

Modeling spatial effects in visual–tactile saccadic reaction time

ADELE DIEDERICH

Jacobs University Bremen, Bremen, Germany

AND

HANS COLONIUS

Oldenburg University, Oldenburg, Germany

Saccadic reaction time (SRT) to visual targets tends to be shorter when nonvisual stimuli are presented in close temporal or spatial proximity, even when subjects are instructed to ignore the accessory input. Here, we investigate visual–tactile interaction effects on SRT under varying spatial configurations. SRT to bimodal stimuli was reduced by up to 30 msec, in comparison with responses to unimodal visual targets. In contrast to previous findings, the amount of multisensory facilitation did not decrease with increases in the physical distance between the target and the nontarget but depended on (1) whether the target and the nontarget were presented in the same hemifield (ipsilateral) or in different hemifields (contralateral), (2) the eccentricity of the stimuli, and (3) the frequency of the vibrotactile nontarget. The time-window-of-integration (TWIN) model for SRT (Colonius & Diederich, 2004) is shown to yield an explicit characterization of the observed multisensory spatial interaction effects through the removal of the peripheral-processing effects of stimulus location and tactile frequency.

Saccades are fast voluntary movements of the eyes that are made in order to align the foveal region of the retina with the spatial position of the stimulus eliciting the eye movement (e.g., Munoz & Schall, 2004). Although saccades are typically made to visual stimuli, they can also be directed to the locations of sounds or tactile stimulation on the body's surface. The saccadic response has become an important paradigm with which to study the effects of multimodal stimulation. In particular, the modulatory effect of an auditory stimulus on saccade onset (saccadic reaction time, SRT) and on saccade trajectory has been studied extensively both in humans and in the monkey. A pervasive result is that SRT for a visual target is reduced by the presence of a spatially and temporally aligned irrelevant auditory stimulus and that this facilitatory effect disappears, or even reverses into inhibition, if the spatiotemporal distance between the stimuli increases (e.g., Colonius & Arndt, 2001; Frens & Van Opstal, 1998; Frens, Van Opstal, & Van der Willigen, 1995; Hughes, Nelson, & Aronchick, 1998; Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994; Lueck, Crawford, Savage, & Kennard, 1990; see Diederich & Colonius, 2004b, and Van Opstal & Munoz, 2004, for reviews).

These behavioral observations can be correlated with the multisensory integration effects measured in neurons of the deep layers of the superior colliculus (SC), a midbrain area involved in saccade generation, in the monkey and the cat (Meredith & Stein, 1986). Information about stimulus location is represented topographically within this struc-

ture by an orderly arrangement of neurons according to the location of their respective receptive fields (RFs). The spatial register among the different sensory maps is formed by multisensory neurons whose different RFs are in register with one another (Stein & Meredith, 1993). Spatially coincident visual–auditory stimuli fall within the overlapping RFs of multisensory SC neurons, producing response enhancements that exceed the individual modality-specific responses (Kadunce, Vaughan, Wallace, & Stein, 2001). In addition, the SC contains a motor map composed of output neurons coding appropriate eye movements (Sparks, 1986; Stanford, 2004). That is, the locus of activity in the motor map encodes a movement command that reorients the eyes (and the head) a given distance in a particular direction. Thus, if the goal is to translate a sensory target into an appropriate motor command, the alignment of the visual and auditory maps with each other and with the motor map is critical. The response enhancement observed in multisensory neurons may serve to facilitate orienting responses to a particular spatial location.

In addition to the visual and auditory maps, the SC also contains a corresponding somatosensory map of the body's surface (Graziano & Gross, 1994; Stein, Magalhães-Castro, & Kruger, 1976). There is increasing evidence (from studies on the monkey or the cat) for the existence of multisensory neurons receptive to both visual and tactile stimuli not only in subcortical areas like the putamen (Graziano & Gross, 1995), but also in several brain areas—for example, in the

A. Diederich, a.diederich@iu-bremen.de

ventral premotor cortex (Graziano & Gross, 1995), the ventral intraparietal area (Duhamel, Colby, & Goldberg, 1998), and the superior temporal sulcus (e.g., Bruce, Desimone, & Gross, 1981). This suggests that not only visual-auditory, but also visual-tactile interactions should occur in saccade generation. Groh and Sparks (1996) found only a slight tendency for SRT reduction when visual and somatosensory stimuli were presented at the same spatial location (in the monkey). In a recent study of visually guided saccades with tactile distractors, Amlôt, Walker, Driver, and Spence (2003) observed (1) that accessory somatosensory stimuli reduced the latency of saccades and (2) that the reduction of SRT was greatest for spatially coincident stimuli. These authors suggested that the enhancement effect reflects a summation at the level of multisensory neurons, although they did not rule out the possibility that a generalized warning signal or alerting effect could be responsible for part of the facilitation. In a similar vein, Diederich, Colonius, Bockhorst, and Tabeling (2003) compared SRTs to unimodal visual targets and to simultaneously presented bimodal stimuli (visual targets and tactile-vibratory nontargets) under various spatial configurations. They found SRTs to bimodal stimuli to be reduced up to 30 msec, in comparison with unimodal visual responses. Contrary to previous results for visual-auditory presentations, however, the magnitude of the speedup did not increase with reductions in the physical distance between the visual and the tactile stimuli. Specifically, with a constant level of visual-tactile distance, SRT facilitation increased with the eccentricity of the stimulus configuration. Obviously, both the absolute and the relative positions of the visual target and the tactile nontarget contribute to the enhancement effect.

In order to make possible a better understanding of these observations, here we will elaborate a recently proposed model of multisensory integration (Colonius & Diederich, 2004) to account for the effects of spatial configuration and will test it on data from our lab. Our time-window-of-integration (TWIN) model distinguishes an early, afferent stage of parallel peripheral processing in the visual and somatosensory pathways (*first stage*), followed by a *second stage* comprising neural integration of the input and preparation of an oculomotor response. In the first stage, visual and tactile activity are processed independently in separate sensory channels.¹ Multisensory integration can occur in the second stage only, manifesting itself through a decrease (or increase) of processing time. Moreover, multisensory integration is contingent upon temporal proximity of the afferent information delivered from the separate modalities: Termination of the afferent peripheral stages for both modalities must occur within a temporal window of fixed duration. Rather than assuming the existence of a joint spatiotemporal window of integration permitting integration to occur only for both spatially and temporally neighboring stimuli, TWIN allows cross-modal interaction to occur even for rather distant stimuli (of different modalities), as long as they fall within the time window. Such interaction will typically be an inhibition or only a small facilitation. This arrangement affords the organism more flexibility in a complex environment. For example, response depression may occur with nearly simultaneous

but distant stimuli, making it easier for the organism to focus on the more important event. The concept of a time window for multisensory integration has been suggested by several authors on the basis of neurophysiological investigations in the human, monkey, and cat (e.g., Corneil, Van Wanrooij, Munoz, & Van Opstal, 2002; Meredith, 2002; Stein & Meredith, 1993).

The TWIN model has been developed for both focused attention (FA) and redundant target (RET) paradigms. In the latter, the subject is instructed to react to stimuli of either modality, whereas in the former a non-target-modality stimulus should be ignored and only the target stimulus responded to. Correspondingly, the TWIN version for FA differs from that for RET with respect to how the time window of integration is "opened." With redundant targets, the winner of the race between the peripheral-processing times in the first stage triggers the opening of the time window. For integration to take place, termination of the second peripheral process must fall within the time window. The RET version of TWIN reflects the symmetry of the role of the modalities in the RET paradigm. It will not be considered any further here, since all the experiments reported below were conducted within the FA paradigm. In the FA paradigm, multisensory integration is assumed to occur only if (1) peripheral processing of the nontarget wins the race and (2) peripheral processing of the target stimulus terminates within a given temporal window of integration opened by the nontarget. This reflects the asymmetry in the role of the modalities: Subjects are not supposed to respond to stimuli of the non-target-modality.²

The window of integration acts like a filter in determining whether the afferent information delivered from different sensory organs is registered close enough in time to elicit multisensory integration. Passing the filter is a necessary, but not a sufficient, condition for multisensory integration to occur, because the latter depends also on the spatial configuration of the stimuli.

