier, in press). The spatial auditory warning signals used in the present study lie closer to the informational end on the inform–alert continuum, as the cues did not simply startle the participants but provided spatial information of the visual target events (though to a lesser extent in Experiment 2). Interface designers should examine the effectiveness of the different warning signals further when specifying the most appropriate ones for the particular applications concerned, both in terms of their ability to alert or arouse an interface operator, and also in terms of their ability to direct their attention in the appropriate internal (i.e., interface display) or external event.

Given that words convey explicit meanings in everyday language, Experiment 4 was designed to use verbal cues (the words “front” and “back”) to examine the effectiveness of verbal instructions for orienting a driver’s visual attention. Selcon et al. (1995) examined the effectiveness of auditorily and/or visually presented verbal and/or spatial missile approach warnings on participants’ ability to respond to the direction of the warnings. These warnings were presented either alone or else in combination. The verbal warnings consisted of a female voice speaking the words “left” or “right,” and the spatial warnings were a pure tone presented from the left or right. The corresponding visual warnings were the words *left* or *right* presented on a monitor, and a row of letter *X*s presented from either side of the display. Selcon and his colleagues reported a significant performance advantage (a reduction in response latencies of up to 30%) when the redundant warnings were presented concurrently from four sources rather than from just a single source. Although Selcon et al. did not compare the relative efficiency of the four different types of information directly, their results did suggest that spatial information can be conveyed either by presenting information from a relevant spatial position or by means of linguistically meaningful spatial directional cues. Although most previous studies of verbal spatial cuing have investigated the effects of the presentation of the words *left* and *right* or *top* and *bottom* on the distribution of spatial attention (e.g., Ho & Spence, in press; Hommel, Pratt, Colzato, & Godijn, 2001; Hunt & Kingstone, 2003), no study has as yet specifically examined the effects of the words *front* and *back* upon perception and/or performance. Nevertheless, recent research suggests that verbal directional cues may, in fact, act in a manner similar to that of social directional cues, such as eye-gaze direction or head orientation, that can automatically capture visual attention and lead to a rapid reflexive or exogenous shifts of attention to the side where a face is looking (e.g., Langton, Watt, & Bruce, 2000).

Method

Twelve new participants (age: $M = 27$, $SD = 6$ years; age range: 20–40 years; 7 men and 5 women) took part in this experiment. All reported normal hearing and normal or corrected-to-normal vision. The average driving experience of the participants was 6.5 years (range: 1–12 years). Eleven of the participants were right-handed and 1 was left-handed by self-report. None of them had taken part in any of the preceding experiments and all received a bar of chocolate in return for their participation.

The apparatus, materials, design, and procedure were exactly the same as in Experiment 3, with the sole exception that the frontal auditory car horn cue was replaced by the word “front” (64 dB[A]), whereas the rear car horn cue was replaced by the word “back” (62 dB[A]) as pronounced by a male voice in the Merriam–Webster Online Dictionary (retrieved from http://www.m-w.com/cgi-bin/audio.pl?front001.wav=front and http://www.m-w.com/cgi-bin/audio.pl?back0001.wav=back; downloaded on 28-02-2004). The verbal cues were always presented from the same loudspeaker cone positioned directly below the participant’s chair (as in Experiment 1). The verbal instructions given to participants again emphasized the fact that the cue would correctly predict the target direction on 80% of the trials, and so they should endeavor to direct their visual attention in the direction indicated by the cue first.

Results

Table 6 shows the statistics for the RT and error data analyses of the driving task in Experiment 4. The results of Experiment 4 were analyzed in a similar manner to those of Experiments 2 and 3; 6.9% of the trials were removed because no response was made within 1,800 ms of the onset of the critical visual driving event. A two-way within-participants ANOVA on the RT data showed no significant main effect of auditory cue direction (see Tables 2 and 6). There was, however, a significant main effect of visual stimulus location with a large effect size. Participants responded more rapidly to visual events occurring in front (i.e., requiring a braking response) than to events occurring in the rearview mirror (i.e., requiring an acceleration response). The interaction between these

Table 6

ANOVA Results for the Response Times (RT) and Error Data Analyses of the Driving Task in Experiment 4

	RT			Error		
Source	*MSE*	*F*(1, 11)	*f*	*MSE*	*F*(1, 11)	*f*
Auditory cue direction (A)	5,895	0.06	0.07	111.4	1.41	0.34
Visual stimulus location (V)	25,913	15.04**	1.12	68.4	0.70	0.24
A × V	5,004	13.47**	1.06	216.8	5.77*	0.69

Note. ANOVA = analysis of variance.
* $p < .05$. ** $p < .01$.

two factors was significant and had a large effect size, revealing that participants responded more rapidly to critical visual driving events cued by an appropriate verbal directional cue ($M = 858$ ms, $SE = 47$ ms) than by an invalid verbal directional cue ($M = 933$ ms, $SE = 51$ ms; see Figure 4). A beneficial effect of valid spatial cuing was found for both the front and the rearview mirror (Cohen's $d = 0.33$ and 0.53, respectively).

