Exp Brain Res (2008) 186:1–22
DOI 10.1007/s00221-007-1197-4

RESEARCH ARTICLE

Crossmodal interaction in saccadic reaction time: separating multisensory from warning effects in the time window of integration model

Adele Diederich · Hans Colonius

Received: 22 March 2007 / Accepted: 19 October 2007 / Published online: 15 November 2007

Abstract In a focused attention task saccadic reaction time (SRT) to a visual target stimulus (LED) was measured with an auditory (white noise burst) or tactile (vibration applied to palm) non-target presented in ipsi- or contralateral position to the target. Crossmodal facilitation of SRT was observed under all configurations and stimulus onset asynchrony (SOA) values ranging from −500 (non-target prior to target) to 0 ms, but the effect was larger for ipsi- than for contralateral presentation within an SOA range from −200 ms to 0. The time-window-of-integration (TWIN) model (Colonius and Diederich in J Cogn Neurosci 16:1000, 2004) is extended here to separate the effect of a spatially unspecific warning effect of the non-target from a spatially specific and genuine multisensory integration effect.

Keywords Multisensory integration · Warning effect · Time-window-of-integration · Saccadic eye movement

Introduction

The sudden occurrence of a peripheral visual or acoustic stimulus usually elicits an overt orienting response involving a coordinated movement of the eyes, head, and body toward the signal source. A similar reaction can also be evoked by a tactile stimulus, for example, a fly settling on your forearm. Gaze orienting will clearly improve vision for the foveated stimulus but, less obviously, the perception of auditory or tactile stimuli in the direction of gaze may be facilitated as well (e.g., Hublet et al. 1976; Schaefer et al. 2005). The existence of such *crossmodal links* has in fact been ascertained in many behavioral and electrophysiological studies (e.g., Bushara et al. 2001; Kennett et al. 2001; McDonald et al. 2000; Spence and Driver 1997; Wallace et al. 2004; Ward 1994).

These findings are consistent with neurophysiological evidence for multisensory integration in structures involved in the control of eye movements, in particular the superior colliculus (SC) (Stein and Meredith 1993). Multisensory neurons in the deep layers of SC in cats and monkeys show an enhanced response to combinations of visual, auditory, and tactile stimuli in spatiotemporal proximity (Meredith and Stein 1986a, b; Wallace et al. 1996; Bell et al. 2001; Frens and Van Opstal 1998). Moreover, multisensory integration properties of most SC neurons, as well as observed orientation behavior, are mediated by influences from two cortical areas—the anterior ectosylvian sulcus (AES) and the rostral aspect of the lateral suprasylvian sulcus (rLS; Jiang et al. 2002; Jiang et al. 2001). In addition, numerous behavioral studies of human eye movements have found a speed-up of saccadic reaction time (SRT) to multiple stimuli from different modalities, compared to responses to each of these modalities alone, following rules of multisensory integration similar to those of neural enhancement (e.g., Amlôt et al. 2003; Corneil and Muñoz 1996; Diederich et al. 2003; Frens et al. 1995; Hughes et al. 1998; Harrington and Peck 1998; Rach and Diederich 2006; for a recent review, see Diederich and Colonius 2004).

A. Diederich (✉)
School of Humanities and Social Sciences,
Jacobs University Bremen, P.O. Box 750 561,
28725 Bremen, Germany
e-mail: a.diederich@jacobs-university.de

H. Colonius
Department of Psychology, Oldenburg University,
P.O. Box 2503, 26111 Oldenburg, Germany
e-mail: hans.colonius@uni-oldenburg.de

From these studies, the temporal contiguity of stimuli from different modalities has been identified as a necessary condition for multisensory integration to occur. Neurophysiological and behavioral findings in human, monkey, and cat have led several authors to suggest the concept of a critical spatiotemporal "window of integration" (e.g., Bell et al. 2005; Meredith 2002; Corneil et al. 2002; Stein and Meredith 1993). Colonius and Diederich (2004) developed a stochastic model of response time in multisensory integration formalizing the notion of a time window of integration (TWIN model, for short). In this paper, we present an extension of TWIN designed to account for data sets that seem to reflect the influence of an additional *warning* mechanism acting over and above multisensory integration and leading to a speed-up of saccadic responses even outside of the alleged time window. A parametric fit of this extended TWIN model to data from a focused attention paradigm with visual targets and auditory and tactile non-targets is presented subsequently.

Crossmodal interaction and warning effects

Two different experimental paradigms have been utilized to measure SRT to a crossmodal stimulus set. In the *redundant target* (aka *divided-attention*) paradigm, stimuli from different modalities are presented simultaneously or with certain stimulus onset asynchrony (SOA), and the participant is instructed to respond by orienting her gaze toward the location of the stimulus detected first. In the *focused attention* paradigm, crossmodal stimulus sets are presented in the same manner, but now participants are instructed to respond only to the onset of a stimulus from a specifically defined target modality, such as the visual, and to ignore the remaining non-target stimulus, the tactile or the auditory, say. In the latter setting, when a stimulus of a non-target modality, a tone, say, appears before the visual target at some spatial disparity, there is no *overt* orienting response toward the tone if the participant is following the task instructions. Nevertheless, the non-target stimulus, though being uninformative with respect to the spatial location of the target, has been shown to modulate the saccadic response to the target: Depending on the exact spatiotemporal configuration of target and non-target, the effect can be a speed-up or an inhibition of SRT (see, e.g., Amlôt et al. 2003; Diederich and Colonius 2007b), and the saccadic trajectory can be affected as well (Doyle and Walker 2002).

As long as the non-target occurs less than about 200–300 ms before the target stimulus, the crossmodal SRT modulation effected by the non-target carries the signature of multisensory integration observed in neurophysiological studies: maximum facilitation at "physiological synchronicity" of both stimuli, a decrease of the speed-up—or even inhibition—at larger target/non-target disparities, and "inverse effectiveness", i.e., larger facilitation effects occur at lower stimulus intensity levels of target and non-target (Rach and Diederich 2006).

Although estimates for the time window of integration vary somewhat across subjects and task specifics, the 200 ms width shows up in several studies (e.g., Eimer 2001). On the other hand, when the non-target occurs at an earlier point in time (stimulus onset asynchrony (SOA) of 200 ms or more before the target), a substantial decrease of SRT compared to the unimodal condition has still been observed in our study (Diederich and Colonius 2007a). This decrease, however, no longer depended on whether target and non-target appeared at ipsi- or contralateral positions, thereby supporting the hypothesis that the non-target plays the role of a *spatially unspecific* alerting or warning cue for the upcoming target whenever the SOA is large enough. Note that the hypothesis of increased cross-modal processing triggered by an alerting or warning cue had already been advanced in Nickerson (1973) who called it "preparation enhancement". In the eye movement literature, the effects of a warning signal have been studied primarily in the context of explaining the gap effect, i.e., the latency to initiate a saccade to an eccentric target is reduced by extinguishing the fixation stimulus prior to the target onset (Reuter-Lorenz et al. 1991; Klein and Kingston 1993). An early study on the effect of auditory or visual warning signals on saccade latency, but without considering multisensory integration effects, was conducted by Ross and Ross (1981).

Here, the dual role of the non-target—inducing multisensory integration that is governed by the above mentioned spatiotemporal rules on the one hand and acting as a spatially unspecific crossmodal warning cue on the other—will be taken into account by a stochastic model that yields an estimate of the relative contribution of either mechanism for any specific SOA value.

Time-window-of-integration-and-warning (TWIN-W) model

Basic assumptions

The anatomical separation of the afferent pathways for the visual and auditory modality suggests at least two serial stages of saccadic reaction time: an early, afferent stage of peripheral processing (*first stage*) followed by a compound stage of converging subprocesses (*second stage*). In the first stage, a race among the peripheral neural excitations in the visual and auditory (or tactile) pathways triggered by a crossmodal stimulus complex takes place. The second stage comprises neural integration of the input and preparation of an oculomotor response. It is hypothesized that crossmodal interaction manifests itself in an increase or decrease of second stage processing time. Moreover, in the redundant target paradigm, the first stage duration is determined by the time of the winner of the race, whereas in the focused attention task—the only case considered here—the first stage duration is determined by the time it takes to process the target stimulus.

Thus, the model retains the classic notion of a race mechanism as an explanation for crossmodal interaction but restricts it to the very first stage of stimulus processing. The assumption of only two stages is certainly an oversimplification. Note, however, that the second stage is defined by default: it includes all subsequent, possibly overlapping, processes that are not part of the peripheral processes in the first stage (for a similar approach, see Van Opstal and Munoz 2004). A recent study by Whitchurch and Takahashi (2006) collecting (head) saccadic reaction times in the barn owl lends further support to the notion of a race between early visual and auditory processes depending on the relative intensity levels of the stimuli. In particular, their data suggest that the faster modality initiates the saccade and the slower modality remains available to refine saccade trajectory.

The TWIN model makes further specific assumptions about the temporal configuration needed for multisensory integration to occur (see also Colonius and Diederich 2004; Diederich and Colonius 2007a):

> (1) *Time-window-of-integration assumptions*: In the focused attention paradigm, crossmodal interaction occurs only if (a) a non-target stimulus wins the race in the first stage, opening a “time window” such that (b) the termination of the target peripheral process falls in the window. The duration of the “time window” is a constant.

The idea here is that the winning non-target will keep the saccadic system in a state of crossmodal reactivity such that the upcoming target stimulus, if it falls into the time window, will trigger crossmodal interaction. At the neural level this might correspond to a gradual inhibition of fixation neurons (in superior colliculus) and/or omnipause neurons (in midline pontine brain stem). In the case of the target being the winner, no discernible effect on saccadic reaction time is predicted, analogous to the unimodal situation.

The window of integration acts as a filter determining whether the afferent information delivered from different sensory organs is registered close enough in time for crossmodal interaction to take place. Passing this filter is necessary for crossmodal interaction to occur. It is not a sufficient condition because interaction also depends on the spatial configuration of the stimuli. Rather than assuming the existence of a joint spatiotemporal window of integration permitting interaction to occur only for both spatially and temporally neighboring stimuli, the TWIN model allows for interaction to occur even for rather distant stimuli of different modalities, as long as they fall exactly within the time window. The notion that crossmodal integration is determined by a window of time has already been suggested in Meredith et al. (1987) recording from SC neurons, and now underlies many studies in this area (for a recent behavioral study, see Navarra et al. 2005).

The two-stage structure of the TWIN model suggests an additional, important assumption about the effects of spatial and temporal factors:

> (2) *Assumption of spatiotemporal separability*: The amount of interaction in second-stage processing time is a function of the spatial configuration of the stimuli, but it does not depend on their (physical) presentation asynchrony (SOA).

Interaction, if it occurs at all, will either be inhibition or facilitation depending on both target and non-target positions. Typically, any facilitation decreases with the distance between the stimuli. More specific hypotheses about the effect of the spatial configuration on the amount of interaction have been studied in Diederich and Colonius (2007b).

These assumptions are part of a more general framework that is based on the distinction between intra- and crossmodal stimulus properties. *Crossmodal* properties are defined when stimuli of more than one modality are present, like spatial distance of target to non-target or similarity between stimuli of different modalities. *Intramodal* properties, on the other hand, refer to properties definable for a single stimulus, no matter whether this property is definable in all modalities (like intensity) or in only one modality (like color or pitch).

Intramodal properties can affect the outcome of the race in the first stage and, thereby, the probability of an interaction. Crossmodal properties may affect the amount of crossmodal interaction (Δ) occurring in the second stage. Note that crossmodal features cannot influence first stage processing time since the stimuli are still being processed in separate pathways. Initial empirical evidence for these predictions has been found in Colonius and Diederich (2004) for visual-tactile stimulation and in Arndt and Colonius (2003) for visual-auditory stimulation.

The final assumption refers to the role of the non-target stimulus as a warning cue:

> (3) *Assumption of warning mechanism*: If the non-target wins the processing race in the first stage by a

margin wide enough for the time window of integration to close before arrival of the target—thereby precluding multisensory integration to take place—then subsequent processing will be facilitated or inhibited without dependence on the spatial configuration of the stimuli.

The occurrence of warning depends on intramodal characteristics of the target and the non-target such as modality or intensity. For instance, an intense auditory non-target may have a higher chance to win the race with a head start compared to a weak tactile non-target.

