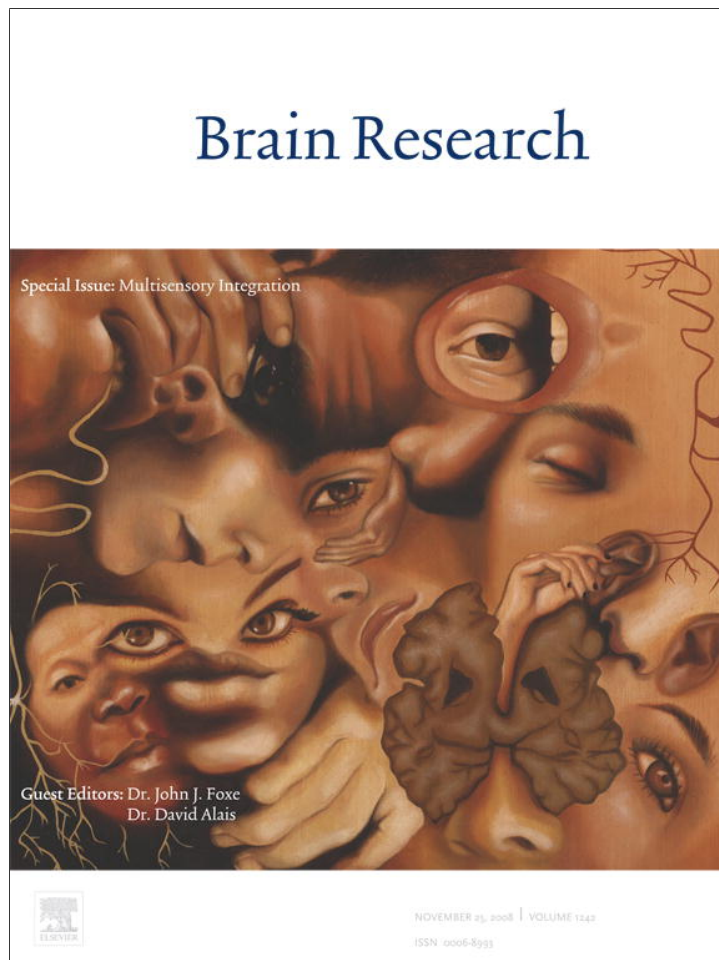


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



ELSEVIER

available at www.sciencedirect.comwww.elsevier.com/locate/brainresBRAIN
RESEARCH

Research Report

When a high-intensity “distractor” is better than a low-intensity one: Modeling the effect of an auditory or tactile nontarget stimulus on visual saccadic reaction time

Adele Diederich^a, Hans Colonius^{b,*}^aSchool of Humanities and Social Sciences, Jacobs University Bremen P.O. Box 750 561, D–28725 Bremen, Germany^bDepartment of Psychology, University of Oldenburg, P.O. Box 2503, D–26111 Oldenburg, Germany

ARTICLE INFO

Article history:

Accepted 29 May 2008

Available online 10 June 2008

Keywords:

Multisensory integration

Saccadic reaction time

Time window of integration

Effect of intensity

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) nontarget 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 –250 ms (nontarget prior to target) to 50 ms. This study specifically addressed the effect of varying nontarget intensity. While facilitation effects for auditory nontargets are somewhat more pronounced than for tactile ones, decreasing intensity slightly reduced facilitation for both types of nontargets. The time course of crossmodal mean SRT over SOA and the pattern of facilitation observed here suggest the existence of two distinct underlying mechanisms: (a) a spatially unspecific crossmodal warning triggered by the nontarget 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. It is shown that the time window of integration (TWIN) model introduced by the authors gives a reasonable quantitative account of the data relating observed SRT to the unobservable probability of integration and crossmodal warning for each SOA value under a high and low intensity level of the nontarget.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Based on neurophysiological and behavioral findings in humans, monkey, and cat, several authors have suggested the existence of a critical spatiotemporal “window” for multisensory integration to occur (e.g., Bell et al., 2005; Meredith, 2002; Corneil et al., 2002). Given that the strength of a multisensory effect is often lawfully related to the spatial and temporal configuration of the stimuli from different modalities (Stein and Meredith, 1993), the notion of a “window” has been proposed to be a critical determinant for multisensory integra-

tion at both neural and the behavioral levels of observation. With respect to the temporal dimension, the idea simply is that the visual and the auditory stimulus, say, must not be presented too far away in time for bimodal integration to occur. This integration may manifest itself in the form of an increased firing rate of a multisensory neuron (relative to unimodal stimulation), an acceleration of saccadic reaction time (Frens et al., 1995; Diederich et al., 2003), an effective audiovisual speech integration (Van Wassenhove et al., 2007), or in an improved or degraded judgment of temporal order of bimodal stimulus pairs (cf. Spence and Squire, 2003). Depending on context, estimates

* Corresponding author. Fax: +49 441 798 5158.

E-mail addresses: a.diederich@jacobs-university.de (A. Diederich), hans.colonius@uni-oldenburg.de (H. Colonius).

of the width of the time window differ widely, however, ranging from 40 to 600 ms. This is not really surprising given the huge differences in the above-mentioned experimental paradigms. In fact, it seems quite questionable that the broad usage of the notion of a time window for multisensory integration reflect a concept that goes beyond a metaphoric level of explanatory power. In this paper, the role of a time window is investigated in the limited context of an explicit model for saccadic reaction time to visual targets in the presence of either an auditory or tactile nontarget (“distractor”).

In this focused attention paradigm, crossmodal stimuli are presented and 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 nontarget stimulus, tactile or the auditory. When a stimulus of a nontarget 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 nontarget 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 nontarget, the effect can be a speed-up or an inhibition of SRT (see, e.g., Amlot et al., 2003; Diederich and Colonius, 2007a), and the saccadic trajectory can be affected as well (Doyle and Walker, 2002). Specifically, as long as the nontarget occurs less than about 200 ms before the target stimulus (stimulus onset asynchrony, SOA, greater or equal to 200 ms), a speed-up of bimodal vs. visual SRT is observed when target and nontarget are presented in spatial proximity, and the speed-up diminishes, or turns into inhibition, with increasing spatial disparity (Van Opstal and Munoz, 2004; Diederich and Colonius, 2007b; Hughes et al., 1998). The spatial effect disappears when the nontarget is presented even earlier, but a spatially unspecific facilitation is sometimes preserved even with an SOA of –500 ms (Diederich and Colonius, 2007a, 2008a). This latter effect may be considered a warning effect, or exogenous shift of attention, rather than a true multisensory integration phenomenon (e.g., McDonald et al., 2000).