A similar analysis of the error data revealed no significant main effect of auditory cue direction or visual stimulus location. The interaction between these two factors was, however, significant and had a large effect size (see Figure 4). Given that the pattern of errors coincides with the RT effects, a speed–accuracy trade-off account of the RT differences reported above can once again be ruled out.

Participants correctly detected 68.5% ($SE = 3.8\%$) of the targets on the concurrent RSVP task, with a mean RT of 597 ms ($SE = 18$ ms). These results were similar to those obtained in the three previous experiments.

Discussion

The results of Experiment 4 suggest that participants could rapidly comprehend the meanings of the verbal warning signals and direct their attention appropriately. Previous studies of auditory verbal warning signals have shown that verbal warning signals can communicate more information than nonspeech warnings about an upcoming event to an operator, especially under situations of high cognitive load, where abstract cues may not convey any explicit meaning (e.g., Graham, 1999; Selcon et al., 1995). We were interested in examining the relative benefits of using verbal versus spatial informative cuing. With reference to Figure 4 (compare 4B and 4C), it can be seen that when cued by a verbal instruction to the likely position of the visual event, participants responded more rapidly than when nonverbal cues were used across all four combinations of auditory cue direction and visual stimulus location. This was, however, accompanied by an increase in errors. Thus, it seems that the spatial and verbal predictive cues may have affected a driver's attention in different ways. Participants were quicker but less accurate with verbal cues than nonverbal cues. Future research should further examine the directional property of verbal cues (as compared with other nonverbal well-learned cues) and interface operators' compliance with verbal instructions produced by an interface (see Simpson, McCauley, Roland, Ruth, & Williges, 1987).

Experiment 5

In our final experiment, we investigated whether combined verbal spatial warning signals (such as the word *front* presented from the front, and the word *back* presented from the back) would be even more effective in orienting a driver's attention than the spatial or verbal cues used in Experiments 3 and 4 (cf. Selcon et al., 1995).

Method

Sixteen new participants (age: $M = 21$, $SD = 3$ years; age range: 19–30 years; 6 men and 10 women) took part in this experiment. All reported normal hearing and normal or corrected-to-normal vision. The average driving experience of participants ranged from 1.5 to 13 years. Fifteen participants were right-handed, and 1 was left-handed by self-report. All of the participants were naïve as to the purpose of the experiment, and all received a chocolate bar in return for their participation. The apparatus, materials, design, and procedure were exactly the same as those used in Experiment 4, with the exception that the verbal cues "front" and "back" were always presented from the front and rear loudspeaker cones, respectively, as in Experiment 3 (rather than from the central loudspeaker as in Experiment 4).

Results

Table 7 shows the statistics for the RT and error data analyses of the driving task in Experiment 5. The results of Experiment 5 were analyzed in a similar manner to those of Experiments 2–4; 4.2% of the trials on average were removed across participants due to no response being made within 1,800 ms of the onset of the critical visual driving event. A two-way within-participants ANOVA on the RT data revealed a significant main effect of visual stimulus location with a large effect size, with participants responding more rapidly to visual driving events presented from the front than from the rear (see Tables 2 and 7). There was, however, no significant main effect of auditory cue direction. The interaction between auditory cue direction and visual stimulus location was significant and had a large effect size. The spatial cuing effects showed a large effect size for targets presented both from the front (mean cuing effect of 170 ms), and from the rear ($M = 190$ ms), thus showing the beneficial cuing effect of an appropriate verbal cue coming from the appropriate (or valid) direction, as opposed to an inappropriate verbal cue coming from an invalid direction (see Figure 4D).

A similar analysis of the error data revealed no significant main effect of auditory cue direction or visual stimulus location. The

Table 7
ANOVA Results for the Response Times (RT) and Error Data Analyses of the Driving Task in Experiment 5

Source	RT			Error		
	MSE	*F*(1, 15)	*f*	*MSE*	*F*(1, 15)	*f*
Auditory cue direction (A)	4,849	0.33	0.14	93.4	3.20	0.45
Visual stimulus location (V)	17,265	28.93**	1.34	127.7	0.63	0.20
A × V	6,601	78.59**	2.22	307.3	9.09**	0.75

Note. ANOVA = analysis of variance.
** $p < .01$.

interaction between auditory cue direction and visual stimulus location was, however, significant and had a large effect size (see Figure 4).

