

# The differential effect of vibrotactile and auditory cues on visual spatial attention

CRISTY HO\*<sup>†</sup>, HONG Z. TAN<sup>‡</sup> and CHARLES SPENCE<sup>†</sup>

<sup>†</sup>Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3UD, UK

<sup>‡</sup>Haptic Interface Research Laboratory, Purdue University, West Lafayette, IN 47907-2035, USA

Previous research has shown that the presentation of spatially predictive auditory and vibrotactile warning signals can facilitate driver responses to driving events seen through the windscreen or rearview mirror. The present study investigated whether this facilitation reflects the priming of the appropriate response (i.e. braking vs. accelerating) or an attentional cuing effect (i.e. a perceptual benefit that facilitates subsequent behavioural responding). In the experiments reported here, participants had to discriminate the colour of a number plate (red vs. blue) following the presentation of either spatially predictive vibrotactile (experiment 1) or auditory (experiment 2) warning signals that indicated the likely location (front or back) of the visual target, while simultaneously performing a highly attention-demanding rapid serial visual presentation task. Numberplate discrimination performance was facilitated following the presentation of valid auditory cues, but not following the presentation of equally informative vibrotactile cues. The use of an orthogonal spatial cuing design enabled with us to rule out of a potential response priming account of these data. The results suggest that whilst directional congruency between a warning signal and a target event may be sufficient to facilitate performance due to the priming of the appropriate response, attentional facilitation effects may also require the co-location of the cue and target within the same functional region of space.

*Keywords:* Warning signal; Driving; Crossmodal spatial attention; Auditory; Vibrotactile; Visual

## 1. Introduction

The rapidly growing body of empirical research on the nature of crossmodal links in spatial attention demonstrates that responses to target events presented in one sensory modality can be facilitated by the prior presentation of a relevant cue in another sensory modality from approximately the same spatial location (see Spence and Driver 2004, for a

---

\*Corresponding author. Email: [cristy.ho@psy.ox.ac.uk](mailto:cristy.ho@psy.ox.ac.uk)

recent review). Such crossmodal spatial facilitation effects have been studied in a variety of different settings, from fundamental laboratory-based research (e.g. Driver and Spence 2004, Spence *et al.* 2004) to more complicated environments in applied research settings (e.g. van Erp and van Veen 2001, 2004). However, it is unclear whether the facilitatory effects reported in many of these previous cuing studies should be attributed to response priming, attentional facilitation or to some unknown combination of these two effects.

According to the response priming account, the facilitation of performance seen following the presentation of a warning signal cue occurs because the response appropriate to the target stimulus is primed by the cue (i.e. performance is facilitated without there necessarily being any change in the perceptual representation of the target itself). According to the attentional account, the presentation of the cue (or warning signal) leads to a crossmodal shift of spatial attention that facilitates the perception of targets subsequently presented near to the cue, and this attentional enhancement of target processing is what leads to improved behavioural performance. The critical research question to be addressed here then is one of understanding how 'early' in information processing the crossmodal cuing effects attributable to the presentation of a warning signal occur. From the applied side, the facilitation of behavioural responses to potential emergency situations is clearly beneficial, no matter how that benefit is elicited (i.e. by response priming, attentional facilitation or by some combination of these two effects). From a theoretical point of view, however, many researchers have undertaken laboratory-based research in order to try and isolate response priming (or compatibility) effects while others have focused on trying to isolate attentional facilitation effects instead (see Spence and McDonald 2004, Proctor *et al.* 2005). Nevertheless, it is only by getting a better understanding of the relative contributions of the two effects (response priming and attentional facilitation) to performance under more realistic conditions that it might, in the future, be possible to design the most effective warning signals – signals that will presumably facilitate performance at both levels.

In an attempt to eliminate the possible contribution of response priming in laboratory-based studies of crossmodal cuing, Spence and Driver (1994) proposed the use of an orthogonal cuing methodology. In a typical orthogonal cuing study, participants are instructed to respond to a dimension, or property, of the target (e.g. its colour or identity) that is orthogonal to the dimension along which cuing occurs. So, for instance, if participants are required to make red vs. blue target discrimination responses for targets appearing on either the left or right, then the dimension of spatial cuing (left vs. right) is orthogonal to the dimension of responding (red vs. blue). The presentation of the cue should therefore not prime a particular response, and so any residual facilitation of performance can unambiguously be attributed to attentional cuing.

In a recent series of laboratory-based simulated-driving experiments, it has been demonstrated that the use of various different types of spatial auditory and vibrotactile warning signals (presented from the front or rear) can facilitate performance when responding to targets presented from the front or rear (e.g. Ho and Spence 2005, Ho *et al.* 2005b). Given that the responses required of participants in these previous studies were either to brake or to accelerate in response to target visual driving events occurring in the front windscreen or the rearview mirror respectively (i.e. as required in a real driving situation), it is unclear whether the facilitation of responding in these previous studies reflects response priming, spatial attentional facilitation or some unknown combination of the two effects. The present study was therefore designed to study whether a purely attentional component to this spatial warning signal effect could be demonstrated by requiring participants to perform an orthogonal discrimination task. Specifically, participants were required to

make speeded discrimination responses regarding the colour of a numberplate (licence plate; i.e. responses were orthogonal to the spatial dimension of interest – the front vs. back cuing of driver attention), with the experiment otherwise being identical to previous research (see Ho and Spence 2005, Ho *et al.* 2005b). The results of the present study should therefore highlight the factors governing the facilitation reported in these previous studies.

It was hypothesized that if the same pattern of results could be replicated in the present study as in the previous research, then this would suggest that the facilitatory crossmodal cuing effects observed previously must reflect, at least in part, a spatial attentional effect. Alternatively, however, if the cues in the previous research simply primed the appropriate responses (i.e. braking vs. accelerating), then no such facilitation would be expected in the present study (given the orthogonal nature of the responses required). The first experiment was designed to replicate experiment 1 of Ho *et al.*'s (2005b) study, where spatially predictive vibrotactile cues (80% valid with regard to the likely location of the upcoming visual targets) were shown to facilitate subsequent responses to target visual driving events in the cued direction. The second experiment replicated Ho and Spence's (2005) experiment 3, where spatial auditory car horn sounds were used as the spatially predictive warning signals.