1.1. Time-window-of-integration (TWIN) model

The distinction between warning and multisensory integration has been taken into account by a stochastic model developed by the authors (Colonius and Diederich, 2004; Diederich and Colonius, 2004a, 2007a,b, 2008a,b) yielding an estimate of the relative contribution of either mechanism for any specific SOA value. The TWIN model distinguishes 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. Thus, the model retains the classic notion of a race mechanism as an explanation for crossmodal interaction (cf. Colonius and Diederich, 2006), 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 temporally overlapping, processes that are not part of the peripheral processes in the first stage (for a similar approach, see Van Opstal and Munoz 2004). Moreover, Whitchurch and Takahashi (2006) collected (head) saccadic reaction times in the barn owl finding further support for 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 specific assumptions about the temporal configuration needed for multisensory integration to occur (see also Colonius and Diederich 2004; Diederich and Colonius 2007a,b, 2008a,b, for further details):

(1) Time-window-of-integration assumptions: In the focused attention paradigm, crossmodal interaction occurs only if ((a)) a nontarget 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 nontarget will keep the saccadic system in a heightened 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 two-stage structure of the TWIN model suggests an additional, important assumption concerning the interplay of spatial and temporal factors:

(2) Assumption of spatiotemporal separability: The amount of interaction ((facilitation or inhibition)) 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)).

These assumptions are part of a more general framework making a 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 nontarget 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 wavelength or frequency).

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 assumptions has been found in [Colonius and Diederich \(2004\)](#) for visual–tactile stimulation and in [Arndt and Colonius \(2003\)](#) for visual–auditory stimulation.

The third assumption refers to the role of the nontarget stimulus as a warning cue. It can be formulated in two different versions, a decision for one version or the other will have to be decided upon empirical evidence:

(3) Assumption of warning mechanism: (Version A) If the nontarget wins the processing race in the first stage by a margin wide enough for the time window of integration to close again before arrival of the target, then subsequent processing will be facilitated or inhibited without depending on the spatial configuration of the stimuli. (Version B) If the nontarget wins the processing race in the first stage by a wide enough margin, then subsequent processing will in part be facilitated or inhibited without dependence on the spatial configuration of the stimuli.

The time margin by which the nontarget may win against the target will be called *headstart*. Version A then stipulates that the headstart is at least as large as the width of the time window. This precludes simultaneous occurrence of warning and multisensory interaction within the same trial. Version B is less restrictive: all that is needed for the nontarget to act as a warning signal is a “large enough” headstart against the target in the race. The actual value of the headstart criterion is a parameter to be estimated in fitting the model under either version of the third assumption.

It follows that the occurrence of warning depends on intramodal characteristics of the target and the nontarget, such as modality or intensity. Assuming that increasing stimulus intensity goes along with decreased reaction time (for auditory stimuli see, e.g., [Frens et al., 1995](#); [Arndt and Colonius, 2003](#); for tactile, [Diederich and Colonius, 2004b](#)), TWIN makes specific predictions regarding the effect of nontarget intensity variation that will be investigated here. For instance, an intense auditory nontarget may have a higher chance to win the race with a headstart compared to a weak tactile nontarget. In general, increasing the intensity of the nontarget (1) increases the probability of it becoming a warning signal, and (2) makes it more likely for the nontarget to win the peripheral race against the processing of the target. Depending on SOA, however, it will be shown below that the latter either results in an increased probability of integration if the target arrives “on time” to fall into the time window, or it may decrease the probability of integration if the nontarget process finishes too early so that the window is already closed before the target arrives.

1.2. Quantitative specification of TWIN

For the race in the first stage of the model, we assign independent nonnegative random variables V and A to the peripheral processing times for the visual target and auditory nontarget stimulus, respectively. For ease of exposition, we

only consider the case of auditory nontargets. Whenever the nontarget 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 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 nontarget is triggered whenever the nontarget wins the race by a certain margin or headstart γ_A . Thus, its occurrence corresponds to the event:

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

The probability of this event, $P(W)$, 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). Under version A of assumption (3), the headstart γ_A is large enough for the integration window to close again, implying $\gamma_A > \omega \geq 0$ 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:

$$RT_{\text{crossmodal}} = S_1 + S_2, \tag{1}$$

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 (under version A of the warning assumption):

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

where $E[S_2|I]$, $E[S_2|W]$, and $E[S_2|I^c \cap W^c]$ denote the expected second stage processing time conditioned on interaction occurring (I), warning occurring (W), or neither of them occurring ($I^c \cap W^c$), respectively (I^c , W^c stand for the complement of events I, W). Setting

$$\begin{aligned} \Delta &\equiv E[S_2|I^c \cap W^c] - E[S_2|I] \\ \kappa &\equiv E[S_2|I^c \cap W^c] - E[S_2|W] \end{aligned}$$

this becomes

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

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

$$E[RT_{\text{unimodal}}] = E[S_1] + E[S_2|I^c \cap W^c],$$

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

$$CI \equiv E[RT_{\text{unimodal}}] - E[RT_{\text{crossmodal}}] = P(I) \cdot \Delta + P(W) \cdot \kappa. \tag{3}$$

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.

Under version B of the warning mechanism assumption, it is assumed for simplicity that the effects on RT of the two events *I* and *W*, integration and warning, combine additively. Interestingly, it can then be shown that the crossmodal interaction prediction of this model version is captured by the same equation as under version A, i.e., Eq. (3). The only difference is in the order restriction for the parameters, $\gamma_A > \omega$, under version A.

1.3. Model predictions

Even without making explicit assumptions about the probability distributions of the peripheral processing times, a number of qualitative predictions from the model can be made. Note, first, that for a given spatial configuration and nontarget 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 nontarget 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 nontarget is presented the smaller are its chances to win the race against the target with some headstart γ . Obviously, there also has to be an upper limit to the headstart for the warning mechanism to become active: if the nontarget occurs several seconds before the target it may no longer act as a warning cue at all. Therefore, in order to fit a parametric version of the model to empirical data, a certain numerical range for each possible parameter value will have to be specified (see Table 1).

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 a 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. The exponential distribution is characterized by a single quantity, the intensity parameter λ (for a full description of the computation of $P(I)$ and $P(W)$ we refer to Diederich and Colonius, 2008a).

To illustrate the mean SRT predictions of TWIN as a function of SOA, we choose a set of arbitrary (but plausible) parameters. Let the parameter for the visual target processing time be $\lambda_V = 0.025$ [1/ms] which amounts to a mean peripheral time for visual unimodal stimuli of 40 ms ($1/\lambda_V$). The parameter for second stage central processing time when neither integration nor warning occurs, $\mu = E[S_2|F \cap Wc]$, 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:

$$E[RT_{\text{unimodal}}] = 1/\lambda_V + \mu.$$

For a prediction of crossmodal SRTs we have to specify the remaining parameters. Let the parameter of nontarget processing time be $\lambda_A = 0.015$ [1/ms], corresponding to a mean peripheral time of 66.7 ms. The time window of integration is set to $\omega = 150$ ms. The parameter for multisensory integration Δ is set to 20 ms. Finally, the warning cue effect also reducing mean SRT to bimodal stimuli, κ , is set to 5 ms. Fig. 1 plots predicted mean SRT as a function of SOA. The horizontal line refers to the unimodal condition, the red curve to the crossmodal condition without warning, and the six blue dashed lines are the predictions for different γ values of the headstart, from $\gamma = 50$ ms (bottom curve) to $\gamma = 300$ ms (top curve) in steps of 50. Without warning, mean SRT approaches the unimodal mean for large enough negative SOA values. For these SOA values, the effect of the headstart is most pronounced, as to be expected. Here, the difference between the two versions of the warning mechanism becomes apparent: for γ values smaller than $\omega = 150$ ms