Due to its two-stage assumption, the TWIN model suggests an important *separability hypothesis*: It allows the separation of effects on SRT based on unimodal stimulus properties (such as intensity and absolute spatial position) from those based on cross-modal properties (such as physical, or perceived, distance between the target and the nontarget). Unimodal stimulus properties modulate the probability that interaction will occur via their influence on the outcome of the race in the first stage, whereas cross-modal stimulus properties affect the duration of the second-stage processing by their influence on multisensory integration. A formal presentation of the model follows below.

The global assumptions of TWIN have been tested under various conditions, including variations of stimulus intensities and stimulus onset asynchrony (Arndt & Colonius, 2003; Colonius & Arndt, 2001; Colonius & Diederich, 2004; Diederich & Colonius, 2004a). So far, the model has not addressed the specifics of the effects of spatial configuration on SRT. Here, the model is extended; several TWIN variants are developed and tested that differ in how the spatial configuration of the stimuli determines the interaction effects.

EXPERIMENT

In an FA paradigm with visual targets and tactile non-targets, SRT effects of visual–tactile spatial stimulus configuration were studied under different eccentricity and laterality conditions.

Method

Subjects. Three students served as paid voluntary subjects. All the subjects had normal vision. They were screened for their ability to follow the experimental instructions (proper fixation, few blinks during trial, and saccades toward the visual target). They gave their informed consent prior to their inclusion in the study. Local ethical approval was obtained for this study, and all experiments were conducted in accordance with the ethical standards described in the 1964 Declaration of Helsinki.

Apparatus and stimulus presentation. Red light-emitting diodes (LED, 5 mm, 3.7 mcd) served as visual targets presented against a black background. An additional LED (red, 5 mm, 0.4 mcd) served as a fixation point. Tactile stimuli were vibrations (50 Hz/0.6 V or 5 Hz/2.4 V) applied to the center of the palm, generated by two silenced oscillation devices. All the stimuli were positioned on top of a table (180 x 130 x 75 cm) 60 cm in front of the subject. The fixation LED was 38.5 cm away from the lower edge of the table. Fifty-six LEDs and 56 holes for the vibrators were placed at various positions measured from the vertex. For each experimental condition, the two vibrators were moved to the respective positions. Vibrators and LEDs were controlled by a PC multifunction card.

Experimental procedure. The experiment was carried out in a completely darkened room, so that the subjects were unable to see their hands during the experiment. Every session began with 10 min of dark adaptation, during which the measurement system was adjusted and calibrated. During this phase, the subjects put their hands at the position used during the entire experimental block. Thus, the subjects were aware of the hand position and, thus, the position of the tactile stimulus. The subjects sat at the longitudinal side (at the vertex), using a chinrest facing the calibrating screen and wearing a video camera frame. Each trial began with the appearance of the fixation point. After a variable fixation time (800–1,500 msec), the fixation point disappeared, and simultaneously, both a visual and a tactile stimulus were presented for 500 msec (no gap). In unimodal trials, only a visual stimulus was presented. The subjects were instructed to move their eyes to the visual target as quickly as possible, whereas the tactile stimulus could be ignored. The interval between stimulus offset and fixation onset for the next trial was 2 sec. Each subject was first trained for 1,000 trials, not included in the data analysis.

Data collection. Saccadic eye movements were recorded by an infrared video camera system (EyeLink system, Sensomotoric Instruments) with a temporal resolution of 250 Hz and a horizontal and vertical spatial resolution of 0.01° (accuracy of 0.25°–0.5°). Saccades were detected on a trial-by-trial basis, using velocity (22°/sec) and acceleration (8,000°/sec²) criteria. SRT (the time between the onset of the visual stimulus and the onset of the saccadic eye movement), start position of the eye, and end position after the saccade (vertical and horizontal positions in degree of visual angle, relative to the straight-ahead fixation point) were calculated from the controlled data samples.

Stimulus arrangement and design. Three different spatial conditions were run. In Condition 1, visual target distance from the fixation point was kept constant (except for left/right presentation), and the tactile stimulus position was varied across experimental blocks. In each trial, a visual target was presented either left or right from fixation. With the visual stimulus appearing always at 10° (left or right) and the tactile stimulus appearing at 10°, 50°, or 70° (left or right), 12 different bimodal configurations (6 ipsi- and 6 contralateral) and two unimodal (visual) conditions were possible, and a total of 100 trials per configuration were recorded. Collapsing over left/right hemifields, this results in 3 ipsi- and 3 contralateral

configurations, plus one unimodal (LED-only) condition, with a total of 200 trials per condition. Since the subjects were required to put their hands at a fixed position, tactile stimulus presentations were blocked for each position (10°, 50°, or 70°), but the order of the positions was randomized over subjects. Moreover, trials were randomized with respect to laterality (ipsi-/contralateral), modality (uni-/bimodal), and frequency of the tactile stimulus (5 Hz/50 Hz). Within each block, 150 trials were presented; a total of 12 blocks were performed by each subject.

Condition 2 was identical to the first one, except for the following: Whereas in the first, the position of the visual target was held constant and tactile stimulus position was changed over blocks, here tactile stimulation was applied at a fixed position (10° left or right from fixation). Thus, the position of the limbs was held constant over the entire condition. The visual stimulus was presented at various spatial locations, but in order to maintain comparability with the first condition, visual stimulus positions were blocked. Thus, within a block, the subjects could anticipate the visual target position, except for the left/right presentation. The two vibrators were positioned at 10° left and right from the fixation point, 55 cm from the vertex. Visual stimulus positions were 10°, 50°, and 70° left and right from fixation, 50 cm from the vertex. As before, bimodal stimuli were presented ipsi- and contralateral. This resulted in 24 different bimodal configurations (6 ipsi-, 6 contralateral, left or right, for both frequencies of the vibration) and six unimodal (visual) conditions, with 100 trials per configuration. Collapsing over left/right hemifields, this amounted to 6 ipsi- and 6 contralateral configurations, plus three unimodal (LED-only) conditions, with a total of 200 trials per condition (per subject).

In Condition 3, both visual and tactile stimuli were presented at positions 10°, 50°, and 70° left and right from the fixation point, with the tactile stimulus 55 cm and the visual stimulus 50 cm away from the vertex. That is, the eccentricity of both stimuli varied; however, the distance between the target and the nontarget was the same for the ipsilateral presentations. As before, visual stimulus positions were blocked, and the bimodal stimuli were presented ipsi- and contralateral, resulting in 24 different bimodal configurations (6 ipsi-, 6 contralateral, left or right, for both frequencies of the vibration) and six unimodal (visual) conditions, with 100 trials per configuration. Collapsing over left/right hemifields, this amounted to 6 ipsi- and 6 contralateral configurations, plus three unimodal (LED-only) conditions, with a total of 200 trials per condition. Table 1 lists the spatial stimulus arrangement for all three conditions.

Results

Anticipatory saccades with reaction times shorter than 80 msec, responses longer than 500 msec, saccades less than 5° and larger than 20° (target at 10°), and gaze direction errors were excluded from the analysis. Overall, this amounted to less than 1% of the total number of responses. The data analysis presented here is across the 3

Table 1
Stimulus Arrangement for the Three Experimental Conditions

		Position		
Stimulus		±10°	±50°	±70°
Condition 1	Visual	●		
	Tactile	↓	↘	↘
Condition 2	Visual	●	●	●
	Tactile	○	↗	↗
Condition 3	Visual	●	●	●
	Tactile	○	○	○

subjects, since their results were qualitatively consistent.³ Moreover, mean SRTs were computed regardless of the specific side (left or right) of stimulus presentation, because there were no systematic hemifield differences. For each of the spatial Conditions 1–3, a three-way ANOVA with factors of *laterality* (with ipsilateral, contralateral, and unimodal [LED] as levels), *eccentricity* (3), and *frequency* (2) resulted in significant main effects throughout ($p < .001$), with the exception of eccentricity in Condition 1 ($p = .116$). Since the spatial configuration of both stimuli being presented at 10° (left or right) occurred in all of them, SRTs for Conditions 1–3 were subsequently combined (both in the figures and in the model fitting). The same holds for the unimodal SRTs in Conditions 2 and 3.