The mean percentage of correct detection responses in the concurrent RSVP task was 79.3% (*SE* = 2.2%), with a mean RT of 571 ms (*SE* = 14 ms). A one-way ANOVA carried out on the performance data from the RSVP task in Experiments 1–5 revealed no significant difference in performance on this concurrent primary task across all the five experiments for either the percentages of correct responses, $F(4, 59) = 2.13$, $MSE = 133.3$, *ns*, Cohen's $f = 0.18$, or the RT data, $F(4, 59) = 0.80$, $MSE = 2558$, *ns*, Cohen's $f = 0.11$. This suggests that the extent to which participants engaged in the primary task was similar across all five of the experiments reported here (see Appendix for an analysis of the RSVP performance as a function of the four critical conditions in the driving task).

Discussion

The results of Experiment 5 demonstrate the potential effectiveness of using informative spatial verbal cues to direct a driver's visual attention. The use of redundant spatial cuing in Experiment 5 led to the fastest overall performance ($M = 868$ ms, $SE = 27$ ms) when compared with the nonspatial-verbal cues used in Experiment 4 ($M = 895$ ms, $SE = 48$ ms) or the spatial-nonverbal car horn sound cues used in Experiment 3 ($M = 1{,}008$ ms, $SE = 26$ ms).

General Discussion

Taken together, the results of the five experiments reported here help to elucidate some of the potential implications of the existence of cross-modal links in spatial attention between audition and vision for interface design. Our results show that a spatially nonpredictive auditory cue coming from an appropriate (or valid) spatial direction can be used to facilitate the overt orientation of visual attention in that direction. The results of Experiments 2–5 show a significant performance advantage (all having a large effect size) in the detection of potential emergency visual driving events (requiring either an acceleration or braking response) when participants were cued by an auditory cue coming from the relevant direction, or else when their attention was directed in the appropriate direction by means of a semantically meaningful verbal cue. Such an effect of a valid auditory spatial cue was enhanced by making the cues spatially predictive (as shown by the comparison of the data from Experiments 2 and 3). These results are consistent with Begault's (1993) previous findings that spatial auditory cues can be used to facilitate the visual search performance of pilots. In our study, spatial auditory warning signals were shown to facilitate visual information detection (and subsequent reactions) in a simulated driving set-up.

Previous laboratory-based studies have demonstrated that the covert orienting of visual and auditory attention influence one another at both the endogenous and exogenous levels (see Driver & Spence, 2004; Spence et al., 2004, for reviews). According to conventional thinking, the spatial cuing effects reported in Experiment 2 would be regarded as purely exogenous (given the spatially nonpredictive nature of the cues), whereas the cuing effects reported in Experiments 3 and 5 would be considered to reflect a combination of both exogenous and endogenous orienting (given that the location of the cue predicted the likely target location on 80% of trials). By contrast, the cuing effects reported in Experiment 4 would be considered as purely endogenous (because the location of the symbolic verbal cue carried no information in-and-of-itself). Cross-modal links in spatial attention between audition and vision have been shown in the present study to draw a driver's visual attention more readily to the validly cued rather than to the invalidly cued spatial direction by the auditory signals, thus demonstrating an overt orienting of visual attention by audition. Overall, our results suggest that the combined use of exogenous and endogenous orienting provides the most effective means of capturing a driver's attention.

On the theoretical side, the successful demonstration of cross-modal links in spatial attention between audition and vision, and the varying effects of the various auditory cues on performance on the same visual task, together argue against traditional modality-specific accounts of independent channels for auditory and visual information processing (see Wickens, 1980, 1991). Our results instead support recent claims that the mechanisms controlling spatial attention in different sensory modalities can interact and thereby influence one another (e.g., Spence & Driver, 2004). Given that we measured the facilitation of responses at the behavioral level, it will be important in future research to investigate whether the performance improvement reported in the present study reflects the consequences of a perceptual enhancement attributable to the spatial aspect of the cues coinciding with that of the targets, a priming of the appropriate responses by the cues (which can be generated by cues regardless of their location relative to the critical driving events), or some unknown combination of the two effects (see Proctor, Tan, Vu, Gray, & Spence, 2005). No matter what the underlying cause(s) turn(s) out to be, we believe that it is important for the applied domain to combine the perceptual and response compatibility aspects of cuing in the design of the most effective multisensory interfaces.

Meaningful Auditory Warnings

Previous studies of cross-modal links in exogenous spatial attention have typically used cues of questionable (or no) ecological validity, such as pure tones or white noise bursts in the case of auditory cuing experiments (see Spence et al., 2004). It is therefore possible that the use of a semantically meaningful sound, such as the car horn used in the present study, might engage attention more effectively than these other arbitrary sounds (cf. Oyer & Hardick, 1963). It remains an interesting question for future research to determine whether the use of a car horn or other semantically meaningful spatial auditory cue is more effective at capturing an interface operator's attention than, say, a white noise burst or pure tone of equivalent intensity and localizability. Note that the semantic meaning conveyed by the car horn sound used in our study might also have activated other attentional mechanisms than those typically thought to control spatial attentional orienting (cf. Langton et al., 2000). For instance, research has shown that the physiological state of arousal of drivers can sometimes determine their driving behaviors, such as modifying their responses to critical emergency driving situations (Collet, Petit, Priez, & Dittmar, 2005).