Furthermore, the magnitude of the warning effect may be influenced by the experimental design. Specifically, presenting non-targets from different modalities in random order within one block of trials may result in a so called "modality switch effect": reaction times are slower when a stimulus is preceded by a stimulus of a different modality in the previous trial. Although the origin of this effect has not yet been clarified, it is not commonly attributed to multisensory interaction but rather to some attention shift mechanism. Whether or not the modality switch effect, which so far has only been observed under the redundant target paradigm, also occurs with non-targets will be tested below.

Note that assumption (3), first, precludes the possibility that both the warning mechanism and multisensory integration occur in one and the same trial and, second, does not rule out that the warning effect may be negative. Intuitively, if target processing terminates within the time window of integration, then the non-target cannot act as a warning cue at the same time. Obviously, there has to be an upper limit to the winning margin of the non-target for the warning mechanism to become active: if the non-target occurs several seconds before the target it may not serve as a warning cue at all. Therefore, in order to fit a parametric version of the model to empirical data, certain numerical ranges for possible parameter values will have to be specified (see Table 3).

The implication of assumption (3) that warning and integration cannot occur simultaneously within a trial, may appear overly restrictive. Let us assume, alternatively, that the non-target winning the race against the target stimulus may trigger a warning mechanism in addition to opening the time window of integration. Note that two details of this assumption have to be specified: (1) Are the two events—integration and warning—statistically independent? and (2) How do the RT facilitation effects for integration and warning combine? Assuming, for simplicity, both statistical independence of these events and additivity of their RT effects, it turns out that the mean RT prediction of this alternative model version is captured by the same equation as the original model (see Eq. 2). Nevertheless, the two versions may differ in their predictions because their probabilities of integration and/or warning may be different. In keeping with the time window concept, we investigated this alternative model version under the assumption that the non-target, if it wins the race opens up a time window of warning, in addition to opening the time window of integration. We refrain from presenting the computational details of this model version here because, in anticipation of the empirical results, it did not improve the model fit compared to the original version (see the section on model fits).

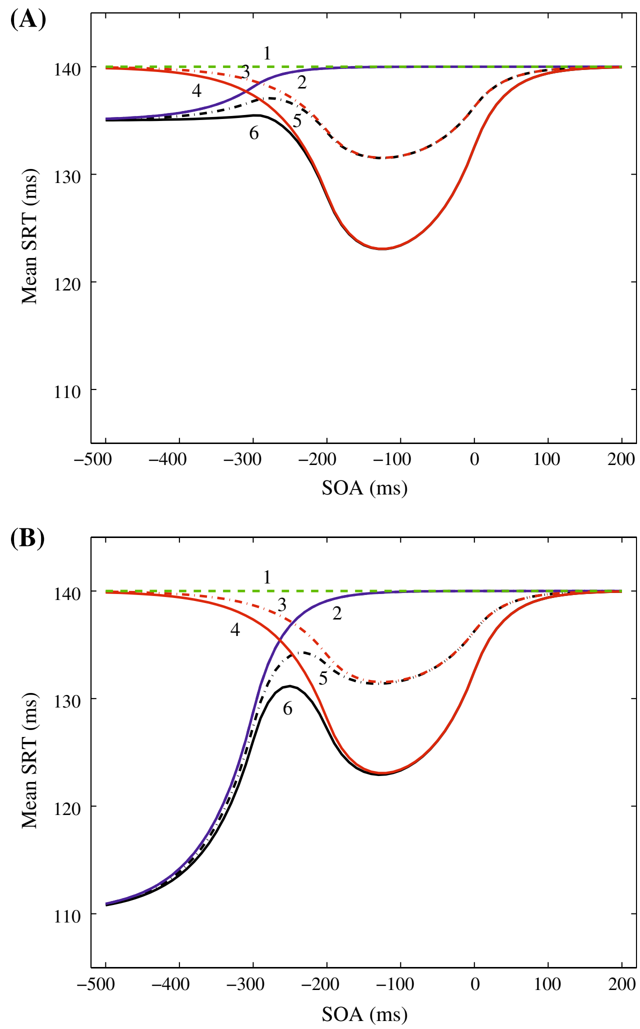

Fig. 1 Predicted mean SRT as a function of SOA. In **a** $\kappa = 5$ ms, in **b** $\kappa = 30$ ms. Line 1 (*horizontal*) indicates the predicted mean SRT to the target only. Lines 2–6 indicate the predicted mean SRT when target and non-target are presented, as a function of SOA. Line 2 refers to the mean SRT when only warning takes place. Lines 3 and 4 refer to the predicted mean SRT when bimodal stimuli were presented contra- and ipsilaterally and only integration occurs. Lines 5 and 6 refer to the predicted mean SRT when bimodal stimuli were presented ipsi- and contralaterally and both integration and warning occur

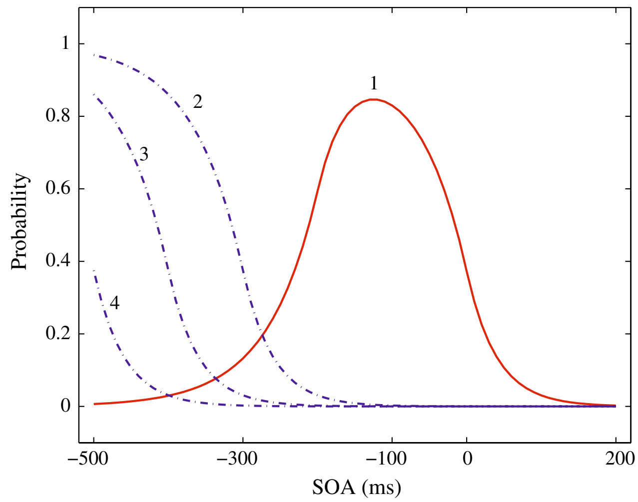

Fig. 2 Probability of integration (Line 1) and probability of warning (Lines 2, 3, and 4) with γ = 300, 400, and 500 ms, respectively

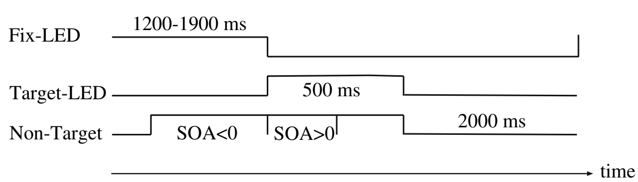

Fig. 3 Time course of a trial

Formal specification of TWIN-W

The race in the first stage of the model is made explicit by assigning independent non-negative random variables V and A to the peripheral processing times for the visual target and auditory non-target stimulus, respectively. For ease of exposition, we only consider the case of auditory non-targets. Whenever the non-target is tactile instead of auditory, symbol A will have to be replaced by symbol T in the expressions below. With τ as SOA value and ω_I as integration window width parameter, the time window of integration assumption is equivalent to the (stochastic) event I, say,

$$I = \{A + \tau < V < A + \tau + \omega\}.$$

Thus, the probability of integration to occur, $P(I)$, is a function of both τ and ω, and it can be determined numerically once the distribution functions of A and V have been specified (see below).

The warning mechanism of the non-target is triggered whenever the non-target wins the race by a certain margin or head start γ_A and, thus, its occurrence corresponds to the event:

$$W = \{A + \tau + \gamma_A < V\}.$$

The probability of warning to occur, $P(W)$, is a function of both τ and γ_A, and its value can be determined numerically as soon as the distribution functions of A and V have been specified (see below). By assumption (3), the head start γ_A is large enough for the integration window to close again, implying

$$\gamma_A > \omega \geq 0 \text{ and, therefore, } P(I \cap W) = 0.$$

The next step is to compute expected total reaction time for the unimodal and crossmodal conditions. From the two-stage assumption, total reaction time in the crossmodal condition can be written as a sum of two random variables:

$$\mathrm{RT}_{\mathrm{crossmodal}} = S_1 + S_2, \tag{1}$$

where S_1 and S_2 refer to the first and second stage processing time, respectively. For the expected saccadic reaction time in the crossmodal condition then follows:

$$\begin{aligned}
\mathrm{E}[\mathrm{RT}_{\mathrm{crossmodal}}] &= \mathrm{E}[S_1] + E[S_2] \\
&= \mathrm{E}[S_1] + P(I)\mathrm{E}[S_2|I] + P(W)\mathrm{E}[S_2|W] \\
&\quad + \{1 - P(I) - P(W)\}\mathrm{E}[S_2|I^c \cap W^c] \\
&= \mathrm{E}[S_1] + \mathrm{E}[S_2|I^c \cap W^c] \\
&\quad - P(I)\{\mathrm{E}[S_2|I^c \cap W^c] - \mathrm{E}[S_2|I]\} \\
&\quad - P(W)\{\mathrm{E}[S_2|I^c \cap W^c] - \mathrm{E}[S_2|W]\},
\end{aligned}$$

where $\mathrm{E}[S_2|I]$, $\mathrm{E}[S_2|W]$, and $\mathrm{E}[S_2|I^c \cap W^c]$ denote the expected second stage processing time conditioned on interaction occurring (I), warning occurring (W), or neither

Table 1 Absolute number of error types (percentage in parentheses) for all participants

Error	Participant				
	P1	P2	P3	P4	P5
Saccades before any signal	29 (0.3)	301 (3.5)	85 (1.0)	127 (1.5)	179 (2.1)
Amplitude not within 3 std	175 (2.0)	318 (3.7)	163 (1.9)	297 (3.4)	109 (1.3)
SRT < 80	18 (0.2)	171 (2.0)	19 (0.2)	84 (1.0)	242 (2.8)
SRT > 500	0	2	0	0	0
Directional	11 (0.1)	22 (0.3)	18(0.2)	54 (0.6)	166 (1.9)
Total	233 (2.7)	814 (9.4)	285 (3.3)	562 (6.4)	696 (8.0)

of them occurring ($I^c \cap W^c$), respectively (I^c, W^c stand for the complement of events I, W). Setting

$$\Delta \equiv \mathrm{E}[S_2|I^c \cap W^c] - \mathrm{E}[S_2|I],$$
$$\kappa \equiv \mathrm{E}[S_2|I^c \cap W^c] - \mathrm{E}[S_2|W],$$

this becomes

$$\mathrm{E}[\mathrm{RT}_{\mathrm{crossmodal}}] = \mathrm{E}[S_1] + \mathrm{E}[S_2|I^c \cap W^c] - P(I)\cdot \Delta - P(W)\cdot\kappa. \tag{2}$$

In the unimodal condition, no integration or warning is possible. Thus,

$$\mathrm{E}[\mathrm{RT}_{\mathrm{unimodal}}] = \mathrm{E}[S_1] + \mathrm{E}[S_2|I^c \cap W^c],$$

and we arrive at a simple expression for the combined effect of multisensory integration and warning, crossmodal interaction (CI),

$$\begin{aligned}\mathrm{CI} &\equiv \mathrm{E}[\mathrm{RT}_{\mathrm{unimodal}}] - \mathrm{E}[\mathrm{RT}_{\mathrm{crossmodal}}]\\ &= P(I)\cdot\Delta + P(W)\cdot\kappa.\end{aligned} \tag{3}$$

Note that Δ and κ can separately take on positive or negative values (or zero) depending on whether multisensory integration and warning have a facilitative or inhibitory effect.

Qualitative predictions

The basic assumptions of TWIN-W imply that for a given spatial configuration and non-target modality there are no sign reversals or changes in magnitude of Δ or κ across all SOA values. In contrast, both the probability of integration $P(I)$ and the probability of warning $P(W)$ do change with SOA. In particular, when the non-target is presented very late relative to the target (large positive SOA), its chances of winning the race against the target and thus opening the window of integration become very small. When it is presented rather early (large negative SOA), it is likely to win the race and to open the window, but the window may be closed by the time the target arrives. Again, the probability of integration, $P(I)$, is small. Therefore, the largest integration effects are expected for some mid-range SOA values. On the other hand, the probability of warning $P(W)$ decreases monotonically with SOA: the later the non-target is presented the smaller are its chances to win the race against the target with some head start γ.

It is interesting to note that this difference in how $P(I)$ and $P(W)$ should depend on SOA is, in principle, empirically testable by manipulating the conditions of the experiment. Specifically, if target and non-target are presented in two distinct spatial conditions, ipsilateral and contralateral, say, one would expect Δ to take on two different values, Δ_i and Δ_c, whereas $P(W)\cdot\kappa$, the expected non-spatial warning effect, should remain the same under both conditions. Subtracting the corresponding crossmodal interaction terms then gives, after canceling the warning effect terms (Eq. 3),

$$\mathrm{CI}_i - \mathrm{CI}_c = P(I)\cdot(\Delta_i - \Delta_c), \tag{4}$$

an expression that should yield the same qualitative behavior, as a function of SOA, as $P(I)$.