Table 1 – Restrictions to model parameters in the estimation routine

Parameters	Restriction limits	
$1/\lambda_V$	20–200	Mean peripheral processing time for visual stimulus
$1/\lambda_{T_H}$	20–200	Mean peripheral processing time for tactile stimulus, higher intensity
$1/\lambda_{A_H}$	20–200	Mean peripheral processing time for auditory stimulus, higher intensity
$1/\lambda_{T_L}$	20–200	Mean peripheral processing time for tactile stimulus, lower intensity
$1/\lambda_{A_L}$	20–200	Mean peripheral processing time for auditory stimulus, lower intensity
μ	>0	Mean central processing time
ω_T	<300	Window of integration for tactile stimulus
ω_A	<300	Window of integration for auditory stimulus
$\Delta_{T_{\text{ipsi}}}$	None	Amount of crossmodal interaction due to tactile stimulus presented ipsilaterally
$\Delta_{T_{\text{contra}}}$	None	Amount of crossmodal interaction due to tactile stimulus presented contralaterally
$\Delta_{A_{\text{ipsi}}}$	None	Amount of crossmodal interaction due to auditory stimulus presented ipsilaterally
$\Delta_{A_{\text{contra}}}$	None	Amount of crossmodal interaction due to auditory stimulus presented contralaterally
γ_{T_H}	Depending on model	Headstart for tactile stimulus, higher intensity
γ_{A_H}	Depending on model	Headstart for auditory stimulus, higher intensity
γ_{T_L}	Depending on model	Headstart for tactile stimulus, lower intensity
γ_{A_L}	Depending on model	Headstart for auditory stimulus, lower intensity
κ	≥ 0	Amount of warning cue effect

(three bottom curves), the warning adds to the multisensory integration effect around SOA = -100 ms because the effects are additive, whereas for larger headstart values warning effects in that SOA range are no longer possible.

A more comprehensive illustration of the effects of warning is given in Fig. 2 depicting mean crossmodal SRT as a function of the headstart parameter, γ , varying from 10 to 300 ms in steps of 10 ms, and of the SOA varying from -300 to 100.

Fig. 3 shows, as a function of SOA, the probability of integration (red line) and the corresponding probability of warning for the six different values of the headstart parameter γ (blue dashed curves, most right hand $\gamma=50$, most left $\gamma=300$). These probabilities are not observable but their time course illustrates the interplay between warning and integration: (i) the probability of warning decreases towards zero as SOA increases, with the headstart parameter controlling the location of decline: the larger γ the smaller the probability of warning at a given SOA; (ii) the probability of integration is a peaked function of SOA:

If target and nontarget 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 $P(W) \kappa$, the expected non-spatial warning effect, should remain the same under both conditions. Subtracting the corresponding cross-modal interaction terms then gives, after canceling the warning effect terms (Eq. (3)),

$$CI_i - CI_c = P(I) \cdot (\Delta_i - \Delta_c), \quad (4)$$

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

Finally, increasing the intensity level of the nontarget peripheral process (λ_A) affects both the probability of warning and of integration as a function of SOA: (i) the higher λ_A the

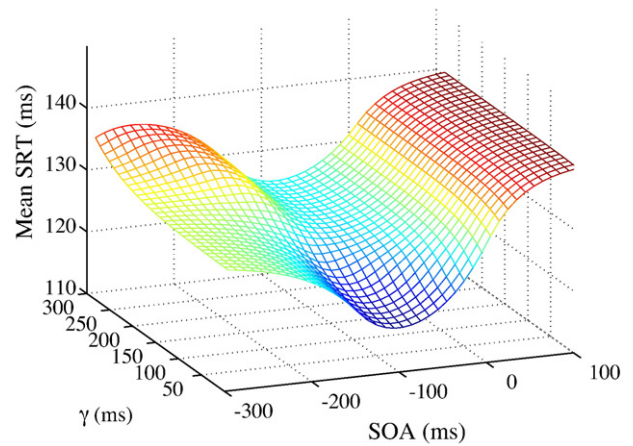


Fig. 2 – TWIN model prediction for mean SRT. Mean SRT as a function of SOA and headstart γ , with other parameters as in Fig. 1.

higher the chances of the nontarget to win against the target with the headstart required for the warning to take place (dashed curves in Fig. 4); (ii) the effect on integration probability is more complex: for large negative SOA, i.e., when the (auditory) nontarget arrives very early – so that warning occurs with high probability – further increasing the auditory intensity makes it more likely for the time window of integration to close before the target arrives and therefore results in lower $P(I)$ values; for SOA values in the range of about -120 to -40 (given the selected parameter values in Fig. 4) turns the effect of around: $P(W)$ is already low and increasing auditory intensity now improves its chances to win against the target and to open the window, as illustrated by the five crossing red lines in Fig. 4).

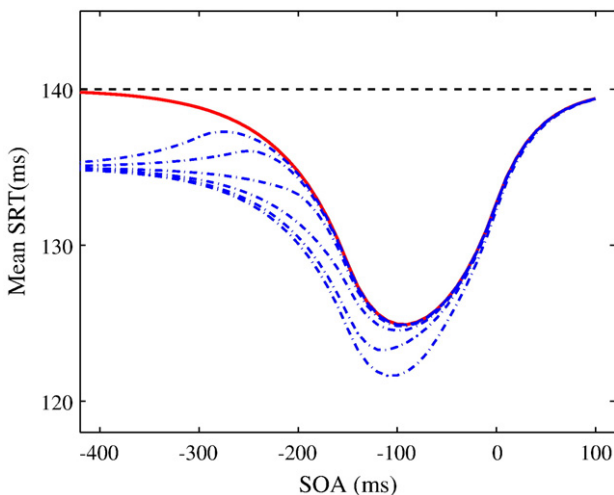


Fig. 1 – TWIN model prediction with and without warning. Mean unimodal SRT (dashed line) and mean crossmodal SRT without warning (red line) as a function of SOA. The blue dashed curves are mean crossmodal SRT for 6 values of the headstart parameter γ , from 50 ms (bottom curve) to 300 ms (top curve) in steps of 50; with $\lambda_V = .025$, $\lambda_A = .015$, $\omega = 150$.

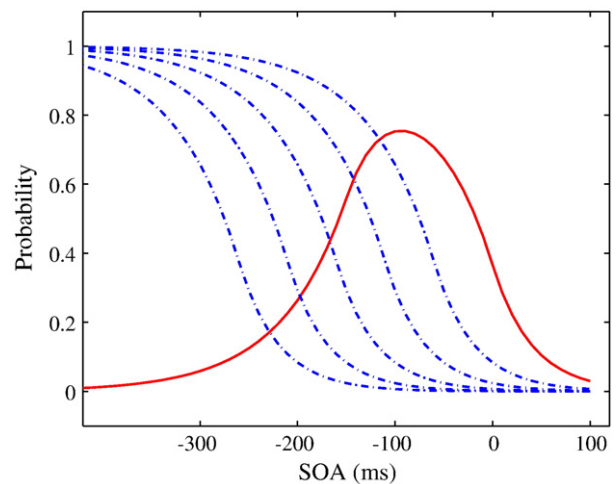


Fig. 3 – TWIN model prediction. Probability of integration (red line) and probability of warning (dashed curves) with $\gamma = 50, 100, 150, 200$ or 250 ms, from right to left. Here, $\lambda_V = .025$, $\lambda_A = .015$, $\omega = 150$.