## 2. Experiment 1

### 2.1. Methods

**2.1.1. Participants.** Sixteen participants (mean age of 19 years, age range from 18–20 years; three males and 13 females) took part in this experiment. All of the participants reported normal tactile sensitivity and normal, or corrected-to-normal, vision. The mean driving experience of participants was 2 years (range from 6 months to 3 years). Twelve participants were right-handed and four were left-handed by self-report. The experiment lasted for approximately 60 min. All of the participants were naïve as to the purpose of the study and they received course credit in return for their participation. The experiments were conducted in accordance with the guidelines laid down by the Department of Experimental Psychology, University of Oxford.

**2.1.2. Apparatus and materials.** The experiment was conducted using the same set-up as that reported in previous research (Ho *et al.* 2005b). The participants were seated in the middle of a 220 × 142 cm experimental booth and were instructed to hold a Logitech<sup>®</sup> MOMO<sup>®</sup> Racing Force Feedback Wheel (Logitech Inc., Fermont, CA, USA) mounted on a desk situated directly in front of them. The foot pedals linked to the steering 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 rapid serial visual presentation (RSVP) stimuli (see figure 1). A monitor showing a video of a road scene taken through the windscreen was positioned 70 cm from the participants (visible over the top of the mirror). A car rearview mirror (6 × 15 cm) 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. The tactors (2.54 × 1.85 × 1.07 cm, VBW32; Audiological Engineering Corp., Somerville, MA, USA) used to present the vibrotactile signals were attached to a Velcro belt fastened around the participant's waist. One tactor was placed on the front in the middle of the participant's stomach and another one was placed in the middle of the participant's back. The belt and the tactors were fastened directly over the top of the

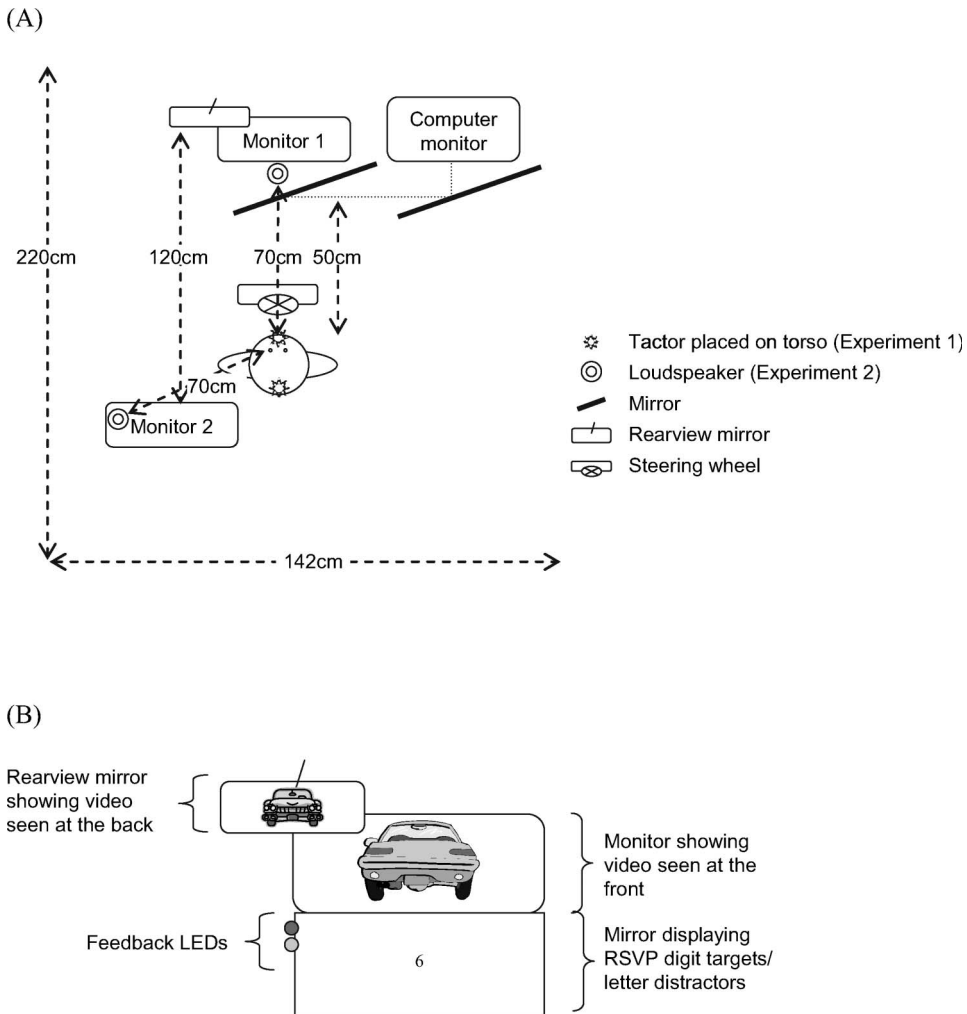


Figure 1. Schematic bird's-eye view (A) and participant's-eye view (B) of the experimental set-up used in experiments 1 and 2. LEDs = light emitting diodes; RSVP = rapid serial visual presentation.

participant's clothing. The tactors were driven with a 290 Hz sinusoidal signal, at an intensity sufficient to deliver clearly perceptible vibrotactile stimuli through clothing. The noise level of the tactors was approximately 42 dB(A) when measured from the participants' ear level with the tactor belt placed around their waist. White noise was delivered through cordless headphones (SBC-HC8355; Philips Electronics, New York, USA) at about 60 dB(A) to mask the noise caused by the operation of the tactors.

The RSVP stimuli consisted of 17 distractor letters and six target digits (cf. Soto-Faraco and 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  $8.3 \times 8.5$  mm in size on the mirror as seen by the participants.

