

Multisensory Interaction in Saccadic Reaction Time:
A Time-Window-of-Integration Model

Hans Colonius*

Adele Diederich

Department of Psychology

School of Humanities and Social Sciences

Universitaet Oldenburg

International University Bremen

September 9, 2003

Ms. #1861, revision

*Corresponding address: Department of Psychology, Universitaet Oldenburg, P.O Box 2503, D-26111 Oldenburg, Germany, E-mail: hans.colonius@uni-oldenburg.de, Phone/Fax: +49-441-7985158

Abstract

Saccadic reaction time to visual targets tends to be faster when stimuli from another modality (in particular, audition and touch) are presented in close temporal or spatial proximity even when subjects are instructed to ignore the accessory input (focused attention task). Multisensory interaction effects measured in neural structures involved in saccade generation (in particular, the superior colliculus) have demonstrated a similar spatio-temporal dependence. Neural network models of multisensory spatial integration have been shown to generate convergence of the visual, auditory, and tactile reference frames and the sensorimotor coordinate transformations necessary for coordinated head and eye movements. However, since these models do not capture the temporal coincidences critical for multisensory integration to occur, they cannot easily predict multisensory effects observed in behavioral data like saccadic reaction times. This paper proposes a quantitative stochastic framework, the time-window-of-integration (TWIN) model, to account for the temporal rules of multisensory integration. Saccadic responses collected from a visual-tactile focused attention task are shown to be consistent with the TWIN model predictions.

INTRODUCTION

Saccades are fast, voluntary movements of the eyes to align the high-resolution fovea with objects and events of interest. In a natural environment saccades are part of a rapid goal-directed orienting response, often including head movements, to stimuli occurring in the periphery. These stimuli are usually crossmodal like, e.g., a rapidly approaching subway train. In addition to visual and auditory inputs, vestibular and somatosensory afferents have access to the saccade-generating mechanism. Thus, it is no surprise that the oculomotor system has become a prominent arena for the analysis of multisensory integration. For example, it has been found that saccadic reaction time to visual targets (the time between the onset of the visual stimulus and the onset of the saccadic eye movement) tends to be faster when auditory stimuli are presented in close temporal or spatial proximity (Colonius & Arndt, 2001; Frens, Van Opstal, & Van der Willigen, 1995; Harrington & Peck 1998; Hughes, Nelson, & Aronchick, 1998). Similar response enhancement effects for saccades have been observed for combining visual and somatosensory stimuli (cf. Groh & Sparks, 1996, for monkeys; Amlôt, Walker, Driver, & Spence, 2003; Diederich, Colonius, Bockhorst, & Tabeling, 2003 for humans).

These psychophysical observations are in line with neurophysiological evidence for multisensory integration in the deep layers of the superior colliculus (SC), an area clearly involved in saccade generation (Robinson, 1972). Multisensory neurons in SC of anesthetized cats (Stein, Magalhães-Castro, & Kruger, 1976; Meredith & Stein, 1986a,b) and monkeys (Wal-

lace, Wilkinson, & Stein, 1996) show an enhanced response to particular combinations of visual, auditory, and tactile stimuli paralleling the spatial-temporal rules observed in behavioral studies. Similar results have recently been obtained for recordings from unanesthetized cats by Wallace, Meredith, & Stein (1998) and from the awake behaving monkey by Bell, Corneil, Meredith, & Munoz (2001) and by Frens & Van Opstal (1998). Information about stimulus location is represented topographically within the structure by an orderly arrangement of neurons according to the location of their respective receptive fields (RFs). The spatial register among the different sensory maps is formed by multisensory neurons whose different RFs are in register with one another (Meredith & Stein, 1986a,b; for a review, see Stein & Meredith 1993). In addition, the SC contains a motor map composed of output neurons coding appropriate eye movements (Sparks, 1986), i.e., the locus of activity in the motor map encodes a movement command that reorients the eyes (and the head) a given distance in a particular direction.

There are a number of explicit neural network models of multisensory spatial integration mechanism describing the integration of visual, auditory, and tactile reference frames and the sensorimotor coordinate transformations necessary for directed head and eye movements (cf. Xing & Andersen, 2000; Deneve, Latham, & Pouget, 2001). While these models capture many aspects of *spatial* rules of multisensory integration (for a recent critical review, see Pouget, Deneve, & Duhamel, 2002), they do not address the effects of *temporal* coincidences of crossmodal stimulus sets. In this paper, we propose a time-window-of-integration (TWIN) model to serve as a framework incorporating the temporal rules of multisensory integration

in saccades. A parametric version of the model is fit to a recent set of data on visual-tactile interaction in saccades.

Time-Window-of-Integration (TWIN) Model

The classic explanation for a speed-up of responses to multisensory stimuli is that subjects are merely responding to the first stimulus detected. Formally, observed reaction time (RT) would then be the minimum of the response time to the visual, auditory, or tactile signal causing a *statistical facilitation* effect (Raab, 1962). However, several studies have clearly shown that statistical facilitation is not sufficient to explain the speed-up in saccadic reaction time (Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994; Corneil & Munoz, 1996; Harrington & Peck, 1998; Hughes et al., 1998). Using Miller's inequality as a benchmark test (see Miller, 1982; Diederich, 1995; Colonius, 1990) responses to bimodal stimuli have been found to be faster than predicted by statistical facilitation, in particular when the stimuli were spatially aligned. Moreover, the race model has no simple explanation for the decrease in facilitation observed with increasing spatial disparity between the stimuli.

While there is increasing evidence for a complex network of largely parallel neural sub-processes underlying performance even in simple crossmodal tasks (Driver & Spence, 2000), the initial separation of the afferent pathways for the different sensory modalities suggests that one can distinguish 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 sub-processes (*second stage*). As shown below, in conjunction with a number of weak additional

assumptions some interesting and empirically testable predictions can be derived from this simple setup.

Even under invariant experimental conditions, saccadic responses typically vary from one trial to the next due to an inherent variability of the underlying neural processes in both ascending and descending pathways. This is taken into account in the time-window-of-integration (TWIN) model by assuming the duration of each of the stages to be a random variable.