Figure 1 shows mean SRTs (with standard errors) in Condition 1 as a function of tactile stimulus position. Bimodal SRTs (squares for the LED/50-Hz condition and circles for LED/5-Hz condition) were shorter than SRTs for unimodal responses (diamonds) under both frequencies and for all eccentricity values of the tactile stimulus; responses to bimodal stimuli with a 5-Hz vibration were slower than responses to those with a 50-Hz vibration (post hoc Tukey HSD tests, $p < .001$). There was a significant interaction of frequency with laterality: Ipsilateral SRTs were shorter than contralateral SRTs for the 50-Hz vibration, but not for the 5-Hz vibration ($p < .001$).

Figure 2 shows mean SRTs in Condition 2 as a function of visual stimulus position. Clearly, SRTs increased with increases in the eccentricity of the visual stimulus both ipsi- and contralateral. Again, responses to bimodal stimuli with a 5-Hz vibration were slower than responses to those with a 50-Hz vibration, and both were faster than unimodal (LED) responses ($p < .001$). Bimodal SRTs are

shorter for ipsilateral than for contralateral presentation, but here, the interaction between frequency and laterality was not significant ($p = .641$).

Figure 3 shows mean SRTs in Condition 3 as a function of joint visual-tactile stimulus position. Again, SRTs increased with increases in eccentricity, bimodal responses were faster than unimodal ones, and ipsilateral responses were faster than contralateral ones. Bimodal 5-Hz responses also took longer than 50-Hz responses, but only for ipsilateral presentations (significant interaction of laterality and frequency, $p = .001$).

To summarize, the simultaneous presentation of a tactile nontarget stimulus and a visual target co-occurred with a reduction of SRT by up to 30 msec, in comparison with unimodal responses, the size of this facilitation effect depending on the spatial configuration. The 5-Hz vibratory nontarget condition (1) exhibited a significant facilitation of bimodal responses, as compared with unimodal responses, but it was smaller than the effect of a 50-Hz nontarget; (2) bimodal 5-Hz SRTs also increased with increases in the eccentricity of the visual target position and were longer for contra- than for ipsilateral presentations (Conditions 2 and 3). Bimodal SRTs did not increase with increases in tactile nontarget eccentricity (Condition 1) either for 5 Hz or 50 Hz, and the ipsi-/contralateral effect found for 50 Hz was not observed in the 5-Hz condition.

TIME-WINDOW-OF-INTEGRATION MODEL

First, the basic assumptions of the TWIN model in the FA version will be presented formally, followed by a series of more specific model versions characterizing the dependence of the facilitation effect on spatial stimulus configuration (cf. Colonius & Diederich, 2004).

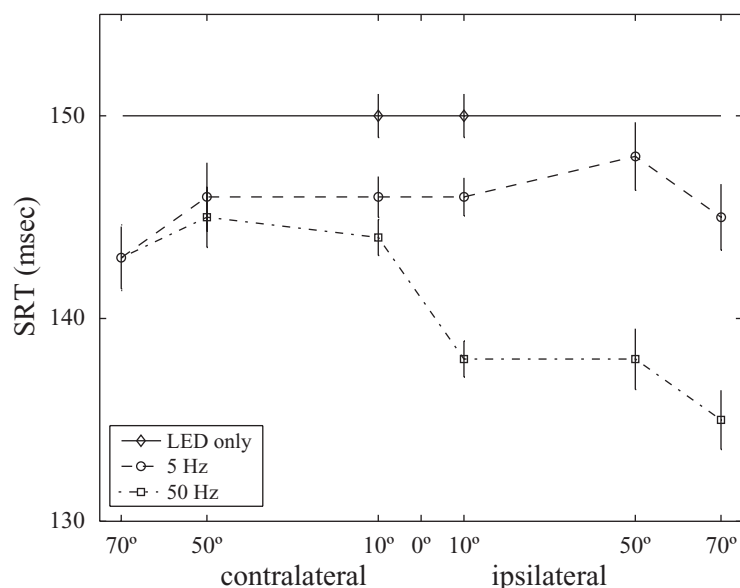


Figure 1. Mean saccade reaction times (SRTs, $\pm SE$ bars) to the visual target at 10° left or right from fixation as a function of tactile stimulus position for 5-Hz and 50-Hz tactile frequency stimuli (Condition 1). The horizontal line refers to unimodal visual stimulus presentation.

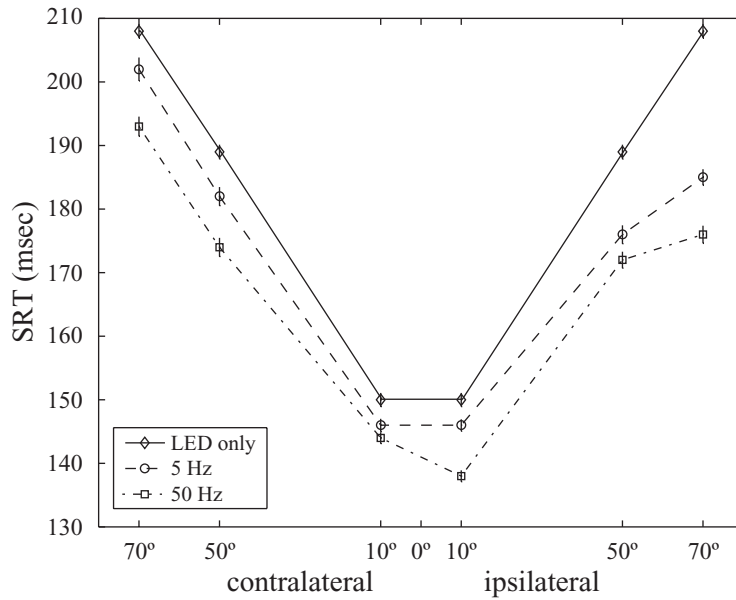


Figure 2. Mean saccade reaction times (SRTs, $\pm SE$ bars) to the visual target as a function of visual stimulus position with tactile stimulus (5-Hz and 50-Hz tactile frequency) at 10° left or right from fixation (Condition 2). The upper graph refers to unimodal visual stimulus presentation.

General Case

SRT is assumed to be the sum of two random variables:

$$SRT = W_1 + W_2, \tag{1}$$

where W_1 and W_2 refer to the first- and second-stage processing times, respectively. Let V denote visual peripheral-

processing time, T tactile peripheral-processing time, and $\Pr(I)$ the probability of the event I that cross-modal interaction will occur according to the model—that is, that T terminates before V and within the time window of length ω :

$$I = \{T < V < T + \omega\}.$$

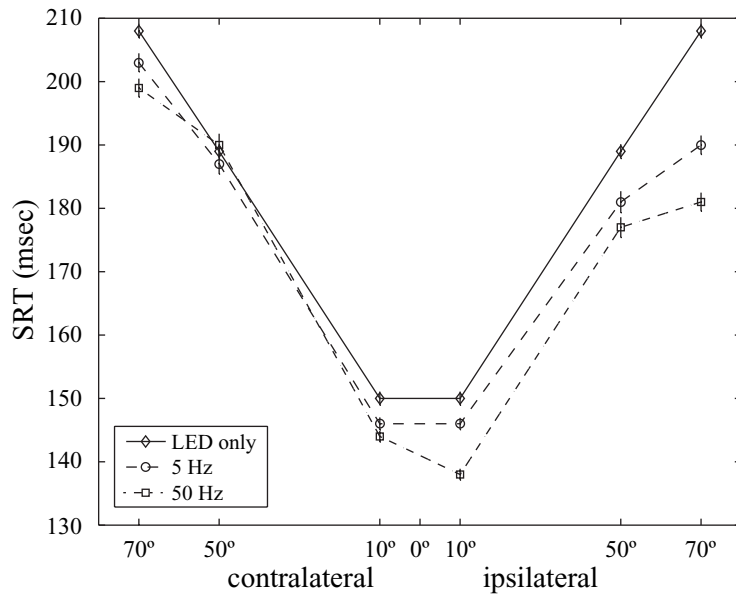


Figure 3. Mean saccade reaction times (SRTs, $\pm SE$ bars) to the visual target as a function of joint visual-tactile stimulus position for 5-Hz and 50-Hz tactile frequency (Condition 3). The upper graph refers to unimodal visual stimulus presentation.