The video recording of the windscreen 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/h. The rear video recording 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 recording included the sudden fast approach (by the car in which the video camera was placed) toward the car in front at approximately 100 km/h or the sudden rapid approach from the car behind. Additional editing was performed using Adobe<sup>®</sup> After Effects<sup>®</sup> 6.5 (Adobe Systems Incorporated, San Jose, CA, USA) to change the colour of the numberplate from yellow (for the car in front) or white (for the car from behind; as required under British law) to either red or blue during the sudden fast approach period (duration of 1800 ms). The non-critical clips included the car in front moving away at normal speed or the car behind retreating, with the colour of the numberplate changing to green instead. To be clear, the numberplate of one of the two cars – either the leading or following car – always changed colour whenever a vibrotactile cue was presented. The critical red and blue colour patches had the same saturation and intensity (set at 100%), with red having a hue of 0° and blue 240°. The non-critical green colour patch had a hue of 120°, with the same saturation and intensity levels as the critical colour patches. Four critical and two non-critical driving clips were created. A red light-emitting diode (LED) was placed on the mirror directly in front of the participants to provide feedback whenever they made an incorrect response in the driving task. An amber LED placed below the red LED was illuminated whenever participants failed to keep the accelerator depressed appropriately during the experiment.

**2.1.3. Design.** The experimental session consisted of eight 6-min blocks of experimental trials. The RSVP task consisted of a continuous stream of distractor letters with target digits periodically embedded within it. The RSVP task was used to simulate a continuously and uniformly highly attention-demanding situation. It has been used extensively in laboratory-based, dual-task attention research to maintain cognitive load on participants and to measure the temporal distribution of attention under conditions of task switching. Essentially, the presentation of the RSVP task to the front of a participant modelled the situation when the visual load of a driver's attention is primarily engaged toward the front. 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. A total of 66 targets were presented in each block, with the temporal gap between successive target digits in the RSVP stream of 2040–6360 ms.

For the driving task, 24 randomized scenarios were presented in each experimental block. Each scenario consisted of a 15 000 ms video clip and a vibrotactile cue lasting for 1060 ms (see figure 2). In a given experimental block, each of the four critical driving clips was presented four times (i.e. the fast approach toward the car in front with either a red or blue target numberplate, or the fast approach of the car from behind with either a red or blue target numberplate) with a vibrotactile cue coming from the appropriate direction, and twice with a vibrotactile cue coming from the opposite direction to the visual driving event (i.e. the cues correctly predicted the direction of the targets on 80% of trials). The onset of the vibrotactile cue coincided with the start of the critical driving event, defined as the initiation of the reduction of the inter-car distance that lasted for 1800 ms (at which time the two cars would have collided). The onset of the vibrotactile cues occurred at the same point in time as in Ho *et al.*'s (2005b) previous study. Each of the four non-critical driving scenarios was presented once (i.e. the ratio of critical to non-critical trials was 83:17). The temporal gap between successive vibrotactile cues was

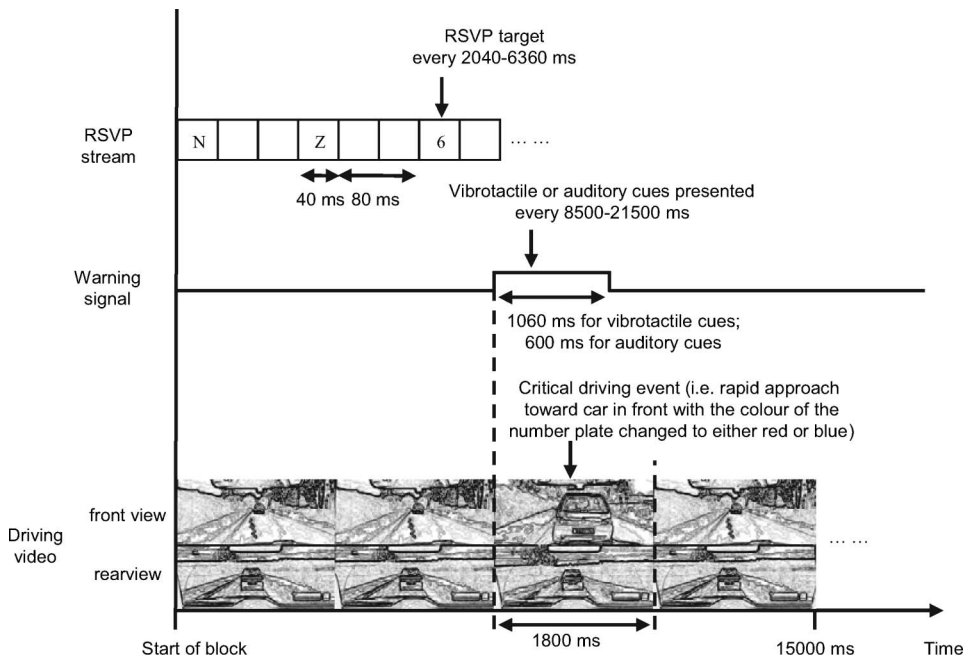


Figure 2. A schematic timeline showing the temporal sequence of events in experiments 1 and 2. RSVP = rapid serial visual presentation.

8500–21500 ms. In total, there were 528 RSVP targets and 192 driving trials (i.e. 160 critical and 32 non-critical trials) in a participant's experimental session.

The participants were given two practice blocks in which to familiarize themselves with the experimental set-up. In the first practice block, they only had to perform the RSVP task, which was initially presented at a slower rate that gradually increased to the experimental rate of stimulus presentation. In the second block, the participants performed both the RSVP and driving tasks as in the subsequent experimental blocks.

**2.1.4. Procedure.** The participants responded to targets in the RSVP task by pressing the right paddle shifter on the steering wheel. For the driving task, half of the participants were instructed to press the upper left button (mounted next to the left paddle shifter on the steering wheel) when they detected a red target numberplate, and to press the lower left button (mounted 2.5 cm vertically below the upper button) for a blue target numberplate. The other half of the participants were instructed to use the opposite response mapping. (They were not instructed to change the pressure of the brake or accelerator pedals in response to the driving events.) For non-critical trials where the numberplate changed to green, the participants simply had to carry on without making any specific response. The red LED was illuminated for 1000 ms following an incorrect response. Participants were also instructed to keep the accelerator depressed slightly throughout the experiment in order to model realistic driving conditions.