Gregory (1998), in his extended analysis of mirrors and perception, mentioned the visual ambiguity in certain mirrored images. In particular, he hints that in an emergency situation, drivers may need to spend time processing what they perceive in the rearview or side mirrors (Gregory, 1998, pp. 208–209). In fact, drivers may not encounter an emergency situation for years, but when they do, the warning signal needs to alert the driver and convey sufficient information for him or her to make the necessary reaction in a cognitively demanding and short time frame (cf. Belz, Robinson, & Casali, 1999). Note that the cuing to the rear in our study seems to be automatic and quick. This is perhaps the first empirical study to have investigated the cuing of attention into the space seen only indirectly via a mirror reflection, an area of growing interest in cognitive neuroscience (e.g., Binkofski, Buccino, Dohle, Seitz, & Freund, 1999; Gregory, 1998; Maravita et al., 2002).

It should be pointed out that it is somewhat unclear whether a spatial auditory cue presented from behind a driver draws his or her attention to the rearview mirror, which is in fact a visual event in the front, or to the actual space behind him or her (i.e., to the rear). If the latter account is correct, then the spatial cues are effective in shifting attention to the relevant locations of the visual elements that represent a certain well-learned spatial location. The learning of the special property of a mirror and how people accommodate that spatial information with other spatial elements in their representation of the environment is a topic deserving of further investigation (e.g., Gregory, 1998; Ramachandran, Altschuler, & Hillyer, 1997).

Reliability of Warning Signals

The issue of warning signal reliability has been investigated in both the aviation and driving literature (e.g., Bliss, 2003; Parasuraman, Hancock, & Olofinboba, 1997). One key point to have emerged from this research is that false alarms can potentially be very distracting to an operator in diverting his or her attention from their primary task. In addition, the evidence suggests that an operator may choose to ignore a warning signal if it proves to be inconsistent and/or unreliable, thus potentially negating its value (e.g., Parasuraman & Riley, 1997; Sorkin, 1988). In a recent study, Bliss and Acton (2003) examined the effect of varying the reliability of collision alarms on performance in a driving simulator. Participants were instructed to drive and to swerve to the left or right depending on the side from which cars approached randomly in the rearview mirror. Bliss and Acton reported that performance improved (i.e., in terms of the frequency of appropriate reactions to the alarm signals) as the reliability of the cue increased (from 50%, through 75%, to 100% valid). Consistent with Bliss and Acton's findings, the present study revealed that responses were both faster and more accurate when the reliability of the auditory cues increased from 50% valid (Experiment 2) to 80% valid (Experiment 3).

Auditory Signals in Driving: Speech Versus Nonspeech

Responses to speech warnings in an emergency situation may also be slower than they would be under normal circumstances, particularly if an operator does not understand the meaning of a speech message until it is finished (e.g., Simpson & Marchionda-Frost, 1984). Given that Experiments 4 and 5 used only two single-syllable verbal cues ("front" and "back"), our results were probably not influenced by any potential problems associated with disambiguating multiple possible speech warnings that may be present in applied settings incorporating multiple verbal warning signals (cf. Chan, Merrifield, & Spence, 2005).

Given that drivers are typically engaged in other tasks involving linguistic elements, such as, for example, listening to the radio, perhaps having a conversation with a passenger or over the mobile phone, or else possibly taking in speech instructions from an in-car navigation system (e.g., Green, 2000; Kames, 1978; Streeter, Vitello, & Wonsiewicz, 1985), the additional verbal cues from the warning systems may inevitably lead to some confusion, misunderstanding, and/or masking (either energetic or informational; see Oh & Lutfi, 1999). Thus, although drivers can easily acquire and react to a verbal instruction presented in isolation (as in the present study), some caution may be needed before implementing verbal cues in a real warning system (cf. Ho & Spence, in press). In a real-world driving situation, it might actually be better to implement nonverbal warning signals, as they are less susceptible to the influence of other concurrent linguistic elements in the environment. It will be particularly interesting in future research to see whether responses to verbal warning signals are particularly adversely affected by concurrent speech signals or speech tasks (cf. Ho & Spence, in press; Strayer & Johnston, 2001).

Some recent research has been designed to investigate the most effective audible icons that can convey the "right degree of urgency" and have a "commonly understood meaning" for in-car systems (e.g., Catchpole, McKeown, & Withington, 1999a, 1999b; Isherwood, McKeown, & Hockey, 2004). Indeed, there is a long history of research on the idea of ecologically valid auditory icons in interface design (see Catchpole, McKeown, & Withington, 2004; Gaver, 1993a, 1993b; Oyer & Hardick, 1963), though that has not been widely implemented in in-car systems to date. Spatial cues may therefore be better than verbal cues or arbitrary nonspatial auditory icons in that they use an orienting response that should be common to everyone, whereas there may be distinct cultural interpretations of the inherent meanings of different auditory icons. A successful warning signal needs to be representational in the sense that any likely operator can intuitively recognize its meaning and be informed of the required actions.