In a similar vein, if target and non-target are presented in two distinct presentation modes, e.g., blocking or mixing the modality of the auditory and tactile non-targets within an experimental block of trials, such that supposedly no changes in the expected amount of multisensory integration should occur, then subtraction of the corresponding CI values yields, after canceling the integration effect terms,

$$\mathrm{CI}_{\mathrm{blocked}} - \mathrm{CI}_{\mathrm{mixed}} = P(W)\cdot(\kappa_{\mathrm{mixed}} - \kappa_{\mathrm{blocked}}), \tag{5}$$

a quantity that should decrease monotonically with SOA because $P(W)$ does.

Quantitative predictions

In the parametric version of the TWIN model, all peripheral processing times are assumed to be exponentially distributed (cf. Colonius and Diederich 2004). Assuming explicit probability distributions is an essential requirement for calculating numerical values of $P(I)$ and $P(W)$, but the particular choice of the exponential is not: it is chosen here for computational simplicity only (see Appendix 1). The exponential distribution is characterized by a single quantity, the *intensity parameter* λ.

To illustrate the predictions of TWIN for mean SRT as a function of SOA we choose a set of arbitrary (but plausible) parameters. Let the parameter for the visual target processing time be $\lambda_{\mathrm{V}} = 0.025\ \mathrm{ms}^{-1}$ which amounts to a mean peripheral time for visual unimodal stimuli of 40 ms ($1/\lambda_{\mathrm{V}}$). The parameter for second stage central processing time when neither integration nor warning occurs, $\mu \equiv \mathrm{E}[S_2|I^c \cap W^c]$, is set to 100 ms. Note that we need not make any distributional assumptions about second stage processing time as long as we restrict our predictions to the level of mean SRT. Neither λ_{V} nor μ are directly observable but the sum of the peripheral and central processing time for the visual target stimulus constitutes a prediction for unimodal mean SRT:

$$\mathrm{E}[\mathrm{RT}_{\mathrm{unimodal}}] = 1/\lambda_{\mathrm{V}} + \mu.$$

Since this expression does not depend on SOA it is indicated as a horizontal line (Line 1) in Fig. 1. For predicting crossmodal SRTs, we have to specify the remaining parameters. Let the parameter of non-target processing time be $\lambda_{\mathrm{NT}} = 0.015\ \mathrm{ms}^{-1}$, corresponding to a mean

Table 2 ANOVA table (*F* values with associated *p* values) for all participants

Effects	Participant				
	P1	P2	P3	P4	P5
laterality ($df = 1$)	9.3 ($p = 0.002$)	48.4 ($p < 0.001$)	48.9 ($p < 0.001$)	53.7 ($p < 0.001$)	151.6 ($p < 0.001$)
modality ($df = 1$)	67.8 ($p < 0.001$)	6.7 ($p = 0.010$)	2.1 ($p = 0.148$)	32.2 ($p < 0.001$)	0.9 ($p = 0.354$)
SOA ($df = 14$)	31.8 ($p < 0.001$)	6.9 ($p < 0.001$)	18.8 ($p < 0.001$)	24.3 ($p < 0.001$)	30.4 ($p < 0.001$)
pres. mode ($df = 1$)	2.6 ($p = 0.106$)	1.7 ($p = 0.192$)	3.0 ($p = 0.082$)	3.3 ($p = 0.071$)	34.4 ($p < 0.001$)
lat × *SOA* ($df = 14$)	2.5 ($p = 0.002$)	2.2 ($p = 0.005$)	2.8 ($p < 0.001$)	1.5 ($p = 0.098$)	6.2 ($p < 0.001$)
modality × *SOA* ($df = 14$)	6.0 ($p < 0.001$)	2.0 ($p = 0.015$)	2.3 ($p = 0.003$)	3.9 ($p < 0.001$)	2.5 ($p = 0.002$)

peripheral time of 66.7 ms. The time window of integration is set to 200 ms, and the head start, γ, for the warning cue is set to 300 ms. The parameters for multisensory integration are set to $\Delta_i = 20$ and $\Delta_c = 10$ ms for ipsilateral and contralateral bimodal stimuli, respectively, implying a facilitation effect. Finally, the warning cue effect also reducing mean SRT to bimodal stimuli, κ, is set to 5 ms (a) or 30 ms (b) in Fig. 1 which presents the mean SRT predictions as a function of SOA for these parameter settings.

Line 2 presents mean SRT when only warning contributes to SRT reduction: its effect is much larger in the right panel ($\kappa = 30$ ms) than in the left panel ($\kappa = 5$ ms), but with increasing SOA it decreases to zero in both cases. When no warning occurs, multisensory integration reduces mean SRT, but the size of the effect depends on the Δ parameter: it is larger for stimuli presented ipsilaterally (Line 4) than contralateral (Line 3). Moreover, as already observed as a qualitative prediction, due to the limited width of the integration time window, the facilitation effect is most pronounced for a certain medium range of SOA values and disappears completely for large enough negative or positive values. Finally, when both integration and warning mechanisms are involved (Lines 5 and 6 for contra- and ipsilateral conditions, respectively), the integration effect remains and the warning mechanism produces a reduction of SRT even for rather large negative SOA values. As seen in the right panel, this reduction can even be more pronounced than the integration effect if the warning parameter is large enough (here: a κ value of 30 ms).

Figure 2 depicts the probability of integration (Line 1) and, for three different values of the head start parameter γ, the corresponding probability of warning as a function of SOA. These probability curves are not observable but their time course illustrates the interplay between warning and integration: (a) the probability of warning decreases towards zero as SOA increases, with the head start parameter controlling the rate of decline; (b) the probability of integration is a peaked function of SOA such that the probability of integration is highest when the probability of warning is (close to) zero.[1]

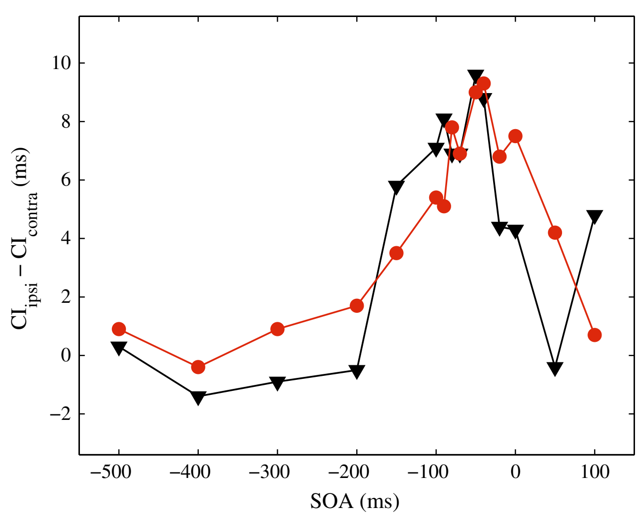

Fig. 4 Qualitative model test: Both functions (circles for auditory, triangles for tactile non-targets) should have a shape similar to $P(I)$

Methods: Experiment

In a focused attention paradigm, saccadic reaction times to a visual target stimulus were measured in the presence or absence of an auditory or tactile accessory stimulus (non-target). The hypothesis is that the non-target triggers a warning mechanism when it is presented early relative to the target, and that it affects saccadic reaction time independent of its spatial position. On the other hand, the multisensory integration mechanism is supposed to occur only if target and non-target fall within a window of integration and that it is most prominent for ipsilateral configurations. In order to tear apart the warning and

[1] This description of how the probability of integration and the probability of warning change with SOA should not conceal the assumption that, for any given trial, warning and integration cannot occur simultaneously.

integration processes a large range of stimulus onset asynchrony values was utilized and the spatial position of the non-target relative to the target was varied as well (ipsi- vs. contralateral).

Participants

Five undergraduate students, aged 18–22, three female, served as paid voluntary participants. All had normal or corrected-to-normal vision and three were right-handed (self-description). They were screened for their ability to follow the experimental instructions (proper fixation, few blinks during trial, saccades towards visual target). They gave their informed consent prior to their inclusion in the study. The experiment was conducted in accordance with the ethical standards described in the 1964 Declaration of Helsinki.

Apparatus and stimulus presentation

The fixation point and the visual stimuli were red light emitting diodes (LEDs) (25 mA, 5.95 mcd and 25 mA, 3.3 mcd, respectively) situated on a table 60 cm in front of the participant, the fixation point in the center and the target LEDs 20° to the left and right of it, respectively. The tactile stimulation (50 Hz, 1 V, 1–2 mm amplitude) was generated by two silenced oscillation exciters (Mini-Shaker, Type 4810, Bruel and Klær) placed on bases situated under the table. Positioned in each shaker was a metal rod extending through a hole in the table approximately 2 cm above the surface. On each rod was a wooden ball of 14 mm diameter, which rested in the palm of the participant and transmitted the vibration to the hand. They were placed 20° to the left and right of the fixation LED and 62 cm in front of the participant. Auditory stimuli were bursts of white noise (59 dbA) generated by two speakers (Canton Plus XS). The speakers were placed at 20° to the left and right of the fixation LED at the same hight as the participants' ear level and 120 cm in front of the participants. The black table was 180 × 130 × 75 cm with a recess to sit in. One PC controlled the stimulus presentation, and two other interlinked PCs controlled the EyeLink program.

Experimental procedure

The participants were seated in a completely darkened, sound attenuated room with the head positioned on a chin rest and the elbows and lower arms resting comfortably on a table. They were unable to see their hands during the experiment. Although the eye movement equipments takes head movements into account, the participants were instructed to leave the head on the chin rest and not to move the head. Every experimental session began with 10 min of dark adaptation during which the measurement system was adjusted and calibrated. During this phase the participants put their hands at the position used during the entire experimental block. Thus, they were aware of the hand position and, thus, the position of the tactile stimulus.

Each trial began with the appearance of the fixation point. After a variable fixation time (1,200–1,900 ms), the fixation LED disappeared and, simultaneously, the visual target stimulus was turned on (i.e., no gap). Participants were instructed to gaze at the visual target as quickly and as accurately as possible ignoring any auditory or tactile non-targets (focused attention paradigm). The visual target appeared alone or in combination with either a tactile or an auditory non-target in ipsi- or contralateral position.

The onset of non-targets was shifted by a stimulus onset asynchrony (SOA) of −500, −400, −300, −200, −150, −100, −90, −80, −70, −50, −40, −20, 0, 50, or 100 ms. Negative values mean that the non-target was

Table 3 Restrictions to model parameters in the estimation routine

Parameters	Restriction limits (in ms)	
$1/\lambda_V$	20–143	Mean peripheral processing time for visual stimulus
$1/\lambda_T$	20–143	Mean peripheral processing time for tactile stimulus
$1/\lambda_A$	20–143	Mean peripheral processing time for auditory stimulus
μ	>0	Mean central processing time
ω_T	≤ 300	Window of integration for tactile stimulus
ω_A	≤ 300	Window of integration for auditory stimulus
$\Delta_{T_{ipsi}}$	None	Amount of crossmodal interaction due to tactile stimulus presented ipsilaterally
$\Delta_{T_{contra}}$	None	Amount of crossmodal interaction due to tactile stimulus presented contralaterally
$\Delta_{A_{ipsi}}$	None	Amount of crossmodal interaction due to auditory stimulus presented ipsilaterally
$\Delta_{A_{contra}}$	None	Amount of crossmodal interaction due to auditory stimulus presented contralaterally
κ	≥ 0	Amount of warning cue effect
γ_T	$>\omega_T$	Head start for tactile stimulus
γ_A	$>\omega_A$	Head start for auditory stimulus

Table 4 Parameters for all participants (estimated after merging mixed and blocked conditions)

Participant	$1/\lambda_V$	$1/\lambda_T$	$1/\lambda_A$	μ	ω_T	ω_A	$\Delta_{T_{ipsi}}$	$\Delta_{T_{contra}}$	$\Delta_{A_{ipsi}}$	$\Delta_{A_{contra}}$	κ	γ_T	γ_A	χ^2
P1	20	143	74	134	173	120	27	25	38	30	12	322	156	83
P2	30	125	49	117	147	100	24	7	16	7	4	147	600	58
P3	34	101	86	139	259	300	17	11	20	12	8	259	470	105
P4	38	143	46	98	187	300	19	12	16	11	6	230	300	100
P5	20	88	85	125	144	109	35	11	43	14	11	263	292	124

χ^2 refers to the expression minimized by the estimation routine (Eq. 7)

presented before the target. The visual stimuli were presented for 500 ms; the auditory and tactile non-targets were turned off simultaneously with the visual stimulus. Thus their duration varied between 1,000 and 400 ms, depending on SOA. Stimulus presentation was followed by a break of 2,000 ms in complete darkness, before the next trial began, indicated by the onset of the fixation LED. The time course of a trial is sketched in Fig. 3.