## 2.2. Results

Trials in which participants made an incorrect response were discarded from the analysis of the reaction time (RT) data. On average, 3.5% of the trials were removed across

participants due to no response being made within 1800 ms of the onset of the critical visual event. False alarms, where participants incorrectly responded when the colour of the numberplate changed to green, occurred on 1.4% of catch trials. An ANOVA was performed on the RT data to assess whether the spatial direction from which the vibrotactile cues were presented had any effect on the participants' ability to respond to the critical visual driving targets presented from either the front or behind (seen via the rearview mirror). The two within-participants factors were: vibrotactile cue location (front vs. back); and visual stimulus location (front vs. back).

This analysis revealed a significant main effect of visual stimulus location,  $F(1,15) = 132.3$ , mean square error (MSE) = 4638,  $p < 0.0001$ , with participants responding more rapidly to critical visual events presented from the front (mean (M) = 987 ms) than from the rearview mirror (M = 1183 ms). The main effect of vibrotactile cue location was, however, not significant,  $F(1,15) = 1.7$ , MSE = 1427,  $p = 0.21$ , nor was the interaction between vibrotactile cue location and visual stimulus location,  $F(1,15) = 1.0$ , MSE = 817,  $p = 0.33$ , suggesting that participants responded equally rapidly regardless of the spatial validity of the vibrotactile cues with regard to the location of the visual targets in the driving videos (i.e. no matter whether the cues were spatially valid or invalid; see figure 3).

A similar analysis of the error data revealed no significant main effect of vibrotactile cue location,  $F(1,15) = 1.1$ , MSE = 19.4,  $p = 0.31$ , or visual stimulus location,  $F(1,15) < 1$ , NS, nor any interaction between these two factors,  $F(1,15) < 1$ , NS.

Performance in the concurrent RSVP task was analysed to ensure that participants performed the central attention-demanding visual task as instructed. Responses occurring 1500 ms or more after the presentation of a target digit were considered as invalid (i.e. false alarms). The mean percentage of correct detection responses in the RSVP task was 74.7 (SE 1.9), with a mean RT of 577 (SE 9) ms and a false alarm rate of 6.2 (SE 1.0)%, which was similar to the results obtained in a previous study (Ho *et al.* 2005b).

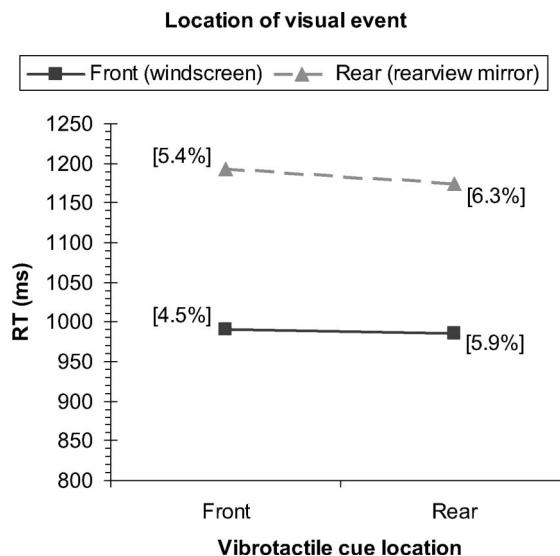


Figure 3. Mean reaction times (RT) for the driving task in experiment 1 as a function of the location of the spatially predictive vibrotactile cue and the location of the critical visual event. Percentages of errors are indicated by numerical values in parentheses.

### 2.3. Discussion

Somewhat surprisingly, the results of experiment 1 revealed no spatial cuing effect following the presentation of the vibrotactile cues. In fact, there was no significant difference in the latency of participants' responses as a function of whether the vibrotactile cues came from the same or opposite direction as the critical visual driving targets. The slightly faster responses following a cue from behind than from in front may reflect the possibility that vibrotactile stimuli are more salient (and/or alerting) when presented to the back than to the front of a participant's waist (see Ho *et al.* 2005a).

The null vibrotactile cuing effects reported in the present study contrast with the robust cuing effects reported in a previous study of the vibrotactile cuing of driver attention (Ho *et al.* 2005 reported mean cuing effects of 66 ms and 6.8% in the RT and error data, respectively). Taken together, these two results suggest that whilst the presentation of a vibrotactile cue from the appropriate spatial direction on the body surface may elicit an automatic (or 'intuitive'; van Erp 2005) response bias (cf. Spence *et al.* 2004, Prinzmetal *et al.* 2005), they do not necessarily lead to a shift of spatial attention (which would have been expected to lead to a perceptual facilitation of colour discrimination responses). One may argue that spatial attention might not have influenced the detection of colour change and the detection of angular size change (as in the detection of a rapidly approaching car reported in Ho and Spence (2005) and Ho *et al.* (2005b)) in the same manner, thus rendering a strict comparison between the present study and the previous studies problematic. However, it should be noted that whether the spatial element is inherent in the task itself (e.g. in angular size change detection) or not (e.g. in colour change detection) is a constant factor within each of the experiments, and that any effects attributable to this factor were cancelled out when the mean cuing effects within each experiment were calculated. Note that the mean cuing effects concern the effectiveness of the spatial cues to orient the attention of the participants to the spatial location of interest. Other researchers have also reported the same pattern of attentional orienting effects across experiments, no matter whether the task was to detect a target presented alone or amongst distractors, or else to discriminate the colour of a target following the presentation of the same visual cues (see e.g. Hommel *et al.* 2001).

The results of experiment 1 show that the presentation of a vibrotactile cue to the front or back of the participant's torso did not lead to a shift of attention in the spatial direction cued by the vibration as indexed by a task involving the discrimination of the colour of a numberplate presented far from the participant's body. It is possible, however, that the vibrotactile cues might have resulted in the participant's attention being oriented toward the source of stimulation on the participant's torso (i.e. to peripersonal space; see Tassinari and Campara 1996, Kitagawa *et al.* 2005), rather than to the extrapersonal visual space outside the car (as seen through the windscreen or rearview mirror). Previous research has shown that crossmodal spatial cuing effects are maximal when the cue and target are presented from the same spatial location and decline as the cue-target separation is increased (Spence *et al.* 2004). (This contrasts with the situation for response priming, where as long as the cue primes the response, it does not really matter so much what the exact spatial relationship between the cue and target is; see Proctor *et al.* 2005.) Indeed, recent neuropsychological research with patients suffering from unilateral spatial extinction has shown far more pronounced interactions between pairs of stimuli when they are both presented within the same region of space (i.e. both within peripersonal space or both within just extrapersonal space) than when they occur in different regions of space (i.e. one in peripersonal space and the other in extrapersonal space; e.g. Làdavas



2002). Such results have been taken to show that the brain represents stimuli originating in these two regions of space in a somewhat distinct manner (cf. Previc 1998, 2000).