————— Figure 1 about here —————

First Stage Assumption: *In the first stage there is a race among the peripheral neural excitations in the visual, auditory, and/or somatosensory pathways triggered by a crossmodal stimulus complex.*

Since the first stage refers to very early sensory processing, random processing times for visual, auditory, and somatosensory stimuli are assumed to be statistically independent.

Second Stage Assumption: *The second stage comprises neural integration of the input and preparation of an ocular motor response. Multisensory integration manifests itself in an increase or decrease of second stage processing time.*

Distinguishing between two stages only is clearly an oversimplification. But note 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 formal description of the model see Methods Section).

Time-Window-of-Integration Assumption: *Multisensory integration occurs only if the peripheral processes of the first stage all terminate within a given time interval, the "window of integration".*

The window of integration acts like a filter determining whether afferent information delivered from different sensory organs is registered close enough in time to allow for multisensory integration. Passing the filter is a necessary, but not a sufficient condition for multisensory integration to occur. The reason is that multisensory integration also depends on the spatial configuration of the stimuli. However, rather than assuming the existence of a joint spatial-temporal window of integration permitting interaction to occur only for both spatially and temporally neighbored stimuli, the TWIN model allows for multisensory integration to occur even for rather distant stimuli (of different modalities) as long as they fall within the time window. Such interaction will typically be an inhibition or only a small facilitation. This arrangement affords more flexibility in a complex environment. For example, response depression may occur with nearly simultaneous but distant stimuli making it easier for the organism to focus attention on the more important event.

Predictions

The TWIN model makes a number of empirical predictions. First, the amount of multisensory integration should depend on the stimulus onset asynchrony (SOA) between the stimuli. Indeed, the effect tends to be most prominent when there is some characteristic temporal

asynchrony between the stimuli (Frens et al., 1995; Colonius & Arndt, 2001). Within the model this simply means that a stimulus with faster peripheral processing has to be delayed in such a way that the arrival times of both stimuli have a higher probability of falling into the window of integration.

Second, the probability of interaction, $\text{Pr}[I]$, should depend on unimodal features that affect the speed of processing in the first stage, like stimulus intensity or eccentricity. For example, if a stimulus from one modality is very strong compared to the other stimulus' intensity, the former stimulus' peripheral processing speed will increase, and the chances that both peripheral processes terminate within the time window will be smaller (assuming simultaneous stimulus presentations). The resulting low value of $\text{Pr}[I]$ is in line with the empirical observation that a very strong target signal will effectively rule out any further reduction of RT by adding a stimulus from another modality (e.g., Corneil, Wanrooij, Munoz, & Van Opstal, 2002).

Finally, the amount of multisensory integration (Δ) and its direction (facilitation or inhibition) occurring in the second stage depend on crossmodal features of the stimulus set, in particular spatial disparity and laterality (laterality here refers to whether or not all stimuli appear in the same hemisphere). Crossmodal features cannot have an influence on first stage processing time since the modalities are yet being processed in separate pathways. More specific predictions require an explication of the rules governing the window-of-integration mechanism in specific task requirements.

Integration Rule Assumptions

Focused Attention Task: *When the task is to orient toward the target stimulus ignoring stimuli from other modalities, the first stage terminates when the target peripheral process ends. Multisensory integration occurs only if the non-target stimulus wins the race in the first stage opening a time window such that the termination of the target peripheral process is enclosed in the window.*

In other words, in the focused attention situation, the window of integration is "opened" only by activity triggered by the non-target stimulus, but first stage duration is determined by the target stimulus processing time regardless of the identity of the winner of the race.

Redundant Target Task: *When the task is to orient toward the first stimulus detected no matter of which modality, the first stage duration is defined by the winner's peripheral processing time, and the window of integration is opened by whichever stimulus wins the race.*

From these assumptions, further predictions concerning the effects of varying stimulus intensity follow. Take, for example, a focused attention task with a visual target and an auditory non-target stimulus. Increasing the intensity of the visual stimulus will speed up visual peripheral processing (up to some minimum level) thereby increasing the chances for the visual target to win the race. Thus, the probability that the window of integration opens decreases predicting less multisensory integration. Increasing the intensity of the non-target auditory stimulus, on the other hand, leads to the opposite prediction: the auditory stimulus will have

a better chance to win the race and to open the window of integration, hence predicting more multisensory integration to occur on average.

Note that the principle of "inverse effectiveness", according to which multisensory facilitation is strongest when stimulus strengths are weak or close to threshold level (Meredith & Stein 1986a), can be accommodated in the model by the following additional hypothesis: for low-intensity stimuli, the window becomes larger so as to increase the likelihood of multisensory integration ¹.

Separating Spatial and Temporal Factors of Integration

Expected multisensory integration is defined as the difference between mean response time in the unimodal and the crossmodal condition. An important property of the TWIN model, formally derived in the model section of the appendix, is the factoring of expected multisensory integration, i.e., expected multisensory integration is simply the product of the probability of interaction $\Pr[I]$ and the amount and sign of interaction (Δ). According to the assumptions, the first factor depends on the temporal configuration of the stimuli (stimulus onset asynchrony, SOA), whereas the second factor depends on their spatial configuration. Note that this separation of temporal and spatial factors is in accordance with the definition of the window of integration: the incidence of multisensory integration hinges upon the stimuli to occur in temporal closeness, whereas the amount and sign of interaction (Δ) is modulated by spatial proximity, reaching from enhancement for neighboring stimuli to possible inhibition for distant stimuli.