Assuming that first-stage processing time is terminated by peripheral-target-processing time V , mean SRT in the bimodal condition, RT_{VT} , is

$$\begin{aligned} E[RT_{VT}] &= E(W_1) + E(W_2) \\ &= E(V) + E(W_2) \\ &= E(V) + \Pr(I) \cdot E(W_2|I) \\ &\quad + [1 - \Pr(I)] \cdot E(W_2|\text{not-}I) \\ &= E(V) + E(W_2|\text{not-}I) - \Pr(I) \\ &\quad \cdot [E(W_2|\text{not-}I) - E(W_2|I)], \end{aligned}$$

where $E(W_2|I)$ and $E(W_2|\text{not-}I)$ denote the expected second-stage processing time conditioned on interaction's occurring (I) or not occurring ($\text{not-}I$), respectively. In the unimodal condition, no interaction is possible. Thus,

$$E(RT_V) = E(V) + E(W_2|\text{not-}I). \quad (2)$$

Putting $\Delta \equiv E(W_2|\text{not-}I) - E(W_2|I)$, bimodal mean SRT becomes

$$E(RT_{VT}) = E(V) + E(W_2|\text{not-}I) - \Pr(I) \cdot \Delta. \quad (3)$$

The amount of expected cross-modal interaction, measured as the difference between expected uni- and bimodal response time, is

$$\begin{aligned} E(RT_V) - E(RT_{VT}) &= \Pr(I) \\ &\quad \cdot \{E(W_2|\text{not-}I) - E(W_2|I)\} \\ &= \Pr(I) \cdot \Delta, \end{aligned}$$

with positive Δ values corresponding to facilitation and negative ones to inhibition. The product form of this last equation reflects the separability hypothesis afforded by the TWIN model, as mentioned in the introduction: Unimodal stimulus properties (such as intensity) affect cross-modal interaction only by modulating the probability of the interaction, $\Pr(I)$, whereas cross-modal properties (such as spatial distance between the target and the nontarget) determine the size (and sign) of the interaction, Δ .

In order to generate quantitative predictions for mean reaction times, peripheral processes are assumed to be exponentially distributed, with a parameter depending on eccentricity (for visual stimuli) or frequency (for tactile stimuli).⁴ Since only mean SRT predictions will be considered, no distributional assumption on second-stage processing time is needed, except for the existence of an expected value (mean) for W_2 .

Thus, besides Δ , the TWIN model requires six basic parameters for this experiment: three for visual peripheral speed depending on eccentricity (λ_{10° , λ_{50° , and λ_{70°), two for tactile peripheral speed depending on frequency ($\lambda_{5\text{Hz}}$ and $\lambda_{50\text{Hz}}$), and one for second-stage processing without cross-modal interaction's occurring, $E(W_2|\text{not-}I) \equiv \mu$. From previous studies, window width was set to $\omega = 200$ msec a priori.

Special Cases

The general model is unspecific with respect to the dependence of the second-stage interaction parameter Δ on the spatial stimulus arrangement. Introducing explicit assumptions regarding this dependence yields a series of

partially nested specialized TWIN models with varying numbers of additional parameters, as follows.

Model 1. The most parsimonious model is the general case itself, where a single value of Δ is to account for all spatial configurations. For Condition 1 with a constant visual target position, it predicts the same mean SRT for all bimodal stimulus conditions of a given tactile stimulus frequency. For Conditions 2 and 3, this model attributes differences in observed SRT to bimodal stimuli to the eccentricity of the visual stimulus and/or the frequency of the tactile stimulus, but not to spatial position. Thus, in addition to the six basic parameters, only one spatial parameter, Δ , is needed.

Model 2. Here, it is assumed that the interaction is most pronounced when both stimuli are presented at a (nearly) identical position. All other configurations have the same effect. That is, interaction may occur in the presence of a nontarget as before, but it makes a difference whether the stimuli are at a close distance or not. This requires two parameters: Δ_s for the interaction when both stimuli are presented at the identical position, and Δ otherwise. Note that for all the models considered, the exact specification of the interaction parameters under all the configuration conditions is collected in Table 2.

Model 3. Alternatively, it is assumed that interaction depends on whether or not the target and the nontarget are presented in the same hemifield. This requires two parameters as well: Δ_i and Δ_c , referring to an ipsi- or a contralateral stimulus arrangement, respectively.

Model 4. This combines the two previous models. It keeps the ipsi- versus contralateral distinction but allows for an even stronger interaction if both stimuli are presented at an identical position. It requires three parameters: Δ_s^i , Δ_i , and Δ_c , for same-location, ipsilateral, and contralateral stimulus arrangements, respectively. Note that "same location" refers to the ipsilateral condition as well.

Table 2
Summary of the Δ Parameters for All Eight Models

Stimulus Position			Model							
Visual	Tactile	Lat.	1	2	3	4	5	6	7	8
10°	10°	i	Δ	Δ_s	Δ_i	Δ_s^i	$i * 10^\circ$	$\Delta_{10^\circ}^i$	Δ_1	Δ_1
10°	50°	i	Δ	Δ	Δ_i	Δ_i	$i * 50^\circ$	$\Delta_{50^\circ}^i$	Δ_3	Δ_3
10°	70°	i	Δ	Δ	Δ_i	Δ_i	$i * 70^\circ$	$\Delta_{70^\circ}^i$	Δ_4	Δ_4
10°	10°	c	Δ	Δ	Δ_c	Δ_c	$c * 10^\circ$	$\Delta_{10^\circ}^c$	Δ_2	Δ_2
10°	50°	c	Δ	Δ	Δ_c	Δ_c	$c * 50^\circ$	$\Delta_{50^\circ}^c$	Δ_4	Δ_4^*
10°	70°	c	Δ	Δ	Δ_c	Δ_c	$c * 70^\circ$	$\Delta_{70^\circ}^c$	Δ_5	Δ_5
50°	10°	i	Δ	Δ	Δ_i	Δ_i	$i * 10^\circ$	$\Delta_{10^\circ}^i$	Δ_3	Δ_3
70°	10°	i	Δ	Δ	Δ_i	Δ_i	$i * 10^\circ$	$\Delta_{10^\circ}^i$	Δ_4	Δ_4
50°	10°	c	Δ	Δ	Δ_c	Δ_c	$c * 10^\circ$	$\Delta_{10^\circ}^c$	Δ_4	Δ_4^*
70°	10°	c	Δ	Δ	Δ_c	Δ_c	$c * 10^\circ$	$\Delta_{10^\circ}^c$	Δ_5	Δ_5
50°	50°	i	Δ	Δ_s	Δ_i	Δ_s^i	$i * 50^\circ$	$\Delta_{50^\circ}^i$	Δ_1	Δ_1
70°	70°	i	Δ	Δ_s	Δ_i	Δ_s^i	$i * 70^\circ$	$\Delta_{70^\circ}^i$	Δ_1	Δ_1
50°	50°	c	Δ	Δ	Δ_c	Δ_c	$c * 50^\circ$	$\Delta_{50^\circ}^c$	Δ_6	Δ_6
70°	70°	c	Δ	Δ	Δ_c	Δ_c	$c * 70^\circ$	$\Delta_{70^\circ}^c$	Δ_7	Δ_7

Note—Lat., laterality; i, ipsilateral; c, contralateral.