Given that vibrotactile cues can only be presented on the body surface (i.e. in near-peripersonal space), the second experiment investigated whether the presentation of spatially predictive auditory warning signals from the same functional region of space as the target visual driving events (i.e. both events presented in extrapersonal space) would facilitate performance on the numberplate colour discrimination task for targets presented from the cued location. (Note that both the rear auditory and visual stimuli originated from the rear, despite the fact that the targets were visually inspected from in front via the rearview mirror.)

### 3. Experiment 2

#### 3.1. Methods

Ten new participants (mean age of 21 years, age range from 18–32; six males and four females) took part in this experiment. All of the participants reported normal hearing and normal, or corrected-to-normal, vision. The mean driving experience of participants was 4 years (range from 1–14 years). All the participants were right-handed by self-report. All of the participants were naïve as to the purpose of the study and none had taken part in experiment 1. Six participants received course credit, the rest received a £5 UK sterling gift voucher for their participation.

The apparatus, materials, design and procedure were exactly the same as those used in experiment 1, with the sole exception that the vibrotactile cues were now replaced by auditory cues. The specification of the auditory stimuli was identical to that reported in experiment 3 of Ho and Spence's (2005) previous study. Specifically, the sound of a real car horn (600 ms duration; 8000 Hz; 66 dB(A)) downloaded from the Internet (retrieved from <http://www.a1freesoundeffects.com/carhornshort.wav>; downloaded on 28 November 2003) was used as the auditory cue. Spatially predictive (80% valid) car horn sounds were presented from one of two loudspeaker cones placed on a virtual circle (70 cm in diameter) centred on the participants' head, one to the front and the other to their rear left-hand side (see figure 1).

#### 3.2. Results

The results of experiment 2 were analysed in a similar manner to those of experiment 1. The mean false alarm rate across participants was 0.3%, and 1.3% of the trials were removed due to no response being made within 1800 ms of the onset of the critical visual event. A two-way, within-participants ANOVA was performed on the RT data with the factors of auditory cue direction (front vs. back) and visual stimulus location (front vs. back). The analysis revealed a significant main effect of auditory cue direction,  $F(1,9) = 8.7$ ,  $MSE = 1793$ ,  $p = 0.02$ , with participants responding more rapidly to critical visual events (i.e. to the colour change of the numberplate) following an auditory cue from the back ( $M = 953$  ms) than from the front ( $M = 992$  ms). The main effect of visual stimulus location was also significant,  $F(1,9) = 34.8$ ,  $MSE = 13225$ ,  $p < 0.001$ , with participants responding more rapidly to critical visual events presented from the front ( $M = 865$  ms) than from the rear ( $M = 1080$  ms), as in experiment 1. However, in contrast to experiment 1, the interaction between these factors was also significant,  $F(1,9) = 25.7$ ,  $MSE = 3105$ ,  $p < 0.001$ . Participants responded more rapidly when the auditory cue and

visual target stimulus were presented from the same direction (i.e. on validly cued trials;  $M = 928$  ms) than when the auditory cue was presented from the opposite direction to the visual stimulus (i.e. on invalidly cued trials;  $M = 1017$  ms). Subsequent paired comparison t-tests revealed that these spatial cuing effects were significant both for targets presented through the front windscreen (mean cuing effect of 50 ms),  $t(9) = 3.3$ ,  $p = 0.01$ , and for those presented in the rearview mirror ( $M = 129$  ms),  $t(9) = 4.7$ ,  $p = 0.001$  (see figure 4).

A similar analysis of the error data again revealed no significant main effects of auditory cue direction,  $F(1,9) = 1.7$ ,  $MSE = 11.8$ ,  $p = 0.22$ , or visual stimulus location,  $F(1,9) < 1$ , NS, nor any interaction between them,  $F(1,9) < 1$ , NS. Analysis of the inverse efficiency data (calculated by dividing the RT data by the proportion of correct responses for each condition for each participant, thus combining the RT and error rate data into a single performance measure; e.g. Townsend and Ashby 1983) also revealed a significant interaction between auditory cue direction and visual stimulus location,  $F(1,9) = 25.9$ ,  $MSE = 2653$ ,  $p < 0.001$ , with significant spatial cuing effects for targets presented both from the front,  $t(9) = 4.4$ ,  $p = 0.002$ , and from the rear,  $t(9) = 3.4$ ,  $p = 0.007$ . Thus, it was possible to rule out a speed-accuracy trade-off account of the cuing effects reported in the RT data (see Duncan 1980).

The mean percentage of correct detection responses in the concurrent RSVP task was 71.9 (SE 4.1), with a mean RT of 585 (SE 13) ms and a false alarm rate of 7.6 (SE 2.5)%. These results were very similar to those reported in experiment 1.

### 3.3. Discussion

Participants in experiment 2 responded significantly more rapidly to target visual driving events (the sudden change in the colour of a car's numberplate) occurring in the same

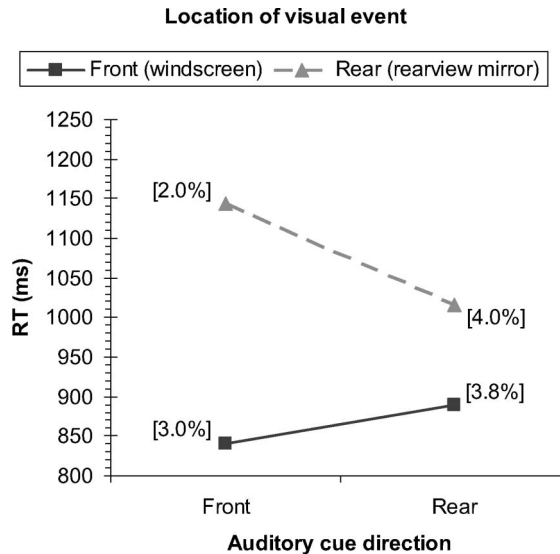


Figure 4. Interaction between the direction of the spatially predictive auditory cue and the location of the critical visual event in the driving task in experiment 2. RT = reaction time. Percentages of errors are indicated by numerical values in parentheses.

direction as the auditory warning signal than when they were presented from opposite directions. These results replicate the crossmodal spatial facilitation effects observed in a previous auditory cuing study, when participants had either to brake or to accelerate in response to the sudden approach toward the car in front or from the car behind (Ho and Spence 2005). The present results allow us to rule out response priming as the sole explanation of the crossmodal effects reported in previous auditory cuing experiments (Ho and Spence 2005). Taken together, these findings support the idea that informative auditory warning signals can effectively orient a driver's spatial attention to a particular region of space (in this case, to front or rear extrapersonal space), leading to a facilitation of responses to visual events occurring subsequently in that direction (e.g. Spence *et al.* 2004).