Validity of Laboratory-Based Studies of Driving

One potential criticism of the present study is that the laboratory set up may not be comparable to a realistic driving environment. In particular, the stimulation was restricted to the visual and auditory stimuli provided in the experiments. Therefore, in order to validate the effects reported in the present study in a more realistic driving situation, future research should investigate whether the same spatial attention effects can be demonstrated in on-road testing or with the use of a driving simulator (see McLane & Wierwille, 1975). It has been argued that, on the whole, driving simulators surpass on-road testing and driving games as being more effective, more ethical, and ultimately providing a safer means of testing (Haigney & Westerman, 2001; cf. Kemeny & Panerai, 2003; Reed & Green, 1999). Note that the large effect sizes associated with all of the significant effects reported in the present study provide a strong basis for further investigation in a more applied setting (i.e., perhaps involving a simulator-based driving task rather than the RSVP task chosen for the present study).

All-in-all, designers need to create the most ecological warning signals that can elicit an intuitive response from drivers to the appropriate spatial location where a critical emergency event may occur. Our investigation into cross-modal interactions in audiovisual spatial attention is suggestive of the generic spatial property in the human information processing system. We have demonstrated that drivers' association of an auditory cue from behind and a visual event seen only indirectly via the rearview mirror is automatic and well-learned. The design of in-car warning signals that carry spatial information is, we believe, a promising direction for future research.

References

Allport, A., & Hsieh, S. (2001). Task-switching: Using RSVP methods to study an experimenter-cued shift of set. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints in human information processing* (pp. 36–64). Oxford, U.K.: Oxford University Press.

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

Belz, S. M., Robinson, G. S., & Casali, J. G. (1999). A new class of auditory warning signals for complex systems: Auditory icons. *Human Factors, 41,* 608–618.

Binkofski, F., Buccino, G., Dohle, C., Seitz, R. J., & Freund, J. (1999). Mirror agnosia and mirror ataxia constitute different parietal lobe disorders. *Annals of Neurology, 46,* 51–61.

Bliss, J. P. (2003). An investigation of alarm related incidents in aviation. *International Journal of Aviation Psychology, 13,* 249–268.

Bliss, J. P., & Acton, S. A. (2003). Alarm mistrust in automobiles: How collision alarm reliability affects driving. *Applied Ergonomics, 34,* 499–509.

Brookhuis, K. A., de Vries, G., & de Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention, 23,* 309–316.

Catchpole, K. R., McKeown, J. D., & Withington, D. J. (1999a). Alerting, informing and localisable auditory warnings. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics, Vol. 4: Job design, product design and human-computer interaction* (pp. 447–454). Aldershot, England: Ashgate.

Catchpole, K. R., McKeown, J. D., & Withington, D. J. (1999b). Localisable auditory warnings: Integral 'where' and 'what' components. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics, Vol. 4: Job design, product design and human-computer interaction* (pp. 439–446). Aldershot, England: Ashgate.

Catchpole, K. R., McKeown, J. D., & Withington, D. J. (2004). Localizable auditory warning pulses. *Ergonomics, 47,* 748–771.

Chan, J. S., Merrifield, K., & Spence, C. (2005). Auditory spatial attention assessed in a flanker interference task. *Acta Acustica, 91,* 554–563.

Collet, C., Petit, C., Priez, A., & Dittmar, A. (2005). Stroop color–word test, arousal, electrodermal activity, and performance in a critical driving situation. *Biological Psychology, 69,* 195–203.

Doyle, M. C., & Snowden, R. J. (2001). Identification of visual stimuli is improved by accompanying auditory stimuli: The role of eye movements and sound location. *Perception, 30,* 795–810.

Driver, J., & Spence, C. (2004). Crossmodal spatial attention: Evidence from human performance. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 179–220). Oxford, U.K.: Oxford University Press.

Duncan, J. (1980). The demonstration of capacity limitation. *Cognitive Psychology, 12,* 75–96.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature, 387,* 808–810.

Edworthy, J., & Hellier, E. (in press). Complex nonverbal auditory signals and speech warnings. In M. Wogalter (Ed.), *Handbook of warnings.* Mahwah, NJ: Erlbaum.

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

Farnè, A., & Làdavas, E. (2002). Auditory peripersonal space in humans. *Journal of Cognitive Neuroscience, 14,* 1030–1043.

Gaver, W. W. (1993a). How do we hear in the world?: Explorations in ecological acoustics. *Ecological Psychology, 5,* 285–313.

Gaver, W. W. (1993b). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Psychology, 5,* 1–29.

Geissler, L. R. (1915). Sound localization under determined expectation. *American Journal of Psychology, 26,* 268–285.