2. Results

Saccades were screened for anticipation errors (SRT < 80 ms, 0.6% of all trials), misses (SRT > 500 ms, none), and accuracy: trials with saccade amplitudes deviating more than three standard deviations from the mean amplitude were excluded from the analysis (0.9% of all trials). Given that apart from individual speed levels all six participants exhibited very similar results across stimulus conditions, all data presented below are averages over participants.

2.1. Qualitative test

We first consider the test implied by Eq. (4),

$$CI_i - CI_c = P(I) \cdot (\Delta_i - \Delta_c). \quad (5)$$

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. Fig. 5 displays the results. Apart from variability due to sampling error, the shapes of both curves (circles for tactile, triangles for auditory nontargets, lines for stronger intensities, dashed lines for weaker intensities) are partially in agreement with the predicted shape of $P(I)$. Specifically, the tactile curve (TH) peaks before the auditory curve (AH) and the two auditory curves cross as predicted by the model.

2.2. Quantitative tests

Table 1 (left column) lists all parameters of TWIN that have to be estimated. The middle column of Table 1 lists the parameter restrictions that were imposed on the estimation routine. The λ values were restricted to fall within a range

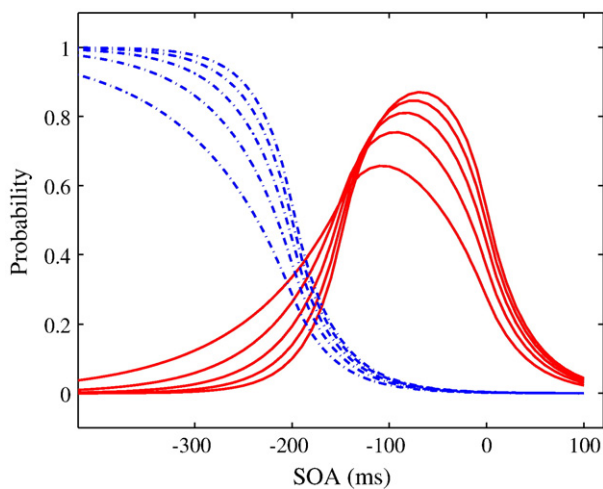


Fig. 4 – TWIN model prediction. Probability of integration (red lines) and probability of warning (dashed curves) for different values of auditory intensity $\lambda_A = .01, .015, .02, .025, \text{ or } .03$. Here, $\gamma = 200$, other parameter values are the same as in Fig. 3.

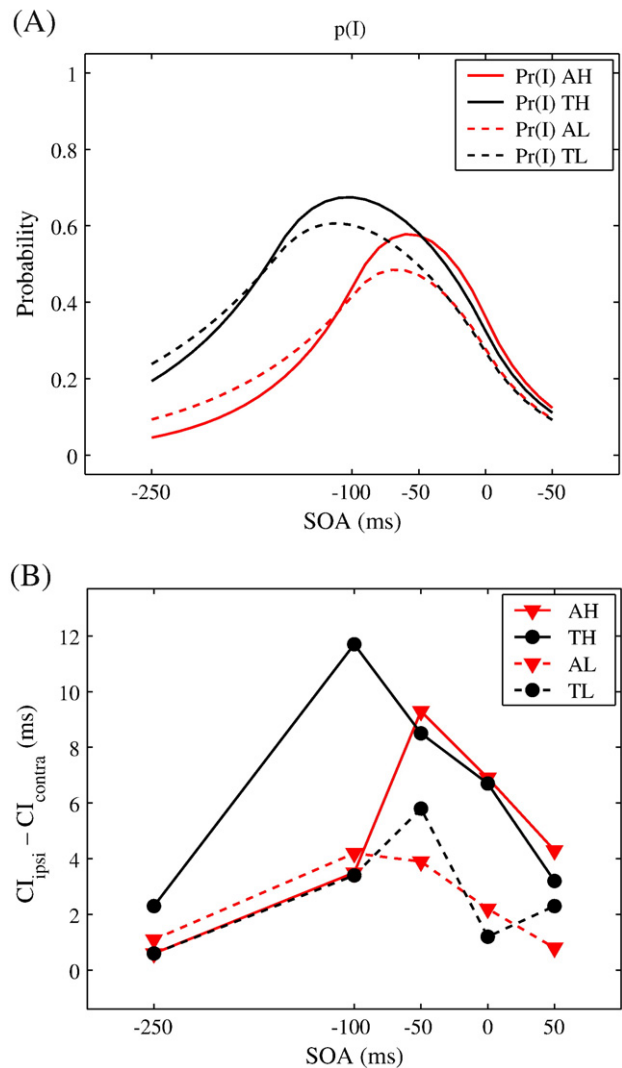


Fig. 5 – Qualitative test of integration and warning mechanism. Panel (A) Probability of integration for auditory and tactile nontargets with high and low intensity level under warning assumption A; Panel (B) Crossmodal integration difference ipsi- vs contralateral as function of SOA that should share qualitative features with curves in (A).

consistent with neurophysiological estimates for peripheral processing times (Stein and Meredith 1993; Groh and Sparks, 1996). The width of the time window of integration for a tactile and an auditory nontarget, ω_T and ω_A , respectively, was limited to an upper bound of 300 ms. These 17 parameters were estimated from 45 means from 25,366 valid observations (40 bimodal and five unimodal mean SRT).

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

$$\chi^2 = \sum_{n=1}^N \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^2 \left(\frac{\overline{\text{SRT}}(i, j, k, n) - \widehat{\text{SRT}}(i, j, k, n)}{\sigma_{\overline{\text{SRT}}(i, j, k, n)}} \right)^2 + \sum_{i=1}^5 \left(\frac{\overline{\text{SRT}}_{\text{uni}}(i) - \widehat{\text{SRT}}_{\text{uni}}(i)}{\sigma_{\overline{\text{SRT}}_{\text{uni}}(i)}} \right)^2 \quad (6)$$

Table 2 – Estimated parameters [in ms] for both versions of the warning mechanism

Estimated parameters	No restriction	$\gamma \geq \omega$
$1/\lambda_V$	58	47
$1/\lambda_{T_H}$	103	92
$1/\lambda_{A_H}$	79	67
$1/\lambda_{T_L}$	131	120
$1/\lambda_{A_L}$	110	101
μ	103	114
ω_T	111	156
ω_A	101	98
$\Delta_{T_{\text{ipsi}}}$	27	28
$\Delta_{T_{\text{contra}}}$	11	16
$\Delta_{A_{\text{ipsi}}}$	17	28
$\Delta_{A_{\text{contra}}}$	5	16
γ_{T_H}	90	175
γ_{A_H}	34	98
γ_{T_L}	102	190
γ_{A_L}	59	104
κ	21	22
χ^2	122	150

using the FMINSEARCH routine of MATLAB. Here $\overline{\text{SRT}}(i, j, k, n)$ and $\text{SRT}(i, j, k, n)$ are, respectively, the observed and the fitted values of the mean SRT to bimodal stimuli (visual–auditory, $i=1$; visual–tactile, $i=2$) with intensities (low, $j=i$, high, $j=2$) presented in spatial positions (ipsilateral, $k=1$; contralateral, $k=2$) with SOA (referred to by n to $N=5$); $\sigma_{\overline{\text{SRT}}(i,j,k,n)}$ and $\sigma_{\overline{\text{SRT}}_{\text{uni}}(i)}$ are the respective standard errors. The second term in Eq. (6) refers to the five unimodal condition. All parameter estimates are collected in Table 2 (left column: no restriction on the ordering of γ and ω , as in Version B of the warning assumption; right column: under the restriction of $\gamma \geq \omega$, as in version A). The minimized χ^2 values ($df=28$) are shown in the last row of Table 2.

