The Spatial Resolution of Crossmodal Attention: Implications for the Design of Multimodal Interfaces

ROB GRAY and RAYKA MOHEBBI
Arizona State University
and
HONG Z. TAN
Purdue University

Previous research on crossmodal attentional orienting has reported speeded reaction times (RT) when the stimuli from the different modalities are in the same spatial location and slowed RTs when the stimuli are presented in very different locations (e.g., opposite sides of the body). However, little is known about what occurs for spatial interactions between these two extremes. We systematically varied the separation between cues and targets to quantify the spatial distribution of crossmodal attention. The orthogonal cueing paradigm [Spence et al. 1998] was used. Visual targets presented above or below the forearm were preceded by either vibrotactile cues presented on the forearm, auditory cues presented below the forearm, or visual cues presented on the forearm. The presentation of both unimodal and crossmodal cues led to a roughly monotonic increase in RT as a function of the cue-target separation. Unimodal visual cueing resulted in an attentional focus that was significantly narrower than that produced by crossmodal cues: the distribution of visual attention for visual cues had roughly half of the lateral extent of that produced by tactile cueing and roughly one fourth of the lateral extent as that produced by auditory cueing. This occurred when both seven (Experiment 1) and three (Experiment 2) cue locations were used suggesting that the effects are not primarily due to differences in the ability to localize the cues. These findings suggest that the location of tactile and auditory warning signals does not have to be controlled as precisely as the location of visual warning signals to facilitate a response to the critical visual event.

Categories and Subject Descriptors: H.1.2 [**User/Machine Systems**]: Human Factors; H.5.2 [**User Interfaces**]: Haptic I/O

General Terms: Experimentation, Human Factors

Additional Key Words and Phrases: Attention, warnings

ACM Reference Format:

Gray, R., Mohebbi, R., and Tan, H. Z. 2009. The spatial resolution of crossmodal attention: Implications for the design of multimodal interfaces. ACM Trans. Appl. Percpt. 6, 1, Article 4 (February 2009), 14 pages. DOI = 10.1145/1462055.1462059 http://doi.acm.org/10.1145/1462055.1462059

1. INTRODUCTION

How is attention distributed across space when directed to a single location? This is an important question not only for theories of attention, but also for applications such as the development of warnings/

This research was supported by a National Science Foundation grant (award #0533908) to authors R.G. and H.Z.T.
Authors' addresses: R. Gray and R. Mohebbi, Department of Applied Psychology, Arizona State University, Sutton 340E, 7001 East Williams Field Road, Mesa AZ, 85212, USA; H. Z. Tan, Purdue University, School of Electrical and Computer Engineering, 1285 Electrical Engineering Building, West Lafayette, Indiana 47907-1285, USA.

alerts in aviation, command and control, and driving. In order to effectively direct an operator's attention to a critical event, one of the first design issues that must be considered is: How close in space must the warning signal be to the critical event?

Results from several experiments using visual cues and visual targets have shown that the facilitatory effects of attention on response performance (e.g., faster reaction times, improved sensitivity) decrease monotonically with increasing distance between the cue and target location (e.g., Downing et al. [1985] and Shulman et al. [1985])—often referred to as the attentional gradient. Furthermore, it has been shown that the spatial distribution of visual attention can be quite broad with the monotonic decrease, continuing up to separations between cue and target as large as 20 deg of visual angle [Sheppard and Muller 1989]. This would suggest that warning signals and alerts could have a positive effect on performance even when not presented in precisely the same spatial location as the critical event.

One limitation of previous research on the distribution of attention is that it has primarily considered only visual cueing of visual attention. Over the past several years, that has been a growing interest in crossmodal attention (i.e., shifting attention in one modality using cues from a different modality). Because the visual system is often overloaded in control tasks such as driving and flying, warning signals presented in other modalities may be more detectable by the operator and therefore may shift visual attention more effectively than visual warnings. The finding that attentional facilitation of response performance can also occur crossmodally has inspired researchers and interface designers alike. For example, it has been shown that reaction times for a visual task can be reduced by roughly 30 ms when preceded by a tactile cue and by roughly 20 ms when preceded by an auditory cue [Spence et al. 1998]. Note that these relatively small magnitude cueing effects are from data collected in a controlled laboratory environment where participants can easily focus most of their attentional resources on the detection task. It is expected that these values would be larger in complex real-world environments, and indeed, there is some support for this proposal [Ferris et al. 2006]. This line of basic research has been paralleled by the development of multimodal human-machine interfaces for aviation (e.g., Rupert [2000a, 2000b]), human-computer-interaction (e.g., Edworthy et al. [1991]), and driving (e.g., Tan et al. [2003], Ho et al. [2005], and Ho and Spence [2005]). Unfortunately, because research on crossmodal attention cueing is in its infancy compared to research on unimodal visual cueing, most previous studies have only compared targets in very different spatial locations (e.g., on opposite sides of the body or opposite sides of the visual field). Therefore, little is known about what occurs for spatial interactions between these two extremes: "It will clearly also be important for future research to address more carefully the spatial distribution of attention following the peripheral presentation of (crossmodal cues)" [Spence 2001]. A systematic parametric investigation of the spatial distribution of crossmodal attention is key to the development of intuitive and efficient multimodal interfaces. Without it, we run the risk of wasting resources creating human-machine interfaces that are well below or well above the capacities of the human sensory system and/or wasting time finding the optimal parameters through trial and error. On a theoretical level, a comparison between the well-known pattern of response facilitation that occurs for unimodal cueing of the visual system (the "spotlight of attention"— Posner [1980]), and the facilitation pattern for crossmodal cueing will have important implications for the ongoing debate between modality-specific and supramodal theories of attention (e.g., Spence and Driver [1996]).

The studies that have performed a more fine grain analysis of the spatial distribution of attention for crossmodal cues have primarily examined audiovisual interactions. Driver and Spence [1998] and Schmitt et al. [2001] measured reaction times in a visual localization task using four possible cue-target locations (two to the left and two to the right of fixation). The main finding from both studies was that RT increased roughly monotonically as a function of cue-target separation for both crossmodal auditory

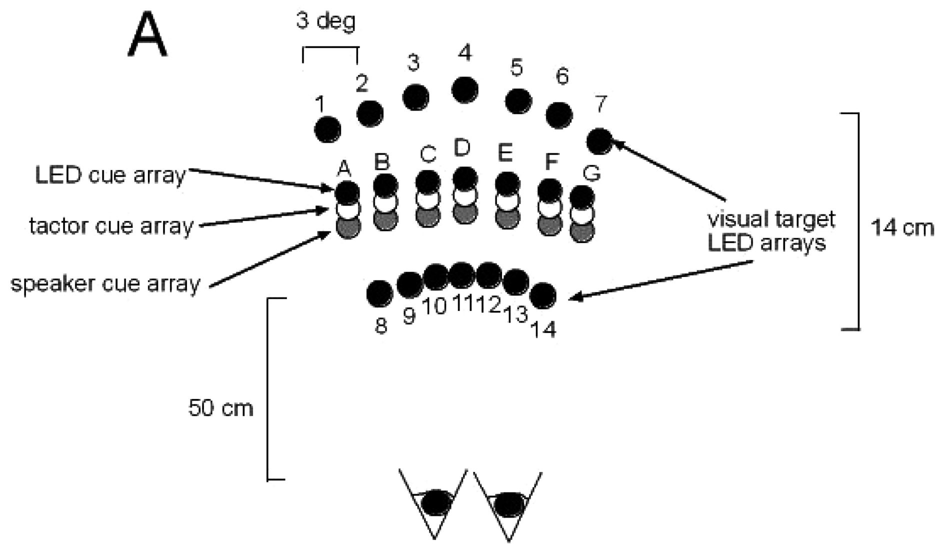

Fig. 1. Cueing apparatus. Visual targets consisted of 14 light-emitting diodes (LEDs) presented in two arrays (#1–7 and #8–14). The participant's task was to indicate whether the illuminated target LED was in the near or far array by pressing one of two response buttons. The presentation of the visual target was preceded by either a tactile, auditory, or visual cue. Tactile cues were vibrotactile pulses delivered via a frontoparallel array of seven tactors (white circles). Auditory cues were burst of white noise presented via a frontoparallel array of seven loudspeaker cones (grey circles). Visual cues were presented via an array of seven LEDs (solid circles).

and visual cueing. Frassinetti et al. [2002] examined the spatial specificity of crossmodal (audio/visual) cueing using eight cue/target locations and a masked visual detection task. Their main finding was that when the auditory cue preceded the visual target by 500 ms there was no effect of cue/target separation on performance. To our knowledge, no previous studies have examined the distribution of attention for crossmodal tactile cues for more than two cue/target locations.

The purpose of the present study was to compare the spatial distribution of attention for crossmodal and unimodal cueing for both auditory and tactile cues using a large range of cue-target separations.

2. EXPERIMENT 1

2.1 Purpose

In Experiment 1, we investigated the spatial distribution of attention for two types of crossmodal attentional orienting (tactile cue/visual target and auditory cue/visual target) and compared this to the spatial distribution for unimodal visual attentional orienting (i.e., visual cue/visual target).

2.2 Methods

2.2.1 *Stimuli and Apparatus.* In the present study, participants were always required to make discriminations about visual targets. As illustrated in Figure 1, visual targets consisted of 14 light-emitting diodes (LEDs) presented in two arrays oriented in semicircular patterns relative to the plane of the participant's eyes. The lateral separation between adjacent LEDs was 3 deg of visual angle (measured at the eye). The two arrays were placed on a table in front of the participant with the center LED of each array located 50 cm and 64 cm from the participant's eyes, respectively. On each trial, one of the 14 LEDs was turned on for a duration of 150 ms.

The presentation of the visual target was preceded by either a tactile, auditory, or visual cue using the apparatus illustrated in Figure 1. Tactile cues were vibrotactile pulses delivered via a frontoparallel array of seven tactors (2.54 × 1.85 × 1.07 cm, VBW32, Audiological Engineering Corp, Somerville, MA) mounted in foam and located between the two LED target arrays. The tactors were driven by a 290 Hz sinusoidal signal. Auditory cues consisted of five 15 ms bursts of white noise presented at 30 ms intervals via a frontoparallel array of seven loudspeaker cones (2.5 cm diameter) that were located below the tactor array. Visual cues were presented via an array of seven LEDs located directly above the tactor array. For all modalities, the cue duration was 195 ms. For the tactile cueing conditions, participants were presented with white noise over headphones to eliminate any sound produced by the tactors.

Participants placed their left forearm between the two LED target arrays such that the underside of their forearm was in contact with the tactor array and the LED cue array was above their forearm. Participants used their right hand to control the two response buttons.

2.2.2 *Procedure.* A modified version of the orthogonal cueing paradigm [Spence et al. 2000] was used. Each trial began with the participant fixating at the center LED in the visual cue array (i.e., location D in Figure 1A) which was illuminated for a duration of 2 sec. Participants were instructed to maintain fixation on this location throughout the trial. On 12 out of 110 trials (i.e., 10.9%), the fixation LED was turned off after a duration of only 500 ms. Participants were instructed to press the left response button whenever they detected this short duration stimulus. These catch trials were used to ensure that our participants were maintaining fixation on the center LED. Within 100 to 300 ms after fixation light offset, a cue signal was presented at one of the seven cue locations for a duration of 195 ms. After an SOA of 200 ms, a visual target signal was presented at one of the 14 visual target locations. The participant's task was to indicate whether the visual target appeared in the near or far array of LEDs by pressing one of two response buttons. Note that the cue location (varying left-to-right) is orthogonal to the target location (varying near to far). Reaction time for each response was recorded. The intertrial interval (ITI) was 2 sec.

All possible combinations of cue and target location occurred with equal probability so that the cues were spatially uninformative to the participant's task (i.e., cues did not reliably indicate the lateral location of the visual target). Data for tactile, auditory, and visual cues were collected in separate blocks.

Each run was comprised of 110 trials (representing all possible combinations of the 14 visual target locations and 7 cue locations plus 12 catch trials) presented in random order. Participants completed 10 repeats for each of the three cue modalities for a total of 30 runs. The order of cue modalities was counterbalanced across participants. Prior to completing these experimental trials, all participants completed three practice runs (one for each of the cue modalities). Participants were informed that the cues would not be spatially informative with regard to the lateral location of the visual target.

2.2.3 *Data Analysis.* To allow for comparison between the different cue locations data were organized in terms of the lateral separation between the cue and target. For example, in Figure 1(a), cue location "B"/target location 9 and cue location "D"/target location 4 both have a lateral separation of 0 deg. This grouping resulted in lateral separations ranging from –18 to +18 deg where negative values indicate that the cue was to the right of the target and positive values indicate the cue was to the left of the target. The grouped data were then analyzed using a 3 × 13 within-subjects ANOVA with cue modality (visual, auditory, and tactile) and cue-target lateral separation (–18, –15, –12, –9, –6, –3, 0, 3, 6, 9, 12, 15, 18 deg) as factors.

2.2.4 *Participants.* Six participants (two male and four female) completed Experiment 1. Participants 1 through 5 were naïve to the aims of the experiment. Participant 6 was author R. M.

2.3 Results

2.3.1 *Reaction Time as a Function of Cue-Target Separation.* Figure 2 shows mean RTs as a function of lateral cue-target separation for the six participants in Experiment 1. Consistent with previous research on unimodal visual cueing, there was a roughly monotonic increase in mean RT as a function of cue-target lateral separation for the visual cue (solid circles). For both the tactile cue (open squares) and auditory cues (solid triangles), it appears that mean RT did not increase as quickly as a function of cue-target separation, as found for the visual cues.

To quantify these data we fit quadratic functions (i.e., of the form $y = ax^2 + bx + c$) to the data in Figure 2. These fits are shown with solid (visual), dashed (tactile), and dotted (auditory) lines. In particular, we were interested in the value of parameter a in the curve fit. This parameter is proportional to the curvature of the quadratic function or in other words, the magnitude of the spatial spread of attention following cueing. Table I shows a and R^2 values for each of the curve fits shown in Figure 2. For all six participants, the visual cues had the largest value of a followed by the tactile cues then by the auditory cues. A one-way ANOVA performed on parameter a values revealed a significant effect of cue modality: $F(2, 10) = 11.4$, $p < 0.01$. As an alternative means of quantifying the spread of attention for the different modalities (that does not rely on the quadratic curve fits), we calculated the mean difference in RT for the 0 deg cue-target separation and the RT for the maximum cue-target separation (18 deg). Mean values for this parameter were as follows: visual cueing (M = 61.3, SE = 12.7 ms), tactile cueing (M = 30.1, SE = 4.6 ms), and tactile cueing (M = 13.8, SE = 3.6 ms). A one-way ANOVA performed on these data also revealed a significant effect of cue modality: $F(2, 10) = 16.9$, $p < 0.001$.

We next turn to an overall analysis of Experiment 1 data. Figure 3 shows the mean RTs as a function of cue-target lateral separation averaged across the six observers. The 3 × 13 within-subjects ANOVA described in the Methods section revealed significant main effects of cue modality ($F[2, 10] = 5.2$, $p < 0.05$) and cue-target lateral separation [$F[12, 60] = 14.1$, $p < 0.001$]. The cue modality × cue-target lateral separation was also significant ($F[24, 120] = 4.2$, $p < 0.001$].

2.3.2 *Reaction Times for Ipsilateral vs. Contralateral Hemisphere Cue-Target Pairs.* Previous research on saccadic generation has shown that the effectiveness of crossmodal orienting depends on whether or not the cue and target are presented in the same or different hemisphere [Colonius and Arndt, 2001; Diederich et al. 2003]. For example, saccades to a visual target are faster when preceded by a tactile cue presented in the same hemisphere (ipsilateral) as compared to when the cue and target are presented in different hemispheres (contralateral). Did the same effect occur for attentional orienting in the present study? To address this question we analyzed RTs for cue target pairs that had the same lateral separation (6 deg) but differed on laterality (ipsilateral vs. contralateral). We choose not to include cues and targets presented on the body midline (i.e., 4 and 11) for this analysis. Referring to Figure 1, ipsilateral cue/target pairs included in this analysis were {C, 1}, {C, 8}, {E, 7}, and {E, 14} and contralateral cue/target pairs included were {C, 5}, {C, 12}, {E, 3}, and {E, 10}. Mean RTs for this analysis are shown in Figure 4. We analyzed these data using a 2 × 3 repeated measure ANOVA with laterality and modality as factors. This analysis revealed significant main effects of modality ($F[2, 10] = 44.6$, $p < 0.001$), laterality ($F[1, 5] = 38.8$, $p < 0.01$), and a significant modality × laterality interaction ($F[2, 10] = 5.5$, $p < 0.05$).

2.3.3 *Response Accuracy.* To ensure that the cueing results described above were not the result of speed-accuracy trade-offs we also examined response accuracy. These data are shown in Figure 5. As shown in this figure, response accuracy was high (>80%) in all conditions. A 3 × 13 within-subjects ANOVA revealed no significant effects of cue modality or cue-target lateral separation on response

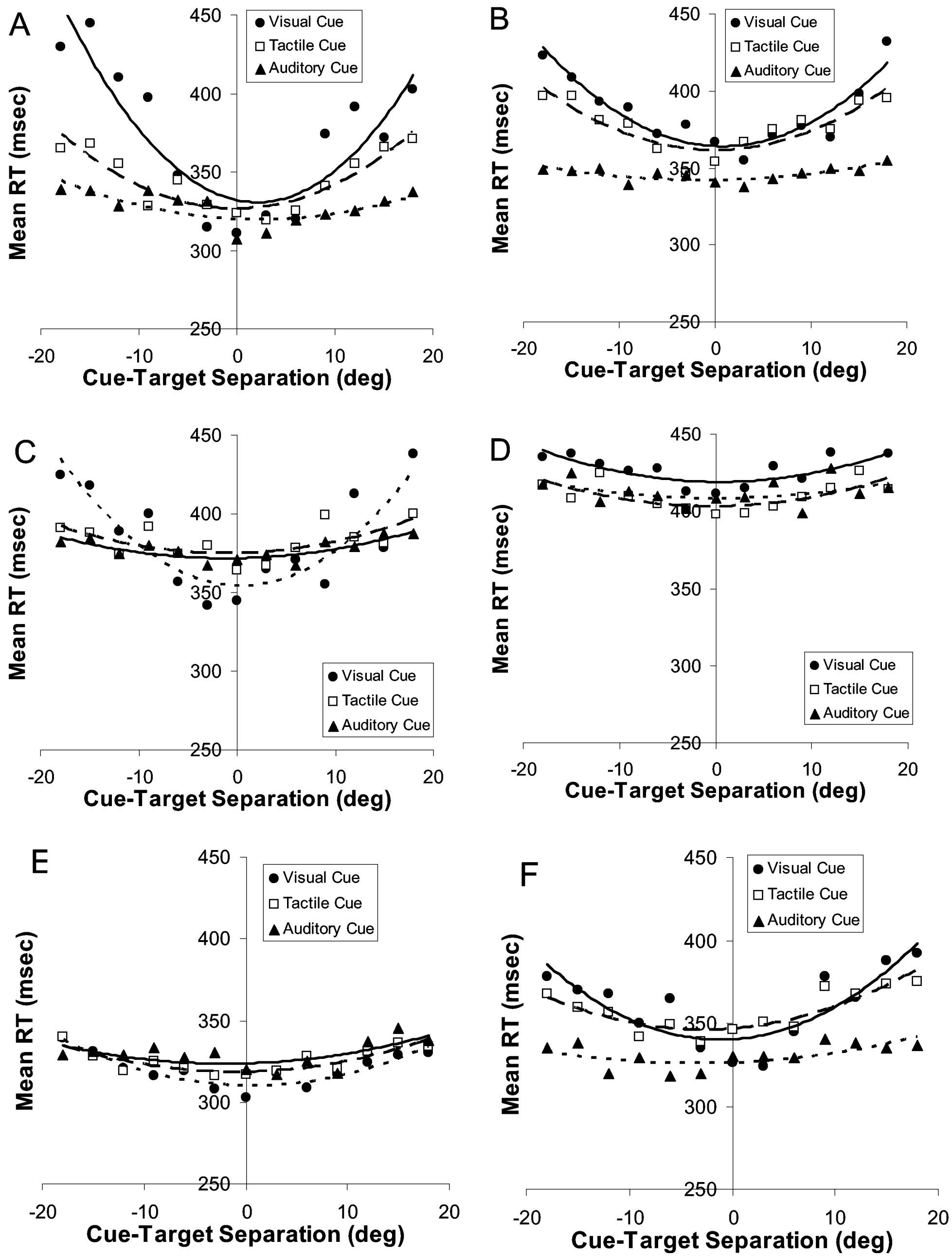

Fig. 2. Mean reaction times as a function of the lateral separation between the cue and target for the three cueing modalities used. Negative values indicate that the cue was to the right of the target and positive values indicate the cue was to the left of the target. Solid lines are quadratic curve fits to the data. Panels A–F are for participants 1–6 respectively.

Table I. Quadratic Curve Fit ($y = ax^2 + bx + c$) Parameters for Experiment 1

Participant	Visual a	Visual R^2	Tactile a	Tactile R^2	Audio a	Audio R^2
1	0.32	0.78	0.15	0.83	0.06	0.55
2	0.18	0.87	0.12	0.77	0.03	0.60
3	0.24	0.75	0.06	0.41	0.05	0.67
4	0.06	0.58	0.05	0.49	0.03	0.29
5	0.08	0.84	0.06	0.74	0.04	0.44
6	0.16	0.72	0.08	0.78	0.03	0.43

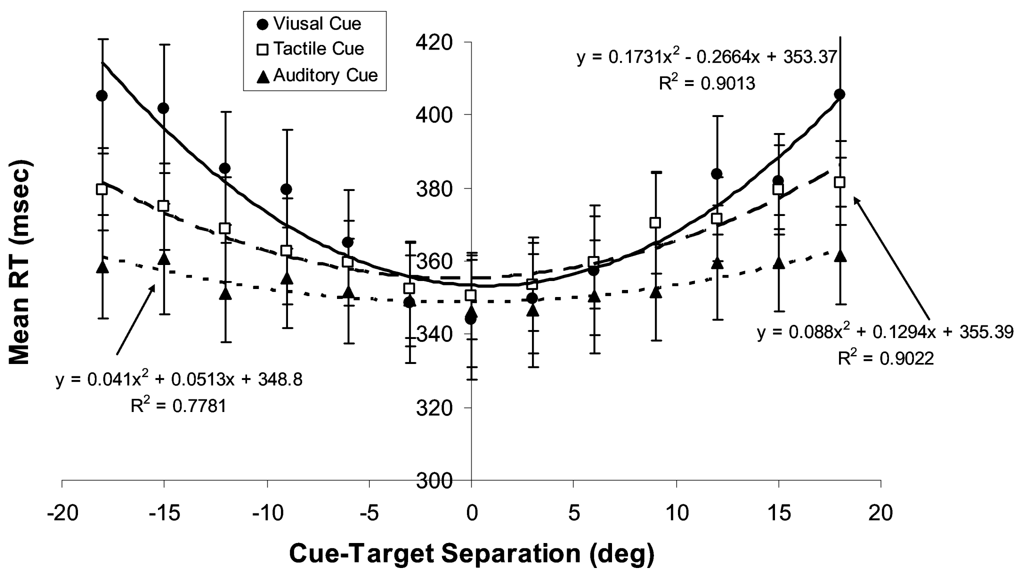

Fig. 3. Mean reaction time as a function of the lateral separation between the cue and target averaged across the six participants. Solid lines, equations, and R^2 values are for quadratic curve fits to the data. Error bars are standard errors.

accuracy (all $p > 0.5$) suggesting that the effects of modality and cue-target separation on RT were not due to a speed-accuracy tradeoff.

2.4 Discussion

The presentation of visual, tactile, and auditory cues had a significant effect on reaction times for the near/far visual judgment in our study. Given that the cues were spatially uninformative about the visual task being performed, these results suggest that the cues used in the present study lead to reorienting of visual attention toward the cued location [Spence, Nicholls, Gilliespie, & Driver, 1998]. Consistent with previous research on unimodal and crossmodal attentional orienting (e.g., Shulman et al. [1986] and Spence and Driver [1997]), RTs were faster when the cue and target were presented in the same spatial location (0 deg) as compared to when they were presented in very different locations for all three cueing modalities. For example, the mean difference between RTs for the 18 deg separation and 0 deg separation conditions in the present study were 61.1, 27.2, and 12.3 ms for the visual, tactile, and auditory cues, respectively. The later two values are consistent with the 30 and 20 ms values previously reported for crossmodal tactile and auditory cues using a 40 deg separation between cue and target [Spence and Driver 1997]. We next turn to the primary goal of the present study—namely to examine what happens when the cue/target separation is systematically increased from zero.

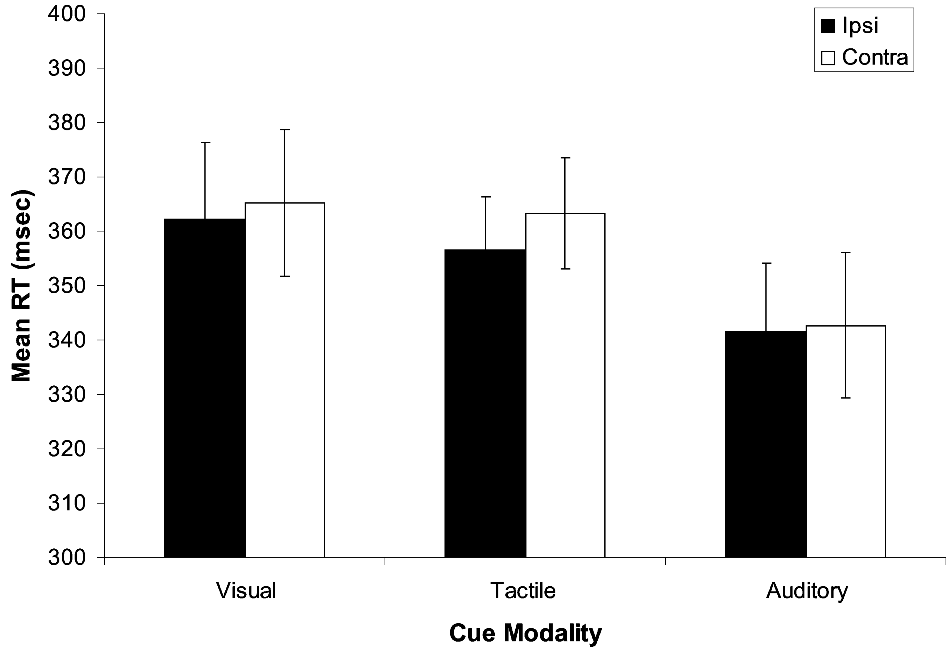

Fig. 4. Mean reaction times for cue/target pairs presented in the same hemisphere (ipsilateral) and different hemispheres (contralateral). Error bars are standard errors. See text for details.

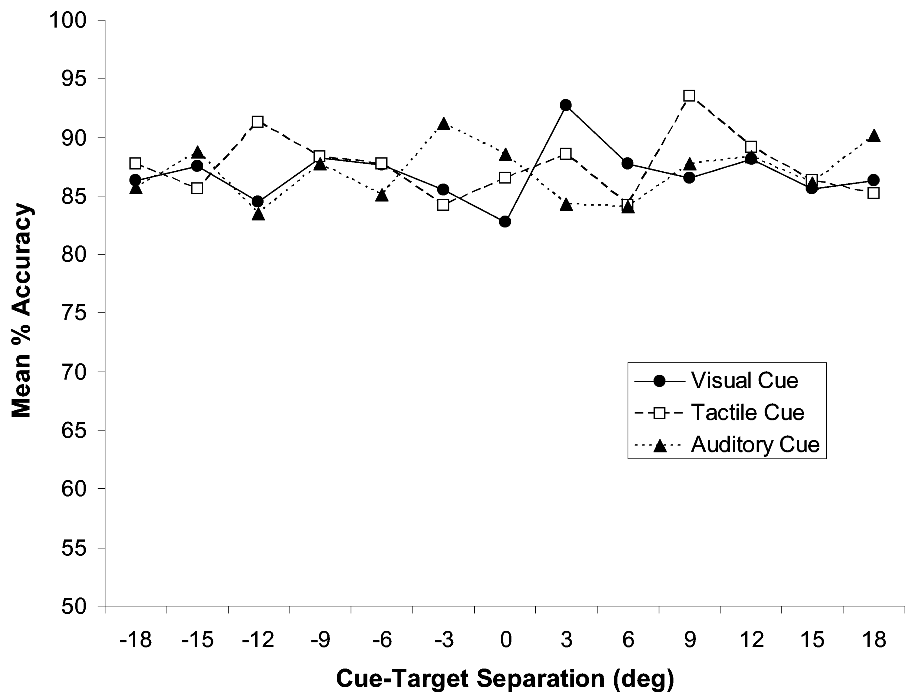

Fig. 5. Mean percentage accuracy for the visual task as a function of the lateral separation between the cue and target averaged across the six participants.

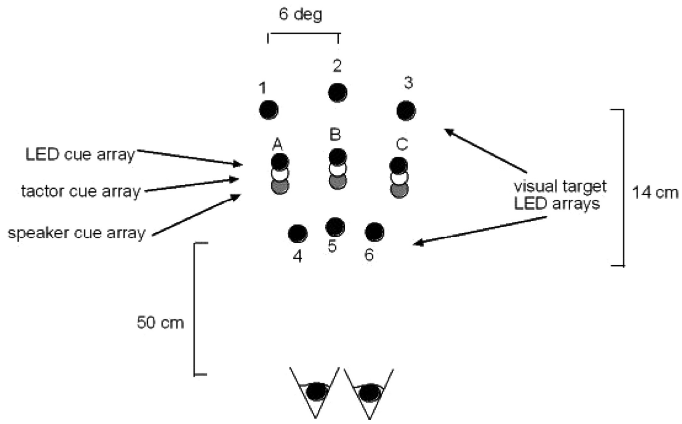

Fig. 6. Cueing apparatus for Experiment 2.

The presentation of both unimodal and crossmodal cues led to a monotonic increase in reaction time as a function of the cue-target lateral separation. These findings provide evidence for an attentional gradient for crossmodal orienting. In other words, tactile and auditory cues shift attention to a relatively small area in visual space rather than distributing it diffusely across an entire hemifield for example. However, it was not the case that the reorienting of visual attention was identical for the visual, tactile, and auditory cues used in the present study. Unimodal visual cueing resulted in an attentional focus that was significantly narrower than that produced by crossmodal cues: the distribution of visual attention for visual cues had roughly half the lateral extent of that produced by tactile cueing and roughly a one quarter of the lateral extent as that produced by auditory cueing. Possible explanations for these differences are discussed below.

Consistent with previous research on eye movements [Diederich et al. 2003], tactile cues that were presented in the same hemisphere as the visual target (ipsilateral) resulted in significantly faster RTs than tactile cues presented in the opposite hemisphere (contralateral) even though the lateral cue-target separation was identical in both cases. This effect may result from the different frames of reference used by vision, touch, and audition, for example, as modeled by Diederich and Colonius [in press]. Unlike in vision and audition, tactile/somatosensory neurons appear to code space in a body-centric frame that is updated as the limbs are moved, and, therefore, multisensory integration differs for ipsilateral and contralateral tactile stimuli (see also Eimer and van Velzen [2005]).

3. EXPERIMENT 2

3.1 Purpose

One possible explanation for the different attentional gradients found for visual, tactile, and auditory cues is that our observers found it more difficult to localize the seven cue locations for auditory and tactile cues then for visual cues. Indeed, for the apparatus used in Experiment 1 the accuracy of cue localization was as follows: visual (M = 99.3%, SE = 0.2), tactile (M = 90.0%, SE = 0.9), and auditory cues (M = 74.2%, SE = 1.1). To partially address this possibility, we used only three cue locations in Experiment 2.

3.2 Methods

3.2.1 *Stimuli & Apparatus.* The stimuli and apparatus were identical to that described in Experiment 1 except that (as shown in Figure 6) several of the cue and target locations were removed. In Experiment 2, there were three cue locations and six possible target locations.

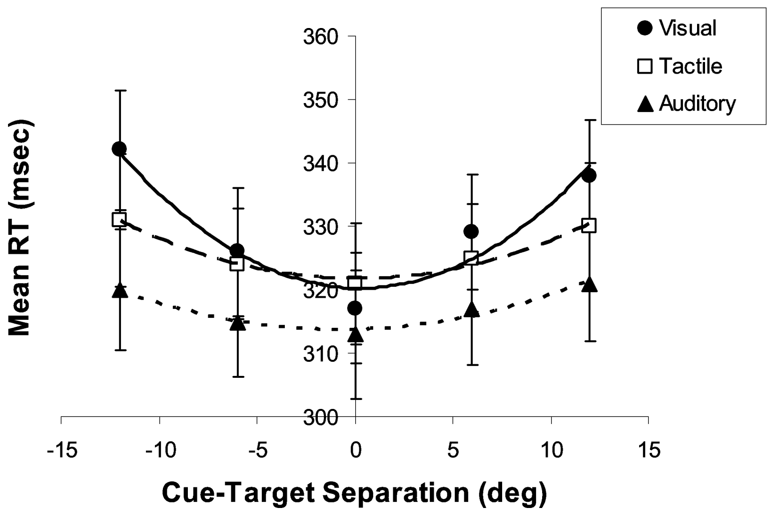

Fig. 7. Mean reaction times (averaged across the six participants) as a function of the lateral separation between the cue and target for the three cueing modalities used in Experiment 2. Negative values indicate that the cue was to the right of the target and positive values indicate the cue was to the left of the target. Solid lines are quadratic curve fits to the data.

3.2.2 *Procedure.* The procedure was identical to that described for Experiment 1 except that each run was comprised of 20 trials (representing all possible combinations of the six visual target locations and three cue locations plus two catch trials) presented in random order. Participants completed 10 repeats for each of the three cue modalities for a total of 30 runs.

3.2.3 *Data Analysis.* To quantify the attentional gradients we calculated the mean difference in RT for the 0 deg cue-target separation and the RT for the maximum cue-target separation (12 deg) for each of the three modalities and performed a one-way ANOVA.

3.2.4 *Participants.* Six participants that did not participate in Experiment 1 took part in Experiment 2.

3.3 Results & Discussion

Figure 7 shows mean RTs as a function of lateral cue-target separation averaged across the six participants in Experiment 1. Visual inspection of this figure suggests that, consistent with the findings of Experiment 1, the attentional gradient was shallower for auditory and tactile cues then for visual cues. For the quadratic curve fits shown in Figure 7, the a (i.e., curvature) and R^2 values were as follows: visual ($a = 0.14$, $R^2 = 0.92$), tactile ($a = 0.06$, $R^2 = 0.91$), and auditory ($a = 0.04$, $R^2 = 0.95$). The mean differences in RT for the 0 deg cue-target separation and the RT for the maximum cue-target separation (18 deg) were as follows: visual cueing (M = 25.4 ms), tactile cueing (M = 10.1 ms), and tactile cueing (M = 7.5 ms). A one-way ANOVA performed on these data revealed a significant effect of cue modality: $F(2, 10) = 17.1, p < 0.001$. Finally, we performed a control experiment to measure localization accuracy for the cues used in Experiment 2. For all six observers, the cue location accuracy was 100% for all three modalities. Together, these findings suggest that the differences in attentional gradients for visual, auditory, and tactile cues found in Experiment 1 were not primarily due to differences in the ability to localize the cues.

4. GENERAL DISCUSSION

4.1 Implications for the Design of Multimodal Warning Systems

The design guidelines for developing effective unimodal visual warnings have been well documented (e.g., Edworthy and Adams [1996]) and in many ways have become common sense. When using a visual warning to draw an operator's attention to the location in space where a critical visual event is occurring (e.g., an oncoming vehicle), motor responses to the critical event (e.g., brake RT) will be fastest when the warning and event occur at the same spatial location. Furthermore, the warning will result in some (albeit less) facilitation of performance if it is not perfectly spatially coincident with the event as long as the spatial separation is kept relatively small, within a range of roughly ± 10 deg [Sheppard and Muller 1989]. Examples of this logic can be found in rear-end collision warnings mounted on the dashboards of commercial trucks and ground enemy approach warnings presented via head-up displays in military aircraft [Wiener and Nagel 1988]. Do the same rules apply for crossmodal warnings? Will an auditory or tactile warning designed to draw attention to a critical visual event be more effective if it is spatially coincident with the event? Is it worth the cost and effort to design crossmodal warnings that can be presented near the critical event (e.g., 3D audio)? In order to answer these questions, we first need to examine how attention is distributed across the visual field when it is reoriented to a specific location in space by crossmodal cues.

In the present study, we found that both tactile and auditory cues lead to a localized distribution of visual attention such that reaction times to a visual target were fastest when the cue and target were spatially coincident and increased roughly monotonically as the cue-target separation was increased. Thus, it appears that crossmodal warning signals should follow similar guidelines to those used for unimodal visual warnings (i.e., for best results, place the warning as close to the critical event as possible). However, there are two important differences between unimodal and crossmodal attentional orienting that should also be considered by the designer.

It is clear from the results of the present study that the spatial tuning of attention was not the same for the three different cueing modalities. If we express the tuning in terms of the half-height of the functions shown in Figure 3, the spatial extents were ± 9, ± 14, and ± 17 deg for the visual, tactile, and auditory cues, respectively. This suggests that the location of tactile and auditory warnings signals does not have to be controlled as precisely as the location of visual warnings to facilitate a response to the critical visual event. It should also be noted that there is some individual variability in the distribution of attention following the presentation of crossmodal cues. As shown in Figure 2, some of our participants produced a relatively narrow tuning for tactile cues (e.g., A, B, and F) while others produced broader tuning functions (e.g., C, D, and E). In terms of multimodal interfaces, the present results suggest that designers should use data from individuals with "narrow tuning" to ensure that attention is effectively shifted to the desired location for all users. If data from individuals with "broad tuning" were used to determine the tolerance for separation between cues and targets, the present results suggest that multimodal warnings may not facilitate responses to a critical event for individuals with "narrow tuning" (i.e., since the event will be outside their attentional focus).

In some cases, tactile cues can be used to effectively reorient visual attention for very large cue-target separations if a surface on the body is mapped onto a set of locations in external space. For example, it has been shown that tactile cues presented to a user's back can be used to reorient visual attention to specific locations on a visual display in front of the user [Tan et al. 2003]. This type of orienting has been shown to occur automatically with minimal training, occurring for low cue validity rates and eliciting eye movements to the cued location [Jones et al. 2007]. Therefore, there appear to be exceptions to the "cue as close as possible to event" rule for tactile crossmodal warnings (although see Ho, Tan, and Spence [2006]).

4.2 Implications for Theories of Attention

Over the years, it has been debated whether people have completely separate pools of attention for the different sensory modalities (e.g., multiple resource theory [Wickens 1980]) or people have a single supramodal system that allocates attention to locations in space regardless of the stimulus modality (e.g., Farah et al. [1989]). Still, others have argued for hybrid attentional systems. For example, Spence and Driver [1996] have proposed a separable-but-linked attentional system in which attention can be directed independently to representations of visual, auditory, and tactile space while at the same time having links that ensure attention in each of these modalities is directed to the same spatial location when an important event occurs. The results from the present study appear to provide further support for this hybrid model. The auditory and tactile cues used in our study caused shifts in visual attention to the cued location even when these shifts lead to slower RTs suggesting that the control of visual attention is not completely separable from audition and touch. But the fact that the attentional distributions found in our study had different lateral extents for visual, auditory, and tactile cues also suggests that attention control operates differently for the three sensory modalities.

Why were the distributions of spatial attention different for the unimodal and crossmodal cues in our study? The results from Experiment 2 suggest that this effect was not primarily due to differences in the ability to localize the cues (which would presumably affect the precision with which attention could be shifted to a specific location in space). Rather, we propose that this effect is most likely due to the alignment of the sensory maps for the different modalities that is required for sensory integration. Since visual space (retinotopic mapping) is coded in a different coordinate system than either auditory space (head-centric mapping) or tactile space (body-centric mapping) these systems must be remapped into a common reference frame before crossmodal interactions can occur [Stein and Meredith 1993]. This remapping process adds noise to the neural signal used to redirect attention and shift gaze (e.g., Groh and Sparks [1996]) that would not be present for unimodal cues (for which remapping is not required). Another possibility is that the differences in attentional gradients were due to the fact that the visual and auditory cues were pulsed whereas the tactile cue was continuous in the present study. We plan to investigate this possibility in future experiments.

In the present study, we used a constant SOA value of 200 ms. It has been shown that the time course of crossmodal cueing can be different for the different cueing modalities. For example, at an SOA of 150 ms, an auditory cue with a tactile target will speed reaction time while at this same SOA a visual cue with a tactile target actually slows reaction time [Spence et al. 1998]. Therefore, it is possible that the attentional distributions reported here might change for different cue-target SOA values. However, for the particular cue/target combinations used in the present study the effect of SOA are similar for values ranging between 150 and 700 ms [Spence and Driver 1997; Spence et al. 1998]; therefore, we would not expect these distributions to change radically. However, this issue needs to be investigated in future research.

4.3 Future Research

There are additional aspects of crossmodal attentional orienting that need to be investigated to aid in the design of multimodal interfaces. In the present study, we found that spatial correspondence between cues and targets required for effective attentional orienting depends on cue modality. Since it is known that the temporal correspondence (i.e., the cue-target SOA) also has a substantial effect on crossmodal attentional orienting (e.g., Spence and Driver [1997]), it will be important for future research to compare a range of different cue-target SOAs for unimodal and crossmodal cues and to investigate how different cue-target temporal separations interact with cue-target spatial separations. As discussed above, our somatosensory system appears to be able to make transformations between proximal locations on the skin surface and distal areas in external space for attentional orienting [Tan et al. 2003]. It will also be important for future research to explore the distribution of attention produced by this special type of cueing.

REFERENCES

CHASTAIN, G. 1992. Analog versus discrete shifts of attention across the visual field. *Psych. Resear. 54*, 175–181.

COLONIUS, H. AND ADNDT, P. 2001. A two-stage model for visual-auditory interaction in saccadic latencies. *Percep. Psychophy. 63*, 126–147.

DIEDERICH, A. AND COLONIUS, H. 2007. Modeling spatial effects in visual-tactile saccadic reaction time. *Percep. Psychophy.*

DIEDERICH, A., COLONIUS, H., BOCKHORST, D., AND TABELING, S. 2003. Visual-tactile spatial interaction in saccade generation. *Exper. Brain Resear. 148*, 328–337.

DOWNING, C. J. AND PINKER, S. 1985. The spatial structure of visual attention. In M. I. POSNER AND O. S. M. MARTIN, Eds. *Attention and Performance XI*. Erlbaum, Hillsdale, NJ, 171–188.

DRIVER, J. AND SPENCE, C. 1998. Crossmodal links in spatial attention. *Philosoph. Trans. the Royal Society Series B 353*, 1319–1331.

EDWORTHY, J. AND ADAMS, A. 1996. *Warning Design: A Research Prospective.* Taylor & Francis, Bristol, PA.

EDWORTHY, J., LOXLEY, S., AND DENNIS, I. 1991. Improving auditory warning design: relationship between warning sound parameters and perceived urgency. *Human Factors 33*, 205–231.

EIMER, M., AND VAN VELZEN, J. 2005. Spatial tuning of tactile attention modulates visual processing within hemifields: an ERP investigation of crossmodal attention. *Exper. Brain Resear. 166*, 402–410.

FARAH, M. J., WONG, A. B., MONHEIT, M. A., AND MORROW, L. A. 1989. Parietal lobe mechanisms of spatial attention: Modality-specific or supramodal? *Neuropsychologica 27*, 461–470.

FERRIS, T., PENFOLD, R., HAMEED, S., AND SARTER, N. 2006. The implications of crossmodal links in attention for the design of multimodal interfaces: a driving simulation study. In *Proceedings of the 50th Annual Meeting on Human Factors and Ergonomics Society*. HFES, Santa Monica, CA, 406–409.

GROH, J. M. AND SPARKS, D. L. 1996. Saccades to somatosensory targets II. Motor convergence in the primate superior colliculus. *J. Neurophysiology 75*, 428–438.

HO, C., AND SPENCE, C. 2005a. Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *J. Exper. Psych.: Applied 11*, 157–174.

HO, C., TAN, H. Z., AND SPENCE, C. 2005. Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transport. Resear. Part F: Traffic Psych. Behav. 8*, 397–412.

HO, C., TAN, H. Z., AND SPENCE, C. 2006. The differential effect of vibrotactile and auditory cues on visual spatial attention. *Ergonomics 49*, 724–738.

JONES, C. M., YOUNG, J. J., GRAY, R., SPENCE, C., AND TAN, H. Z. 2007. An eyetracker study of the haptic cuing of visual attention, peer reviewed extended abstract. In *Proceedings of the World Haptics Conference (WHC07): The 2nd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, Los Alamitos, CA.

KENNETT, S., SPENCE, C., AND DRIVER, J. 2002. Visuo-tactile links in covert exogenous spatial attention remap across changes in unseen hand posture. *Percep. Psychophy. 64*, 1083–1094.

RUPERT, A. H. 2000a. An instrumentation solution for reducing spatial disorientation mishaps—A more "natural" approach to maintaining spatial orientation. *IEEE Engineering Med. Biology Mag. 19*, 71–80.

RUPERT, A. H. 2000b. Tactile situation awareness system: Proprioceptive prostheses for sensory deficiencies. *Aviation Space Environ. Medicine 71*, A92–A99.

SHEPHERD, M. AND MULLER, H. J. 1989. Movement versus focusing of visual attention. *Percept. Psychophy. 46*, 146–154.

SHULMAN, G. L., REMINGTON, R. W., AND MCLEAN, J. P. 1979. Moving attention through physical space. *J. Exper. Psych.-Hum. Percep. Perform. 5*, 522–526.

SHULMAN, G. L., WILSON, J., AND SHEEHEY, J. B. 1985. Spatial determinants of the distribution of attention. *Percep. Psychophy. 37*, 59–65.

SPENCE, C. 2001. Crossmodal attentional capture: A controversy resolved? In C. FOLK AND B. GIBSON, Eds. Attention, distraction and action: Multiple perspectives on attentional capture. *Advances in Psychology*, *133*. Elsevier Science, Amsterdam, 231–262.

SPENCE, C. AND DRIVER, J. 1996. Audiovisual links in endogenous covert spatial attention. *J. Exper. Psych. Hum. Percep. Perform. 22*, 1005–1030.

SPENCE, C. AND DRIVER, J. 1997. Cross-modal links in attention between audition, vision, and touch: Implications for interface design. *Int. J. Cognitive Ergonomics 1*, 351–373.

SPENCE, C. NICHOLLS, M. E. R., GILLESPIE, N., AND DRIVER, J. 1998. Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. *Percep. Psychophy. 60*, 544–557.

SPENCE, C., PAVAINI, F., AND DRIVER, J. 2000. Crossmodal links between vision and touch incovert endogenous spatial attention. *J. Exper. Psych.-Hum. Percep. Perform. 26*, 1298–1319.

STEIN, B. M. AND MEREDITH, M. A. 1993. *The Merging of the Senses*. MIT Press, Cambridge, MA

TAN, H. Z., GRAY, R., YOUNG, J. J., AND TRAYLOR, R. 2003. A haptic back display for attentional and directional cueing. *Haptics-e*: *Electron. J. Haptics Resear.* 3.

WICKENS, C. D. 1980. The structure of attentional resources. In Nickerson, R. S., Ed. *Attention and Performance VIII*. Lawrence Erlbaum Associates, Hillsdale, NJ, 239–254.

WIENER, E. L. AND NAGEL, D. C. 1988. *Human Factors in Aviation.* Academic Press, New York

Received January 2007; revised June 2007, October 2007; accepted January 2008