Graham, R. (1999). Use of auditory icons as emergency warnings: Evaluation within a vehicle collision avoidance application. *Ergonomics, 42,* 1233–1248.

Green, P. (2000). Crashes induced by driver information systems and what can be done to reduce them. *SEA Convergence 2000 Conference Proceedings,* 27–36.

Gregory, R. (1998). *Mirrors in mind.* London: Penguin Books.

Haigney, D., & Westerman, S. J. (2001). Mobile (cellular) phone use and driving: A critical review of research methodology. *Ergonomics, 44,* 132–143.

Higashiyama, A., & Shimono, K. (2004). Mirror vision: Perceived size and perceived distance of virtual images. *Perception & Psychophysics, 66,* 679–691.

Higashiyama, A., Yokoyama, Y., & Shimono, K. (2001). Perceived distance of targets in convex mirrors. *Japanese Psychological Research, 43,* 13–24.

Ho, C., & Spence, C. (in press). Verbal interface design: Do verbal directional cues automatically orient visual spatial attention? *Computers in Human Behavior.*

Ho, C., Tan, H. Z., & Spence, C. (2005). Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*, doi:10.1016/j.trf.2005.05.002.

Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science, 12,* 360–365.

Horowitz, A. D., & Dingus, T. A. (1992). Warning signal design: A key human factors issue in an in-vehicle front-to-rear-end collision warning system. *Proceedings of the Human Factors and Ergonomics Society 36th Annual Meeting,* 1011–1013.

Hublet, C., Morais, J., & Bertelson, P. (1976). Spatial constraints on focused attention: Beyond the right-side advantage. *Perception, 5,* 3–8.

Hublet, C., Morais, J., & Bertelson, P. (1977). Spatial effects in speech perception in the absence of spatial competition. *Perception, 6,* 461–466.

Hunt, A. R., & Kingstone, A. (2003). Covert and overt voluntary attention: Linked or independent? *Cognitive Brain Research, 18,* 102–105.

Isherwood, S., McKeown, J. D., & Hockey, G. R. J. (2004). *A positive correlation between rated unpleasantness and perceived urgency of auditory icons.* Unpublished manuscript.

Kames, A. J. (1978). A study of the effects of mobile telephone use and control unit design on driving performance. *IEEE Transactions on Vehicular Technology, VT-24,* 282–287.

Kemeny, A., & Panerai, F. (2003). Evaluating perception in driving simulation experiments. *Trends in Cognitive Sciences, 7,* 31–37.

Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction, 12,* 391–438.

Kitagawa, N., Zampini, M., & Spence, C. (2005). Audiotactile interactions in near and far space. *Experimental Brain Research,* doi:10.1007/s00221-005-2393-8.

Klein, R. M., & Dick, B. (2002). Temporal dynamics of reflexive attention shifts: A dual-stream rapid serial visual presentation exploration. *Psychological Science, 13,* 176–179.

Klein, R. M., & Shore, D. I. (2000). Relations among modes of visual orienting. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 195–208). Cambridge, MA: MIT Press.

Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences, 4,* 50–59.

Lansdown, T. C. (2002). Individual differences during driver secondary task performance: Verbal protocol and visual allocation findings. *Accident Analysis and Prevention, 34,* 655–662.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences, 9,* 75–82.

Müller, H. J., & Findlay, J. M. (1987). Sensitivity and criterion effects in the spatial cuing of visual attention. *Perception & Psychophysics, 42,* 383–399.

Maravita, A., Spence, C., Sergent, C., & Driver, J. (2002). Seeing your own touched hands in a mirror modulates cross-modal interactions. *Psychological Science, 13,* 350–355.

McLane, R. C., & Wierwille, W. W. (1975). The influence of motion and audio cues on driver performance in an automobile simulator. *Human Factors, 17,* 488–501.

Mondor, T. A., & Amirault, K. J. (1998). Effect of same- and different-modality spatial cues on auditory and visual target identification. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 745–755.

Oh, E. L., & Lutfi, R. A. (1999). Informational masking by everyday sounds. *Journal of the Acoustical Society of America, 106,* 3521–3528.

Oviatt, S. (1999). Ten myths of multimodal interaction. *Communications of the ACM, 42,* 74–81.

Oviatt, S. (2002). Multimodal interfaces. In J. A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 286–304). Mahwah, NJ: Erlbaum.

Oviatt, S., DeAngeli, A., & Kuhn, K. (1997). Integration and synchronization of input modes during multimodal human-computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems: CHI '97* (pp. 415–422). New York: ACM Press.

Oyer, J., & Hardick, J. (1963). *Response of population to optimum warning signal* (Rep. No. SHSLR163). Washington, D.C.: Office of Civil Defence.

Parasuraman, R., Hancock, P. A., & Olofinboba, O. (1997). Alarm effectiveness in driver-centred collision-warning systems. *Ergonomics, 40,* 390–399.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39,* 230–253.