Figs. 6 and 7 plot the observed mean SRT values (\pm standard error) and model predictions across SOAs for all nontarget conditions for version A and B, respectively. Upper panels show auditory nontargets with high (left panel) and low (right panel) intensities. Lower panels show the respective tactile nontarget conditions. Circles (squares) 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 estimated mean unimodal (visual) SRT.

Figs. 8 and 9 plot the probability of integration (lines) and the probability of warning (dashed lines) computed from the model parameter estimates for version A and B, respectively. The differences between the dashed and the solid lines express the effect of changing the intensity of the nontarget: the warning probability decreases with decreasing intensity, whereas integration probability first increases and then, for larger SOA, decreases with intensity, as predicted (see Figs. 3 and 4).

3. Discussion

Inspecting the five intensity parameters (λ values) in Table 2, the estimates are found not to be in line with the expectation

that the visual peripheral latencies are longer than the auditory or somatosensory ones. However, peripheral latencies are also known to depend on the relative intensity levels; here, the stimuli were not matched so as to achieve a crossmodal subjective equivalence in intensity. Reassuringly, the order of high vs. low intensity within the auditory and the somatosensory modality, respectively, was reflected correctly in the parameter values. Second, by definition “peripheral latency” in TWIN is the processing time before any crosstalk between the senses occurs; it is not entirely clear (to us) whether the estimates for peripheral latencies from the neurophysiology literature are in one-to-one correspondence to this definition. Finally, some of the discrepancy may simply result from problems in estimating these parameter values. In future work, the stability and reproducibility of the parameter estimates should be investigated systematically.

Visual inspection of Figs. 6 and 7 indicates that both TWIN model versions reflect the main qualitative properties of the data relatively well: (i) for large negative SOA (-250 ms), there is facilitation but no difference between ipsi- and contralateral presentation of target and nontarget, (ii) the difference between the ipsi- and contralateral condition increases until a maximum around SOA $= -100$ to -50 is achieved, and (iii) facilitation decreases toward zero with increasing SOA further. This study specifically addressed the effect of varying nontarget intensity. The plots indicate that decreasing intensity somewhat flattens the increase of the SOA curve for both the auditory and the tactile nontargets such that facilitation becomes lesser. This is in line with the decrease of integration probability in this SOA range predicted by the model (see Fig. 4). Comparing the nontarget modality conditions suggests that facilitation effects for auditory nontargets are slightly more pronounced than for tactile ones. Again, this could be predicted from the corresponding probability of integration/warning plots (Figs. 8, 9): the probability of warning levels off much earlier for the tactile than for the auditory nontargets (under both model versions).

The prediction alluded to in the title of this paper may be interpreted as being in conflict with the well known principle of “inverse effectiveness” according to which a decrease of stimulus intensity should go along with an enhancement of multisensory interaction, i.e. in this context, a reduction of SRT (Stein and Meredith 1993). As explicated above (cf. Fig. 4), the effect of decreasing nontarget intensity on integration probability critically depends on the SOA range. In particular, inverse effectiveness would only be expected for rather large negative SOAs, i.e., when the nontarget leads the target by a sufficient amount of time. For SOA values of -120 to -40 , where the largest facilitation effects occur, however, decreasing intensity would have the opposite effect since it would diminish the nontarget’s chance of winning the race and opening the time window. In this context, it should be noted that the validity of the principle of inverse effectiveness, as well as its precise definition, have become a matter debate recently (Holmes 2007; Ross et al. 2007). These predictions of TWIN certainly need further empirical testing, with a higher number of intensity levels, before a definite resolution of this issue is possible.

The crossmodal spatial interaction effects in saccades reported here are in line with findings from other labs using