EXPERIMENT

The alignment of visual, auditory, and tactile topographical maps in SC (cf. Meredith & Stein, 1986a,b) suggests that, in analogy to visual-auditory interaction, visual-tactile interaction in saccade generation should depend on the spatial configuration of the stimuli. Here we examined the effect of a tactile stimulus (vibration applied to palm) on response time for saccades toward a visual target as a function of the spatial-temporal visual-tactile stimulus configuration. Visual and tactile stimuli were presented 20° or 70° left or right of the fixation point. Subjects were asked to make a saccade as quickly and as accurately as possible toward the visual stimulus appearing randomly left or right. They were instructed to ignore a tactile accessory stimulus presented at the corresponding hand position ipsi- or contralateral to the visual target either before the visual stimulus (100 or 50 ms), simultaneous to, or 50 ms after the visual. In unimodal trials, only a visual stimulus was presented. They were randomly mixed with bimodal trials.

RESULTS

Mean RTs and standard errors over all subjects as a function of eccentricity and laterality are shown in Figure 2. If tactile stimulation had no effect on RT, all means within a given eccentricity (20° or 70°) should be about the same. However, presentation of the tactile accessory stimulus had a facilitating effect on saccadic RT ($p < .01$). Ipsilateral bimodal stimulation had a stronger effect than contralateral presentation ($p < .01$), and responses to

the 20° position were faster than to the 70° position ($p < .01$). Furthermore, all SOAs had a significant effect on saccadic RT ($p < .01$).

————— Figure 2 about here —————

To better illustrate the effects of eccentricity and SOA on multisensory integration, the data are presented in terms of multisensory response enhancement (MRE) in analogy to Meredith & Stein (1986a) (cf. Anastasio et al. 2000; Colonius & Diederich, 2002):

$$\text{MRE} = \frac{RT_{unimodal} - RT_{bimodal}}{RT_{unimodal}} \times 100. \quad (1)$$

The following MRE values obtained:

————— Table 1 about here —————

Multisensory integration, as measured by MRE, is a monotonically decreasing function of SOA. That is, in the given range of SOA values, multisensory integration is the larger the earlier the tactile stimulus is presented. Note that this is consistent with the TWIN model's prediction: the earlier the tactile non-target is presented relative to the target, the higher its chance to win the peripheral race and to open the integration window. At SOA= 50 ms the tactile peripheral process terminates too late to open the window.

The fact that saccadic RT typically increases with eccentricity of the visual target (up to 30° the increase is about .4 ms per 1°, e.g., Findlay & Walker 1999) could be a confounding factor for the effect of eccentricity on facilitation. However, the relative enhancement as measured by the MRE introduced above would not be affected by this. Nonetheless, MRE

increases with increasing eccentricity of the visual stimulus separately for the ipsi- and the contralateral condition (Table 1).

The error rates for moving the eyes to the non-target direction are listed in parentheses in Table 1. As expected, this rate increases the earlier the non-target is presented contralateral to the target, but for ipsilateral presentations erroneous responses are below 5% at 20° and below 8% at 70°.

TWIN Model Fit

According to the TWIN model, saccadic RT in the bimodal condition equals the sum of peripheral visual target processing time and the duration of the neural integration stage. Multisensory integration in the second stage occurs if (i) the race in the first stage is won by the tactile (non-target) modality and (ii) termination of the visual peripheral processing time falls within the time window of integration.

To allow for quantitative predictions, independent exponential peripheral processing times are assumed in the TWIN model. In total, there were 9 parameters to be estimated from 18 data points (corresponding to the means observed in the different experimental conditions).

————— Figure 3 —————

The fits, obtained separately for each subject, are illustrated in Figure 3 for two of the 6 subjects, the other fits being similar².

————— Table 2 about here —————

Obviously, the model captures the main features of the data at the level of the means. Mean saccadic reaction time is increasing with SOA for all stimulus configurations, except for the 70° contralateral condition in participant ML where monotonicity is reversed due to inhibition, consistent with the model. Moreover, the parameter estimates are ordered as expected for all participants (including those not reported here): (a) visual peripheral processing is faster for the 20° position than for the 70° position; (b) peripheral tactile processing is slower than visual processing; (c) the amount of facilitation is the larger the closer the tactile stimulus is to the visual. The optimal width for the time window was about 200 ms (cf. Table 2).

DISCUSSION

Experiment

Saccadic reaction time to a visual target presented randomly left or right from fixation was reduced by up to 22% when accompanied by a spatially non-predictive tactile stimulus. This facilitation effect was larger for spatially aligned visual-tactile stimulus configurations than for contralateral presentation, and it increased with stimulus eccentricity (20° vs. 70°). Moreover, responses were the faster the earlier the tactile stimulus was presented (in a range of the tactile preceding the visual by 100 ms to following it by 50 ms). These results extend previous findings in our lab to non-simultaneous visual-tactile stimulus presentations

(Diederich et al., 2003).

A partial explanation for the observed speed-up not involving any multisensory integration is that the tactile stimulus acts as a warning signal about the forthcoming visual target onset. In terms of the underlying neural oculomotor circuitry, a warning signal would lead to a decrease in activity of the fixation neurons in SC thereby facilitating saccade generation (cf. Frens et al. 1995). Note, however, that this would not account for the observation that ipsilateral stimulation was uniformly more effective than contralateral stimulation, since the tactile stimulus was not predictive with respect to the location of the target. Moreover, the differential effect of eccentricity on multisensory response enhancement cannot be explained with an invariant warning signal influence.

While in line with recent results in Amlôt et al. (2003), the fact that even contralateral distractors had a facilitating effect on saccadic RT seems to be at odds both with certain empirical findings and with models of saccade target selection. A case in point is the "remote distractor" effect that consists in an increase of RT when the target and a distractor are presented at remote positions (Walker, Deubel, Schneider, & Findlay, 1997), in accordance with a model by Trappenberg, Dorris, Munoz, & Klein (2001) holding that each stimulus produces an increase of collicular neuronal activity in separate populations in the collicular salience map. Target selection is accomplished by the existence of a single salience peak in the collicular map. This requires an inhibitory process taking a certain amount of time that is not needed in the single target case, contrary to the facilitation observed in our contralateral distractor conditions. It should be noted here, however, that the "remote distractor" effect

seems to be confined to the visual modality. For example, both Frens et al. (1995) and Colonius & Arndt (2001) did find a facilitating effect of remote auditory distractors as well. While the size of these effects obviously depends on the relative intensity of target and distractor, it could be that visual distractors generate a more salient representation on the collicular activity map. This may be due in part to a more spread-out representation of auditory or tactile stimuli relative to visual stimuli, possibly resulting from differences in the corresponding receptive field sizes in the periphery.