#### 4. General discussion

The facilitatory effects of crossmodal spatial attentional cuing were observed following the presentation of spatially predictive auditory warning signals (experiment 2), but not following the presentation of spatially predictive vibrotactile warning signals (experiment 1). These findings contrast with the significant crossmodal spatial facilitation effects evidenced following the presentation of both spatially predictive auditory and vibrotactile warning signals in previous research (see Ho and Spence 2005, Ho *et al.* 2005b; figure 5). There are at least two possible explanations for the differential effects of auditory and vibrotactile cuing on performance in the experiments reported here. First, it may be that audition and touch affect visual spatial attention in somewhat different ways. For although it has been argued previously that spatial attention may function in the same manner across the various sensory modalities (see Spence and Driver 2004), the translation (or switching of attention) from touch to vision has been shown to be less efficient (i.e. more time-consuming) than from audition to vision (see Spence *et al.* 2001). The change in the colour of the numberplate (which was superimposed on a rapidly approaching, i.e. looming, car) in the present study may have been more attention-capturing (cf. Turatto and Galfano 2001, Franconeri *et al.* 2004) than the rapid approach toward the car in front or the rapid approach of the car from behind in previous studies. Consequently, the change of the colour of the numberplate may have more automatically (and/or rapidly) captured the participants' attention in the two experiments reported here. Given that tactile information is transduced at the receptor surface (i.e. skin) more slowly than auditory stimuli at the basilar membrane of the ear (see Spence and Squire 2003, Zampini *et al.* 2005), it might be suggested that participants in experiment 1 may not have had time to localize the vibrations and to shift their spatial attention to the visual driving task from the RSVP task prior to their attention being captured so automatically by the onset of the visual target (i.e. by the colour change of the numberplate itself).

However, the results of another study, which investigated the relative speed with which people can discriminate the direction (front vs. back) from which warning signals are presented in the different sensory modalities, reveals this explanation to be incorrect (Ho *et al.* 2005a). There, it was found that participants could discriminate the direction of vibrotactile warning signals significantly more rapidly than the direction of auditory warning signals, presumably because of the difficulty people have in localizing auditory stimuli within confined spaces (e.g. Moore and King 1999, Catchpole *et al.* 2004).

Alternatively, however, it could also be argued that the colour discrimination task itself may simply have been less sensitive to the distribution of spatial attention than the looming (i.e. braking/accelerating) task used in previous studies, with the latter task in