To investigate the influence of experimental settings on the size of the warning effect two different presentation modes were employed. In one set of trials, hereafter called *blocked presentation mode*, the non-target was either auditory or tactile. Sixteen trials were visual targets only (unimodal), 60 bimodal trials were presented ipsilaterally and 60 bimodal trials were presented contralaterally. In another set of 136 trials, hereafter called *mixed presentation mode*, bimodal stimuli could be visual-auditory and visual-tactile. That is, a mixed block consisted of 30 visual-auditory and 30 visual-tactile bimodal stimuli, presented ipsilaterally as well as contralaterally, plus 16 unimodal stimuli. All trial types were presented in random order. Two blocks of mixed and two blocks of blocked presentation modes were conducted within one hour. A participant's first two hours were considered as training and the results not included in the data analysis. Afterwards, each participant was engaged for 6 h completing a total of 8,704 trials (136 trials per condition).

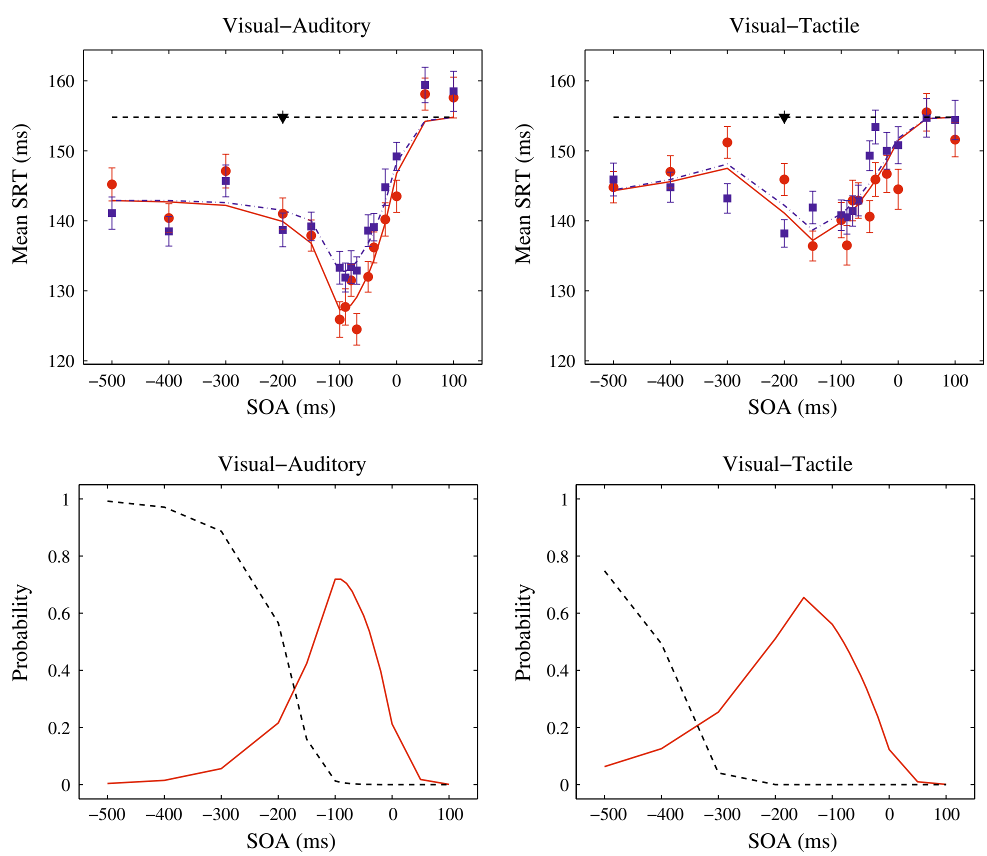

Fig. 5 Participant P1. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

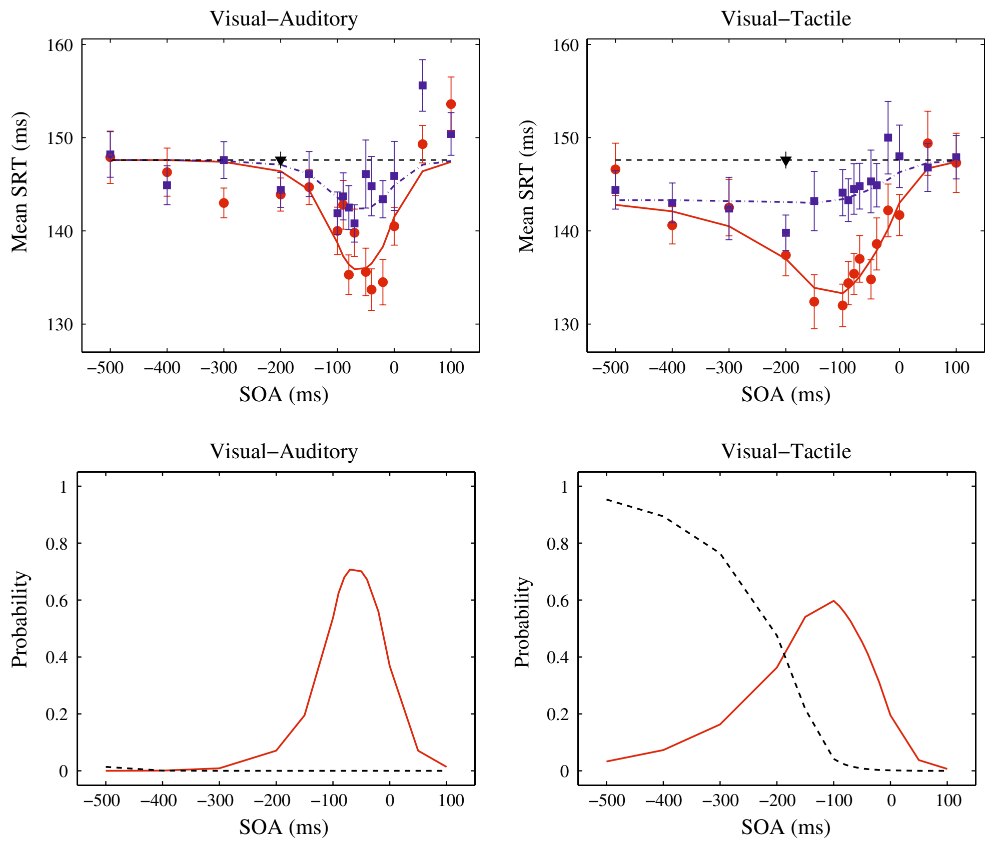

Fig. 6 Participant P2. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

Data collection

Saccadic eye movements were recorded by an infrared video camera system (EyeLink II, SR Research) with a temporal resolution of 500 Hz and a horizontal and vertical spatial resolution of 0.01°. Criteria for saccade detection on a trial by trial basis were velocity (35 deg s^{-1}) and acceleration (9,500 deg s^{-2}). The recorded eye movements from each trial were checked for proper fixation at the beginning of the trial, eye blinks and correct detection of start and end point of the saccade.

Results

Saccades were screened for anticipation errors (SRT < 80 ms), misses (SRT > 500 ms), and accuracy: trials with saccade amplitudes deviating more than three standard deviations from the mean amplitude were excluded from the analysis. Table 1 presents the frequency of different error types for all five participants. The mean amplitude (standard deviation) for each participant was 16.5° (2.8°), 18.1° (2.7°), 18.1° (1.9°), 17.1° (2.4°), and 17.0° (4.3°), respectively.

Mean SRTs of all participants under all uni- and bimodal conditions are listed in Tables 5, 6, 7, 8, 9 and 10 in Appendix 2. We defined four factors for an analysis of variance as follows: *laterality* with levels ipsilateral and contralateral, *modality* with levels auditory (A) and tactile (T), *SOA* with levels −500, −400, −300, −200, −150, −100, −90, −80, −70, −50, −40, −20, 0, 50, 100 ms, and *presentation mode* with levels blocked and mixed. Four-way (1 × 3 × 15 × 2) ANOVAs were conducted for each participant separately, such that the individual trials were the random unit.

The ANOVA results for all participants are shown in Table 2. Except for participant P5, the factor *presentation mode* had no effect on mean SRT. For participants P3 and P5 the factor *modality* showed no significant results on mean SRT. However, post-hoc tests (Tukey) revealed significant differences between unimodal stimuli (i.e., no non-target) and bimodal stimuli, for both auditory and tactile non-targets. The remaining main effects were all significant. For all participants the two-way interaction *modality* × *SOA* was significant at the $p \leq 0.02$ level. For four participants (P1, P2, P3 and P5) the interaction *laterality* × *SOA* was significant at the $p \leq 0.005$ level. None of the remaining two-way interactions,

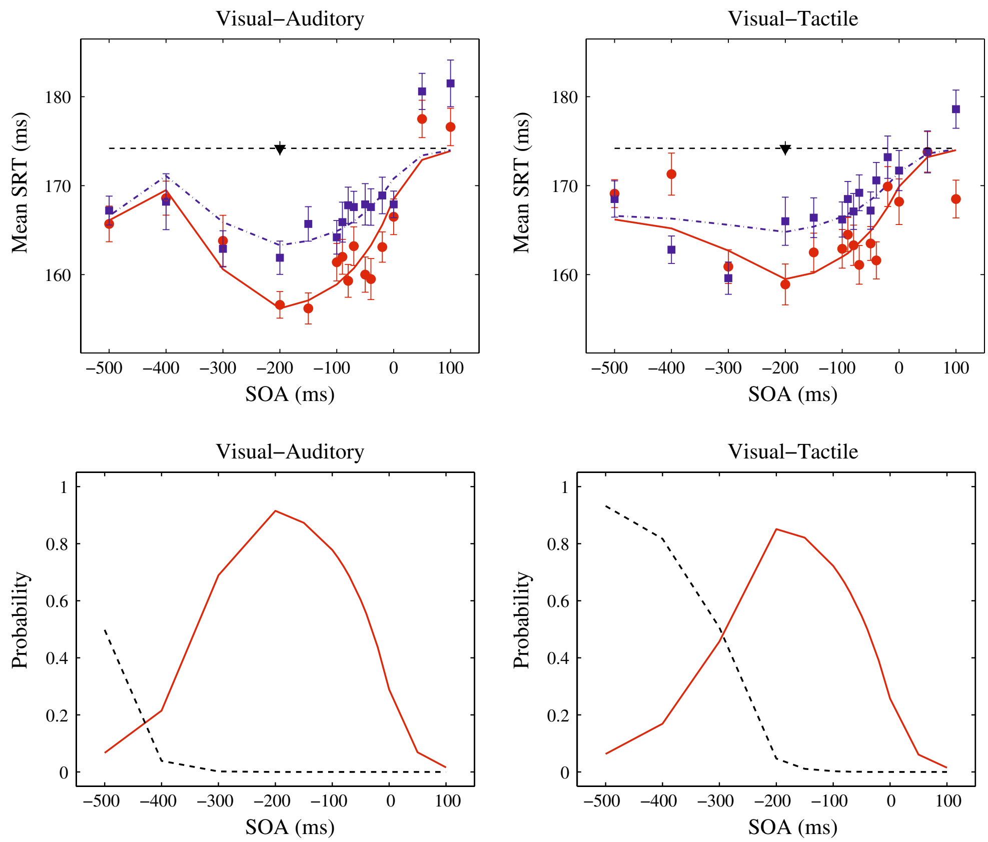

Fig. 7 Participant P3. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

or higher interactions, was significant at the $p \leq 0.01$ level.

Post-hoc tests (Tukey) gave the following:

(1) For all participants mean SRT to a visual target was shorter in the presence of non-targets ($p < 0.001$). For participants P1 and P4 mean SRT to visual-auditory stimuli was shorter than to visual-tactile stimuli ($p < 0.001$). For participant P2 mean SRT to visual-auditory stimuli was longer than to visual-tactile stimuli ($p = 0.027$). For participants P4 and P5 the difference between visual-auditory and visual-tactile mean SRT was not significant.
(2) For all participants mean SRTs were shorter when target and non-targets were presented ipsilaterally rather than contralaterally ($p < 0.005$).