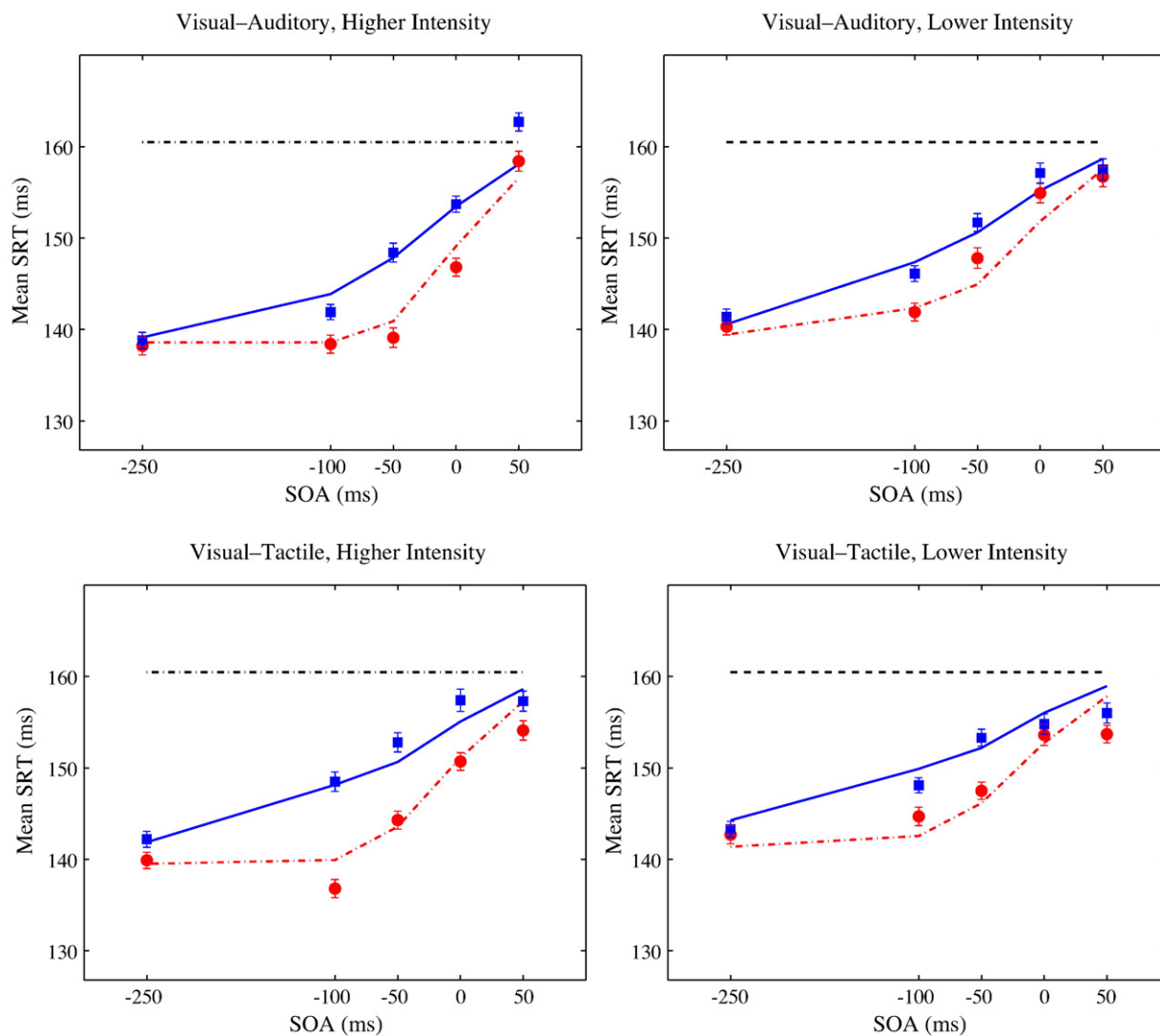


Fig. 6 – Model fit for TWIN with warning version A. Observed and predicted mean SRT for model with warning mechanism version A as a function of SOA for auditory (top panels) and tactile (bottom panels) nontargets with high (left panels) and low (right panels) intensity. Horizontal dashed line: unimodal visual observed data; blue lines and points: contralateral condition, red lines and points: ipsilateral condition.

visual targets and either auditory or tactile nontargets (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, 2008a). The time course of cross-modal mean SRT over SOAs and the pattern of facilitation observed here suggest the existence of two distinct underlying mechanisms: (a) a spatially unspecific crossmodal warning triggered by the nontarget 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.

On the other hand, due to the large number of observations (about 100 trials/condition), a χ^2 goodness of fit test has large power and clearly rejects both versions of the model ($p < .001$).

There are several, not mutually exclusive explanations for this failure: First, the model, although not really simple, may not capture all relevant factors affecting the timing of the saccadic response. Second, there may be idiosyncratic effects of individual participants that have not been averaged out given their small number. Third, since the number of parameters is high, the estimation procedure may not yet have delivered the best estimates possible.

Consequently, it appears sensible to consider the χ^2 value more as a descriptive measure of model appropriateness than as a criterion to reject the model outright. In this respect version B, i.e., no restriction of γ relative to ω , accounts slightly better for the data as reflected in the smaller χ^2 value in Table 2. This suggests that the restriction that in any given trial only one mechanism can be operative – warning or integration – may be too strong. In other words, the

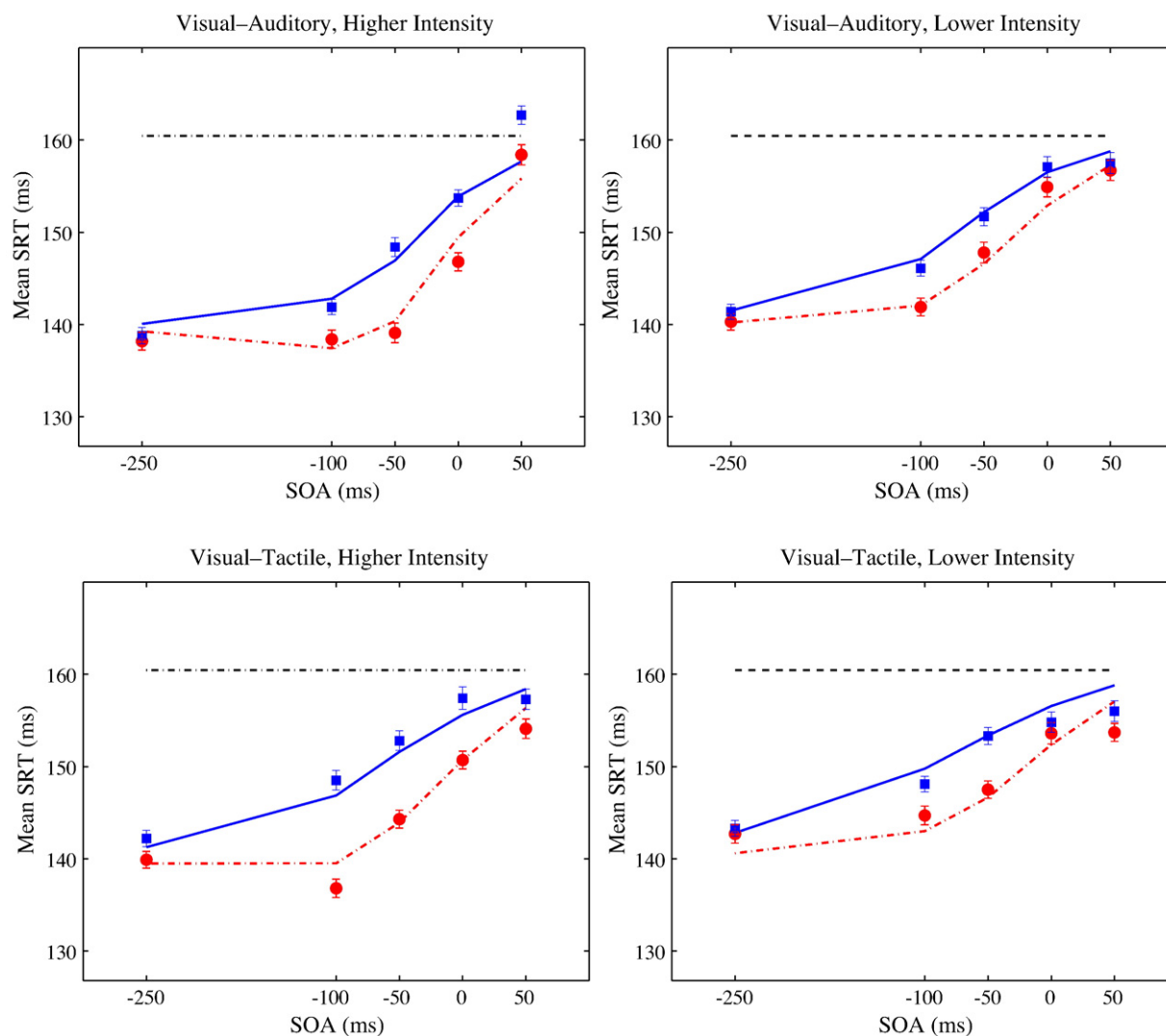


Fig. 7 – Model fit for TWIN with warning version B. Same as Fig. 6 for model under warning mechanism version B.

nontarget may be effective both as a warning signal and as an element of the bimodal stimulus configuration inducing multisensory integration, and its influence in either role depends on SOA in a manner depicted in Fig. 3. Thus, this finding goes beyond our previous study (Diederich and Colonius, 2008a) where no advantage of either version was discernable.

The time window of integration model introduced in Colonius and Diederich (2004) and extended by a warning mechanism in Diederich and Colonius (2008a) has by now been tested in a variety of ways: it accounts for the spatial configuration of the stimuli (Diederich and Colonius 2007b), for the effect of increasing the number of nontargets presented together with the target (Diederich and Colonius 2007a), for the warning effect occurring with large negative SOAs (Diederich and Colonius, 2008a) and, as shown here, for the effects of increasing the intensity of the nontarget. The qualitative test presented here, based on Equation (4) and depicted in Figure 5, showing that the peakedness of the crossmodal integration curves over SOA can be recovered constitutes further strong

support for the central assumption of spatiotemporal separability in TWIN, independent of the specific distributional assumptions for the peripheral latencies.

Finally, it should be noted that the idea of multisensory integration being determined by a time window had already been suggested in Meredith, Nemitz, and Stein (1987) recording from SC neurons. It, together with the concept of a race among peripheral processes, now underlies many studies of crossmodal temporal interaction (e.g., Van Opstal and Munoz, 2004; Lewald and Guski, 2003; Morein-Zamir, Soto-Faraco, Kingston, 2003; Spence and Squire, 2003; see Whitchurch and Takahashi, 2006, for head saccades in the barn owl).

4. Experimental procedures

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 (nontarget) stimulus.

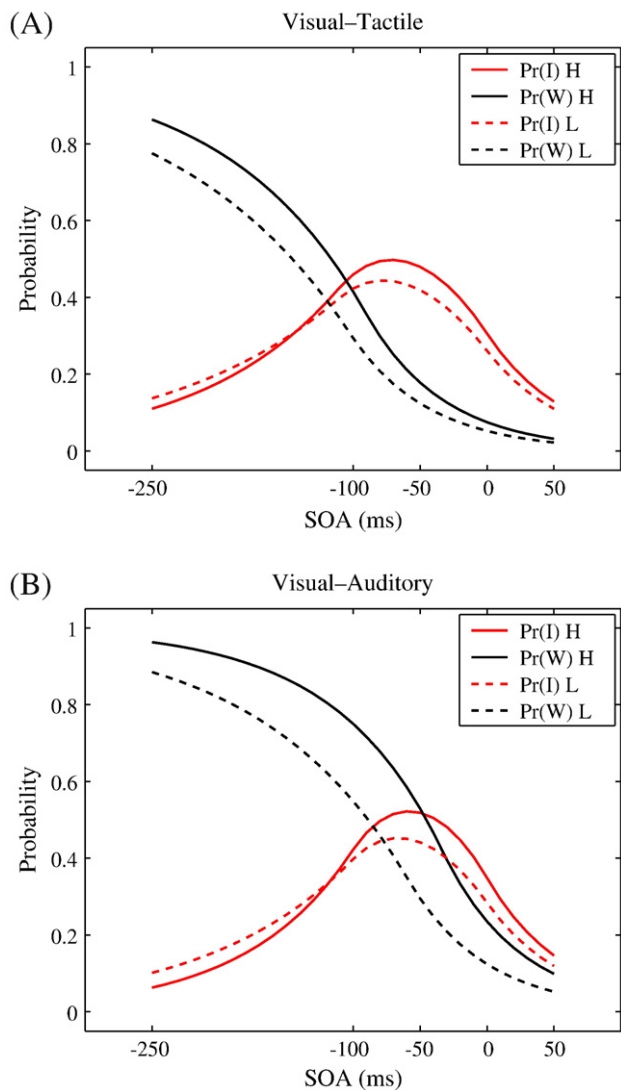


Fig. 8 – Probability of warning and integration. Probability of integration and of warning for high and low intensity of nontarget (top panel: tactile, bottom panel: auditory) with parameters from Table 2 for model under warning mechanism version A.

4.1. Participants

Six undergraduate students, aged 20 to 23, one female, served as paid voluntary participants. All had normal or corrected to normal vision. 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.

4.2. Apparatus and stimulus presentation

Two red light-emitting diodes (LED, 25 mA, 3.3 mcd) presented against a black background served as visual targets (V). They were placed on a table 600 mm in front of the participant, 20° to the left or right of a central fixation point (fixation LED, red,

25 mA, 5.95 mcd). The tactile stimulation (50 Hz, 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 20 mm 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 620 mm in front of the participant. Two intensities were employed: a lower intensity with 0.25 V and denoted (t) and a higher intensity with 0.75 V, denoted (T). Auditory stimuli were bursts of white noise with two levels of intensity, a lower intensity with 25.5 dBA SPL, (a), and a higher intensity with 45 dBA SPL). They were generated by two speakers (Canton Plus XS). The speakers were placed at 20° to the left and right of the fixation

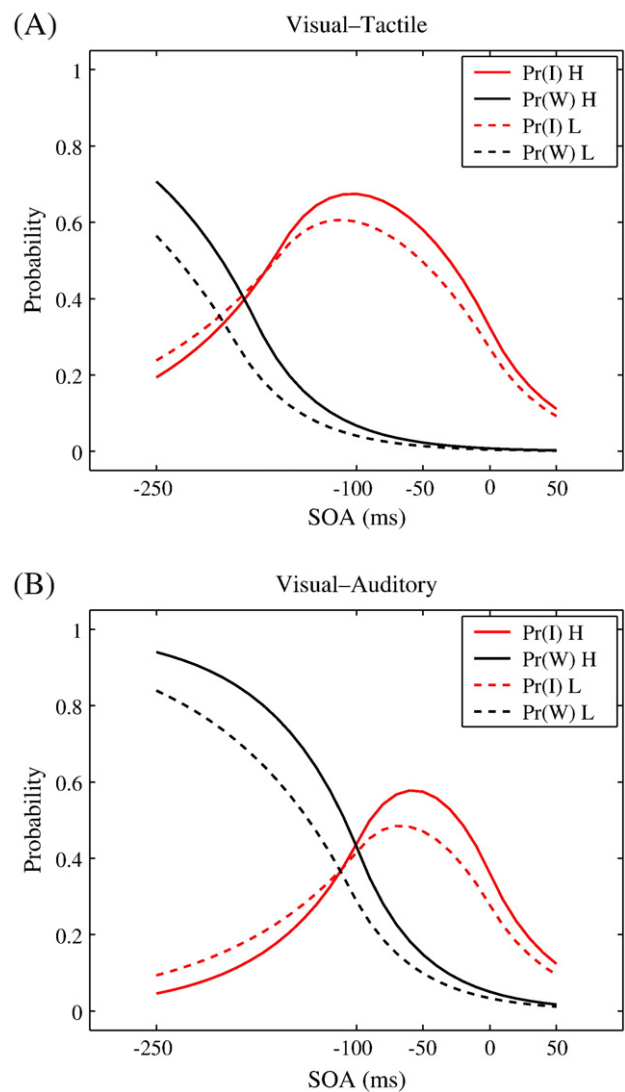


Fig. 9 – Probability of warning and integration. Probability of integration and of warning for high and low intensity of nontarget (top panel: tactile, bottom panel: auditory) with parameters from Table 2 for model under warning mechanism version B.