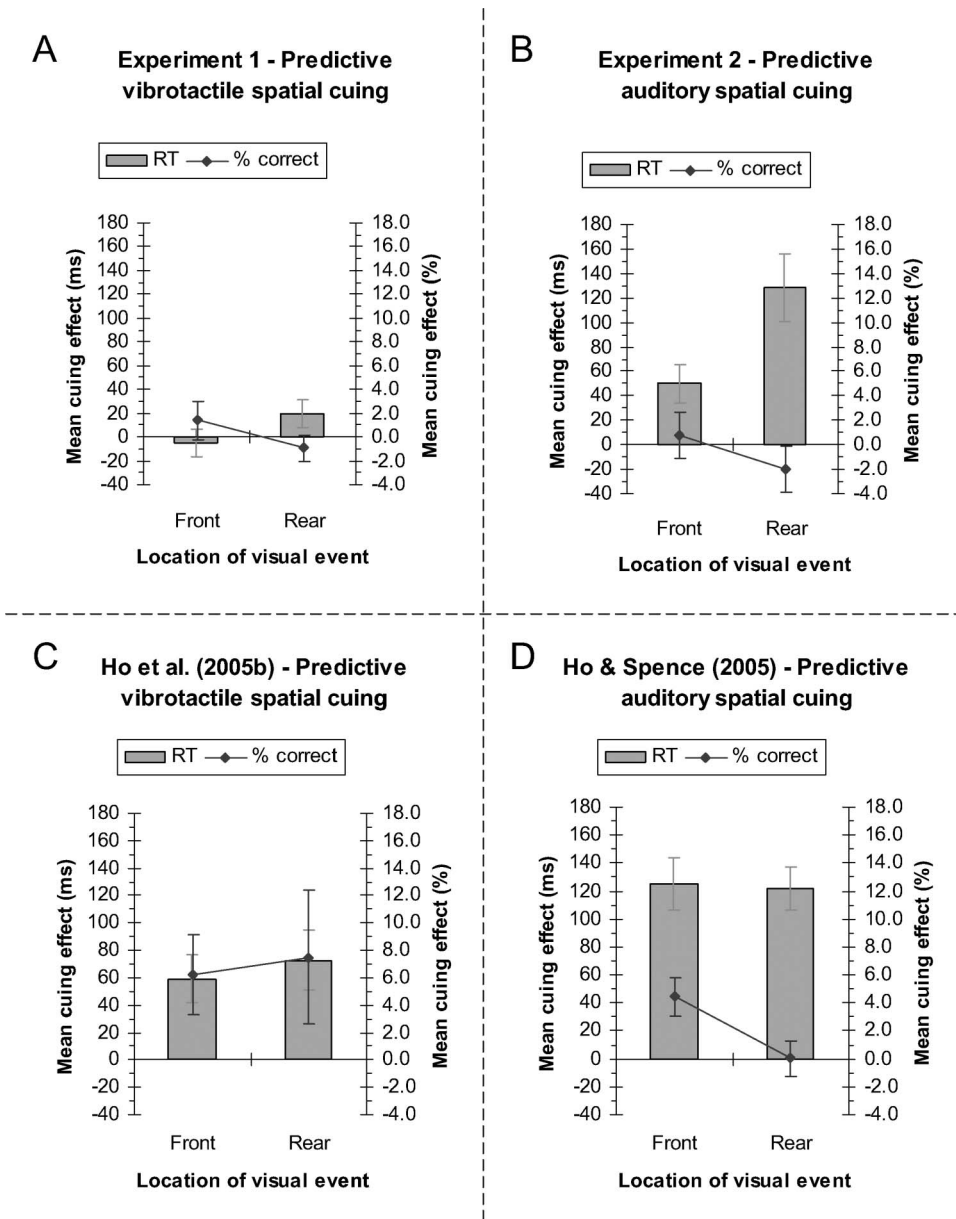


Figure 5. Summary of the mean cuing effects following vibrotactile (experiment 1) and auditory (experiment 2) cuing in the present study (A and B, respectively), spatially predictive vibrotactile cuing in Ho *et al.*'s (2005b) study (C) and spatially predictive auditory cuing in Ho and Spence's (2005) study (D). RT = reaction times. The error bars show the standard errors of the means.

some sense involving the integration of spatial information over time (cf. Spence *et al.* 2004). There are two reasons why the authors believe that such an account can be ruled out. First, several researchers have shown reliable spatial attentional orienting effects using colour discrimination tasks (e.g. see Hommel *et al.* 2001). Second, the reliable

spatial cuing effects observed with auditory cuing in experiment 2 argue against the colour discrimination task simply being insensitive to the distribution of spatial attention. Hence, it is very unlikely that the use of the colour discrimination task represents a major factor contributing to the null vibrotactile cuing effect observed in experiment 1.

Several researchers have suggested that the use of vibrotactile cues may provide an 'intuitive' means of presenting directional information in interface displays, such as in situation awareness and navigation systems for pilots and drivers (e.g. Rupert 2000, van Erp and van Veen 2004, van Erp 2005). For instance, van Erp and van Veen showed that drivers could respond more rapidly following navigational messages presented in a bimodal (tactile and visual) display than when the messages were presented unimodally (i.e. via touch or vision only). Participants also reported significantly lower subjective mental workload ratings with the touch only display, as compared to when using the visual only or bimodal displays.

Van Erp (2005) recently reported that people seem to have no trouble in localizing vibrotactile stimuli presented to their torsos and can indicate the associated direction in the external environment accurately (albeit with a consistent mislocalization bias toward the navel or spine). However, the results of the present study suggest that vibrotactile cues are in some sense limited. Indeed, when assessing the effectiveness of vibrotactile cues in conveying information it may be useful to break the role of such vibrotactile stimuli into distinct categories, namely, their ability to direct spatial attention to the location of the critical target event, their ability to prime the response appropriate to that target event and, finally, their ability to act as a means of information transfer to an interface operator.

The most likely explanation for the difference in the effectiveness of auditory and vibrotactile cuing in the present study comes from recent evidence suggesting that the brain represents stimuli occurring in peripersonal space somewhat differently from those occurring in extrapersonal space (e.g. Rizzolatti *et al.* 1997, Previc 1998, 2000, Spence and Driver 2004). It may be that the remapping of a front or back directional cue on the torso (i.e. in peripersonal space) to a target visual event happening in a location farther away in extrapersonal space is less efficient than when the cue and target both occur within the same functional region of space (i.e. both within peripersonal space or both within extrapersonal space). In this regard, it would be interesting in future research to examine how people subjectively represent extrapersonal and peripersonal space in driving situations (e.g. does the peripersonal vs. extrapersonal distinction equate to events happening within the car vs. outside the car, respectively, cf. Kitagawa and Spence 2005, Spence and Ho 2005). Understanding the representation of near-peripersonal space may therefore assist in interpreting the present findings (see Lådavas and Farnè 2004, Kitagawa *et al.* 2005).

Future research should determine whether the effectiveness of auditory cues changes as a function of their distance from the body, such as when they are presented from locations very close to the participants' head or waist vs. from locations further away (cf. Warren 1970, Moore and King 1999). If, as recent cognitive neuroscience data suggest, the brain really does handle peripersonal and extrapersonal space in qualitatively different ways (Previc 1998, Lådavas 2002, Spence and Driver 2004), then it might be predicted that the crossmodal spatial facilitation effects observed in experiment 2 would be eliminated if the spatial auditory cues were actually presented from sources that were much closer to the participants' body (just as was the case for the vibrotactile cues in experiment 1). If, however, crossmodal spatial facilitation effects were still to be present

in this situation, then it would suggest that for the differential effects of auditory and vibrotactile cues on visual spatial attention in the driving situation, there must be another cause. Taking things one step further, it might also be predicted that auditory cues could be made even more effective if it was somehow possible to fool the brain into thinking that they came from outside the car (i.e. from far extrapersonal space; see Previc 1998, 2000, Alais and Carlile 2005). The authors are currently extending their research to address precisely this question.

From an applied perspective, the research outlined here suggests that the use of vibrotactile and auditory warning signals in automobile interface design to inform drivers of the occurrence of time-critical events outside the car is both effective and practical. The present findings also imply that response compatibility is an important factor in multimodal interface design and may facilitate performance whenever the warning signal and the relevant target event are presented from approximately the same direction. However, the additional benefits attributable to attentional facilitation may only occur when the warning signal and target event locations match much more closely (i.e. when they are both presented within the same region of space).

### Acknowledgements

This research was supported in part by scholarships from the Clarendon Fund, Oxford UK, and from the Overseas Research Student Awards Scheme (ORS) to C.H. and a Network Grant from the Oxford McDonnell-Pew Centre for Cognitive Neuroscience to H.Z.T. and C.S.