Model 5. It is assumed that both laterality of the bimodal stimulus and position of the tactile stimulus determine the interaction. Each Δ value is a product of one of two parameter values indicating laterality (ipsi-/contralateral) and one of three parameter values indicating the eccentricity of the tactile stimulus (10°, 50°, and 70°). This generates six different Δ s from five parameters.

Model 6. Here, we retain the assumptions of Model 5 on the role of laterality and eccentricity but drop the multiplicativity property. This yields parameters $\Delta_{10^\circ}^i, \Delta_{10^\circ}^c, \Delta_{50^\circ}^i, \Delta_{50^\circ}^c, \Delta_{70^\circ}^i, \Delta_{70^\circ}^c$, increasing the number of parameters by one.

Model 7. Here, we drop the perspective of the subject as the reference point for determining laterality and eccentricity. Instead, each physical distance between visual and tactile stimulus is presented by one parameter. For Condition 1, a visual stimulus at +10° is combined with a tactile stimulus at +10°, +0°, +10°, -10°, -50°, and -70°, creating six distances, and the same holds for a visual stimulus presented at -10°. However, for symmetry and because the distance from +10° to +70° equals the distance from +10° to -50°, only five different parameters are needed. For Condition 2, this also results in five different parameters, whereas Condition 3 adds another two. For convenience, they are simply numbered by order of distance magnitude (cf. Table 2).

Model 8. This generalizes Model 7 in allowing for an effect of laterality even if physical distance is identical. One more parameter must be added.

Figure 4 depicts the partially nested family of TWIN model versions for the effects of spatial configuration.

Model Fits

Table 2 summarizes the specific parameter sets for all the model variants.⁵ These parameters and the model-

specific parameters listed in Table 2 were estimated simultaneously from 31 data points by minimizing the Pearson χ^2 statistic⁶

$$\chi^2 = \sum_i \frac{(\overline{\text{SRT}}_i - \widehat{\text{SRT}}_i)^2}{\sigma_{\overline{\text{SRT}}_i}^2}. \tag{4}$$

$\overline{\text{SRT}}_i$ and $\widehat{\text{SRT}}_i$ are the observed and predicted values, respectively, of the mean SRTs to the i th stimulus condition, and $\sigma_{\overline{\text{SRT}}_i}$ is the associated standard error.

Parameters $\lambda_{10^\circ}, \lambda_{50^\circ}, \lambda_{70^\circ}, \lambda_{5\text{Hz}}, \lambda_{50\text{Hz}}$, and μ were assumed to be fixed under all eight models. Their estimates were $\lambda_{10^\circ} = 0.063, \lambda_{50^\circ} = 0.017, \lambda_{70^\circ} = 0.014, \lambda_{50\text{Hz}} = 0.035, \lambda_{5\text{Hz}} = 0.009$, and $\mu = 133.99$. They determine the peripheral and second-stage processing time estimates shown in Table 3. The estimates of visual peripheral-processing times were ordered as expected; that is, they increased with increases in the eccentricity of the target position at a rate of about 1 msec/degree of eccentricity, in accordance with an estimate from an earlier visual-auditory study (Colonius & Arndt, 2001). Estimated peripheral-tactile-processing time was substantially shorter for the 50-Hz vibration than for 5 Hz. This is in line with the fact that the 50-Hz stimulus was reported by the subjects as being much more salient than the 5-Hz stimulus. In terms of the model, this means that winning the race against the visual target and, thereby, the occurrence of interaction is more likely for the 50-Hz vibration stimulus than for the 5-Hz vibration stimulus.

Estimates for the Δ parameters, specific for each model, are shown in Table 4. Multiplied by the probability of an interaction, they determine the amount of facilitation/inhibition (in milliseconds). The larger the Δ s, the larger the interaction for a given probability. All Δ values were larger than zero, indicating a facilitation of SRT for bimodal

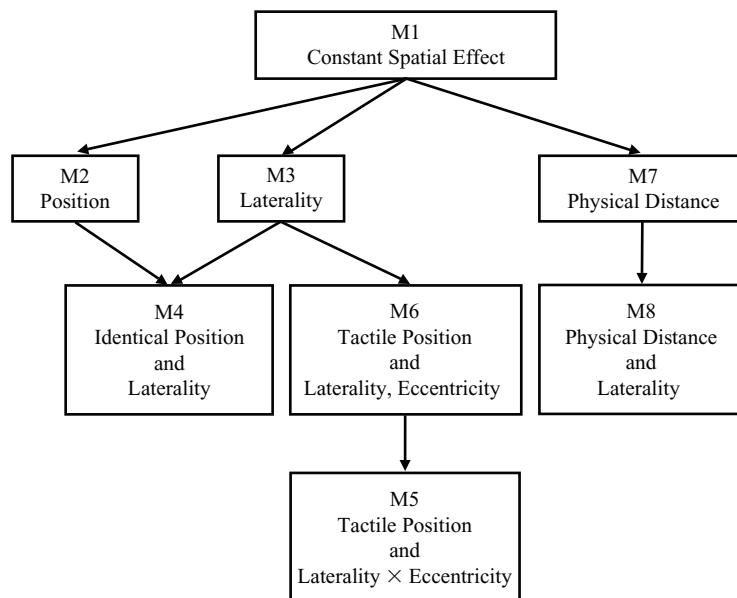


Figure 4. TWIN model hierarchy for the effects of spatial configuration features. Lower level models refer to more specific assumptions.

Table 3
Processing Time Estimates (in Milliseconds) for the
Peripheral Processes and the Second Stage

	Condition	Time
Visual peripheral	10°	16
	50°	58
	70°	71
Tactile peripheral	50 Hz	29
	5 Hz	111
Second stage		134

stimuli. Throughout, the model parameters predict larger amounts of facilitation when the target and the nontarget are presented ipsilaterally, rather than contralaterally.

Turning to a comparison of the models with respect to their overall fit, the observed and predicted values for mean SRTs for Conditions 1–3 and LED only are shown in Tables 5–8, respectively.

For Condition 1, all the models predict mean SRTs for the 5-Hz tactile accessory stimulus to be the same for all the conditions. Indeed, observed SRTs do not vary much as a function of stimulus positions. For the 50-Hz accessory stimulus, Model 6 seems to describe the data best. This model takes into account the laterality of stimulus presentation (ipsilateral or contralateral) and the eccentricity of the stimuli, but not the physical distance between the stimuli, as does Model 7.

In Condition 2, for the 5-Hz tactile accessory stimulus, all the models predict SRTs to be the same within a given eccentricity of the visual stimulus, regardless of the laterality. This resulted in a rather poor fit when the target was presented 50° or 70° left or right from fixation. As before, the models gave a better account for the 50-Hz than for the 5-Hz tactile nontarget condition. Model 8, with the highest number of parameters, provided the best fit.

In Condition 3, the estimated SRTs for the 5-Hz tactile accessory stimulus are identical to the those in Condition 2. Note that both conditions differ with respect to the position of the tactile nontarget, but not with respect to the visual target. Obviously, ignoring laterality leads, as before, to a poor fit of the models. For the 50-Hz stimulus condition, Models 7 and 8 provided the best fits.

For all the models, it was easier to account for the observed patterns of data from Condition 1 than for those from Conditions 2 and 3. In particular, all the models had problems accounting for the observed data in Conditions

2 and 3 when the nontarget was the 5 Hz, rather than the 50-Hz, vibration. This may have been due to the relatively long peripheral-processing time for the 5-Hz stimulus (111 msec; see Table 3), leading to low probabilities of interaction (.34 and .39) for the visual stimuli of eccentricity (50° and 70°, respectively): Since the spatial effect is the product of that probability times the Δ value, the ability of the different models to make differing predictions is somewhat limited. Indeed, an inspection of Tables 6 and 7 reveals that model predictions under the 5-Hz condition vary very little for any given stimulus configuration and do not cover the whole range of observed mean SRT values.