Our results are compatible with the notion that the observed saccadic RT facilitation results from genuine visual-tactile integration taking place in the SC, an area involved in eye movement control. Stein & Meredith (1993, pp. 134-140) specify that an optimal interactive window for most visual-somatosensory neurons in cat SC appears to be less than 250 ms, which accords with our estimate of 200 ms. They estimate about 26 ms for a stimulus touching the skin to activate the neuron, while peripheral processing for a visual stimulus requires 65-100 ms. If similar numbers hold for the human brain, then maximal temporal overlap of discharges would be expected for SOAs from -100 to -50 ms, the range where we actually did observe maximal response enhancement.

Note that the observed lack of contralateral inhibition on saccade generation discussed above is also at odds with the general finding of multisensory (visual-auditory) "response depression" for spatially disparate stimuli measured both in SC neurons and overt orientation behavior in the cat (Stein, Meredith, Huneycutt, & McDade, 1989; Wilkinson, Meredith, & Stein, 1996). Interestingly, response enhancement under spatial disparity in an orienting

task has also been found in peripheral sensory space of the cat (visual target at $\pm 30^\circ$ or $\pm 45^\circ$) when an auditory stimulus was more peripheral to the visual target (Jiang, Jiang, & Stein, 2002). Moreover, multisensory integration properties of most SC neurons are mediated by influences from two cortical areas, the anterior ectosylvian sulcus (AES) and the rostral aspect of the lateral suprasylvian sulcus (rLS) (Jiang, Wallace, Jiang, Vaughan, & Stein, 2001). In the focused attention paradigm employed here, where target and non-target modalities must be distinguished to accomplish the task, the contribution of further cortical areas making use of the warning signal function of the tactile stimulus cannot be ruled out as a partial explanation of the lack of contralateral inhibition.

Finally, note that the observed increase of multisensory response enhancement with the eccentricity of the visual target makes sense from a functional point of view. Orienting the eyes toward more eccentric targets takes more time (cf. Bell, Everling, & Munoz, 2000), and thus a facilitating effect of the tactile accessory stimulus would be more important for more eccentric targets (see also Bell, Corneil, Meredith, & Munoz, 2001). This differential contribution of multisensory neurons is another instantiation of the principle of *inverse effectiveness* (see also Anastasio, Patton, & Belkacem-Boussaid, 2000; Colonius & Diederich, 2002).

TWIN Model

The TWIN model mechanisms postulated to generate the behavioral data are not meant to mirror the processes at the level of an individual neuron. There are many different

types of multisensory convergence occurring in individual neurons (for a recent review, see Meredith, 2002) and some of their activities are consistent with the TWIN assumptions while others are not. For example, for certain neurons in cat SC, given a visual stimulus is always presented in the visual receptive field of the bimodal (visual-auditory) neuron, an auditory stimulus presented outside its receptive field will generate only response depression (or no interaction) regardless of the SOA between the stimuli (Stein & Meredith, 1993, p. 140). This accords nicely with the separation of temporal and spatial factors postulated by the TWIN model. However, in other visual-auditory neurons changing the temporal order and interval between the stimuli can change enhancement to depression even if their spatial arrangement is left invariant (Meredith, Nemitz, & Stein, 1987). This latter type of behavior could only be accounted for in TWIN if two stimuli, after having passed the filter of the temporal window of integration, could still carry over information about their temporal distance to the subsequent convergence stage. Moreover, besides the common excitatory-excitatory type of multisensory convergence there is evidence for an excitatory-inhibitory type of neural circuit by which inputs from one modality inhibit those from the other (cf. Meredith, 2002). These circuits could play a specific role in focused attention situations (ibid, p. 37). In any event, at this stage of development the TWIN model operates at the level of behavioral data resulting from the combined activity of a possibly large number of neurons and from the specific task instructions, so that the existence of different types of multisensory convergence in individual neurons does not provide a strong model constraint.

Since stimulation from different modalities like vision and touch cannot interact (e.g.,

on the retina), the main assumption of the two-stage model, namely the existence of a first stage of parallel independent modality-specific activations in the afferent pathways, seems uncontroversial. It refers to a very early stage of processing where detection of the stimuli, but possibly no "higher" processes like localization and identification, take place. Note that the two-stage assumption does not preclude the possibility of interaction between modality-specific pathways, nor between modality-specific and multisensory areas, at a later stage. In fact, there is increasing evidence that multisensory processing does not take place entirely in feed-forward convergent pathways but that it can also modulate early cortical unisensory processing, as suggested by recent functional magnetic resonance imaging (fMRI) and ERP studies (see Macaluso, Frith, & Driver, 2000, and Laurienti, Burdette, Wallace, Yen, Field, & Stein, 2002, for fMRI; Molholm, Ritter, Murray, Javitt, Schroeder, & Foxe, 2002, and McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003, for ERP).

Even if certain assumptions of the TWIN model turn out to be inconsistent with an experimental result, one advantage of our model framework is that it facilitates the statement of clear-cut hypotheses about multisensory integration rules. For example, consider the "hypothesis of restricted influence". According to TWIN, unimodal stimulus properties, like intensity or spatial position, have a direct influence on first-stage processing time. Moreover, unimodal properties may *indirectly* influence the resulting multisensory integration taking place in the second stage by modulating the opening of the integration window. The "restricted influence hypothesis" then denies existence of a direct influence of unimodal stimulus properties upon second stage processing. For example, increasing or decreasing stimulus in-

tensity may not affect multisensory integration directly but only via changing properties of the window of integration, e.g., its probability of opening or its width. A saccadic eye movement study in the visual-auditory domain lends some support to this hypothesis. In a focused attention study Arndt & Colonius (2003) found that saccadic response time decreased with increasing the intensity of a non-target auditory stimulus while they could rule out any direct effect of intensity on second stage processing³.