LED at the same height 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.

4.3. 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°/s) and acceleration (9500°/s²). 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.

4.4. Spatial and temporal stimulus configuration

The visual target was presented (*unimodally*) or in the presence of an auditory or a tactile nontarget. For bimodal presentations, two spatial and five temporal stimulus configurations were used. **Spatial configuration:** (i) target (V) and high- or low intensity nontarget (A_H , A_L , T_H , T_L) were presented in the same hemifield (20° left or right from fixation) (*ipsilateral*), or (ii) target and nontarget were presented in opposite hemispheres (20° left and right, or vice versa) (*contralateral*). **Temporal configuration:** five different stimulus onset asynchrony levels (SOAs) were used between targets and nontargets. The nontarget was presented 250 ms, 100 ms, or 50 ms before the target appeared (SOAs = -250, -100, -50), or target and nontarget were presented simultaneously (SOA = 0), or the nontarget was presented 50 ms after the target appeared (SOA = 50).

4.5. 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. Every experimental session began with 10 min dark adaptation during which the measurement system was adjusted and calibrated. During this phase, participants put the hands at the position used during the entire experimental block.

Each bimodal trial began with the appearance of the fixation point. After a variable fixation time (800–1500 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 nontargets (focused attention paradigm). The visual target appeared in combination with either a tactile or an auditory nontarget in ipsi- or contralateral position. The visual stimuli were presented for 500 ms; the auditory and tactile nontargets were turned off simultaneously with the visual stimulus. Thus their duration varied between 750 and 450 ms, depending on SOA. Stimulus presentation was followed by a break of 2000 ms in complete darkness, before the next trial began, indicated by the onset of the fixation LED. 240 bimodal trials were presented within one block of trials randomized across all experimental

conditions (two nontarget modalities, two intensities, five SOA, two lateralities). Two blocks of trials were conducted within 1 h. After extensive training (60–120 min), each participant completed a total of 3840 bimodal trials (96 trials per condition). Unimodal blocks consisted of 250 trials and were conducted within 30 min. One block was presented before and one after the entire bimodal presentation, resulting in a total of 500 unimodal trials (100 trials per condition).

Acknowledgments

This research was supported by grants from Deutsche Forschungsgemeinschaft Di 506/8–1 and /8–3. We are grateful to Dr. Annette Schomburg, Rike Steenken, and Stefan Rach for helpful discussions of this work.

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, A.H., Meredith, A., Van Opstal, A.J., Munoz, D.P., 2005. Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. *J. Neurophysiol.* 93, 3659–3673.
- Colonius, H., Diederich, A., 2004. Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *J. Cogn. Neurosci.* 16, 1000–1009.
- Colonius, H., Diederich, A., 2006. Race model inequality: Interpreting a geometric measure of the amount of violation. *Psychol. Review* 113 (1), 148–154.
- Cornell, B.D., Van Wanrooij, M., Munoz, D.P., Van Opstal, A.J., 2002. Auditory-visual interactions subserving goal-directed saccades in a complex scene. *J. Neurophysiol.* 88, 438–454.
- Diederich, A., Colonius, H., 2004a. Modeling the time course of multisensory interaction in manual and saccadic responses. In: Calvert, G., Spence, C., Stein, B.E. (Eds.), *Handbook of Multisensory Processes*. MIT Press, Cambridge, MA.
- Diederich, A., Colonius, H., 2004b. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Percept Psychophys.* 66 (8), 1388–1404.
- Diederich, A., Colonius, H., 2007a. Why two “distractors” are better than one: modeling the effect of nontarget auditory and tactile stimuli on visual saccadic reaction time. *Exp. Brain Res.* 179, 43–54.
- Diederich, A., Colonius, H., 2007b. Modeling spatial effects in visual-tactile saccadic reaction time. *Percept. Psychophys.* 69 (1), 56–67.
- Diederich, A., Colonius, H., 2008a. Crossmodal interaction in saccadic reaction time: separating multisensory from warning effects in the time window of integration model. *Exp. Brain Res.* 186, 1–22.
- Diederich, A., Colonius, H., Schomburg, A., 2008b. Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia*.
- Diederich, A., Colonius, H., Bockhorst, D., Tabeling, S., 2003. Visual-tactile spatial interaction in saccade generation. *Exp. Brain Res* 148, 328–337.

- Doyle, M.C., Walker, R., 2002. Multisensory interactions in saccade target selection: Curved saccade trajectories. *Exp. Brain Res.* 142, 116–130.
- 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. *Percept. Psychophys.* 57, 802–816.
- Groh, J.M., Sparks, D.L., 1996. Saccades to somatosensory targets. III. Eye-position-dependent somatosensory activity in primates superior colliculus. *J. Neurophysiol.* 75, 439–453.
- Harrington, L.K., Peck, C.K., 1998. Spatial disparity affects visual–auditory interactions in human sensorimotor processing. *Exp. Brain Res.* 122, 247–252.
- Holmes, P.N., 2007. The law of inverse effectiveness in neurons and behavior: Multisensory integration versus normal variability. *Nicholas P. Holmes. Neuropsychologia* 45, 3340–3345.
- Hughes, H.C., Nelson, M.D., Aronchick, D.M., 1998. Spatial characteristics of visual–auditory summation in human saccades. *Vis. Res.* 38, 3955–3963.
- 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, J.J., Teder-Sälejärvi, W.A., Hillyard, S.A., 2000. Involuntary orienting to sound improves visual perception. *Nature* 407, 906–908.
- Meredith, M.A., Nemitz, J.W., Stein, B.E., 1987. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J. Neurosci* 10, 3215–3229.
- Meredith, M.A., 2002. On the neural basis for multisensory convergence: a brief overview. *Cogn. Brain Res.* 14, 31–40.
- Morein-Zamir, S., Soto-Faraco, S., Kingstone, A., 2003. Auditory capture of vision: examining temporal ventriloquism. *Cogn. Brain Res.* 17, 154–163.
- Ross, L.A., Saint-Amour, D., Leavitt, V., Javitt, D.C., Foxe, J.J., 2007. Do you see what I'm saying? Optimal visual enhancement of speech recognition in noisy environments. *Cereb. Cortex.* 17, 1147–1153.
- Spence, C., Squire, S., 2003. Multisensory integration: maintaining the perception of synchrony. *Curr. Biol.* 13, R519–R521.
- Stein, B.E., Meredith, M.A., 1993. *The merging of the senses.* The MIT Press, Cambridge MA.
- Van Opstal, A.J., Munoz, D.P., 2004. Auditory–visual interactions subserving primate gaze orienting. In: Calvert, G., Spence, C., Stein, B.E. (Eds.), *Handbook of Multisensory Processes.* MIT Press, Cambridge, MA, pp. 373–393.
- Van Wassenhove, V., Grant, K.W., Poeppel, D., 2007. Temporal window of integration in auditory–visual speech perception. *Neuropsychologia* 45, 598–607.
- Whitchurch, E.A., Takahashi, T.T., 2006. Combined auditory and visual stimuli facilitate head saccades in the barn owl (*Tyto alba*). *J. Neurophysiol.* 96, 730–745.