### References

- ALAIS, D. and CARLILE, S., 2005, Synchronizing to real events: Subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proceedings of the National Academy of Sciences*, **102**, 2244–2247.
- CATCHPOLE, K.R., MCKEOWN, J.D. and WITHINGTON, D.J., 2004, Localizable auditory warning pulses. *Ergonomics*, **47**, 748–771.
- DRIVER, J. and SPENCE, C., 2004, Crossmodal spatial attention: Evidence from human performance. In *Crossmodal Space and Crossmodal Attention*, C. Spence and J. Driver (Eds.), pp. 179–220 (Oxford: Oxford University Press).
- DUNCAN, J., 1980, The demonstration of capacity limitation. *Cognitive Psychology*, **12**, 75–96.
- FRANCONERI, S.L., SIMONS, D.J. and JUNGE, J.A., 2004, Searching for stimulus-driven shifts of attention. *Psychonomic Bulletin and Review*, **11**, 876–881.
- HO, C. and SPENCE, C., 2005, Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Psychology: Applied*, **11**, 157–174.
- HO, C., SPENCE, C. and TAN, H.Z., 2005a, Warning signals go multisensory. *Proceedings of the 11th Annual Conference on Human-Computer Interaction, Vol. 9* (Mahwah, NJ: Erlbaum).
- HO, C., TAN, H.Z. and SPENCE, C., 2005b, Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, **8**, 397–412.
- HOMMEL, B., PRATT, J., COLZATO, L. and GODIJN, R., 2001, Symbolic control of visual attention. *Psychological Science*, **12**, 360–365.
- KITAGAWA, N. and SPENCE, C., 2005, Investigating the effect of a transparent barrier on the crossmodal congruency effect. *Experimental Brain Research*, **161**, 62–71.
- KITAGAWA, N., ZAMPINI, M. and SPENCE, C., 2005, Audiotactile interactions in near and far space. *Experimental Brain Research*, **166**, 528–537.
- LÁDAVAS, E., 2002, Functional and dynamic properties of visual peripersonal space. *Trends in Cognitive Sciences*, **6**, 17–22.
- LÁDAVAS, E. and FARNÉ, A., 2004, Neuropsychological evidence for multimodal representations of space near specific body parts. In *Crossmodal Space and Crossmodal Attention*, C. Spence and J. Driver (Eds.), pp. 69–98 (Oxford: Oxford University Press).

- MOORE, D.R. and KING, A.J., 1999, Auditory perception: The near and far of sound localization. *Current Biology*, **9**, R361–R363.
- PREVIC, F.H., 1998, The neuropsychology of 3-D space. *Psychological Bulletin*, **124**, 123–164.
- PREVIC, F.H., 2000, Neuropsychological guidelines for aircraft control stations. *IEEE Engineering in Medicine and Biology Magazine*, **19**, 81–88.
- PRINZMETAL, W., MCCOOL, C. and PARK, S., 2005, Attention: Reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General*, **134**, 73–92.
- PROCTOR, W., TAN, H.Z., VU, K.-P.L., GRAY, R. and SPENCE, C., 2005, Implications of compatibility and cuing effects for multimodal interfaces. *Proceedings of the 11th Annual Conference on Human-Computer Interaction, Vol. 11* (Mahwah, NJ: Erlbaum).
- RIZZOLATTI, G., FADIGA, L., FOGASSI, L. and GALLESE, V., 1997, The space around us. *Science*, **277**, 190–191.
- RUPERT, A.H., 2000, An instrumentation solution for reducing spatial disorientation mishaps. *IEEE Engineering in Medicine and Biology Magazine*, **19**, 71–80.
- SOTO-FARACO, S. and SPENCE, C., 2002, Modality-specific auditory and visual temporal processing deficits. *Quarterly Journal of Experimental Psychology*, **55A**, 23–40.
- SPENCE, C.J. and DRIVER, J. (Eds.), 1994, Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 555–574.
- SPENCE, C. and DRIVER, J., 2004, *Crossmodal Space and Crossmodal Attention* (Oxford: Oxford University Press).
- SPENCE, C. and HO, C., 2005, Multisensory warning signals for event perception and safe driving (in press).
- SPENCE, C. and McDONALD, J., 2004, The cross-modal consequences of the exogenous spatial orienting of attention. In *The Handbook of Multisensory Processes*, G.A., Calvert, C. Spence and B.E. Stein (Eds.), pp. 3–25 (Cambridge, MA: MIT Press).
- SPENCE, C., McDONALD, J. and DRIVER, J., 2004, Exogenous spatial cuing studies of human crossmodal attention and multisensory integration. In *Crossmodal Space and Crossmodal Attention*, C. Spence and J. Driver (Eds.), pp. 277–320 (Oxford: Oxford University Press).
- SPENCE, C., NICHOLLS, M.E.R. and DRIVER, J., 2001, The cost of expecting events in the wrong sensory modality. *Perception and Psychophysics*, **63**, 330–336.
- SPENCE, C. and SQUIRE, S., 2003, Multisensory integration: Maintaining the perception of synchrony. *Current Biology*, **13**, R519–R521.
- TASSINARI, G. and CAMPARA, D., 1996, Consequences of covert orienting to non-informative stimuli of different modalities: A unitary mechanism? *Neuropsychologia*, **34**, 235–245.
- TOWNSEND, J.T. and ASHBY, F.G., 1983, *Stochastic Modeling of Elementary Psychological Processes* (Cambridge: Cambridge University Press).
- TURATTO, M. and GALFANO, G., 2001, Attentional capture by color without any relevant attentional set. *Perception and Psychophysics*, **63**, 286–297.
- VAN ERP, J.B.F., 2005, Presenting directions with a vibrotactile torso display. *Ergonomics*, **48**, 302–313.
- VAN ERP, J.B.F. and VAN VEEN, H.A.H.C., 2001, Vibro-tactile information processing in automobiles. *Proceedings of EuroHaptics 2001* (Birmingham, UK: University of Birmingham), pp. 99–104.
- VAN ERP, J.B.F. and VAN VEEN, H.A.H.C., 2004, Vibrotactile in-vehicle navigation system. *Transportation Research Part F: Traffic Psychology and Behaviour*, **7**, 247–256.
- WARREN, D.H., 1970, Intermodality interactions in spatial localization. *Cognitive Psychology*, **1**, 114–133.
- ZAMPINI, M., BROWN, T., SHORE, D.I., MARAVITA, A., RÖDER, B. and SPENCE, C., 2005, Audiotactile temporal order judgments. *Acta Psychologica*, **118**, 277–291.