To compare the models while taking into account the different numbers of parameters, root-mean squared error (RMSE; Browne & Cudeck, 1992) was calculated for each fit. RMSE is the square root of the χ^2 value, divided by the number of data points (N) minus the number of the parameters (k) of the specific model; that is, $RMSE = \sqrt{\chi^2 / (N - k)}$. The model with the smallest RMSE is considered to be the best fitting. Note that RMSE penalizes each model by the number of parameters required (RMSE increases as the number of parameters increases). Table 9 shows the number of model-specific parameters, χ^2 values (degrees of freedom in parentheses), and RMSE values for each model.

There is a definite effect of spatial configuration on bimodal SRT, since Model 1, which assumes a constant Δ value, fares worst. Attributing a special effect to coincident presentation of visual and tactile stimulus (Model 2) does not improve the fit substantially. However, distinguishing between same- and different-hemifield presentation (laterality; Model 3) reduces the RMSE value by 22%. An additional parameter tracking whether visual and tactile positions are identical or not (Model 4) leads only to a marginal improvement. Further reduction of RMSE is achieved by Model 6, which postulates a simultaneous effect of both the eccentricity of the tactile stimulus and laterality. Assuming that this effect is multiplicative (Model 5) seems not be warranted. Model 7, which claims that physical distance is the only relevant parameter, is inferior to Model 6, but providing an additional parameter for laterality in Model 7 (Model 8) leads to a fit comparable to that of Model 6.

In terms of relative fit measured by RMSE, Model 6 features the lowest RMSE value, having two parameters less than Model 8. The following picture emerges. The

Table 4
 Δ Parameters for Each of the Eight Models

Model							
1	2	3	4	5	6	7	8
$\Delta = 26$	$\Delta_s = 32$	$\Delta_i = 33$	$\Delta_s^i = 32$	$i = 1$	$\Delta_{10^\circ}^i = 32$	$\Delta_1 = 32$	$\Delta_1 = 32$
	$\Delta = 22$	$\Delta_c = 16$	$\Delta_c = 16$	$c = 0.46$	$\Delta_{10^\circ}^c = 12$	$\Delta_2 = 16$	$\Delta_2 = 16$
			$\Delta_i = 37$	$10^\circ = 32$	$\Delta_{50^\circ}^i = 31$	$\Delta_3 = 31$	$\Delta_3 = 31$
				$50^\circ = 29$	$\Delta_{50^\circ}^c = 24$	$\Delta_4 = 34$	$\Delta_4^* = 24$
				$70^\circ = 38$	$\Delta_{70^\circ}^i = 43$	$\Delta_5 = 18$	$\Delta_4 = 43$
					$\Delta_{70^\circ}^c = 18$	$\Delta_6 = 2$	$\Delta_5 = 18$
						$\Delta_7 = 9$	$\Delta_6 = 3$
							$\Delta_7 = 9$

Table 5
Condition 1: Observed and Predicted Saccadic Reaction Times (SRTs, in Milliseconds)
for 5- and 50-Hz Nontargets

Stimulus Position		Laterality	Observed SRT		Predicted SRT for Model							
Visual	Tactile		<i>M</i>	<i>SE</i>	1	2	3	4	5	6	7	8
5 Hz												
10°	10°	i	146	0.94	146	146	146	146	146	146	146	146
10°	10°	c	146	1.01	146	146	146	146	146	146	146	146
10°	50°	i	148	1.68	146	146	146	146	146	146	146	146
10°	50°	c	146	1.69	146	146	146	146	146	146	146	146
10°	70°	i	145	1.63	146	146	146	146	146	146	146	146
10°	70°	c	143	1.62	146	146	146	146	146	146	146	146
50 Hz												
10°	10°	i	138	0.89	140	139	138	139	139	138	139	139
10°	10°	c	144	0.89	140	142	144	144	145	144	144	144
10°	50°	i	138	1.50	140	142	138	137	139	139	139	139
10°	50°	c	145	1.49	140	142	144	144	141	145	138	141
10°	70°	i	135	1.46	140	142	138	137	135	136	138	135
10°	70°	c	143	1.53	140	142	144	144	143	143	143	143

Note—i, ipsilateral; c, contralateral.

spatial effect is not due primarily to the physical distance between the stimuli but, rather, to laterality—that is, whether or not the target and the nontarget are presented in the same hemifield. Moreover, there is a tendency for the effect to increase with increases in the eccentricity of the tactile stimulus, especially with ipsilateral configurations, but laterality and eccentricity seem not to combine according to a multiplicative rule.

It should be noted that all the models must be rejected, from a statistical-hypothesis-testing point of view, given the large χ^2 values. This is due to the large sample size (600 observations per condition), leading to very small standard errors and, consequently, to an excessive power of the χ^2 goodness-of-fit test. On the other hand, all the models can account for a high percentage of the observed variability in the data, ranging from 94% (Model 1) to 97% (Model 6) to 98% (Model 8).

Finally, given that the model hierarchy was developed post hoc, further empirical testing of the TWIN model versions for spatial effects is required.

SUMMARY AND DISCUSSION

The experiment reported here demonstrated a reliable facilitatory effect of a vibrotactile accessory stimulus on SRTs to a visual target. The size of the reduction of bimodal SRT, relative to unimodal SRT, was shown to depend on (1) whether both stimuli were presented ipsi- or contralateral, (2) the eccentricity of the stimuli, and (3) the frequency of the vibrotactile signal. This extends and complements previous findings in Diederich et al. (2003), where it was demonstrated in addition that the effect was not due to a speed-accuracy trade-off in saccade execution. Our results are in line with those in the study

Table 6
Condition 2: Observed and Predicted Saccadic Reaction Times (SRTs, in Milliseconds)
for 5- and 50-Hz Nontargets

Stimulus Position		Laterality	Observed SRT		Predicted SRT for Model							
Visual	Tactile		<i>M</i>	<i>SE</i>	1	2	3	4	5	6	7	8
5 Hz												
10°	10°	i	146	0.93	146	146	146	146	146	146	146	146
10°	10°	c	146	1.01	146	146	146	146	146	146	146	146
50°	10°	i	176	1.45	183	181	181	181	181	181	181	181
50°	10°	c	182	1.46	183	181	181	181	181	181	181	181
70°	10°	i	185	1.28	196	194	193	194	194	193	194	194
70°	10°	c	202	1.85	196	194	193	194	194	193	194	194
50 Hz												
10°	10°	i	138	0.89	140	139	138	139	139	138	139	139
10°	10°	c	144	0.89	140	142	144	144	145	144	144	144
50°	10°	i	172	1.29	175	177	170	168	172	173	172	172
50°	10°	c	174	1.47	175	177	182	182	176	183	169	176
70°	10°	i	176	1.40	188	190	183	181	176	179	182	176
70°	10°	c	193	1.55	188	190	195	195	193	193	193	193

Note—i, ipsilateral; c, contralateral.