Since factor *presentation mode* was not significant for participants P1, P2, P3, and P4, their data were collapsed across presentation mode for model fitting, separately for each individual. P5's data for blocked and mixed presentation modes were collapsed as well to simplify the data fitting procedure.

Model fits

Qualitative test

We consider the test implied by Eq. 4,

$$\mathrm{CI}_i - \mathrm{CI}_c = P(I) \cdot (\Delta_i - \Delta_c).$$

Since according to the model the difference $\Delta_i - \Delta_c$ does not depend on SOA, this expression should have the same qualitative shape as $P(I)$, i.e., it should go to zero for small and for large enough SOA values and peak in between, at around SOA = −100 ms. Figure 4 displays the results from averaging the individual curves across all participants. Apart from variability due to sampling errors and some individual differences in the parameters (see data vs. model prediction figures in the next section), the shapes of both curves (circles for auditory, triangles for tactile non-targets) are roughly in agreement with the predicted shape of $P(I)$.

The second qualitative test that would be possible, capitalizing on the difference between blocked and mixed presentations, was not performed since the *presentation mode* factor did not reach significance in the analysis of variance (except for one participant).

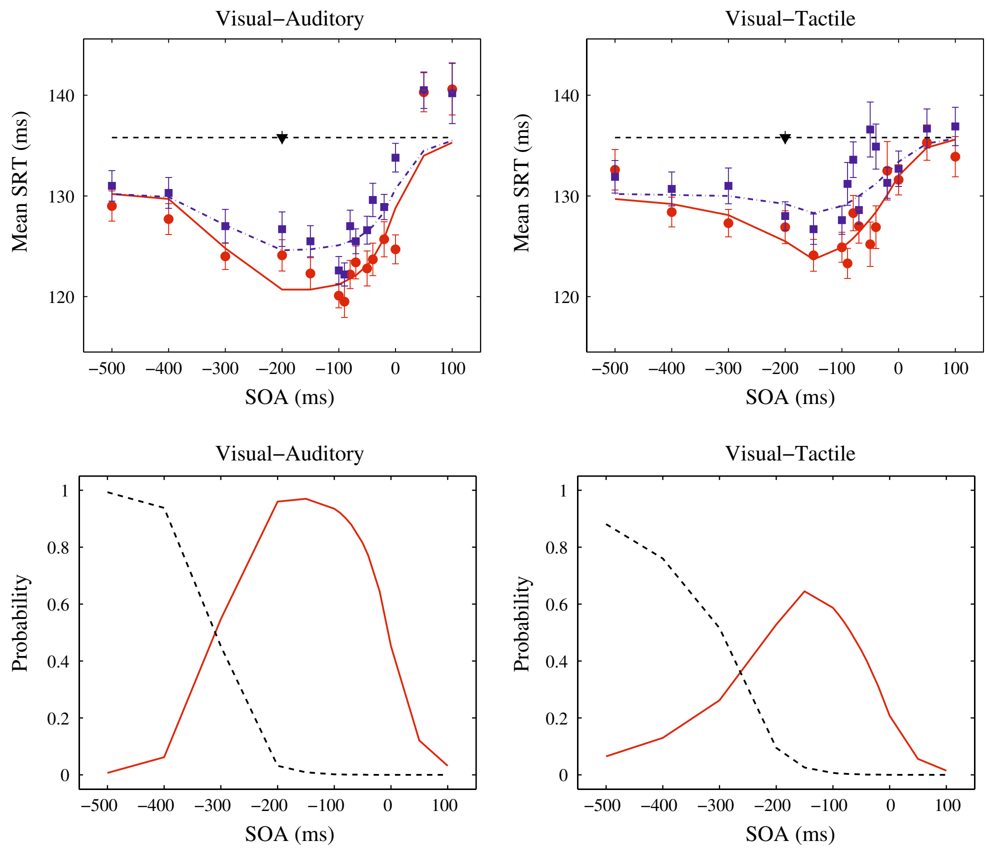

Fig. 8 Participant P4. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

Quantitative tests

Table 3 (left column) lists all parameters of TWIN-W that have to be estimated for each participant separately. For each participant there are 13 parameters to be estimated from 61 data points (i.e., mean SRT values for visual-auditory and visual-tactile stimulus pairs presented ipsi- and contralaterally in 15 SOA steps, plus one unimodal stimulus presentation).

The parameters were estimated by minimizing the Pearson χ^2 statistic

$$\chi^2 = \sum_{n=1}^{N}\sum_{i=1}^{2}\sum_{j=1}^{2}\left(\frac{\overline{\text{SRT}}(i,j,n) - \widehat{\overline{\text{SRT}}}(i,j,n)}{\sigma_{\overline{\text{SRT}}(i,j,n)}}\right)^2 + \left(\frac{\overline{\text{SRT}}_{\text{LED}} - \widehat{\overline{\text{SRT}}}_{\text{LED}}}{\sigma_{\overline{\text{SRT}}_{\text{LED}}}}\right)^2 \quad (6)$$

using the FMINSEARCH[2] routine of MATLAB. Here $\overline{SRT}(i,j,n)$ and $\widehat{\overline{SRT}}(i,j,n)$ are, respectively, the observed and the fitted values of the mean SRT to bimodal stimuli (visual-auditory, $i = 1$; visual-tactile, $i = 2$) presented in spatial positions (ipsilateral, $j = 1$; contralateral, $j = 2$) with SOA (referred to by n to $N = 15$); $\sigma_{\overline{T}_{ij(k,n)}}$ are the respective standard errors. The second term in Eq. 7 refers to the unimodal visual target condition.

The right column of Table 3 lists the parameter restrictions that were imposed on the estimation routine. The λ values were restricted to fall within a range consistent with neurophysiological estimates for peripheral processing times (Stein and Meredith 1993; Groh and Sparks 1996). The width of the the time window of integration for a tactile and an auditory non-target, ω_T and ω_A, respectively, was limited to an upper bound of 300 ms.

All parameter estimates are shown in Table 4. The minimized χ^2 values are listed in the rightmost column of the table. Except for P2 (a participant highly trained in earlier experiments) having a p value of 0.15 ($\chi^2_{(48)} = 58$), all other participants show statistically significant deviations from the model predictions ($p < 0.01$).

The upper panels in Figs. 5, 6, 7, 8 and 9 plot the observed mean SRT values (±standard error) and model predictions across SOAs for all participants, separately for target and non-target presented in the same hemifield

[2] FMINSEARCH uses the simplex search method of Lagarias et al. (1998).

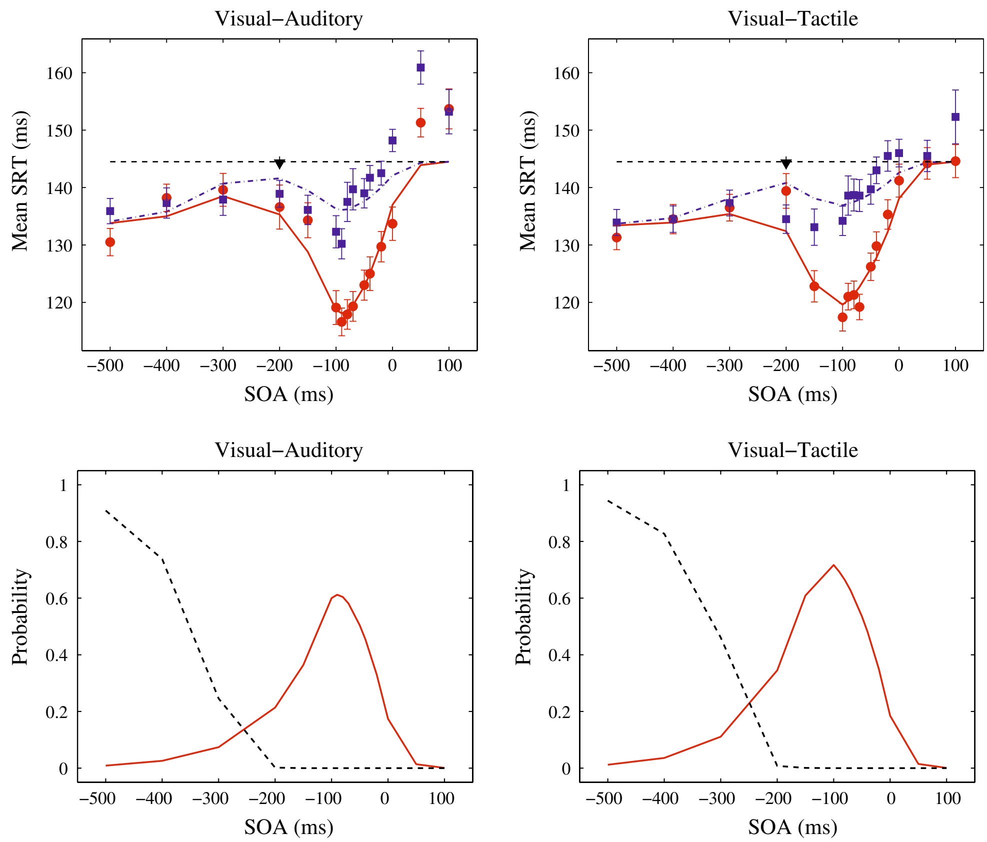

Fig. 9 Participant P5. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

(ipsilateral) or in different hemifields (contralateral). The lower panels depict the probability of integration (line) and the probability of warning (dashed line) computed from the model parameter estimates specific to the participant. In the upper panels, squares (circles) refer to mean SRTs to bimodal stimuli presented ipsilaterally (contralateral), corresponding model predictions are presented by a dashed line (ipsilateral) and a line (contralateral). The horizontal dashed line marks the mean unimodal (visual) SRT. In Figs. 5, 6, 7, 8 and 9 left panels depict results for auditory non-targets, right panels for tactile non-targets.

Visual inspection of Figs. 5, 6, 7, 8 and 9 suggests that the TWIN model does reflect the main qualitative properties of the data rather well: (a) for large negative SOA, there is some facilitation but no difference between ipsi- and contralateral presentation of target and non-target, (b) facilitation, and the difference between ipsi- and contralateral, increase until a maximum around SOA = −200 to −100 is achieved, and (c) facilitation decreases toward zero with increasing SOA further. Calculating the amount of variance accounted for by the model (adjusted R^2) yields values of more than 80% for all subjects.

The relatively high number of observations underlying the observed mean SRTs (136 per data point) allows for a rather powerful model test demonstrating that the model does not capture many of the finer details of the SOA curves. However, there are no deviations that occur consistently across all subjects, suggesting that these deviations may be idiosyncratic and not due to factors the model is supposed to take into account. There is one exception, however: For all participants and under many conditions, for the positive SOA values 50 and 100 ms mean SRT values lie above the unimodal SRT mean. This “late inhibition” phenomenon cannot be predicted by the TWIN model in its current form (see discussion section below).