Further experimental tests of the TWIN model should include a larger range of stimulus onset asynchronies. There is some empirical evidence that the assumed symmetry of the crossmodal effects with respect to the temporal order in which the modalities are presented may not hold up in general (e.g., Corneil, et al., 2002). Moreover, the second stage mechanisms should be specified in more detail with respect to the spatial stimulus configuration effects, possibly integrating aspects of the neural network models mentioned in the introduction. There is a large data base on receptive field properties of multisensory neurons available now (cf. Kadunce et al., 2001), and connecting these with behavioral data via an appropriate elaboration of the TWIN model should be a challenging task.

METHODS

Participants

Six students (4 female) served as paid voluntary participants in the experiments. All participants had 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.

Apparatus and stimulus presentation

Red light-emitting diodes (LED, 5 mm, 3.7 mcd) served as visual targets presented against a black background. An additional LED (red, 5 mm, 0.4 mcd) served as fixation point. Tactile stimuli were vibrations (50 Hz, 0.6 V, 1–2 mm amplitude) transmitted through wooden balls applied to the center of the palm, generated by two silenced oscillation devices (Mini-shaker, Type 4810, B & K). All stimuli were positioned on top of a table (180 cm × 130 cm × 75 cm) with a recess to sit in (referred to as *vertex*). The fixation LED was 38.5 cm away from the lower edge of the table. For each experimental condition the two vibrators were moved to the respective positions.

Experimental procedure

The experiment was carried out in a completely darkened room so that participants were unable to see their hands during the experiment. Every session began with 10 min of dark adaptation during which the measurement system was adjusted and calibrated. During this phase the participants put their hands at the position used during the entire experimental block. Thus, the participants were aware of the hand position and, thus, the position of the tactile stimulus. Participants were sitting at the longitudinal side (at the vertex) using a chin rest facing the calibrating screen and wearing a video camera frame. Each trial began with the appearance of the fixation point. After a variable fixation time (800-1500 ms), the fixation point disappeared and both a visual and a tactile stimulus were presented for

500 ms (no gap) at positions 20° or 70° left or right of the fixation point, the tactile stimulus 55 cm and the visual stimulus 50 cm away from the vertex. In unimodal trials, only a visual stimulus was presented. Subjects were instructed to move their eyes to the visual target as quickly as possible, while the tactile stimulus could be ignored. The interval between stimulus offset and fixation onset for the next trial was 2 seconds. Each participant was first trained for 1000 trials not included in the data analysis.

Data collection

Saccadic eye movements were recorded by an infrared video camera system (EyeLink system, Sensomotoric Instruments) with temporal resolution of 250 Hz and horizontal and vertical spatial resolution of 0.01° . Eye position data from each trial were inspected for proper fixation at the beginning of the trial, for blinks, and for correct detection of start and endpoint of the saccade. Saccadic reaction time, start position of the eye, and end position after the saccade (vertical and horizontal positions in degree of visual angle relative to the straight ahead fixation point) were calculated from the controlled data samples.

Experimental Design

There were three levels of factor *laterality*: ipsilateral, contralateral, and visual (LED) only. The second factor, *eccentricity*, refers to the eccentricity of the visual and tactile stimuli with levels 20° and 70° . The third factor, SOA, had levels -100 , -50 , 0 , and 50 ms. In each block of 270 trials stimulus position (20° or 70°) was held fixed resulting in 18 different conditions per block (4 SOA levels, visual-tactile stimulus pairs presented ipsi-

or contralateral, left or right, plus 2 unimodal visual conditions). Thus, except for the laterality, the participants knew where the stimuli would appear. There were 7 blocks for either stimulus position resulting in 105 replications for each condition. RTs were then pooled over left/right presentation.

Statistical Analysis

A three-way ANOVA ($3 \times 2 \times 4$) yielded significant main effects ($p < .01$). Moreover, all interactions were significant ($p < .01$ for the pairs, $p < .05$ for the triple). Post-hoc tests showed that the presence of a tactile accessory stimulus had a significant effect ($p < .01$). Ipsi- vs. contralateral bimodal stimulation was significant ($p < .01$). Furthermore, all SOAs had a significant effect ($p < .01$) on RT as well as the eccentricity ($p < .01$).

Quantifying multisensory integration in the TWIN Model

According to the two-stage assumption, total reaction time in the crossmodal condition can be written as a sum of two random variables:

$$RT_{crossmodal} = W_1 + W_2, \quad (2)$$

where W_1 and W_2 refer to the first and second stage processing time, respectively. Let I denote the event that multisensory integration occurs, having probability $\Pr(I)$. For the expected saccadic reaction time in the crossmodal condition then follows:

$$E[RT_{crossmodal}] = E[W_1] + E[W_2]$$

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

where $E[W_2|I]$ and $E[W_2|\text{not-}I]$ denote the expected second stage processing time conditioned on interaction occurring (I) or not occurring ($\text{not-}I$), respectively. Putting $\Delta \equiv E[W_2|\text{not-}I] - E[W_2|I]$, this becomes

$$E[RT_{crossmodal}] = E[W_1] + E[W_2|\text{not-}I] - \Pr[I] * \Delta. \quad (3)$$

The term $\Pr[I] * \Delta$ can be interpreted as a measure of the expected saccadic RT speed-up in the second stage with positive Δ values corresponding to facilitation, negative ones to inhibition. In the unimodal condition, no interaction is possible. Thus,

$$E[RT_{unimodal}] = E[W_1] + E[W_2|\text{not-}I], \quad (4)$$

and

$$E[RT_{unimodal}] - E[RT_{crossmodal}] = \Pr[I] * \Delta$$

TWIN Model for Visual-Tactile Interaction in Focused Attention

With V denoting visual peripheral processing time, T tactile peripheral processing time, τ stimulus onset asynchrony, and ω the width of the integration window, multisensory integration depends on the event I that the tactile stimulus wins the race in the first stage and that visual peripheral processing terminates within the window of integration,

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

With W_2 denoting second stage processing time, expected saccadic response time to a visual target stimulus with tactile accessory presented at τ is

$$E[RT_{VT,\tau}] = E[V] + E[W_2 | \text{not-}I] - \Pr[I] * \Delta. \quad (5)$$

As before, expected multisensory integration, measured as difference between expected uni- and bimodal response time, is

$$\begin{aligned} E[RT_V] - E[RT_{VT,\tau}] &= \Pr[I] * \{ E[W_2 | \text{not-}I] - E[W_2 | I] \} \\ &= \Pr[I] * \Delta, \end{aligned}$$

where $E[RT_V] = E[V] + E[W_2 | \text{not-}I]$, the unimodal response time to the visual target only.