Table 7
Condition 3: Observed and Predicted Saccadic Reaction Times (SRTs, in Milliseconds)
for 5- and 50-Hz Nontargets

Stimulus Position		Laterality	Observed SRT		Predicted SRT for Model							
Visual	Tactile		<i>M</i>	<i>SE</i>	1	2	3	4	5	6	7	8
5 Hz												
10°	10°	i	146	0.93	146	146	146	146	146	146	146	146
10°	10°	c	146	1.01	146	146	146	146	146	146	146	146
50°	50°	i	181	1.70	183	181	181	181	181	181	181	181
50°	50°	c	187	1.61	183	181	181	181	181	181	181	181
70°	70°	i	190	1.49	196	194	193	194	194	193	194	194
70°	70°	c	203	1.47	196	194	193	194	194	193	194	194
50 Hz												
10°	10°	i	138	0.89	140	139	138	139	139	138	139	139
10°	10°	c	144	0.89	140	142	144	144	145	144	144	144
50°	50°	i	177	1.61	175	171	170	171	171	173	171	171
50°	50°	c	190	1.78	175	177	182	182	184	183	189	190
70°	70°	i	181	1.48	188	184	183	184	184	179	184	184
70°	70°	c	199	1.46	188	190	195	195	197	193	199	199

Note—i, ipsilateral; c, contralateral.

by Amlôt et al. (2003) cited above. In that study, facilitation was found only when the tactile stimulus preceded the visual target, but this discrepancy may well have been due to the different intensity levels of the vibrotactile stimuli being used in the two studies.

Part of the reduction of bimodal SRT might be due to the nontarget stimulus’s providing a “warning signal”: Although the tactile stimulus was not a reliable predictor of which hemifield the target would occur in (contralateral presentations were as frequent as ipsilateral ones) the nontarget—in particular, the fast 50-Hz stimulus—could have given temporal information about the forthcoming target onset.⁷ It is difficult to assess the exact amount of a possible warning effect, but it is clear that it would not account for the effects of spatial configuration. Another potential explanation often considered for cross-modal facilitation is probability summation: a reduction of SRT due simply to statistical facilitation in a race model situation (cf. Colonius, 1990; Colonius & Diederich, 2006). But again, it is not obvious how such a mechanism would be able to account for the spatial effects observed here.⁸

In order to further probe the nature of the spatial effects, several versions of the TWIN model were fit to the data. The effects of peripheral-processing speed depended on the eccentricity of the visual stimulus and the frequency of the vibrotactile stimulus and were clearly reflected in the parameter estimates (cf. Table 3). The TWIN model

framework can be seen as a filter removing these confounding effects of peripheral processing to yield a true estimate of the amount of multisensory spatial interaction. The remaining effect of spatial configuration is entirely represented in the model’s Δ parameters, and the different TWIN model versions tested differ only in their assumptions on how Δ depends on the configuration. Although testing the basic assumptions of the TWIN model was not the primary goal of this study, the separability of the peripheral-processing parameters from the parameters capturing the spatial effects and the plausible quantitative ordering of the parameter estimates lend further support to the general framework underlying the TWIN model.

In comparison with the visual-auditory domain, very little seems to be known about the coordinate transformations required to determine that, for example, a tactile stimulus applied to the palm is in spatial register with a visual stimulus (cf. Andersen & Buneo, 2002). Our modeling results may help researchers to pursue such an approach. The finding that laterality of the stimulus arrangement plays the dominant role in determining the amount of interaction and that the eccentricity of the tactile stimulus position also matters is consistent with a representation of space in eye-centered response fields that are gain modulated by limb position signals (cf. Andersen, Snyder, Bradley, & Xing, 1997; Pouget & Snyder, 2000).

Recent brain imaging (fMRI) studies with subjects receiving simple visual or tactile stimuli in either the right or the left hemifield have identified multimodal areas in the intraparietal sulcus and in the frontal eye fields (Macaluso & Driver, 2001). In a related FA (fMRI) study, subjects received either left or right visual targets near to the hand at which vibrotactile nontargets were also delivered on half of the trials (Macaluso, Frith, & Driver, 2000a). Neural activity in the occipital visual areas was significantly amplified when visual and tactile stimuli were presented at the same location, as compared with when tactile stimulation was applied to the opposite hand. Moreover, it was

Table 8
Observed and Predicted Saccadic Reaction Times
(SRTs, in Milliseconds) for the LED-Only Condition
As a Function of Eccentricity

LED Eccentricity	SRT	
	Observed	Predicted
10°	150	150
50°	189	192
70°	208	205

Table 9
Number of Parameters, χ^2 (*df*), and Root-Mean Squared Errors (RMSEs) for Models 1–8

	Model							
	1	2	3	4	5	6	7	8
Number of parameters	1	2	2	3	5	6	7	8
χ^2 (<i>df</i>)	505 (24)	432 (23)	260 (22)	244 (21)	231 (20)	192 (19)	219 (18)	175 (17)
RMSE	4.10	3.86	2.99	2.95	2.98	2.77	3.02	2.76

shown that this multisensory interaction depended on the spatial alignment of the cross-modal stimuli in external space, rather than on initial projection of the visual and tactile stimuli to a common hemisphere (Macaluso, Frith, & Driver, 2002). Note how this squares with our results in Condition 3: In ipsilateral trials, the visual target and the tactile nontarget were always presented spatially close to each other but, over different blocks, at widely different locations, from 10° eccentricity to 50° and to 70°. Nevertheless, in the best-fitting Models 6 and 8, the corresponding Δ values were nearly (Model 6) or perfectly (Model 8) invariant over these configurations.

These imaging results suggest that the spatial location of the stimuli are first represented cortically in modality-specific maps and, at a later stage of processing, are integrated into a multimodal spatial representation. Note that this assumption of a feedforward convergence from unimodal to multimodal brain areas is also consistent with the basic TWIN architecture, where, however, no specific assumptions are made with respect to the underlying circuitry (subcortical or cortical). Another important finding of these imaging studies is the following: Tactile stimulation was found to influence activity in unimodal visual occipital areas when tactile attention was directed to one hemifield (Macaluso, Frith, & Driver, 2000b) or with unpredictable tactile stimulation (Macaluso et al., 2000a). This spatially specific influence is suggested to be due to back-projections from multimodal parietal areas and has also been shown to depend on gaze direction, regardless of whether the stimulated hand was seen or not (Macaluso et al., 2002). This feedback system may be part of a mechanism operating to solve the coordinate transformations required for cross-modal sensory integration (Andersen et al., 1997).⁹ There is no evidence at present, however, that this activity modulation in unimodal cortical areas will also affect the early peripheral-processing stage that is represented as the first stage in the TWIN model.

Finally, a general warning concerning the interpretation of these fMRI results seems in order.¹⁰ Given that many brain areas contain both uni- and multisensory cells, an enhanced BOLD signal could be the result of a recruitment of unimodal cells, rather than of an increased response of multisensory cells. Moreover, in a recent study, the authors above found cross-modal congruency effects (in lateral occipital and parietal operculum regions—i.e., in the visual and somatosensory cortex), irrespective of which modality was defined as the target and of whether the stimuli were used to guide eye movements or were received passively (Macaluso, Frith, & Driver, 2005). This independence of the BOLD response from task context and target/nontarget modality makes it unlikely that fMRI

data speak powerfully to the issues raised by the specific assumptions of the TWIN model.

AUTHOR NOTE

This research was supported by Deutsche Forschungsgemeinschaft (DFG) Grants Di 506/8–1 and 506/8–2 to both authors. Correspondence concerning this article should be addressed to A. Diederich, School of Humanities and Social Sciences, International University Bremen, P.O. Box 750 561, D-28725, Bremen, Germany (e-mail: a.diederich@iu-bremen.de).

REFERENCES

- AMLÔT, R., WALKER, R., DRIVER, J., & SPENCE, C. (2003). Multimodal visual–somatosensory integration in saccade generation. *Neuropsychologia*, *41*, 1-15.
- ANDERSEN, R. A., & BUNEO, C. A. (2002). Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience*, *25*, 189-220.
- ANDERSEN, R. A., SNYDER, L. H., BRADLEY, D. C., & XING, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience*, *20*, 303-330.
- ARNDT, A., & COLONIUS, H. (2003). Two separate stages in crossmodal saccadic integration: Evidence from varying intensity of an auditory accessory stimulus. *Experimental Brain Research*, *150*, 417-426.
- BROWNE, M. W., & CUDECK, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*, 230-258.
- BRUCE, C., DESIMONE, R., & GROSS, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, *46*, 369-384.
- COLONIUS, H. (1990). Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology*, *34*, 253-275.
- COLONIUS, H., & ARNDT, P. (2001). A two-stage model for visual–auditory interaction in saccadic latencies. *Perception & Psychophysics*, *63*, 126-147.
- COLONIUS, H., & DIEDERICH, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *Journal of Cognitive Neuroscience*, *16*, 1000-1009.
- COLONIUS, H., & DIEDERICH, A. (2006). Race model inequality: Interpreting a geometric measure of the amount of violation. *Psychological Review*, *113*, 148-154.
- CORNEIL, B. D., VAN WANROOIJ, M., MUNOZ, D. P., & VAN OPSTAL, A. J. (2002). Auditory–visual interactions subserving goal-directed saccades in a complex scene. *Journal of Neurophysiology*, *88*, 438-454.
- DENÈVE, S., LATHAM, P. E., & POUGET, A. (2001). Efficient computation and cue integration with noisy population codes. *Nature Neuroscience*, *4*, 826-831.
- DIEDERICH, A., & COLONIUS, H. (2004a). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, *66*, 1388-1404.
- DIEDERICH, A., & COLONIUS, H. (2004b). Modeling the time course of multisensory interaction in manual and saccadic responses. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes* (pp. 395-408). Cambridge, MA: MIT Press.
- DIEDERICH, A., COLONIUS, H., BOCKHORST, D., & TABELING, S. (2003). Visual–tactile spatial interaction in saccade generation. *Experimental Brain Research*, *148*, 328-337.
- DUHAMEL, J. R., COLBY, C. L., & GOLDBERG, M. E. (1998). Ventral intraparietal area of the macaque: Congruent visual and somatic response properties. *Journal of Neurophysiology*, *79*, 126-136.

- FRENS, M. A., & VAN OPSTAL, A. J. (1998). Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Research Bulletin*, **46**, 211-224.
- FRENS, M. A., VAN OPSTAL, A. J., & VAN DER WILLIGEN, R. F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception & Psychophysics*, **57**, 802-816.
- GRAZIANO, M. S. A., & GROSS, C. G. (1994). The representation of extrapersonal space: A possible role for bimodal, visual-tactile neurons. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1021-1034). Cambridge, MA: MIT Press.
- GROH, J. M., & SPARKS, D. L. (1996). Saccades to somatosensory targets: I. Behavioral characteristics. *Journal of Neurophysiology*, **75**, 412-427.
- HUGHES, H. C., NELSON, M. D., & ARONCHICK, D. M. (1998). Spatial characteristics of visual-auditory summation in human saccades. *Vision Research*, **38**, 3955-3963.
- HUGHES, H. C., REUTER-LORENZ, P. A., NOZAWA, G., & FENDRICH, R. (1994). Visual-auditory interactions in sensorimotor processing: Saccades versus manual responses. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 131-153.
- KADUNCE, D. C., VAUGHAN, J. W., WALLACE, M. T., & STEIN, B. E. (2001). The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Experimental Brain Research*, **139**, 303-310.
- KIRCHNER, H., & COLONIUS, H. (2005). Interstimulus contingency facilitates saccadic responses in a bimodal go/no-go task. *Cognitive Brain Research*, **25**, 261-272.
- LUECK, C. J., CRAWFORD, T. J., SAVAGE, C. J., & KENNARD, C. (1990). Auditory-visual interaction in the generation of saccades in man. *Experimental Brain Research*, **82**, 149-157.
- MACALUSO, E., & DRIVER, J. (2001). Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia*, **39**, 1304-1316.
- MACALUSO, E., FRITH, C. D., & DRIVER, J. (2000a). Modulation of human visual cortex by crossmodal spatial attention. *Science*, **289**, 1206-1208.
- MACALUSO, E., FRITH, C. D., & DRIVER, J. (2000b). Selective spatial attention in vision and touch: Unimodal and multimodal mechanisms revealed by PET. *Journal of Neurophysiology*, **83**, 3062-3075.
- MACALUSO, E., FRITH, C. D., & DRIVER, J. (2002). Crossmodal spatial influences of touch on extrastriate visual areas take current gaze direction into account. *Neuron*, **34**, 647-658.
- MACALUSO, E., FRITH, C. D., & DRIVER, J. (2005). Multisensory stimulation with or without saccades: fMRI evidence for crossmodal effects on sensory-specific cortices that reflect multisensory location-congruence rather than task-relevance. *NeuroImage*, **26**, 414-425.
- MEREDITH, M. A. (2002). On the neuronal basis for multisensory convergence: A brief overview. *Cognitive Brain Research*, **14**, 31-40.
- MEREDITH, M. A., & STEIN, B. E. (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research*, **365**, 350-354.
- MUNOZ, D. P., & SCHALL, J. D. (2004). Concurrent, distributed control of saccade initiation in the frontal eye field and superior colliculus. In W. C. Hall & A. Moschovakis (Eds.), *The superior colliculus: New approaches for studying sensorimotor integration* (pp. 55-82). Boca Raton, FL: CRC Press.
- POUGET, A., & SNYDER, L. H. (2000). Computational approaches to sensorimotor transformations. *Nature Neuroscience*, **3**, 1193-1198.
- SPARKS, D. L. (1986). Translation of sensory signals into commands for control of saccadic eye movements: Role of primate superior colliculus. *Physiological Review*, **66**, 118-171.
- STANFORD, T. R. (2004). Signal coding in the primate superior colliculus revealed through the use of artificial signals. In W. C. Hall & A. Moschovakis (Eds.), *The superior colliculus: New approaches for studying sensorimotor integration* (pp. 35-53). Boca Raton, FL: CRC Press.
- STEIN, B. E., MAGALHÃES-CASTRO, B., & KRUGER, L. (1976). Relationship between visual and tactile representations in cat superior colliculus. *Journal of Neurophysiology*, **39**, 401-419.
- STEIN, B. E., & MEREDITH, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
- VAN OPSTAL, A. J., & MUNOZ, D. P. (2004). Auditory-visual interactions subserving primate gaze orienting. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 373-393). Cambridge, MA: MIT Press.

NOTES

1. Independence can certainly be assumed at the receptor level. The first stage can be defined by this independence assumption: It lasts up to a level of processing at which independence breaks down.
2. Although nontargets can be ignored, subjects are certainly aware of them and evoke an inhibitory mechanism to suppress saccades directed to nontargets (see, e.g., Kirchner & Colonius, 2005).
3. Part of the data in the 50-Hz vibration condition have previously been published in Diederich et al. (2003). An investigation of saccade accuracy presented there allowed us to exclude the possibility of any speed-accuracy trade-offs in saccade execution.
4. This implies

$$\Pr(I) = \Pr(T < V < T + \omega) \\ = \frac{\lambda_T}{\lambda_T + \lambda_V} \{1 - \exp[-\lambda_V \omega]\},$$

where λ_T and λ_V refer to the tactile and visual intensity (peripheral speed) parameters, respectively.

5. As in Colonius and Diederich (2004), window width (ω) was set to 200 msec.
6. The FMINSEARCH routine of MATLAB was used. Merging the data from the 10° conditions reduced the original number of 39 data points to 31.
7. At the neural level, this warning effect could be effected by facilitating the release of fixation-related activity.
8. An assessment of statistical facilitation would have required inclusion of unimodal tactile trials in the experiments, as was suggested by one of the reviewers. In our view, introducing unimodal nontarget (tactile) trials in an FA paradigm is problematic: Here, subjects are explicitly instructed to ignore the tactile stimulus and to respond to the visual target only. This asymmetry in the role of the two modalities is specific to the FA paradigm; it does not apply to the RET paradigm, where subjects respond to whichever stimulus they detect first.
9. For a neural network model with forward and backward connections, using known properties of cortical neurons, see Denève, Latham, and Pouget (2001).
10. We are grateful to a reviewer for pointing this out.

(Manuscript received November 1, 2004;
revision accepted for publication February 24, 2006.)