Test of alternative model

The alternative model allowing warning and integration to occur in one and the same trial was also fitted to the data. Assuming (a) that the events of warning and integration are statistically independent and (b) that their SRT effects

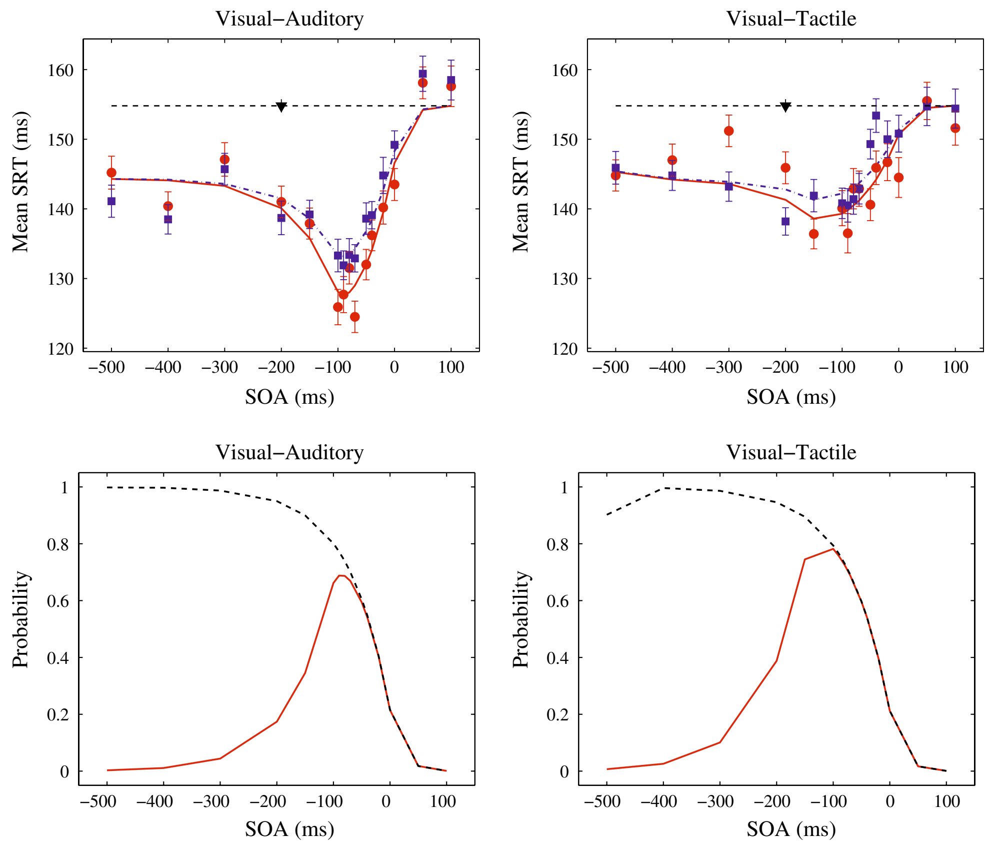

Fig. 10 Participant P1. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

combine in an additive fashion, the SOA functions predicted by this model are depicted in the upper panels of Figs. 10, 11, 12, 13, 14 together with the data, separately of each participant. Although the predictions differ slightly from the ones of the original model, there is no marked difference in overall fit between the models. This is corroborated by the corresponding χ^2 values (not presented here) showing no improvement consistently across all participants. The most conspicuous difference between the models shows up in the lower panels of the figures presenting the probabilities of integration and warning as a function of SOA: In the original model, the probability of warning quickly decreases with SOA and the probability of integration increases, whereas in the alternative model the probability of warning stays up high and decreases in line with the probability of integration. In other words, the alternative model predicts an effect, albeit small, of warning even for simultaneous presentation of target and non-target.

Discussion

The crossmodal spatial interaction effects in saccades reported here are in line with findings from other labs using visual targets and either auditory or tactile non-targets (e.g., Amlôt et al. 2003; Bell et al. 2005; Frens et al. 1995; Harrington and Peck 1998) and also with our own results (Colonius and Diederich 2004; Diederich et al. 2003; Diederich and Colonius 2007a, b). None of the previous studies, however, applied stimulus asynchrony values over such a broad range and in such large numbers. The time course of crossmodal mean SRT over SOAs and the pattern of facilitation observed here suggested the existence of two distinct underlying mechanisms: (a) a spatially unspecific crossmodal warning triggered by the non-target being detected early enough before the arrival of the target plus (b) a spatially specific multisensory integration mechanism triggered by the target processing time terminating within the time window of integration. Note the two roles played by the non-target in this conceptual setup: It can act either as a warning cue or become an element of a bimodal stimulus configuration inducing multisensory integration.

Whether, at any given trial, only one mechanism can be operative or whether both—warning and integration—can co-occur could not be settled in this study because they yielded very similar fits to the data. In both versions of the TWIN model the likelihood for the events of warning and integration —$P(W)$ and $P(I)$—can be calculated as a function of SOA, and these functions turned out to differ

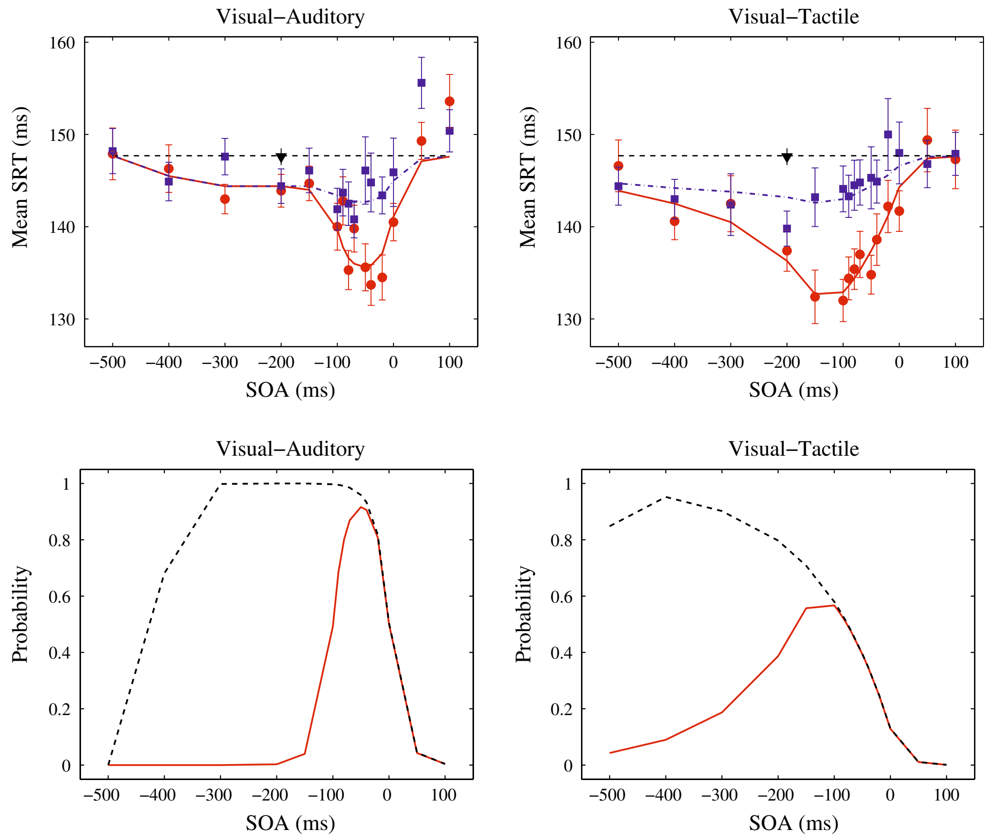

Fig. 11 Participant P2. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

quite substantially. This may potentially offer an opportunity for an experimental test between these models in future investigations.

The extended TWIN model gave a rather satisfactory account of the mean observed SRTs over the wide range of stimulus onset asynchrony values, except for the positive SOA values. The "late inhibition" exhibited by all participants in both ipsi-and contralateral conditions could not be predicted by the model, for a simple reason: the warning parameter κ is assumed to be invariant over SOA; in particular, it cannot change its sign. A facilitative warning effect for very early non-target presentations forces κ to be greater than zero, thus it cannot simultaneously switch to a negative value to account for the late inhibition effect. In a similar vein, the multisensory integration parameter Δ is either positive or negative over all SOA values, thus it cannot contribute to the late inhibition either as long as there is facilitative multisensory interaction for some SOA values. Since this invariance of Δ over SOA is essential for the two-stage assumption and the separation of crossmodal and intramodal stimulus properties in TWIN, it should not be given up. On the other hand, it would be straightforward to extend the warning mechanism assumption in such a way that the non-target will have a *negative* warning effect whenever peripheral target processing has finished before the non-target processing or, formally, when $V < A + \tau$ (for auditory non-targets). The probability of this event will be large for large SOA (τ) leading to the late inhibition behavior observed. For some of the participants the absolute amount of positive and negative warning appears to be about equal (cf. Figs. 5, 6, 7, 8, 9) so that no new (negative) κ parameter value would be necessary, whereas for others an additional parameter may have to be estimated. In any event, we refrain from fitting such an extended warning mechanism model here since it would only be an a posteriori explanation. A more comprehensive exploration of the late inhibition phenomenon using additional levels of positive SOA values seems to be worthwhile for future investigations.

Several aspects of the time window of integration model introduced in Colonius and Diederich (2004) have now been tested, including the effect of the number of non-targets presented (Diederich and Colonius 2007a) and of the spatial configuration of the stimuli (Diederich and Colonius 2007b). The present study demonstrates that the TWIN model can be extended in a simple way to account

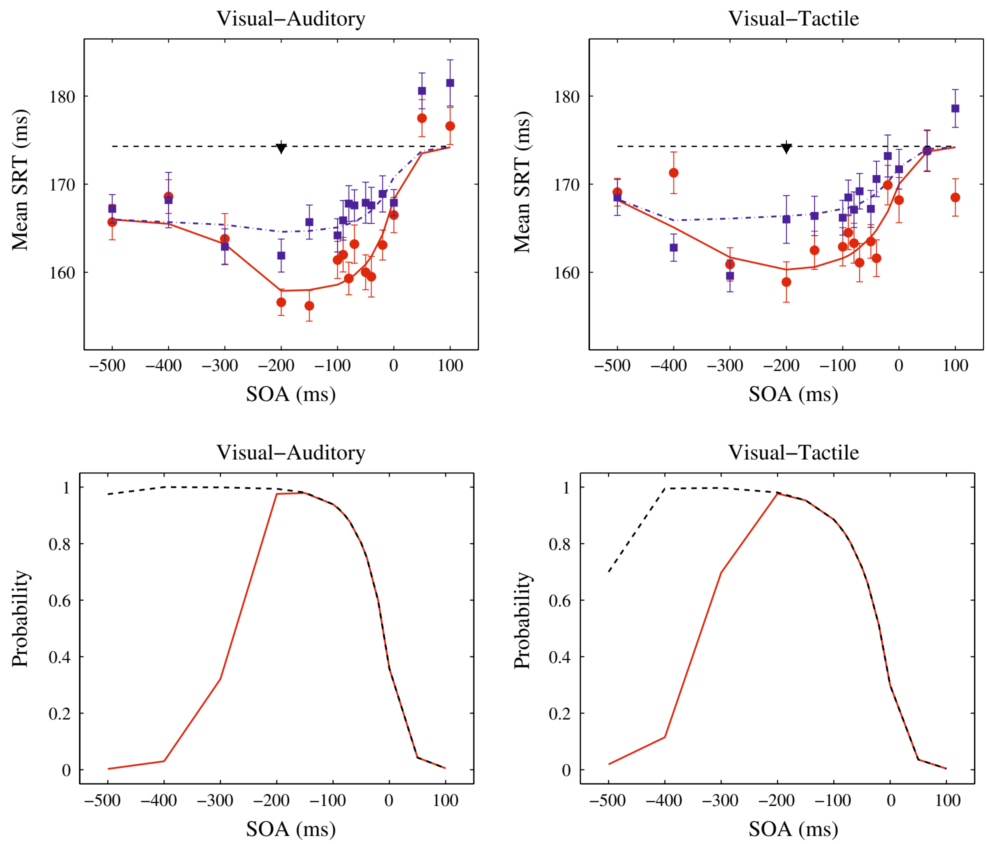

Fig. 12 Participant P3. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

for the warning effect observable whenever the non-target is presented more than about 200–300 ms before the target. Importantly, this extension is in keeping with the general structure of the model, i.e., the race among the peripheral processes and the time window of integration. The only additional parameter (for the model version claiming exclusiveness of warning and integration) is the head start, γ, the time it takes for the time window to close before the arrival of the target stimulus. Estimates of this head start vary from 150 to 350 ms depending on subject and stimulus modality.[3]

Concluding, it should be noted that the idea of multisensory integration being determined by a time window had already been suggested in Meredith et al. (1987) recording from SC neurons. It, together with the concept of a race among peripheral processes, now underlies many studies of crossmodal interaction (e.g., Van Opstal and Munoz 2004; Lewald and Guski 2003; Morein-Zamir et al. 2003; Spence and Squire 2003; see Whitchurch and Takahashi 2006, for head saccades in the barn owl) and also occurs in recent studies on the integration of audiovisual speech (Navarra 2005; Van Wassenhove et al. 2007; Van Atteveldt et al. 2007). Separation of multisensory integration processes from warning mechanism at the neural level seems to be quite intricate at this point, however, given that the detection of temporal synchrony–asynchrony of visual-auditory stimulus configurations is obviously not limited to the SC but, rather, is mediated by a large-scale neural network of insular, posterior parietal, prefrontal, and cerebellar areas possessing a functional interaction with the SC (e.g., Bushara et al. 2001).

[3] There were also 2 “outlier” estimates beyond that range most likely due to some problem in the difficult optimization task.

Acknowledgment This research was supported by grants from Deutsche Forschungsgemeinschaft Di 506/8-1 and /-3. We are grateful to Rike Steenken, and Stefan Rach for several discussions of this work and and to Dr. Annette Schomburg for her help in setting up the experiment.

Appendix 1

Here we derive the probability of the events I (multisensory integration) and W (warning) under the exponential distribution assumption. Specifically, the probability distributions for peripheral processing times V for a visual target and A for an auditory non-target are, respectively,

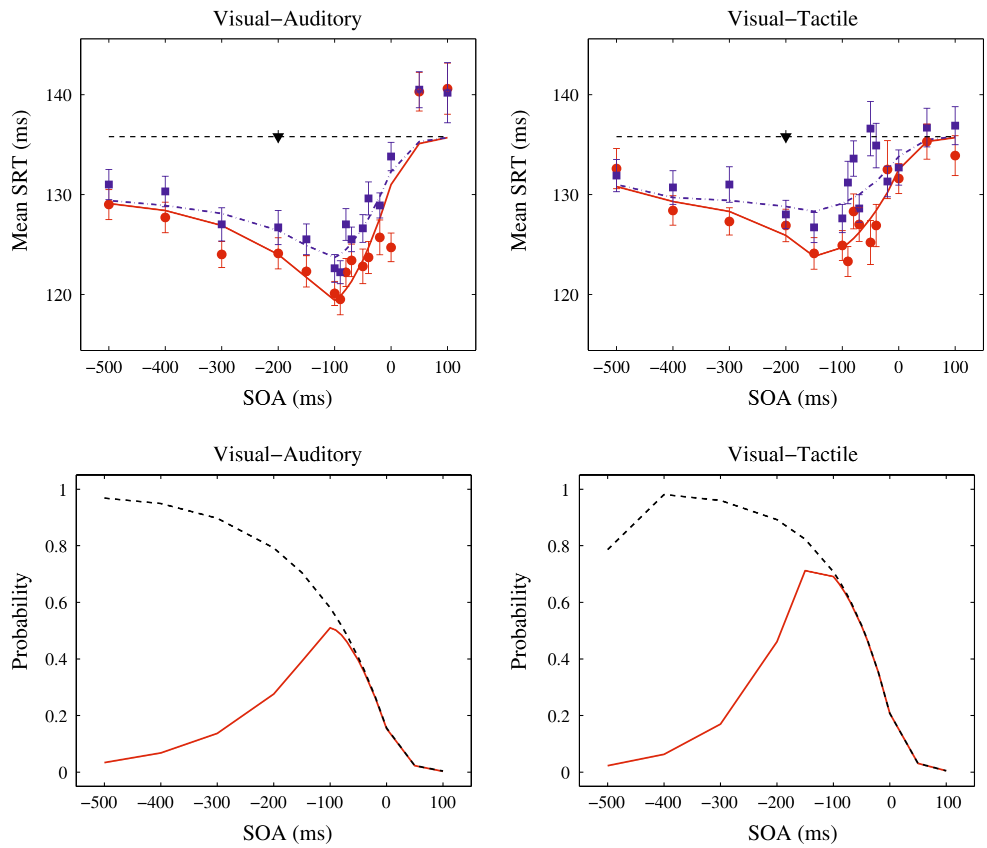

Fig. 13 Participant P4. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

$$f_V(t) = \lambda_V e^{-\lambda_V t},$$
$$f_A(t) = \lambda_A e^{-\lambda_A t}$$

for $t \geq 0$, and $f_V(t) = f_A(t) \equiv 0$ for $t < 0$. The corresponding distribution functions are referred to by $F_V(t)$ and $F_A(t)$.

Probability of integration: *P*(*I*)

By definition,

$$P(I) = Pr(A + \tau < V < A + \tau + \omega) = \int_0^\infty f_A(a)\{F_V(a + \tau + \omega) - F_V(a + \tau)\}da,$$

where τ denotes the SOA value and ω is the width of the integration window. Computing the integral expression requires that we distinguish between three cases for the sign of $\tau + \omega$:

(a) $\tau < \tau + \omega < 0$

$$P(I) = \int_{-\tau-\omega}^{-\tau} \lambda_A e^{-\lambda_A a}\{1 - e^{-\lambda_V(a+\tau+\omega)}\}da + \int_{-\tau}^{\infty} \lambda_A e^{-\lambda_A a}\{e^{-\lambda_V(a+\tau)} - e^{-\lambda_V(a+\tau+\omega)}\}da = \frac{\lambda_V}{\lambda_V + \lambda_A} e^{\lambda_A \tau}(-1 + e^{\lambda_A \omega});$$

(b) $\tau < 0 < \tau + \omega$

$$P(I) = \int_0^{-\tau} \lambda_A e^{-\lambda_A a}\{1 - e^{-\lambda_V(a+\tau+\omega)}\}da + \int_{-\tau}^{\infty} \lambda_A e^{-\lambda_A a}\{e^{-\lambda_V(a+\tau)} - e^{-\lambda_V(a+\tau+\omega)}\}da = \frac{1}{\lambda_V + \lambda_A}\{\lambda_A(1 - e^{-\lambda_V(\omega+\tau)}) + \lambda_V(1 - e^{\lambda_A \tau})\};$$

(c) $0 < \tau < \tau + \omega$

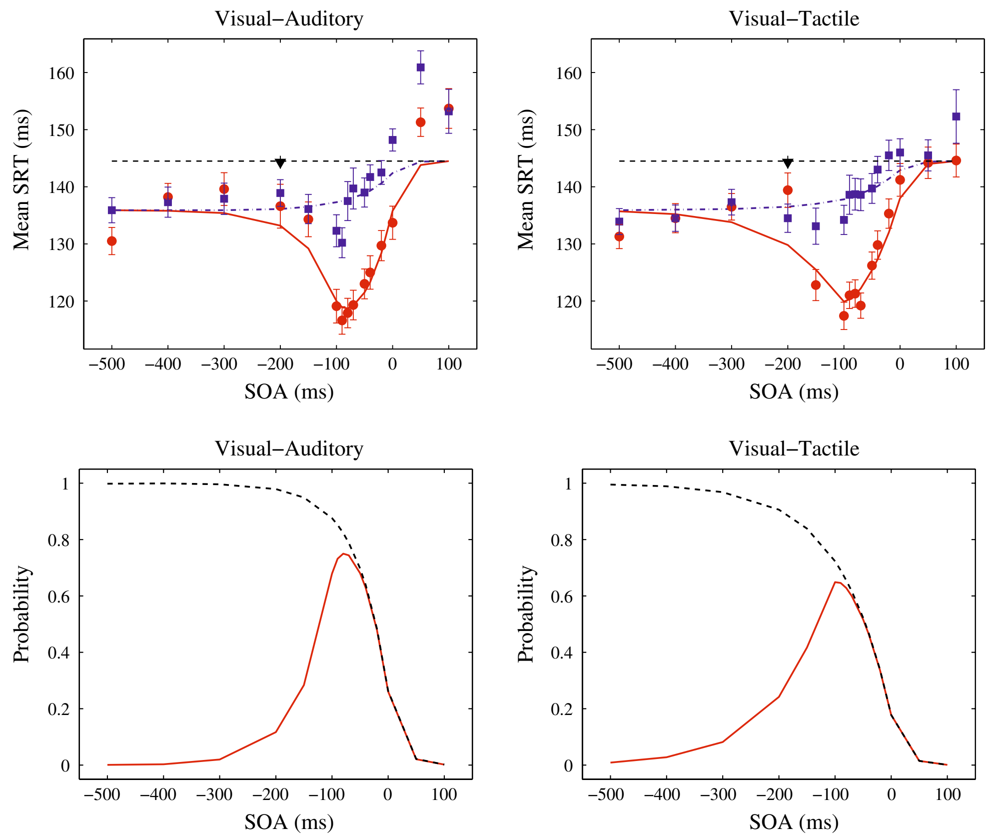

Fig. 14 Participant P5. *Upper panels* Observed and predicted mean SRT for visual-auditory stimuli (*left*) and visual-tactile stimuli (*right*). *Squares* refer to observed mean SRT to bimodal stimuli presented contralaterally, *circles* refer to observed mean SRT to bimodal stimuli presented ipsilaterally. *Lower panels* Corresponding probability of integration (*line*) and probability of warning (*dashed line*)

$$P(I) = \int_0^\infty \lambda_A e^{-\lambda_A a}\{e^{-\lambda_V(a+\tau)} - e^{-\lambda_V(a+\tau+\omega)}\}da = \frac{\lambda_A}{\lambda_V+\lambda_A}\{e^{-\lambda_V\tau} - e^{-\lambda_V(\omega+\tau)}\}.$$

Probability of warning: *P*(*W*)

By definition,

$$P(W) = \Pr(A+\tau+\gamma_A < V) = \int_0^\infty f_A(a)\{1 - F_V(a+\tau+\gamma_A)\}da = 1 - \int_0^\infty f_A(a)F_V(a+\tau+\gamma_A)da.$$

Again, we need to consider different cases:

(i) $\tau + \gamma_A < 0$

$$P(W) = 1 - \int_{-\tau-\gamma_A}^\infty \lambda_A e^{-\lambda_A a}\{1 - e^{-\lambda_V(a+\tau+\gamma_A)}\}da = 1 - \frac{\lambda_V}{\lambda_V+\lambda_A} e^{\lambda_A(\tau+\gamma_A)};$$

(ii) $\tau + \gamma_A \geq 0$

$$P(W) = 1 - \int_0^\infty \lambda_A e^{-\lambda_A a}\{1 - e^{-\lambda_V(a+\tau+\gamma_A)}\}da = \frac{\lambda_A}{\lambda_V+\lambda_A} e^{-\lambda_V(\tau+\gamma_A)}.$$

Appendix 2

Tables 5, 6, 7, 8, 9 and 10 for mean SRTs under all uni- and bimodal conditions are listed, by participant.

Table 5 Mean SRT to unimodal visual stimuli, averages across all experimental conditions

P1	P2	P3	P4	P5
154.8	147.6	174.2	135.8	144.3

Table 6 Participant P1: Mean SRT to bimodal stimuli with auditory non-target (left four columns) and tactile non-target (right four columns) presented ipsi- or contralaterally in a blocked or mixed trial condition

SOA	Auditory non-target				Tactile non-target			
	Ipsilateral		Contralateral		Ipsilateral		Contralateral	
	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed
−500	146.4	143.9	142.9	139.9	147.9	141.7	147.9	143.5
−400	140.1	140.7	139.1	138.4	149.3	144.6	147.6	142.4
−300	145.1	149.8	143.9	147.5	155.2	147.2	144.5	141.8
−200	141.8	140.3	139.8	137.5	146.1	145.7	140.4	135.9
−150	136.2	139.6	141.8	137.3	137.1	135.5	143.8	140.0
−100	126.0	125.8	132.8	133.9	141.0	139.2	142.1	139.9
−90	127.3	128.4	130.2	133.9	139.0	133.9	140.9	139.9
−80	132.0	131.0	134.5	132.3	141.0	145.0	137.4	145.9
−70	122.0	126.7	135.8	130.0	144.6	141.2	142.7	143.2
−50	128.7	135.3	138.4	138.9	140.4	142.4	147.7	150.8
−40	137.0	135.5	137.2	141.2	148.4	143.5	150.1	156.7
−20	138.8	141.6	140.2	149.3	147.7	145.8	153.4	146.7
0	144.1	142.9	150.9	147.6	145.3	143.7	153.2	148.5
50	158.2	158.1	159.1	159.7	156.3	154.0	154.9	154.3
100	160.0	155.2	157.5	159.5	150.0	153.1	157.7	151.1

Table 7 Participant P2: Mean SRT to bimodal stimuli with auditory non-target (left four columns) and tactile non-target (right four columns) presented ipsi- or contralaterally in a blocked or mixed trial condition

SOA	Auditory non-target				Tactile non-target			
	Ipsilateral		Contralateral		Ipsilateral		Contralateral	
	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed
−500	146.9	149.0	147.0	149.6	148.6	144.7	143.5	145.2
−400	143.8	149.0	143.4	146.4	145.1	136.4	144.0	141.9
−300	145.0	140.9	149.5	145.6	146.1	139.0	149.2	136.6
−200	147.2	140.7	144.6	144.2	135.8	138.9	138.9	140.6
−150	148.3	141.2	144.0	148.3	131.8	132.9	140.0	146.2
−100	140.5	139.5	141.7	142.0	133.1	130.9	142.7	145.5
−90	147.8	137.4	141.9	145.5	129.8	139.1	141.2	145.3
−80	135.2	135.4	143.5	141.5	133.1	137.7	144.1	144.9
−70	144.2	134.9	143.1	138.4	135.7	138.3	141.5	147.7
−50	134.2	136.6	138.3	153.9	134.0	135.6	142.1	148.3
−40	133.0	134.4	145.5	144.1	137.9	139.2	142.3	147.3
−20	131.3	137.9	143.9	143.3	142.9	141.4	144.4	155.5
0	139.8	141.2	144.3	147.6	141.1	142.4	145.4	150.6
50	146.3	152.4	149.1	162.2	148.1	150.7	145.7	147.9
100	151.3	155.1	151.0	149.8	148.1	146.4	144.8	150.3

Table 8 Participant P3: Mean SRT to bimodal stimuli with auditory non-target (left four columns) and tactile non-target (right four columns) presented ipsi- or contralaterally in a blocked or mixed trial condition

SOA	Auditory non-target				Tactile non-target			
	Ipsilateral		Contralateral		Ipsilateral		Contralateral	
	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed
−500	167.5	163.8	166.0	168.4	169.5	168.7	168.2	168.7
−400	170.5	166.5	165.2	171.2	167.1	175.7	161.3	164.3
−300	163.5	164.2	162.8	162.9	160.6	161.2	158.6	160.5
−200	156.9	156.2	163.3	160.5	158.0	159.8	163.6	168.4
−150	158.0	154.4	165.6	165.7	159.6	165.3	165.3	167.5
−100	162.1	160.6	165.0	163.4	161.4	164.4	166.0	166.6
−90	161.6	162.4	166.7	165.0	165.3	163.8	168.7	168.4
−80	161.0	157.6	167.2	168.4	164.0	162.6	168.0	166.3
−70	165.8	160.6	168.5	166.7	158.6	163.7	168.1	170.3
−50	156.5	163.5	167.3	168.5	164.3	162.6	166.0	168.4
−40	164.3	154.8	168.3	166.9	160.7	162.6	167.1	174.0
−20	164.5	161.7	169.4	168.4	172.2	167.7	172.8	173.5
0	166.6	166.4	166.1	169.7	165.9	170.6	172.0	171.4
50	176.8	178.0	178.5	182.7	173.5	174.1	174.5	173.0
100	174.9	178.3	179.6	183.5	165.0	172.0	175.8	181.4

Table 9 Participant P4: Mean SRT to bimodal stimuli with auditory non-target (left four columns) and tactile non-target (right four columns) presented ipsi- or contralaterally in a blocked or mixed trial condition

SOA	Auditory Non-target				Tactile Non-target			
	Ipsilateral		Contralateral		Ipsilateral		Contralateral	
	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed
−500	129.5	128.5	132.4	129.7	130.7	134.6	132.9	130.9
−400	128.5	126.8	129.4	131.2	129.8	126.9	132.2	129.4
−300	125.0	123.0	125.9	128.1	126.9	127.7	129.8	132.2
−200	120.9	127.6	124.6	128.8	128.5	125.3	126.6	129.3
−150	126.0	118.5	122.9	128.0	121.9	126.4	128.2	125.3
−100	121.4	118.8	124.4	120.6	124.4	125.4	131.4	124.0
−90	120.3	118.7	122.3	122.1	124.2	122.3	130.9	131.4
−80	121.8	122.6	128.7	125.4	126.0	130.4	134.6	132.7
−70	122.3	124.6	125.1	125.9	127.3	126.6	127.9	129.2
−50	120.8	124.8	126.0	127.2	123.4	127.0	133.7	139.5
−40	122.7	124.8	130.2	129.1	127.6	126.3	132.7	137.2
−20	124.1	127.3	128.1	129.7	131.5	133.4	128.2	134.2
0	124.3	125.1	132.5	135.1	128.7	134.4	129.4	135.9
50	138.3	142.4	137.8	143.3	133.9	136.8	138.1	135.2
100	143.2	138.0	137.4	142.9	130.8	136.9	137.2	136.7

Table 10 Participant P5: Mean SRT to bimodal stimuli with auditory non-target (left four columns) and tactile non-target (right four columns) presented ipsi- or contralaterally in a blocked or mixed trial condition

SOA	Auditory Non-target				Tactile Non-target			
	Ipsilateral		Contralateral		Ipsilateral		Contralateral	
	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed	Blocked	Mixed
−500	128.5	132.4	135.5	136.3	130.8	131.9	133.3	134.5
−400	136.6	140.0	133.8	140.9	132.8	136.1	136.5	132.5
−300	138.7	140.5	136.6	139.2	134.0	139.0	133.8	141.1
−200	135.5	137.7	136.6	141.4	135.8	143.3	133.6	135.5
−150	138.1	130.2	132.7	139.6	122.0	123.7	127.0	140.0
−100	120.0	118.1	131.9	131.9	114.2	121.0	137.4	131.0
−90	113.4	119.6	127.5	132.9	117.5	124.5	132.4	144.7
−80	116.5	119.4	135.4	135.8	124.8	117.6	135.5	142.1
−70	118.7	119.8	141.0	138.4	117.6	120.8	137.2	140.0
−50	120.4	125.4	138.3	139.7	126.5	125.8	139.2	140.2
−40	121.2	128.8	140.8	141.9	128.9	130.8	143.0	143.1
−20	122.3	137.1	138.5	146.5	132.6	138.1	145.9	145.1
0	127.4	137.1	147.4	148.9	137.6	144.8	146.1	145.9
50	153.1	149.5	157.4	165.0	142.3	146.1	139.6	152.9
100	148.1	159.5	148.9	157.3	141.8	147.8	144.3	160.7

References

Amlôt R, Walker R, Driver J, Spence C (2003). Multimodal visual-somatosensory integration in saccade generation. Neuropsychologia 41:1–15

Arndt A, Colonius H (2003) Two separate stages in crossmodal saccadic integration: evidence from varying intensity of an auditory accessory stimulus. Exp Brain Res 150:417–426

Bell AH, Corneil BD, Meredith MA, Munoz DP (2001) The influence of stimulus properties on multisensory processing in the awake primate superior colliculus. Can J Exp Psychol 55(2):123–132

Bell AH, Meredith A, Van Opstal AJ, Munoz DP (2005) Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. J Neurophysiol 93:3659–3673

Bushara KO, Grafman J, Hallett M (2001) Neural correlates of auditory-visual stimulus onset asynchrony detection. J Neurosci 21:300–304

Colonius H, Diederich A (2004) Multisensory interaction in saccadic reaction time: a time-window-of-integration model. J Cogn Neurosci 16:1000–1009

Corneil BD, Munoz DP (1996) The influence of auditory and visual distractors on human orienting gaze shifts. J Neurosci 16:8193–8207

Corneil BD, Van Wanrooij M, Munoz DP, Van Opstal AJ (2002). Auditory-visual interactions subserving goal-directed saccades in a complex scene. J Neurophysiol 88:438–454

Diederich A, Colonius H (2004) Modeling the time course of multisensory interaction in manual and saccadic responses. In: Calvert G, Spence C, Stein BE (eds) Handbook of multisensory processes. MIT, Cambridge

Diederich A, Colonius H (2007a) Why two "distractors" are better than one: modeling the effect of non-target auditory and tactile stimuli on visual saccadic reaction time. Exp Brain Res. doi: 10.1007/s00221-006-0768-0

Diederich A, Colonius H (2007b) Modeling spatial effects in visual-tactile saccadic reaction time. Percept Psychophys 69(1):56–67

Diederich A, Colonius H, Bockhorst D, Tabeling S (2003) Visual-tactile spatial interaction in saccade generation. Exp Brain Res 148:328–337

Doyle MC, Walker R (2002) Multisensory interactions in saccade target selection: curved saccade trajectories. Exp Brain Res 142:116–130

Eimer M (2001) Crossmodal links in spatial attention between vision, audition, and touch: evidence from event-related brain potentials. Neuropsychologia 39:1292–1303

Frens MA, Van Opstal AJ (1998) Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. Brain Res Bull 46(3):211–224

Frens MA, Van Opstal AJ, Van der Willigen RF (1995) Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. Percept Psychophys 57:802–816

Groh JM, Sparks DL (1996) Saccades to somotosensory targets. III. Eye-position-dependent somatosensory activity in primates superior colliculus. J Neurophysiol 75:439-453

Harrington LK, Peck CK (1998) Spatial disparity affects visual-auditory interactions in human sensorimotor processing. Exp Brain Res 122:247–252

Hublet C, Morais J, Bertelson P (1976) Spatial constraints on focussed attention: beyond the right side advantage. Perception 5:3–8

Hughes HC, Nelson MD, Aronchick DM (1998) Spatial characteristics of visual-auditory summation in human saccades. Vis Res 38:3955–3963

Jiang W, Wallace MT, Jiang H, Vaughan JW, Stein BE (2001) Two cortical areas mediate multisensory integration in superior colliculus neurons. J Neurophysiol 85:506–522

Jiang W, Jiang H, Stein BE (2002) Two corticotectal areas facilitate orientation behavior. J Cogn Neurosci 14:1240–1255

Kennett S, Eimer M, Spence C, Driver J (2001) Tactile-visual links in exogenous spatial attention under different postures: convergent evidence from psychophysics and ERPs. J Cogn Neurosci 13:462–478

Klein R, Kingstone A (1993) Why do visual offsets reduce saccadic latencies? Behav Brain Sci 16(3):583–584

Lagarias JC, Reeds JA, Wright MH, and Wright PE (1998) Convergence properties of the Nelder-Mead simplex method in low dimensions. SIAM J Optim 9(1):112–147

Lewald J, Guski R (2003) Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. Cogn Brain Res 16:468–478

McDonald JJ, Teder-Sälejärvi WA, Hillyard SA (2000). Involuntary orienting to sound improves visual perception. Nature 407:906–908

Meredith MA (2002) On the neural basis for multisensory convergence: a brief overview. Cogn Brain Res 14:31–40

Meredith MA, Stein BE (1986a) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. J Neurophysiol 56:640–662

Meredith MA, Stein BE (1986b) Spatial factors determine the activity of multisensory neurons in cat superior colliculus. Brain Res 365:350–354

Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. J Neurosci 10:3215–3229

Morein-Zamir S, Soto-Faraco S, Kingstone A (2003) Auditory capture of vision: examining temporal ventriloquism. Cogn Brain Res 17:154–163

Navarra J, Vatakis A, Zampini M, Soto-Faraco S, Humphreys W, Spence C (2005) Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. Cogn Brain Res 25:499–507

Nickerson RS (1973) Intersensory facilitation of reaction time: energy summation or preparation enhancement. Psychol Rev 80:489–509

Posner MI (1980) Orienting of attention. Quart J Exp Psychol 32:3–25

Rach S, Diederich A (2006) Visual-tactile integration: Does stimulus duration influence the relative amount of response enhancement? Exp Brain Res 173:514–520

Reuter-Lorenz PA, Hughes HC, Fendrich R (1991) The reduction of saccadic latency by prior offset of the fixation point: an analysis of the gap effect. Percept Psychophys 49(2):167–175

Ross SM, Ross LE (1981) Saccade latency and warning signals: effects of auditory and visual stimulus onset and offset. Percept Psychophys 29(5):429–437

Schaefer M, Heinze HJ, Rotte M (2005) Viewing touch improves tactile sensory threshold. Neuroreport 16(4):367–70

Spence C, Driver J (1997) Audiovisual links in exogenous covert spatial orienting. Perc Psychophys 59(1):1–22

Spence C, Squire S (2003) Multisensory integration: maintaining the perception of synchrony. Curr Biol 13:R519–R521

Stein BE, Meredith MA (1993) The merging of the senses. MIT, Cambridge

Van Atteveldt NM, Formisano E, Blomert L, Goebel R (2006) The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. Cereb Cortex. doi: 10.1093/cercor/bhl007

Van Opstal AJ, Munoz DP (2004) Auditory-visual interactions subserving primate gaze orienting. In: Calvert G, SpenceC , Stein BE (eds) Handbook of multisensory processes. MIT, Cambridge, pp 373–393

Van Wassenhove V, Grant KW, Poeppel D (2007) Temporal window of integration in auditory-visual speech perception. Neuropsychologia 45:598–607

Wallace MT, Wilkinson LK, Stein BE (1996) Representation and integration of multiple sensory inputs in primate superior colliculus. J Neurophysiol 76:1246–1266

Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA (2004) Unifying multisensory signals across time and space. Exp Brain Res 158:252–258

Ward LM (1994) Supramodal and modality-specific mechanisms for stimulus-driven shifts of auditory and visual attention. Can J Exp Psychol 48:242–259

Whitchurch EA, Takahashi TT (2006) Combined auditory and visual stimuli facilitate head saccades in the barn owl (Tyto alba). J Neurophysiol 96:730–745