Acknowledgments

We thank B. E. Stein and two anonymous reviewers for helpful comments. This research was supported by grants from Deutsche Forschungsgemeinschaft SFB 517/C3 “Neurokognition” (HC) and Di 506/7-1, /7-2, /8-1 (AD).

Reprint requests should be sent to Hans Colonius, Department of Psychology, Oldenburg University, P.O.Box 2503, D-26111 Oldenburg (Germany), or via email: hans.colonius@uni-oldenburg.de.

Notes

¹In the focused attention task, the occurrence of inverse effectiveness should only depend on the intensity of the *target* stimulus. While not implausible, this explanation of "inverse effectiveness" is post-hoc and needs further empirical testing.

²Given the large number of parameters and the relatively small data set we refrain from reporting quantitative fit statistics at this exploratory state of model testing.

³A probability inequality test revealed that the amount of multisensory integration (as measured by comparison with the prediction of a simple, one-stage race model) was a function of the spatial visual-auditory configuration, but not of auditory intensity. Within the logic of the model, an auditory non-target with higher auditory intensity level has a greater probability to open the time window predicting a response speed-up. Since the auditory stimulus was not the target, this does not contradict the "inverse effectiveness" hypothesis.

References

- Amlôt, R., Walker, R., Driver, J., & Spence, C. (2003). Multimodal visual-somatosensory integration in saccade generation. *Neuropsychologia*, *41*, 1-15.
- Anastasio, T. J., Patton, P. E., Belkacem-Boussaid, K. (2000). Using Bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Computation*, *12*, 1165-1187.

- Arndt, A., & Colonius, H. (2003) Two separate stages in crossmodal saccadic integration: Evidence from varying intensity of an auditory accessory stimulus. *Experimental Brain Research*, *148*, 328-337.
- Bell, A. H., Corneil, B. D., Meredith, M. A., & Munoz, D. P. (2001). The influence of stimulus properties on multisensory processing in the awake primate superior colliculus. *Canadian Journal of Experimental Psychology*, *55*, 123-132.
- Bell, A. H., Everling, D., & Munoz, D. P. (2000). Influence of stimulus eccentricity and direction on characteristics of pro- and antisaccades in non-human primates. *Journal of Neurophysiology*, *84*, 2595-2604.
- Colonius, H. 1990. Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology*, *34*, 253-275.
- Colonius, H., & Arndt, P. (2001). A two-stage model for visual-auditory interaction in saccadic latencies. *Perception and Psychophysics*, *63*, 126-147.
- Colonius, H., Diederich, A. (2002). A maximum-likelihood approach to modeling multisensory enhancement. In: Dietterich, T.G., Becker, S., Ghahramani, Z. (eds) *Advances in Neural Information Processing Systems 14*, The MIT Press: Cambridge.
- Corneil, B. D., & Munoz, D. P. (1996). The influence of auditory and visual distractors on human orienting gaze shifts. *Journal of Neuroscience*, *16*, 8193-8207.

- Corneil, B. D., Van Wanrooij, M., Munoz, D. P., & Van Opstal, A. J. (2002). Auditory-visual interactions subserving goal-directed saccades in a complex scene. *Journal of Neurophysiology*, *88*, 438-454.
- Deneve, S., Latham, P., & Pouget, A. (2001). Efficient computation and cue integration with noisy population codes. *Nature Neuroscience*, *4*, 826-831.
- Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology*, *39*, 197-215.
- Diederich, A., Colonius, H., Bockhorst, D. & Tabeling, S. (2003). Visual-tactile spatial interaction in saccade generation. *Experimental Brain Research*, *148*, 328-337.
- Driver, J., & Spence, C. (2000). Multisensory perception: beyond modularity and convergence. *Current Biology*, *10*, R731-R735.
- Findlay, J. M., Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, *22*, 661-674.
- Frens, M. A., & Van Opstal (1998). Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Research Bulletin*, *46*, 211-224.
- Frens, M. A., Van Opstal, A. J., Van der Willigen, R. F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception and Psychophysics*, *57*, 802-816.

- Groh, J. M., & Sparks, D. L. (1996). Saccades to somatosensory targets. I. Behavioral characteristics. *Journal of Neurophysiology*, *75*, 412-427.
- Harrington, L. K., & Peck, C. K. (1998). Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Experimental Brain Research*, *122*, 247-252.
- Hughes, H. C., Nelson, M. D., & Aronchick, D. M. (1998). Spatial characteristic of visual-auditory summation in human saccades. *Vision Research*, *38*, 3955-3963.
- Hughes, H. C., Reuter-Lorenz, P. A., Nozawa, G., & Fendrich, R. (1994). Visual-auditory interactions in sensorimotor processing: Saccades versus manual responses. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 131-153.
- Jiang, W., Jiang, H., & Stein, B. E. (2002). Two corticotectal areas facilitate orientation behavior. *Journal of Cognitive Neuroscience*, *14*, 1240-1255.
- Jiang, W., Wallace, M. T., Jiang, H., Vaughan, J. W., & Stein, B. E. (2001). Two cortical areas mediate multisensory integration in superior colliculus neurons. *Journal of Neurophysiology*, *85*, 506-522.
- Kadunce, D. C., Vaughan, J. W., Wallace, M. T., Benedek, G., & Stein, B. E. (1997). Mechanisms of within- and cross-modality suppression in the superior colliculus. *Journal of Neurophysiology*, *78*, 2834-2847.
- Kadunce, D. C., Vaughan, J. W., Wallace, M. T., & Stein, B. E. (2001). The influence of visual and auditory receptive field organization on multisensory integration in the