Perrott, D. R., Cisneros, J., McKinley, R. L., & D'Angelo, W. R. (1996). Aurally aided visual search under virtual and free-field listening conditions. *Human Factors, 38,* 702–715.

Perrott, D. R., Saberi, K., Brown, K., & Strybel, T. Z. (1990). Auditory psychomotor coordination and visual search performance. *Perception & Psychophysics, 48,* 214–226.

Perrott, D. R., Sadralodabai, T., Saberi, K., & Strybel, T. Z. (1991). Aurally aided visual search in the central visual field: Effects of visual load and visual enhancement of the target. *Human Factors, 33,* 389–400.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Erlbaum.

Proctor, R. W., Tan, H. Z., Vu, K.-P. L., Gray, R., & Spence, C. (2005). Implications of compatibility and cuing effects for multimodal interfaces. *Proceedings of the HCI International 2005, 11.*

Ramachandran, V. S., Altschuler, E. L., & Hillyer, S. (1997). Mirror agnosia. *Proceedings of the Royal Society of London Series B, Biological Sciences, 264,* 645–647.

Reed, M. P., & Green, P. A. (1999). Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task. *Ergonomics, 42,* 1015–1037.

Rees, G., & Lavie, N. (2001). What can functional imaging reveal about the role of attention in visual awareness? *Neuropsychologia, 39,* 1343–1353.

Selcon, S. J., Taylor, R. M., & McKenna, F. P. (1995). Integrating multiple information sources: Using redundancy in the design of warnings. *Ergonomics, 38,* 2362–2370.

Shapiro, K. (Ed.). (2001). *The limits of attention: Temporal constraints in human information processing.* Oxford, U.K.: Oxford University Press.

Shinn-Cunningham, B., & Ihlefeld, A. (2004). Selective and divided attention: Extracting information from simultaneous sound sources. *Proceedings of the 10th Meeting of the International Conference on Auditory Displays.*

Simpson, C. A., & Marchionda-Frost, K. (1984). Synthesized speech rate and pitch effects on intelligibility of warning messages for pilots. *Human Factors, 26,* 509–517.

Simpson, C. A., McCauley, M. E., Roland, E. F., Ruth, J. C., & Williges, B. H. (1987). Speech controls and displays. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 1490–1525). New York: Wiley.

Sinnett, S., Costa, A., & Soto-Faraco, S. (in press). Inattentional blindness across sensory modalities. *Quarterly Journal of Experimental Psychology: A Human Experimental Psychology.*

Sorkin, R. D. (1988). Why are people turning off our alarms? *Journal of the Acoustical Society of America, 84,* 1107–1108.

Soto-Faraco, S., Morein-Zamir, S., & Kingstone, A. (2005). On audiovisual spatial synergy: The fragility of an effect. *Perception & Psychophysics, 67,* 444–457.

Soto-Faraco, S., & Spence, C. (2002). Modality-specific auditory and visual temporal processing deficits. *Quarterly Journal of Experimental Psychology,* 55A, 23–40.

Spence, C. (2001). Crossmodal attentional capture: A controversy resolved? In C. Folk & B. Gibson (Eds.), *Attraction, distraction and action: Multiple perspectives on attentional capture* (pp. 231–262). Amsterdam: Elsevier Science.

Spence, C., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 555–574.

Spence, C., & Driver, J. (1997a). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics, 59,* 1–22.

Spence, C., & Driver, J. (1997b). Cross-modal links in attention between audition, vision, and touch: Implications for interface design. *International Journal of Cognitive Ergonomics, 1,* 351–373.

Spence, C., & Driver, J. (1999). A new approach to the design of multimodal warning signals. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics, Vol. 4: Job design, product design and human-computer interaction* (pp. 455–461). Aldershot, England: Ashgate.

Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention.* Oxford, U.K.: Oxford University Press.

Spence, C., McDonald, J., & Driver, J. (2004). Exogenous spatial-cuing studies of human crossmodal attention and multisensory integration. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 277–320). Oxford, U.K.: Oxford University Press.

Spence, C., Nicholls, M. E. R., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics, 63,* 330–336.

Spence, C., Ranson, J., & Driver, J. (2000). Cross-modal selective attention: On the difficulty of ignoring sounds at the locus of visual attention. *Perception & Psychophysics, 62,* 410–424.

Spence, C., & Read, L. (2003). Speech shadowing while driving: On the difficulty of splitting attention between eye and ear. *Psychological Science, 14,* 251–256.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General, 130,* 799–832.

Stevens, S. S., & Newman, E. B. (1936). The localization of actual sources of sound. *American Journal of Psychology, 48,* 297–306.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science, 12,* 462–466.

Streeter, L. A., Vitello, D., & Wonsiewicz, S. A. (1985). How to tell people where to go: Comparing navigational aids. *International Journal of Man–Machine Studies, 22,* 549–562.

Von Fieandt, K. (1966). Perception of the self directly and as a mirror image. In K. von Fieandt (Ed.), *The world of perception* (pp. 322–335). Homewood, IL: Dorsey.

Wallace, J. S., & Fisher, D. L. (1998). Sound localization information theory analysis. *Human Factors, 40,* 50–68.

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1984). *Engineering psychology and human performance.* Columbus, OH: Merrill.

Wickens, C. D. (1991). Processing resources and attention. In D. L. Damos (Ed.), *Multiple-task performance* (pp. 3–24). London: Taylor & Francis.

van Winsum, W., & Brouwer, W. (1997). Time headway in car following and operational performance during unexpected braking. *Perceptual and Motor Skills, 84,* 1247–1257.

van Winsum, W., & Heino, A. (1996). Choice of time-headway in car following and the role of time-to-collision information in braking. *Ergonomics, 39,* 579–592.

Zeigler, B. L., Graham, F. K., & Hackley, S. A. (2001). Cross-modal warning effects on reflexive and voluntary reactions. *Psychophysiology, 38,* 903–911.

Appendix

Further Analysis of RSVP Performance

Performance in the RSVP task (i.e., probability of correct detection of RSVP targets) was analyzed for each participant in each of the five experiments reported in the present study as a function of the four critical driving conditions in the driving task. Target stimuli in the RSVP task presented after the onset of one auditory cue and before the onset of the next auditory cue were classified as belonging to the driving condition associated with the first auditory cue. Given that there were only eight trials for each of the four noncritical driving conditions, we did not include the data from these conditions in our analyses.

Table A1 shows the statistics for the data analyses as a function of the probability of correct detection of the RSVP targets in Experiments 1–5. Paired-comparison *t* tests performed on the data from Experiment 1 (nonspatial–nonpredictive car horn sounds) revealed no significant difference in participants' performance in the RSVP task as a function of whether the critical visual driving trials occurred at the front or back.

Separate analyses of variance were performed on the data from Experiments 2–5 to assess whether the spatial location from which the auditory cues and critical visual scenes were presented had any effect on participants' ability to respond to the RSVP targets. The two variables in the within-participants design were auditory cue direction (front vs. back) and visual stimulus location (front vs. back). The analyses of the probability data from Experiments 2 (spatially nonpredictive car horn sounds) and 3 (spatially predictive car horn sounds) revealed no significant effect on the RSVP task as a function of the critical driving conditions (see Table A1). It is interesting to note that the analysis of the data from Experiment 4 (predictive verbal cues) showed a significant interaction between auditory cue direction and visual stimulus location. In particular, participants detected more RSVP targets when the critical visual driving event was validly cued by the verbal cue (probability of correct detection: $M = 0.69$, $SE = 0.04$) instead of invalidly cued ($M = 0.64$, $SE = 0.04$). There was no significant main effect of either Auditory cue direction or Visual stimulus location.

The analysis of the data from Experiment 5 (spatially presented verbal cues) again revealed a significant interaction between the auditory cue direction and visual stimulus location, with participants detecting more RSVP targets when the driving cue and target were presented from the valid ($M = 0.79$, $SE = 0.02$) rather than from the invalid ($M = 0.74$, $SE = 0.02$) directions. The main effects of auditory cue direction and visual stimulus location were also significant, with participants detecting more RSVP targets when the driving trials contained a "back" auditory cue and/or a visual critical target from in front.

Taken together, these analyses of performance on the RSVP monitoring task as a function of the conditions in the driving task seem to suggest that participants were more distracted by the spatial elements inherent in the driving trials when verbal rather than nonverbal spatial auditory cues were used. These observations also coincide with the higher error rates in verbal over nonverbal spatial cuing.

(*Appendix continues*)

Table A1

ANOVA and t Test Results for the Probability of Correct Detection of RSVP Targets (as a Function of the Conditions in the Driving Task) Data Analyses in Experiments 1–5

Experiment	Source	$t(11)$	d	
1	Visual stimulus location (V)	−1.30	0.22	
		MSE	$F(1, 11)$	f
2	Auditory cue direction (A)	0.0029	2.65	0.47
	Visual stimulus location	0.0038	0.84	0.26
	A × V	0.0025	0.00	0.01
3	Auditory cue direction	0.0035	0.02	0.04
	Visual stimulus location	0.0027	0.01	0.03
	A × V	0.0028	0.00	0.02
4	Auditory cue direction	0.0057	0.11	0.09
	Visual stimulus location	0.0041	0.18	0.12
	A × V	0.0024	12.60**	1.03
		MSE	$F(1, 15)$	f
5	Auditory cue direction	0.0037	10.60**	0.82
	Visual stimulus location	0.0016	83.70**	2.29
	A × V	0.0012	38.60**	1.55

Note. RSVP = rapid serial visual presentation.
** $p < .01$.

Received February 28, 2005
Revision received May 16, 2005
Accepted May 16, 2005 ■