- superior colliculus. *Experimental Brain Research*, 139, 303-310.
- Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y., Field, A. S., & Stein, B.E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *Journal of Cognitive Neuroscience*, 14, 420-429.
- Macaluso, E., Frith, C. D., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, 289, 1206-1208.
- McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F. & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, 15, 10-19.
- Meredith, M. A. (2002). On the neuronal basis for multisensory convergence: a brief overview. *Cognitive Brain Research*, 14, 31-40.
- Meredith, M. A., & Stein, B. E. (1986a). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, 56, 640-662.
- Meredith, M. A., & Stein, B. E. (1986b). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research*, 365, 350-354.
- Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience*, 7, 3215-3229.

- Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology, 14*, 247-279.
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive Brain Research, 14*, 115-128.
- Pouget, A., Deneve, S., & Duhamel, J. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews, 3*, 741-747.
- Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences, 24*, 574-590.
- Robinson, D. A. (1972). Eye movements evoked by collicular stimulation in the alert monkey. *Vision Research, 12*, 1795-1808.
- Sparks, D. L. (1986). Translation of sensory signals into commands for control of saccadic eye movements: Role of primate superior colliculus. *Physiological Review, 66*, 116-177.
- Stein, B. E., Meredith, M. A. (1993). *The merging of the senses*. The MIT Press: Cambridge, MA.
- Stein, B. E., Magalhães-Castro, & Kruger, L. (1975). Relationship between visual and tactile representations in cat superior colliculus. *Journal of Neurophysiology, 39*, 401-419.

- Stein, B. E., Meredith, M. A., Huneycutt, W. S., & McDade, L. (1989). Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience*, *1*, 12-24.
- Trappenberg, T. P., Dorris, M. C., Munoz, D. P., & Klein, R. M. (2001). A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. *Journal of Cognitive Neuroscience*, *13*, 256-271.
- Wallace, M. T., Meredith, M. A., & Stein, B. E. (1998). Multisensory integration in the superior colliculus of the alert cat. *Journal of Neurophysiology*, *80*, 1006-1010.
- Wallace, M. T., & Stein, B. E. (1994). Cross-modal synthesis in the midbrain depends on input from association cortices. *Journal of Neurophysiology*, *71*, 429-432.
- Wallace, M. T., Wilkinson, L. K., & Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology*, *76*, 1246-1266.
- Walker, R., Deubel, H., Schneider, W. X., & Findlay, J. M. (1997). Effect of remote distractors on saccade programming: evidence for an extended fixation zone. *Journal of Neurophysiology*, *78*, 1108-1119.
- Wilkinson, L. K., Meredith, M. A., & Stein, B. E. (1996). The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Experimental Brain Research*, *112*, 1-10.

Xing, J., & Andersen, R. A. (2000). Models of the posterior parietal cortex which perform multimodal integration and represent space in several coordinate frames. *Journal of Cognitive Neuroscience*, 12, 601-614.

Table Captions

Table 1: Multisensory response enhancement (MRE) (in ms) as a function of eccentricity, SOA, and laterality. For example, an MRE value of 10 means that saccadic reaction time to the visual target is reduced by 10% when a tactile stimulus is present.

Table 2: Parameter estimates (in ms) for TWIN model processing stages, window width, and facilitation for participants EN and ML.

Figure Captions

Figure 1: Two-stage schema for TWIN model. Multisensory integration occurs only if peripheral processes terminate within the temporal integration window (see text).

Figure 2: Mean SRT (in ms) across all 6 participants for 20° position (upper panel) and 70° position (lower panel) of visual target, as a function of laterality and SOA. Unimodal visual RT (LED only) is indicated by the horizontal lines.

Figure 3 Fit of TWIN model to mean SRTs for participants EN and ML. Symbols refer to data, lines to predictions. Squares = observed ipsilateral 20°; triangles = observed contralateral 20°; circles = observed ipsilateral 70°; and stars = contra 70°. Horizontal lines refer to predicted unimodal SRTs.

SOA	Eccentricity			
	20°		70°	
	Laterality		Laterality	
	ipsi	contra	ipsi	contra
-100	16.2(3.9)	7.8(4.4)	22.4(5.8)	8.9(17.4)
-50	10.3(1.9)	6.3(3.2)	14.8(7.1)	5.4(12.3)
0	5.3(3.9)	6.2(2.7)	12.4(6.1)	2.9(9.3)
50	-0.3(2.9)	1.3(2.4)	2.5(5.2)	-0.2(8.3)

Table 1:

Processing time estimates (ms)	EN	ML
visual peripheral (20°)	38	13
visual peripheral (70°)	75	52
tactile peripheral	103	100
second stage	118	186
integration window width	201	200
facilitation ipsilateral	49	51
facilitation contralateral 20°/70°	33/39	2/-16

Table 2:

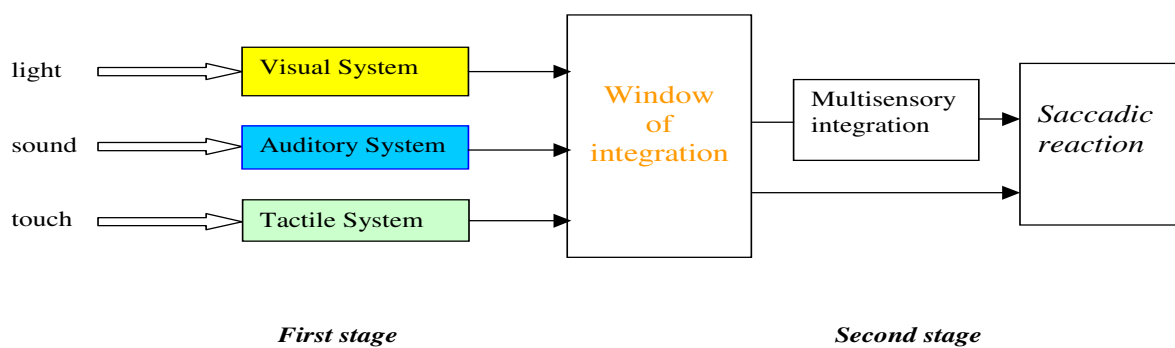


Figure 1:

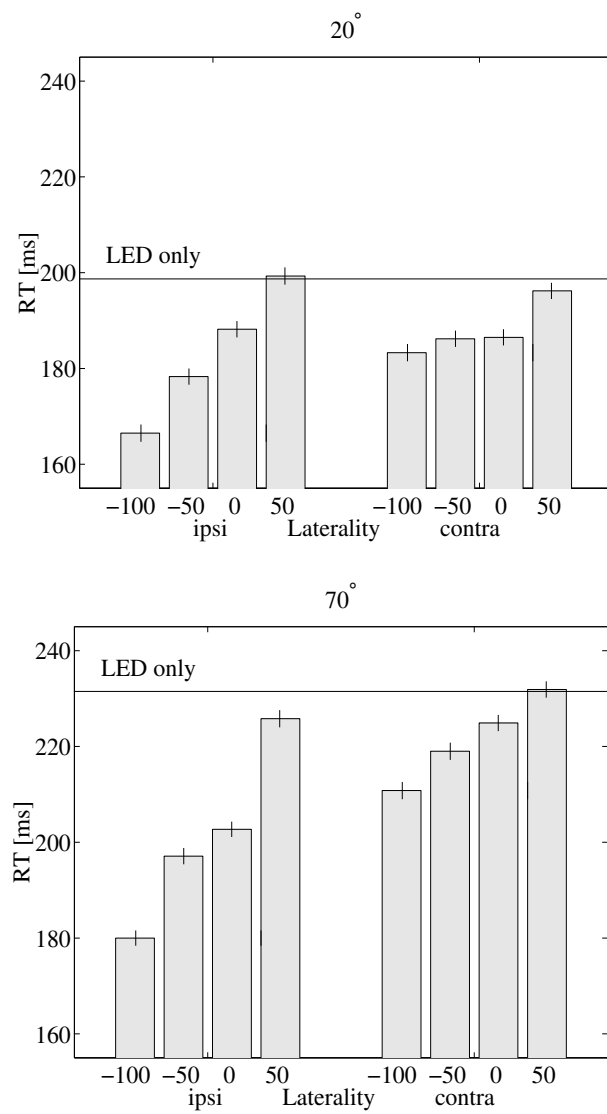


Figure 2:

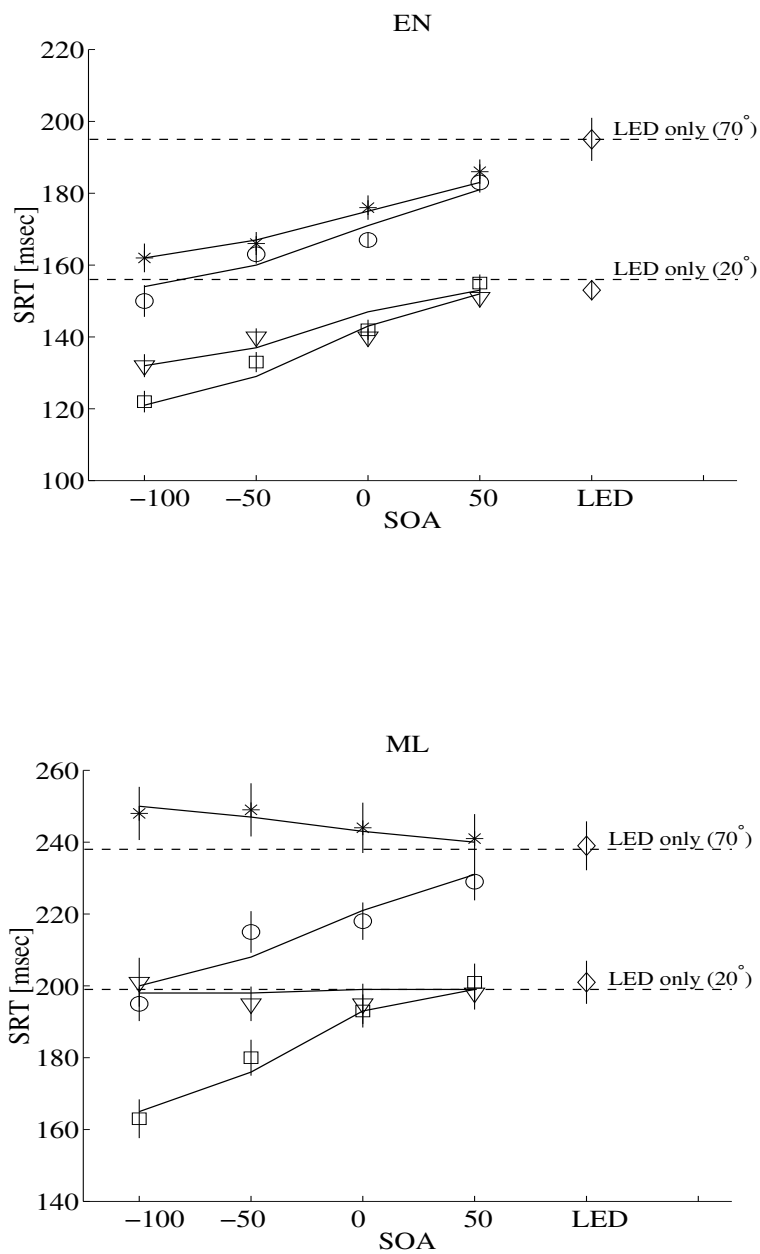


Figure 3: