

A two-stage model for visual–auditory interaction in saccadic latencies

HANS COLONIUS and PETRA ARNDT
University of Oldenburg, Oldenburg, Germany

In two experiments, saccadic response time (SRT) for eye movements toward visual target stimuli at different horizontal positions was measured under simultaneous or near-simultaneous presentation of an auditory nontarget (distractor). The horizontal position of the auditory signal was varied, using a virtual auditory environment setup. Mean SRT to a visual target increased with distance to the auditory nontarget and with delay of the onset of the auditory signal relative to the onset of the visual stimulus. A stochastic model is presented that distinguishes a peripheral processing stage with separate parallel activation by visual and auditory information from a central processing stage at which intersensory integration takes place. Two model versions differing with respect to the role of the auditory distractors are tested against the SRT data.

In a bimodal *focused attention* task, subjects are asked to make a speeded response to the onset of a stimulus (called the *imperative* stimulus) from one modality—for example, a visual stimulus—in the presence of a stimulus from another modality (e.g., an auditory stimulus), called the *accessory*. Since the early study by Todd (1912), the general finding for manual responses is that reaction time (RT) in bimodal stimulus trials is shorter, on average, than that in unimodal trials, in which only stimuli from the modality of the imperative stimulus are presented. Specifically, for medium-level intensities, manual RT to a visual imperative stimulus is 20–80 msec shorter if it is paired with an accessory auditory stimulus. This *intersensory facilitation effect* (IFE) can be observed under a wide variety of experimental conditions and is commonly attributed to some kind of multisensory integration (see Welch & Warren, 1986, for a review). In this paper, a quantitative model for the IFE is proposed in the context of saccadic responses, and its qualitative and quantitative properties are tested against saccadic RTs (SRTs) to visual–auditory stimuli at varying spatial positions and with different stimulus onset asynchronies (SOAs). Although this two-stage model makes predictions only at the behavioral level, it is compatible with an interpretation in physiological terms in which the first stage represents peripheral sensory processing and the second stage comprises neural multisensory convergence processes as found in the deep layers of the superior colliculus (SC) and in other, cortical areas (see Meredith & Stein, 1986).

THE INTERSENSORY FACILITATION EFFECT

Traditionally, IFEs have been studied for manual RTs only. Accordingly, the early approaches to the IFE that will be sketched here are based primarily on results from simple or choice RT paradigms. More recently, however, the IFE has also been observed in studies measuring SRTs. Even though a systematic investigation of potential dissimilarities between saccadic and manual intersensory interaction effects seems worthwhile, this is outside the objectives of this study. Rather, the model proposed here is geared toward the description of saccadic responses. It is not unlikely, however, that the basic structure of the model could be shown to apply to both manual and saccadic responses.

Theoretical Approaches to the Intersensory Facilitation Effect

To our knowledge, the literature provides no “stringent testable processing models” to account for the IFE (see Giray & Ulrich, 1993, p. 1287). Nonetheless, a number of different theoretical explanations for the facilitating effect of an irrelevant stimulus (i.e., the accessory) were advanced early on. According to the *arousal hypothesis* (Bertelson & Tisseyre, 1969), the accessory stimulus causes an unspecific activation of the motor system that does not trigger execution but can still speed up the response (e.g., Posner, Nissen, & Klein, 1976; Sanders, 1983). In a similar vein, *energy integration* or *summation* was suggested as an explanation for the IFE early on (e.g., Bernstein, Rose, & Ashe, 1970). It assumes that stimulus energy of both the visual and the auditory stimuli is somehow combined in bimodal trials so that the effective stimulus intensity of the imperative stimulus is increased, leading to shorter RTs. In his review, Nickerson (1973) forcefully argued that energy summation falls short of explaining the entire set of observations. In particular,

This research was supported by a grant from Deutsche Forschungsgemeinschaft to H.C. (SFB 517/C3, Neurokognition). We are grateful to Howard C. Hughes and two anonymous reviewers for helpful comments and suggestions. Correspondence concerning this article should be addressed to H. Colonius or P. Arndt, Institut für Kognitionsforschung, Universität Oldenburg, FB 5, D-26111 Oldenburg, Germany (e-mail: hans.colonius@uni-oldenburg.de or petra.arndt@uni-oldenburg.de).

evidence against total stimulus energy as the sole determinant of bimodal RT comes from early experiments in which the spatial configuration of the stimuli was varied (Bernstein & Edelstein, 1971; Simon & Craft, 1970). In these studies, RT was shorter when the accessory stimulus and the imperative stimulus were presented on the same side (left/right) than when they were presented on opposite sides.

Statistical facilitation has also been discussed as a possible source of the IFE. According to *separate activation* models (Raab, 1962), each signal is processed within its channel, and a response is initiated as soon as an activation level is exceeded in either channel. Since the mean of two random processing times is smaller than (or equal to) the smaller of the two unimodal means, the RT reduction in bimodal trials is thus explained by a statistical facilitation, or probability summation, effect. Since in many of the early studies, there were no catch trials (on which the accessory is presented alone and the subject is to withhold his or her response), the subjects were in fact free to respond to either the accessory or the imperative stimulus. However, in their study without catch trials, Gielen, Schmidt, and van den Heuvel (1983) could show that the RT speed-up for visual targets (with auditory or kinesthetic accessory) was larger than that predicted by statistical facilitation.¹ Although it is difficult to rule out the possibility that at least part of the RT speed-up may be due to some statistical facilitation mechanism, experimental paradigms in which the position of the auditory accessory is completely independent of the position of the visual target make the presence of statistical facilitation rather implausible. As was observed in Frens, Van Opstal, and Van der Willigen (1995), since the auditory cue is irrelevant for the correct performance of the task, the subjects cannot directly benefit from simply responding to the onset of the auditory stimulus. In addition, recent studies (Harrington & Peck, 1998; Hughes, Nelson, & Aronchick, 1998; Nozawa, Reuter-Lorenz, & Hughes, 1994) using the so-called race-model inequality test come to the conclusion that the saccadic RT speed-up in visual-auditory trials exceeds levels predictable by probabilistic summation between unimodal pathways (for a further scrutiny of these results, see below).

A promising line for modeling the IFE arises from the concept of *preparation enhancement* suggested in Nickerson (1973). Here, the accessory stimulus is assumed to provide an *alerting* role on one or more of the stages of processing by decreasing temporal stimulus uncertainty, so that the affected stages are terminated more quickly. In analogy to Raab's (1962) separate activation model, the preparation enhancement model assumes independent central arrival times for the visual and the auditory signals. According to Nickerson, these arrival times include "both afferent transmission time and the central processing time that is required for the nervous system to do whatever it does before a response is evoked. It excludes the time involved in response execution" (p. 501). Thus, whereas in the separate activation model the first stimu-

lus to arrive evokes the response, here "the first stimulus to arrive may simply increase the subject's preparedness to respond to, or to process further, the second" (Nickerson, 1973, p. 501). It should be noted here that this approach takes for granted the assumption of these subprocesses' being organized in sequential, additive stages. In fact, an experimental study using Sternberg's additive-factors logic (Sternberg, 1969) led Schmidt, Gielen, and van den Heuvel (1984) to conclude that the site of the effect of the accessory—that is, the preparation enhancement—was the late response-programming stage. More recent studies using lateralized readiness potentials (LRPs) suggest that an early phase of response selection is influenced by the accessory stimulation (Hackley & Valle-Inclán, 1998, 1999).

According to the preparation enhancement model, the effectiveness of the auditory accessory signal as an alerting cue varies randomly from one trial to the next, since it depends on the auditory arrival time's finishing before the visual arrival time. However, its average effectiveness will depend on the shape and the relative location of the two arrival time distributions. In particular, any experimental manipulation that affects the probability that the auditory arrival time will finish before the visual arrival time will also influence the amount of the IFE. As was noted by Nickerson (1973), the most obvious way to change this probability is to vary the SOA between the imperative and the accessory signals. For example, the more the accessory signal is delayed with respect to the presentation of the imperative signal, the smaller is its chance to modify the subsequent processing of the imperative signal and, thus, to produce the IFE.

The effects of the spatial (or directional) disparity between the imperative and the accessory are more difficult to reconcile with the preparation enhancement idea. Why should there be a spatial effect if the role of the accessory is simply that of an alerting cue? In the model to be outlined in the next section, this issue will be taken up again in a framework that, while incorporating important aspects of the preparation enhancement model, extends it in a way that seems compatible with recent evidence from neurophysiological studies.

Modeling the Generation of Saccadic Response Times

Several models for the initiation of visually guided saccades have been proposed since the late 1970s, but only a few have a direct bearing on the study of visual-auditory interaction. A prominent feature of most models is the idea of two separate pathways concerned, respectively, with the spatial and the temporal programming of the saccades (e.g., Becker & Jürgens, 1979; Findlay, 1983). In their comprehensive presentation of this framework, Findlay and Walker (1999) described a hierarchy of processing levels through which information and command streams run in parallel in the *where*- and the *when*-pathways. An important aspect of the *where*-pathway is the involvement of spatially distributed coding and the

selection of the saccade target from a *saliency map*, where the point of maximum activity in the map determines the metric of the saccade. Although not explicitly mentioned in that paper, the concept of a saliency map can obviously be extended to include both visual and auditory spatial information (Colnius & Arndt, 1999; Kopecz, 1995). In terms of the Findlay–Walker framework, the exact time point of the initiation of a saccade is determined by the resolution of a conflict between a *fixate center* in the *when*-pathway and a *move center* in the *where*-pathway. On the basis of electrical stimulation studies in the deep layers of the SC, Van Opstal and Van Gisbergen (1989) formulated one of the most detailed models of saccade metrics involving nonlinear collicular spatial interaction mechanisms. Although neurophysiological studies by Munoz, Wurtz, and collaborators (e.g., Munoz & Wurtz, 1993a, 1993b) have shown that fixation cells in the monkey SC play an important role in the operation of the fixate system, the activity of frontal eye field neurons also highly correlates with saccadic latencies (Hanes & Schall, 1996).

At the behavioral level, Carpenter and Williams (1995) proposed a simple but efficient model for visual saccade initiation. At stimulus onset, a decision signal rises with a constant rate r , say, and upon exceeding a fixed threshold, a saccade is initiated. To account for the variability of saccadic responses, the increase rate r is drawn from a Gaussian distribution in each trial. From this it follows that the *reciprocal* latency follows a Gaussian distribution as well. Although no severe empirical violations of the model have been reported, it should be noted that the model makes the implausible prediction that latencies of infinite duration occur with nonzero probability.²

Before the two-stage model proposed in this paper is presented in detail, we will report data from Experiment 1, which was performed to find out whether intersensory facilitation effects would occur under our specific setup, in accordance with previous studies. First, the general setup employed in both experiments is laid out. In particular, the use of the virtual auditory environment will be described in detail, and a study of its effectiveness will be reported.

GENERAL METHOD

Apparatus

The subjects were seated in a sound-attenuated chamber ($1.0 \times 1.2 \times 1.9$ m), with the head fixed by a bite bar. The visual stimuli were presented on a 37-in. monitor (XP37, NEC), which could be seen through a window (74.5×59 cm) in front of the subject. The monitor was placed directly outside of the window, resulting in a viewing distance of 57 cm. Presentation of visual and auditory stimuli was controlled by a personal computer (PC). A second PC was employed for data acquisition. The temporal arrangement of stimulus presentation and data acquisition were synchronized, with the exact presentation time of the visual stimulus being determined by the monitor update rate.

Visual Stimuli

White dots with a diameter of 0.1° served as stimuli. They were presented with a luminance of 70 cd/m^2 against a dark background

(less than 0.01 cd/m^2) at eccentricities of 15° or 25° to the left or to the right of the fixation point (Figure 1). The monitor was driven with a vertical frequency of 75000 Hz. Monitor update rate and vertical and horizontal movement time of the cathode rays to different positions on the screen were taken into account in the calculation of the onset and offset times of the visual stimuli.

Auditory Stimuli

Auditory stimuli were generated prior to the experiments on the basis of a noise signal with a spectrum equivalent to the spectrum of human speech (bandwidth 500–14000 Hz). The noise signal was convolved with head-related transfer functions (HRTF) of a dummy head. This method was suitable because all the auditory stimuli employed had zero elevation. The auditory direction discrimination in the horizontal plane is determined by time and level differences between sounds detected by the left ear and the right ear. These time and level differences are easily produced by a convolution of an auditory signal with the HRTFs of a dummy head (Middlebrooks & Green, 1991). In contrast to judgments of the elevation of auditory stimuli, for which the individual shape of the outer ear produces individual direction cues, the direction discrimination in the horizontal plane does not depend on differences between individual HRTFs. Since all the auditory stimuli had zero elevation, the application of the HRTFs of a dummy head led to a good approximation of the exact auditory position. This effect was supported by the choice of a noise signal with a strong representation of low frequencies and a continuous decay of energy between 6 and 14 kHz. The low-frequency portion of the auditory signal is the main carrier for the interaural time difference cue. Signals for five different stimulus positions were produced (straight ahead 0°), 15° and 30° to the left, and 15° and 30° to the right), digitized and stored for display during the experiments. For stimulus presentation, the signals were played back by a high-precision sound card (Tahiti, Turtle Beach), amplified and presented via headphones (Sennheiser HD 580) with an intensity of 76 dB SPL.

A direct comparison of free-field (loudspeaker) and nonindividual virtual acoustics was carried out by Wenzel, Arruda, Kistler, and Wightman (1993), using oral reports of the perceived stimulus positions. They showed that the main errors in nonindividualized virtual acoustics are front–back and up–down confusions: “With confusions resolved, there were high correlations between the judged and actual source locations and a close correspondence between free-field and virtual-source conditions. When performance degradations did occur with virtual stimuli, they were in the dimension of elevation” (Wenzel et al., 1993). Although front–back confusions cannot occur with the stimuli in the horizontal plane used in our experiments, up–down confusions might occur. However, in tests with 3 of the subjects (M.R., R.V., and K.S.), performed after completion of the experiments reported below, no up–down confusions were found.

To demonstrate the localizability of the auditory stimuli, a control study (Heuermann & Colnius, 1999) with 4 additional subjects was carried out. For the stimulus positions used in the main experiment, mean eye positions for free-field stimuli, as compared with stimuli presented over virtual acoustics, were very much alike: For stimuli at 30° , eccentricity was 32.3° in the free-field condition and 30.3° in the virtual acoustics condition, and for stimulus eccentricity of 15° , mean eye position was 21.1° in the free-field condition and 15.3° in virtual acoustics condition. Mean absolute angle of error in the azimuthal plane was 6.7° for free-field stimuli and 7.4° for virtual acoustics stimuli.

Response Recording

Eye movements were registered with an infrared light reflecting system (IRIS, Skalar Medicals) providing an analog signal of the eye position, with a spatial resolution of up to 2 min of arc. The signal is linear in the range between -25° and $+25^\circ$. Eye position data were sampled with a rate of 1 kHz and stored in a PC for subse-

quent analysis. Owing to the calibration procedure and the digitalization of the signal, we achieved a spatial accuracy of up to 12 min of arc. Eye position was controlled on line during the complete session. Data recording was carried out on a trial-by-trial basis. The data acquisition PC was triggered by the stimulus PC. For each trial, 1,500 msec were recorded.

Procedure

Each trial started with the presentation of a fixation point. After a truncated exponentially distributed random time interval (with a minimum of 800 msec and a maximum of 2,500 msec), the fixation point disappeared, and simultaneously, the visual target was presented either alone (unimodal condition) or together with the auditory nontarget (bimodal condition). Stimulus duration was 500 msec for both stimuli. Trials were separated by time periods of 1,500 msec, during which neither a fixation point nor a visual target was presented. The random fixation time was applied in order to prevent the subjects from preparing the temporal aspects of saccade initiation in advance on the basis of expected stimulus presentation time. Such preparation processes could interfere with changes in processing times elicited by bimodal stimulation.

Response Detection

Saccades were detected automatically from the calibrated eye position signals. Saccadic onset and offset were defined by velocity and acceleration criteria. The eye position data of each trial were checked visually for proper fixation at the beginning of the trial, for blinks, and for the correct detection of start and endpoint of the detected saccade. Trials with improper fixation or with blinks during fixation or stimulus presentation were excluded from further analysis. If necessary, onset and end of the saccade were marked manually. RTs (defined as the time between the onset of the visual stimulus and the onset of the saccadic eye movement, in milliseconds), 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. RTs of less than 100 msec or more than 500 msec were discarded as anticipations and misses, respectively. For the bimodal stimulus condition, only primary saccades directed toward the respective visual target were included in further analysis. We utilized the mean postsaccadic eye position in the unimodal condition for defining the bimodal threshold criterion. That is, the postsaccadic end position of the eye in bimodal trials had to be within an area of 4° around the mean end position for unimodal trials.

EXPERIMENT 1

In the first experiment, we investigated how the spatial arrangement of the stimuli affected the visual–auditory interaction in saccadic eye movements. Visual stimuli were presented on a monitor, whereas auditory stimuli were presented in a virtual auditory environment setup (see the General Method section above). The main purpose was to ensure that existing findings concerning the IFE could be confirmed with our setup, using a virtual auditory environment.

Method

Subjects. Four female subjects (ages 20, 22, 24, and 25 years) and three male subjects (ages 22, 39, and 42 years) participated in the experiment. All had normal auditory and oculomotor function and normal or corrected-to-normal vision. The subjects were naive with respect to the specific goals of the study; they were paid for their participation.

Design/stimulus arrangement. The auditory stimuli were presented as described in the General Method section. Four different horizontal positions were chosen for the presentation of the visual stimuli: 15° and 25° to the left or right of the central fixation point. Thus, visual as well as auditory stimuli were within 30° from the fixation point. Each visual stimulus was combined with each spatial position of the auditory stimuli (Figure 1, panel a). The combination of only partly coincident visual and auditory stimuli resulted in 20 bimodal stimulus combinations, with nine different spatial distances between visual and auditory stimulus positions. A large number of different spatial distances was desirable, given that the dependence of intersensory interaction on the spatial distance between the auditory and the visual stimulus was to be investigated.

With the head fixed, saccadic eye movements with an amplitude of up to 30° usually show an increase of both peak velocity and duration, whereas for larger movement amplitudes, velocity reaches an upper limit, so that further growth is accompanied by a large increase in movement duration (Becker, 1991). Saccadic eye movements beyond 30° may be considered nonphysiological under head-fixed conditions and are thus not studied in this experiment.

Stimulus onset times in bimodal trials were in temporal alignment (Figure 1, panel b). The bimodal stimulus combinations, together with the four unimodal visual stimuli, were presented four times in 1 block in pseudorandom order. The presentation of 1 block took about 8 min, and 3 to 4 blocks were presented in one session, with 2-min breaks between blocks. The subjects were free to ask for longer breaks or to end the session if they felt fatigue or discomfort. Usually, only one session was performed per day. Ten to 14 blocks were measured per subject.

Results

Mean saccadic reaction times. In Figure 2, mean saccadic latencies (averaged over all subjects) are plotted separately for each of the four visual positions as a function of the position of the auditory stimulus. Mean unimodal latencies for each of the visual positions are indicated by the rightmost points.

For all the spatial positions, latencies were clearly reduced under bimodal stimulation, as compared with unimodal visual stimulation, and this IFE held for each subject separately ($p < .001$). The IFE (defined quantitatively as the mean unimodal visual latency minus the mean bimodal latency) varied with the spatial distance between the target and the accessory auditory stimulus position. Specifically, for large distances between stimuli—that is, auditory and visual stimuli in different hemifields and an interstimulus distance of 45° or 55° of visual angle—IFEs of 4.8–41.2 msec were found. IFE for neighboring stimuli (5° distance) and for coincident stimuli varied between 19 and 63.3 msec. To test for the specifics of the bimodal interaction, a three-way between-subjects analysis of variance (ANOVA) was performed, with repeated measures on target position (four levels) and auditory position (five levels) as within-subjects factors. There were significant main effects for the within-subjects factors of target position [$F(3,645) = 38.02, p < .001$] and auditory position [$F(4,860) = 34.84, p < .001$]. The target position \times subject and the auditory position \times subject two-way interactions were also significant [$F(18,645) = 7.03, p < .001$, and $F(24,860) = 2.043, p = .002$]. For the two-way target position \times auditory position interaction, we

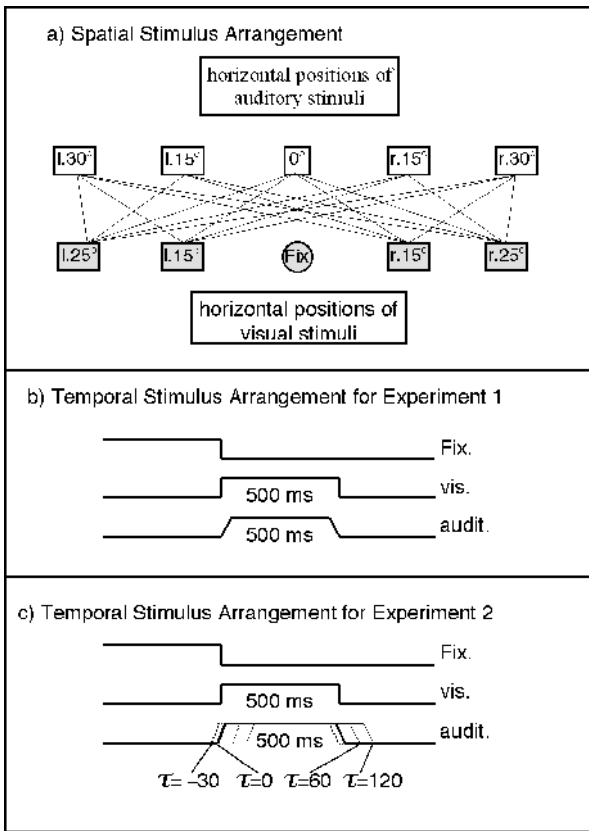


Figure 1. (a) Spatial stimulus arrangement: Visual and auditory stimuli were located along the horizontal meridian. Horizontal separation differed between 0° and 55° of visual angle. (b) Temporal stimulus arrangement for Experiment 1: A step paradigm was used, with the visual target and the auditory accessory being presented immediately after the fixation point was extinguished. (c) Temporal stimulus arrangement for Experiment 2: The temporal arrangement of visual stimulation was maintained, whereas the presentation time of the auditory stimulus was varied between -30 msec (auditory first) and 120 msec (visual stimulus first).

found that $F(12,2580) = 1.726, p = .058$, whereas no significant effect was found for the three-way target position \times distractor position \times subject interaction.

Saccade amplitudes. In bimodal trials, 94%–99% of the saccades met the criterion for correctly directed saccades (see the General Method section) and were included in the RT analysis. No influence of the accessory auditory stimulus on amplitude was found for any spatial combination of visual and auditory stimuli. In bimodal trials, at most 1.7% of the saccades were directed toward the auditory stimulus. Here, the direction of the eye movement at the onset of the saccade (to the left or to the right) was used as a criterion when the visual target and the auditory stimulus were located in opposite hemifields. For bimodal trials with both visual and auditory stimuli in the same hemifield, the saccade was judged to be directed to the auditory stimulus if (1) the end position did not meet

the criterion for correctly directed saccades and (2) the amplitude was either too small to meet the criterion in trials with a low eccentricity of the auditory stimulus (15°) or too large in trials with a large eccentricity of the auditory stimulus (30°). In bimodal trials, only 0.1%–4% of the saccades were executed to directions in which no stimulus was presented at all. A detailed analysis for each subject is given in Table 1.

There was no evidence for a center-of-gravity effect on saccadic amplitudes, as is typically found for the presentation of two visual targets (e.g., Findlay, 1982; Walker, Deubel, Schneider, & Findlay, 1997)—that is, saccades did not tend to be directed to the center of gravity of the target configuration. The end positions of the primary saccades remained unchanged for all bimodal stimulus combinations and did not differ significantly from those to unimodal visual targets, as can be seen from Figure 3.

Discussion

It has been shown that an accessory auditory stimulus presented over a virtual auditory environment significantly reduced mean saccadic latencies toward visual targets for interstimulus distances of up to 55° . Only saccades obviously directed toward the visual target were included in our data analysis. Therefore, this effect cannot be attributed to the incorporation of purely auditory saccades with, possibly, shorter latencies in the calculation of the mean SRTs.³

The general reduction in latency for bimodal versus unimodal stimuli is consistent with previous free-field studies of the IFE, involving both simple manual responses and saccades (e.g., Engelken & Stevens, 1989; Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994). Moreover, the dependency of the size of this reduction on the spatial distance between the auditory and the visual stimuli is similar to a result in Frens et al. (1995) in human subjects and to that in a more recent study by the same authors (Frens & van Opstal, 1998) in the monkey. Indirectly, our results vindicate the use of a virtual auditory environment. Note, however, that whether or not a virtual environment is being used, it cannot be deduced from our data that the spatial effect is indeed an effect of geometric distance per se. It is possible, for example, that the IFE is some decreasing function of the difference in the *directions* of the visual and the auditory signal sources and that the eccentricity of the stimuli also has some influence on the effect.

Results from Lueck, Crawford, Savage, and Kennard (1990) support this complex pattern of the influence of interstimulus arrangement. They measured the amplitudes of horizontal saccadic eye movements toward *auditory targets* combined with accessory *visual nontargets*. Saccadic amplitudes changed systematically with the position of the visual nontarget if auditory and visual stimuli were presented in the same hemifield. However, no effect on saccadic amplitude was found for stimuli presented on opposite sides of the central fixation point. The fact that neither in our data nor in the results from other

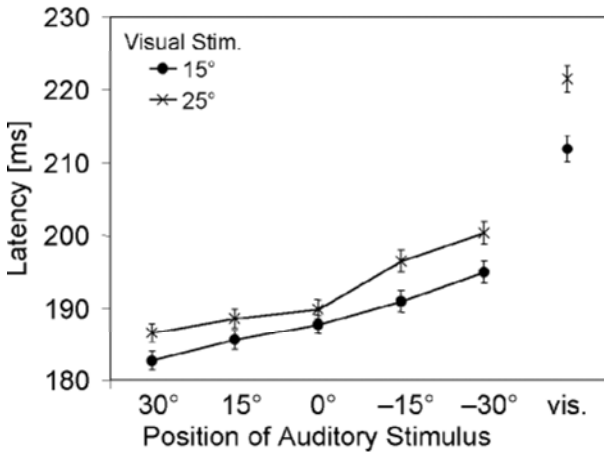


Figure 2. Mean latencies (with standard errors) for different spatial stimulus conditions averaged over all subjects. Data are plotted separately for the two eccentricities of the visual target stimuli. The position of the auditory stimulus is indicated on the abscissa, with positive values referring to an auditory position ipsilateral to the visual stimulus and negative values referring to a position contralateral to the visual stimulus. The symbols on the right represent the mean reaction time to purely visual stimulation.

groups (Engelken & Stevens, 1989) have modifications of saccadic amplitude under bimodal visual–auditory stimulation with visual targets, as compared with unimodal visual stimulation, been found suggests an asymmetric effect of visual versus auditory stimuli on the generation of saccades. This asymmetry is reflected in the SRTs as well. Whereas Lueck et al. found a pronounced reduction of SRTs for stimuli in opposite hemifields, we did not observe inhibitory effects by the accessory auditory stimulus in any condition (see Corneil & Munoz, 1996). Asymmetric effects of visual–auditory stimuli in eye movement programming are in line with experiments on explicit, conscious localization judgments (Pick & Warren, 1969)—that is, conscious perception and the programming of saccadic eye movements are in good correspondence under bimodal stimulation.

Our data support the hypothesis of a separate programming of the decision to initiate a saccade (*when*) and of the spatial parameters (*where*) of the saccade (Becker & Jürgens, 1979; Findlay & Walker, 1999): Accessory auditory stimuli changed the SRT (i.e., the onset of sac-

cade initiation), but not the amplitude. In a similar vein, Frens et al. (1995) did not find any influence of auditory stimuli on saccadic trajectories for stimuli with intensities in the range used in our experiments. In their data, an influence of the accessory auditory stimulus on trajectories (i.e., the *where* pathway) appeared at low-intensity stimuli, however. To sum up, accessory auditory stimulation at the intensity levels used in our experiment seems to be involved in the process of saccade initiation, but not in the computation of the spatial characteristics of saccadic eye movements.

TWO-STAGE-MODEL

In this section, a stochastic model for the mechanism of saccade initiation in the focused attention situation is proposed. It is based on the simple but fundamental assumption that given the initial separation of the peripheral pathways, the entire processing time must consist of at least two stages arranged in series.⁴ This two-stage model distinguishes an early stage of processing, at which the stimuli trigger peripheral visual and/or auditory sensory processes, from a central stage of processing, at which stimulus information from different modalities is integrated and an ocular motor response is prepared. The duration of each of the (sub)stages is considered a random variable. Although the model postulates that saccadic onset time depends on the specific spatial configuration of visual and auditory stimuli, it does not specify this relationship in any detail. Thus, the primary purpose of the model at this stage of development is to capture the dynamics of the *when*-pathway under co-occurrence of both visual and auditory stimulation.

Basic Assumptions

First, the basic features of the model are outlined. Later, the assumptions will be specified in more and more detail. Of course, the higher the level of specification, the more—more or less arbitrary—choices have to be made. On the other hand, more specific assumptions permit a more rigorous empirical test of the model. For each level of specification, the predictions from the model will be discussed. Finally, quantitative predictions of the model at the level of the means and standard deviations of the SRTs are tested in Experiment 2, to be reported below.

Table 1
Percentages of Correct and Misdirected Saccades for Individual Subjects

Saccades	Stimulus	A.M.	A.W.	B.G.	K.S.	M.R.	O.V.	R.V.
Correctly directed	unimodal	99	100	95	99	98	99	100
	bimodal	99	94	96	99	98	99	99
Anticipatory	unimodal	—	—	1	—	—	—	—
	bimodal	—	1	1	—	—	—	—
Toward auditory stimulus		—	1	2	1	1	1	—
Toward no stimulus	unimodal	1	—	4	1	2	1	—
	bimodal	1	4	1	4	—	—	1

Note—All saccades with latencies of up to 500 msec, with proper fixation, and without blinks were analyzed. Saccades with reaction times under 100 msec were considered anticipatory.

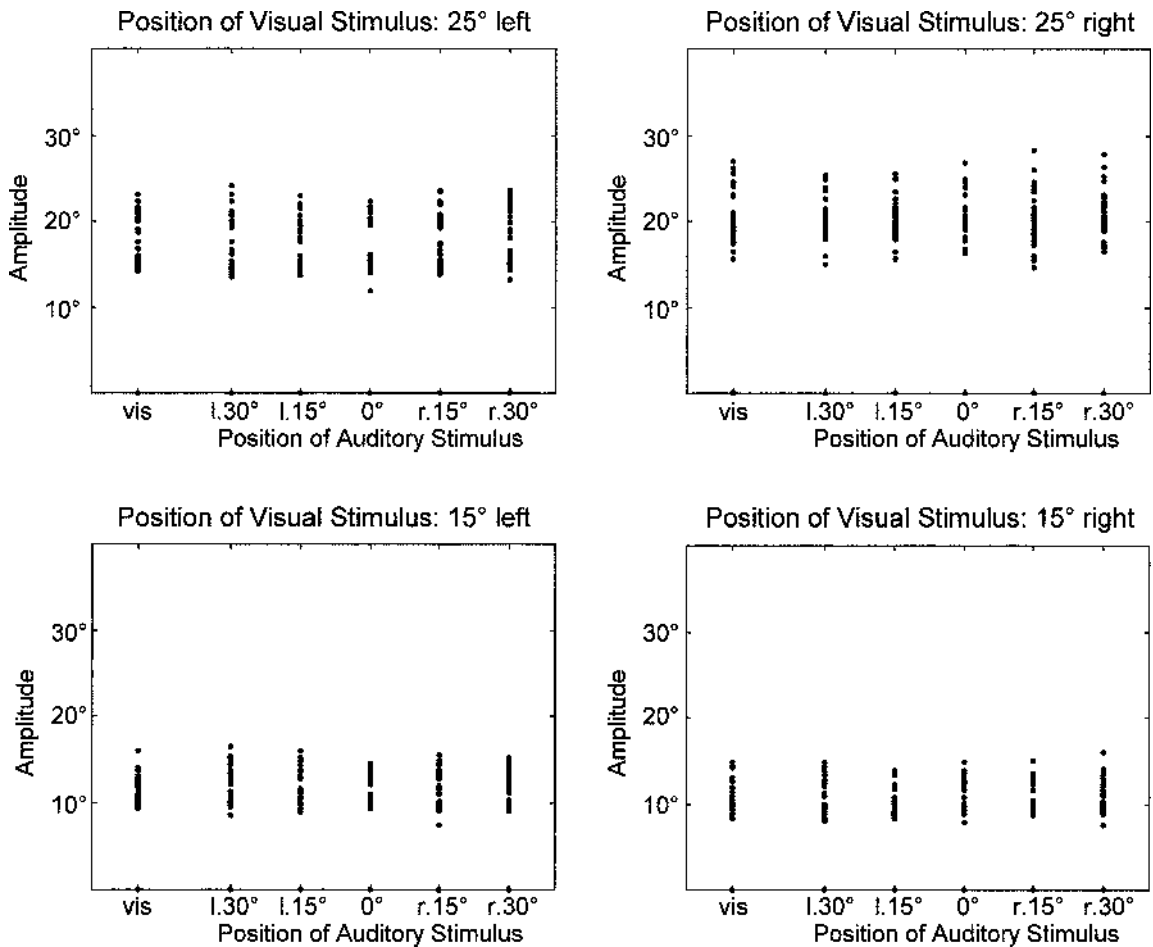


Figure 3. Amplitudes of primary saccades for 1 subject. Each diagram refers to one position of the visual stimulus. Each dot indicates the horizontal amplitude of an individual primary saccade. The points on the left in each diagram represent amplitudes of saccadic eye movements toward unimodal visual stimuli. For bimodal stimulus conditions, the position of the auditory signal is given on the abscissa; l./r. indicates an auditory stimulus left or right of the fixation point, respectively.

Assumption 1: Independent peripheral processes.

Any presentation of a bimodal signal triggers independent peripheral neural processes in both sensory pathways.

Although not strictly necessary for developing the model, it seems plausible that the variables representing the time lengths of these processes do not interact with each other in a statistical sense at this early stage. Although these trigger processes are analogous to those in the preparation enhancement model by Nickerson (1973), the interpretation of these processes is different. The *arrival times* in the preparation enhancement model comprise all stages except response evocation. In the two-stage model suggested here, the peripheral processes only represent mechanisms involved in detecting the presence of the signals, visual or auditory.

Assumption 2: The termination rule. Two different termination rules for the first stage of processing will be considered: (1) (*MIN-version*) the first stage is terminated by the first of the two peripheral processes to arrive

at a hypothetical decision center that registers the detection of the imperative and the accessory stimuli, and (2) (*V-version*) the first stage is terminated only at the time the visual peripheral process is registered at the decision center. As will be discussed shortly, the difference between these two versions is quite fundamental, and testing between them is one of the goals of the experiment below.

Assumption 3: The interaction rule. Visual-auditory interaction occurs only if the auditory accessory stimulus is the winner of the race between the two peripheral processes.

This rather strong assumption deserves some comment. First, the assumption is stated for a focused attention situation with the visual stimulus as target and the auditory stimulus as accessory, and it will have to be modified in an obvious way when their roles are interchanged. Second, for the accessory to win the race is a necessary, but not a sufficient, condition for the occurrence of interaction. The spatiotemporal configuration may be such that no interaction takes place even if the au-

ditory has been detected first. For example, the distance between the visual and the auditory stimuli in time or in space may be too large, so that they come to lie outside of a *window of integration*. Finally, the outcome of the race is not observable. However, there is always a nonzero probability that the accessory will win the race, as long as the probability distribution for the target peripheral process does not lie entirely below (i.e., to the left of) the probability distribution of the accessory.⁵

Assumption 4: Intersensory interaction in the second stage. The interaction comes about as a speed-up (facilitation) or prolongation (inhibition) of the second-stage processing duration. The amount of interaction (neural summation) only depends on the visual–auditory configuration. The functional form of this dependence is left unspecified at this point.

One important feature of this model is that the spatial configuration between the stimuli does not affect the duration of the first stage, but only that of a subsequent—postdetection—stage at which some kind of neural summation takes place and the saccadic eye movement toward the visual target is being preprogrammed. This does not mean that the location or the intensity of the stimuli has no influence on the outcome of the first stage. For example, it is known that RT to visual targets and to auditory targets varies with the eccentricity of the stimuli in specific ways (Yao & Peck, 1997). However, the model assigns the specific effects of spatial coincidence or disparity of the bimodal stimulus configuration to a later processing stage.

If the auditory accessory is detected prior to the visual target, it is assumed in the MIN-version that it triggers some early part of central processing—that is, it serves “to prepare” the saccadic response while the target location of the saccade is not yet amenable to the system. At the neurophysiological level, this can be interpreted as the detection of the auditory signal to start inhibiting the activation of the fixation neurons in the SC and/or that of the omnipause neurons in the brainstem (see Wurtz, Basso, Paré, & Sommer, 2000). Thus, in the MIN-version, the auditory accessory acts both as a nonspecific warning signal for the onset of the visual target stimulus and as a component in the bimodal convergence taking place in the later stage.

Although the absence of any catch trials⁶—which is presupposed here—should make the detection of the accessory a completely valid cue for the occurrence of a visual target, it cannot be ruled out a priori that the beginning of the second stage always depends on the visual target to be detected and/or its location to be identified. This possibility was referred to as the *V-version* of the model in Assumption 2. Note that the two versions make different predictions about how the IFE is generated. Although both versions attribute facilitatory or inhibitory effects to a neural summation mechanism in the central stage of processing, the MIN-version allows for an additional effect of statistical facilitation. As was discussed

in the Theoretical Approaches to the Intersensory Facilitation Effect section, statistical facilitation has been found not to be sufficient as the sole explanation for the observed intersensory effects (e.g., Hughes et al., 1998). However, the MIN-version of the two-stage model attributes only part of the entire amount of facilitation to statistical facilitation. Moreover, as will be specified below in more detail, since statistical facilitation acts only in the direction of speeding up RT no matter what the specific characteristics of the race are (see Colonius, 1990), in the MIN-version, but not in the V-version, a possible inhibitory IFE can partly be offset by statistical facilitation.

Formal Description of the Two-Stage Model

In order to derive testable predictions from the model, some more formal description of the underlying assumptions is needed. First, no specific assumptions concerning the parametric form of the distributions for the various processing times will be made. It turns out that, nonetheless, certain empirical predictions follow from the model that are, in principle, amenable to testing. Later, specific distributional assumptions will be added that allow a quantitative fit of the model. The following notation will be used:

V	visual peripheral processing time,
A	auditory peripheral processing time,
W_1	first stage (peripheral) processing time,
W_2	central processing time,
τ	SOA,
RT_V	observed SRT when only the visual stimulus is presented,
$RT_{VA,\tau}$	observed SRT to the visual stimulus with the auditory accessory presented at SOA τ (in milliseconds),
I	the event $\{A + \tau < V\}$,
π	the probability of event I .

Taking V , A , W_1 , and W_2 to be (nonnegative) random variables with finite means and variances, the observed SRT to a visual stimulus with auditory accessory presented with $SOA = \tau$ ($-\infty < \tau < \infty$) is assumed to be

$$RT_{VA,\tau} = W_1 + W_2, \quad (1)$$

with, for the MIN-version,

$$W_1 = \min(V, A + \tau),$$

and for the V-version,

$$W_1 = V.$$

The simple additive form of the model equation (Equation 1) is somewhat deceptive. Whereas according to Assumption 1 above, the peripheral processing times, V and A , are stochastically independent, the stage processing

times, W_1 and W_2 , are not independent, because W_2 depends on the outcome of the race in the first stage—that is, I or its complementary event. Technically, W_1 and W_2 are *conditionally independent, given event I*. This is another way of saying that all (stochastic) dependence between W_1 and W_2 disappears, given that the outcome of the race is known. Moreover, the second stage, W_2 , includes a number of substages, such as motor preparation and execution, that will be given specific consideration at this level of modeling.

A complete description of the model requires specification of how the random variables in Equation 1 depend on a manipulation of various experimental factors. An inherent part of the two-stage assumption is that all of the stimulus parameters that determine their detectability, such as auditory and visual intensity, eccentricity of the visual stimulus, and so forth, have a direct impact on the first stage of processing only. Nevertheless, since these factors have an effect on the outcome of the race in the first stage, they also indirectly influence the second stage.⁷ Finally, in order to derive a quantitative expression for the IFE, the unimodal visual SRT is also assumed to be decomposable into a peripheral and a more central component,

$$RT_V = V' + W'_2,$$

where V' is a random variable equal in distribution to V and W'_2 is equal in distribution to W_2 given I^c —that is, given that no (spatial) interaction occurs.

Distribution-Free Predictions

Although fitting the model to data quantitatively requires specification of the (joint) probability distribution for all the random variables involved, the structure of the model is such that a number of distribution-free predictions are possible. The advantage of such tests is, of course, that a failure of the model cannot simply be attributed to choosing inappropriate distribution types but, rather, would point to a more fundamental inadequacy of one or more of the model assumptions.

An important feature of the model that opens up a number of routes to empirical testing is that it predicts the observable SRT distribution to be a binary mixture of two distributions. These two component distributions are generated by the realization of either the event

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

or its complement. More specifically, by conditioning on I and its complement, the (cumulative) bimodal RT distribution can be written as

$$P[RT_{VA, \tau} \leq t] = P[W_1 + W_2 \leq t] = \pi P[W_1 + W_2 \leq t | I] + (1 - \pi) P[W_1 + W_2 \leq t | I^c]. \quad (2)$$

Although neither π nor the two conditional distributions in Equation 2 can be estimated directly from the data, mixture distributions have several distinctive properties that lead to empirically testable predictions. Many tests

have been developed (see, e.g., Yantis, Meyer, & Smith, 1991), but here only the following two will be discussed.

First, the mixture property is applied to the expected SRT:

$$\begin{aligned} E[RT_{VA, \tau}] &= E[W_1] + E[W_2] = E[W_1] + \pi E[W_2 | I] \\ &\quad + (1 - \pi) E[W_2 | I^c] = E[W_1] + E[W_2 | I^c] \\ &\quad - \pi (E[W_2 | I^c] - E[W_2 | I]). \end{aligned}$$

Writing $\Delta \equiv E[W_2 | I^c] - E[W_2 | I]$, the above becomes

$$E[RT_{VA, \tau}] = E[W_1] + E[W_2 | I^c] - \pi * \Delta, \quad (3)$$

where $\pi * \Delta$ can be interpreted as a measure of the expected SRT speed-up in the second stage. Specifically, positive Δ values correspond to (spatial) facilitation, negative ones to (spatial) inhibition. According to the model assumptions, Δ depends on the spatial arrangement of the auditory and the visual stimuli, but not on τ , whereas π depends on τ , but not on the spatial arrangement of the two stimuli. Thus, with respect to Experiment 1, varying the auditory stimulus position for a fixed visual target position should affect only the Δ value in the above expression.⁸ Specifically, if the SRT speed-up in the second stage (Δ) is a decreasing function of the distance between the visual and the auditory stimuli, observed mean SRTs should be increasing⁹ with that distance. This prediction has been borne out, although not perfectly, in the data presented in Figure 2 (see also the discussion in the Results section of Experiment 1, above).

Second, a statistical test for the presence of a mixture distribution developed in Yantis et al. (1991) that relies on the entire shape of the RT distribution is applicable to the bimodal SRTs. In this multinomial maximum-likelihood mixture (MMLM) analysis, each experimental condition yielding a set of (RT) observations is hypothesized as producing either a mixture distribution or one of the two basis distributions from which the mixtures are generated. The observations from each condition are grouped into bins and treated as coming from a family of multinomial distributions. The iterative algorithm then searches for maximum-likelihood estimates for the bin probabilities, both in the basis distributions and in the mixture distributions. The asymptotically chi-square-distributed G^2 goodness-of-fit statistic is calculated,¹⁰ providing a test of the null hypothesis that a mixture distribution is present in the data. In an earlier version of this paper, this MMLM analysis was applied to parts of the data of Experiment 1. Although no significant evidence against the presence of mixture distributions under bimodal conditions was found, estimates of the power of these tests were rather low, between .10 and .40, so that those mixture tests will not be considered further for these data.

An interesting distribution-free prediction distinguishing between the two model versions derives from the dependence (via π) of mean SRT on SOA. For the MIN-version, Equation 3 becomes

$$\begin{aligned} E[RT_{VA, \tau}] &= E[\min(V, A + \tau)] + E[W_2] \\ &= E[\min(V, A + \tau)] + E[W_2 | I^c] \\ &\quad - \pi * \{E[W_2 | I^c] - E[W_2 | I]\}. \end{aligned}$$

According to the model, Δ depends on the spatial arrangement of the auditory and the visual stimuli, but not on τ . It is revealing to write down the expression for the IFE emanating from this model version in the following way:

$$\begin{aligned} \text{IFE} &= E[RT_V] - E[RT_{VA, \tau}] = E[V'] + E[W_2'] \\ &\quad - \{E[\min(V, A + \tau)] + E[W_2]\} \\ &= \{E[V] - E[\min(V, A + \tau)]\} + \pi * \Delta. \end{aligned} \quad (4)$$

Thus, the IFE can be additively decomposed into a statistical facilitation effect occurring in the first stage (the difference in braces) plus the spatial interaction term. Furthermore, it is easy to see that $E[\min(V, A + \tau)]$ is increasing in τ and that π is decreasing in τ . This implies the following:

PREDICTION 1: *In the MIN-version of the model, given that facilitation occurs in the second processing stage ($\Delta > 0$), the SOA function is increasing; or, equivalently, the IFE is decreasing in τ .*

On the other hand, inhibition in the second stage does not necessarily imply a monotonic IFE function. To see this, an inspection of Equation 4 reveals that, with τ increasing, the first term decreases, whereas the last term, $\pi * \Delta$, increases for negative Δ , implying a possible non-monotonicity of the IFE function. Whether or not a non-monotonicity occurs depends on the relative size of the two additive components of the IFE, and further predictions require more specific parametric assumptions (see below).

For the V-version of the model, predictions turn out to be simpler. The SOA function for the V-version,

$$E[RT_{VA, \tau}] = E[V] + E[W_2],$$

immediately yields

$$\text{IFE} = E[RT_V] - E[RT_{VA, \tau}] = \pi * \Delta. \quad (5)$$

Note that now Δ and the IFE must have the same sign, implying the following:

PREDICTION 2: *In the V-version of the model, the SOA function is either increasing (in the case of facilitation) or decreasing (in the case of inhibition); or, equivalently, the IFE is either decreasing or increasing in τ , respectively.*

In other words: *nonmonotone SOA (or IFE) functions rule out the V-version, but not the MIN-version.* The following experiment was set up to allow the computation of the SOA function for various onset asynchronies.

A model close in spirit to the one considered here was proposed by Nozawa et al. (1994), but their parallel-serial model differs in a number of important aspects. Both models consider serially arranged stages with parallel visual-auditory peripheral pathways. However, the activities in the first stage of that model combine into a neural

summation process, rather than into an all-or-none race, as is suggested here. The choice of neural summation over the race concept was based on finding consistent violations of the following inequality first proposed by Miller (1982; cf. Colonius, 1990; Diederich, 1992):

$$P[\min(V, A) \leq t] \leq P[V \leq t] + P[A \leq t], \quad (6)$$

where the right-hand side of Equation 6 presents an upper bound for the distribution function under bimodal stimulation compatible with a race between visual and auditory processing. Note that for computing this upper bound, the distribution functions for both visual and auditory unimodal responses must be obtained. In the divided attention paradigm employed by Nozawa et al., subjects are free to respond to either the visual or the auditory stimulus; thus, there is an inherent symmetry between the modalities, whereas in the focused attention paradigm that is considered in this paper, one modality is assigned to be the imperative stimulus (to which a response is required), and the other has only an accessory status—that is, can be ignored by the subject. Nonetheless, the distinction between focused and divided attention paradigms becomes somewhat blurred when both the visual and the auditory stimuli are presented at the same location without catch trials. Interestingly, it is easy to show that the two-stage model for focused attention can be made to violate the above inequality by choosing a value large enough for the interaction parameter δ .

EXPERIMENT 2

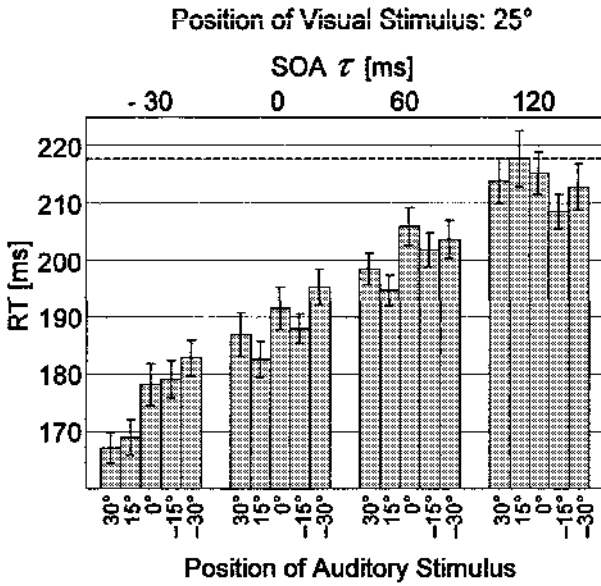
One important factor affecting the amount of response speed-up is the SOA between the presentations of the visual and the auditory signals. Specifically, this amount is often found to be maximal when the visual signal is followed by the auditory signal with a delay that equals the difference between the mean RTs for the unimodal latencies (*physiological synchrony*; see, e.g., Colonius & Townsend, 1997; Diederich, 1992, 1995; Diederich & Colonius, 1987; Miller, 1986). As was discussed in the previous section, the finding of a nonmonotone SOA function would lead one to reject the V-version of the model. In this experiment, in addition to varying spatial visual-auditory positions, an SOA is applied in the range of -30 msec (auditory onset 30 msec prior to visual onset) to 120 msec (auditory onset 120 msec after visual onset).

Method

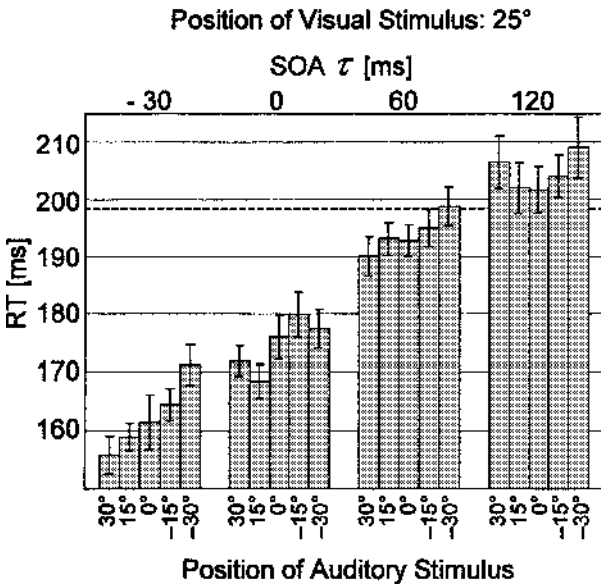
Subjects. Six of the 7 subjects from Experiment 1 participated in Experiment 2. In addition, 1 female subject with normal visual, auditory, and oculomotor functions took part in the experiment. The subjects were paid for their participation.

Stimulus arrangement. The same spatial arrangement as that in Experiment 1 was used for the visual and auditory stimuli, orthogonally combined with four different SOAs. The auditory stimulus onset was 30 msec prior to the visual stimulus (SOA = -30), simultaneous (SOA = 0), as in Experiment 1, or 60 or 120 msec after the visual stimulus (see Figure 1, panel c). Each spatiotemporal configuration between visual and auditory stimuli and the unimodal visual stimuli was presented once within each block. The presenta-

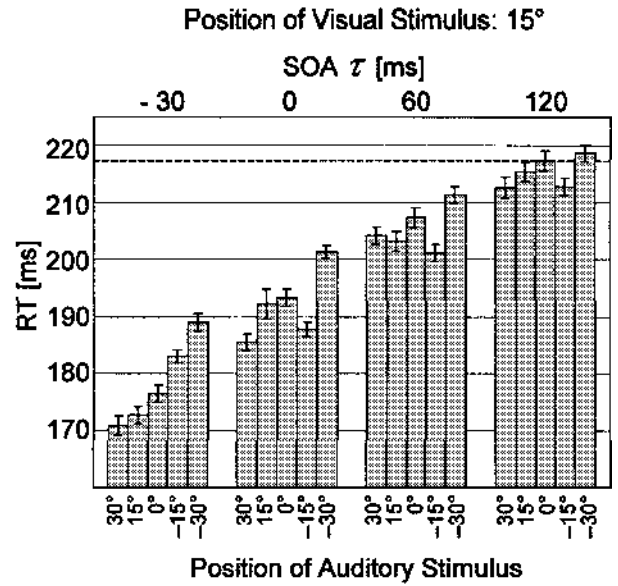
KS



MR



KS



MR

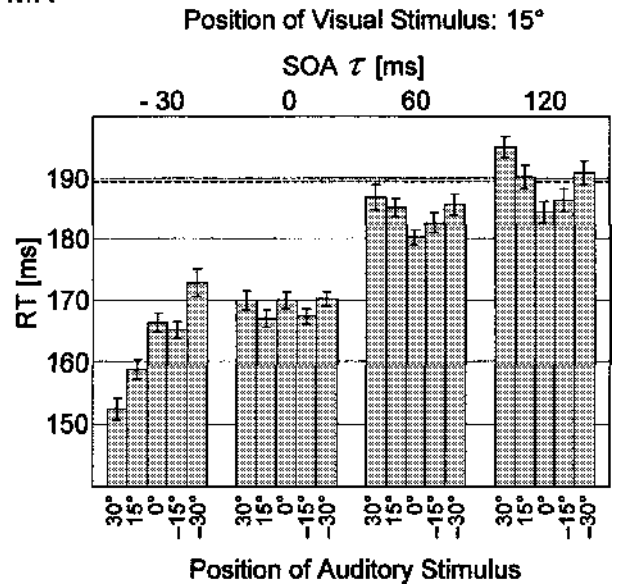


Figure 4. Mean saccadic latencies of 2 subjects (K.S. and M.R.) to bimodal targets for all spatiotemporal stimulus conditions used in Experiment 2. Each of the four sets of columns in every diagram refers to one of the four temporal relationships used in the experiment. The respective stimulus onset asynchronies (SOAs) are given above the sets. Each column within a set represents the reaction time (RT) for a specific spatial arrangement. On the abscissa, the position of the auditory stimulus is given with respect to the visual stimulus; positive values indicate an auditory position ipsilateral to the visual stimulus, and negative values a position contralateral to the visual stimulus. The left diagrams refer to visual stimulus eccentricity of 25°, the right diagrams of 15°. Mean unimodal RTs are indicated by dashed lines in each diagram.

tion of one block took about 8 min. Three blocks were presented in one session, with 2-min breaks in between. Only one session was performed per day. Twenty blocks were measured with each subject.

Results

In Figure 4, mean SRTs for 2 subjects (K.S. and M.R.) are plotted as a function of spatial stimulus combination

and SOA. The effect of the auditory stimulus is facilitatory for nearly all conditions, as is reflected by the reduction of RTs, as compared with the unimodal condition. For all the subjects, mean RT increases monotonically with the length of the SOA. Thus, at least in the temporal range covered here, the auditory stimulus is the more effective the earlier it is given. This holds true for all spatial com-

binations of the visual stimulus with the auditory stimulus. Moreover, for each level of SOA, there is a tendency to replicate the spatial effect observed in Experiment 1.

A four-way ANOVA was computed, with repeated measures on target eccentricity (two levels), auditory stimulus position (with respect to the target, five levels), and SOA (four levels) as within-subjects factors. There were significant main effects for all within-subjects factors: target eccentricity [$F(1,196) = 32.33, p < .001$], distractor position [$F(4,784) = 29.86, p < .001$], and SOA [4 levels; $F(3,588) = 904.51, p < .001$]. All two-way interactions, as well as the three-way interaction of within-subjects factors, were significant: target eccentricity \times auditory stimulus position [$F(4,784) = 2.79, p < .026$], target eccentricity \times SOA [$F(3,588) = 8.20, p < .001$], auditory stimulus position \times SOA [$F(12,2352) = 5.38, p < .001$], and target eccentricity \times auditory stimulus position \times SOA [$F(12,2352) = 2.82, p = .001$]. Two-way interactions with the subject factor were found for target eccentricity [$F(5,196) = 21.88, p < .001$] and SOA [$F(7,588) = 7.04, p < .001$]. The three-way auditory stimulus position \times SOA \times subject interaction was significant also [$F(60,2352) = 1.35, p = .039$].

Significant effects of interstimulus distance were found in a three-way ANOVA, with repeated measures on interstimulus distance (10 levels) and SOA (4 levels) as within-subjects factors [interstimulus distance: $F(9,1953) = 17.41, p < .001$; SOA: $F(3,651) = 1,002.79, p < .001$]. All two-way interactions were significant: interstimulus distance \times SOA [$F(9,1953) = 17.41, p < .001$], interstimulus distance \times subject [$F(54,1953) = 4.34, p < .001$], and SOA \times subject [$F(18,651) = 7.00, p < .001$], whereas the three-way interstimulus distance \times SOA \times subject interaction was not significant. The interaction between interstimulus distance and SOA was analyzed in more detail by separate ANOVAs for each SOA. The effect of interstimulus distance remained significant for all SOAs [SOA = -30 msec: $F(9,2133) = 8.33, p < .001$; SOA = 0 msec: $F(9,2097) = 5.13, p < .001$; SOA = 60 msec: $F(9,2034) = 4.82, p < .001$; SOA = 120 msec: $F(9,2052) = 6.69, p < .001$]. A post hoc test (Bonferroni) showed significant differences between small and large interstimulus distances ($p < .05$). For example, RTs for the interstimulus distance of 55° differed significantly from those measured with distances of up to 30° (SOA = -30 msec, and SOA = 0 msec) or 15° (SOA = 60 msec). The pattern is less clear for the 120-msec SOA.

Discussion

The data from this experiment display a clear effect of the SOA on mean SRTs. The IFE is larger the earlier the auditory stimulus is presented, at least up to an SOA value of -30 msec. For SOA = -30 msec, a small percentage of express saccades contributes to the RT reduction. Although the effect of interstimulus distance could be validated, the presence of significant two-way (and higher) interactions makes the interpretation of the spatial effects less clear-cut. In particular, for the large SOA values, spatial effects disappear or are no longer in the hypothesized

direction for many configurations and/or subjects. Note, however, that according to the two-stage model, the probability of visual-auditory interaction should decrease with increasing SOA. Whether or not these data are compatible with the model can be more appropriately tested by fitting the data to a variant of the model with specific distributional assumptions to which we now turn. Moreover, it should be noted that since the SOA function is monotone for all the spatial stimulus conditions and for all the subjects tested, there is no evidence to exclude either of the two versions on the basis of this qualitative test.

EX-GAUSSIAN TWO-STAGE MODEL

Distribution-free predictions are useful, since their validity does not depend on certain parametric assumptions' being true. On the other hand, introducing distributional assumptions permits more specific predictions that can be tested, at least to some degree, against the data. In particular, one would hope that fitting the data to both versions of the model will lead to favoring one version over the other.

First, all the peripheral processes will be assumed to have an exponentially distributed duration, with one specific (intensity) parameter for each modality. Although as a first approximation, this may not be completely implausible, we do not try to motivate this assumption from neurophysiological observations. The main reason for choosing the exponential distribution is its simplicity and its ensuing mathematical tractability. Second, the central processing stage is assumed to be normally distributed, with mean and standard deviation depending on whether or not interaction in this second stage occurs. The choice of the normal distribution is again somewhat arbitrary, but there are two pragmatic arguments in favor of this decision. First, the additive combination of exponential plus normal distributions has been quite successful in describing RT distributions previously (e.g., Van Zandt & Ratcliff, 1995). Second, in the normal distribution, there is no functional relationship between its mean and its variance, which makes it more flexible in fitting the data than, for example, the gamma distribution.¹¹

A nonnegative random variable T is exponentially distributed with intensity parameter λ ($\lambda > 0$) if

$$P[T \leq t] = 1 - \exp[-\lambda * t] \text{ or, for short, } T \sim \text{EXP}(\lambda).$$

For the expected value of T , it is well known (e.g., Ross, 1983) that $E[T] = 1/\lambda$.

The following parameters will be used in the ex-Gaussian model:

λ_V	intensity parameter for V, V' ,
λ_A	intensity parameter for A ,
μ	mean of the normally distributed central stage processing time when no interaction occurs,
δ	visual-auditory interaction parameter,
$\sigma(I)$	standard deviation of central stage processing time when interaction (I) occurs,
$\sigma(I^c)$	same when no interaction (I^c) occurs.

Specifically, λ_V and λ_A are the intensity parameters for the peripheral processes. An attractive property of the exponential is that under independence the distribution of the minimum of V and A is $\text{EXP}(\lambda_V + \lambda_A)$. Moreover, μ is the mean of W_2 , given that no interaction takes place—that is, if $I^c = \{A + \tau > V\}$ occurs—and $\mu - \delta$ is the mean of W_2 if interaction occurs. Thus, $\delta > 0$ corresponds to (spatial) facilitation and $\delta < 0$ to inhibition.

Mean Saccadic Response Time

In deriving mean SRT for the MIN-version, positive SOAs ($\tau > 0$) and negative SOAs ($\tau < 0$) have to be considered separately. Straightforward calculation yields, for $P[I] = \pi$,

$$\pi = \begin{cases} \exp[-\lambda_V \tau] \lambda_A / (\lambda_V + \lambda_A) & \text{for } \tau \geq 0 \\ 1 - \exp[\lambda_A \tau] \lambda_V / (\lambda_V + \lambda_A) & \text{for } \tau \leq 0, \end{cases}$$

and for the SOA function in the MIN-version, for $\tau \geq 0$,

$$\begin{aligned} E[\text{RT}_{VA, \tau}] &= 1/\lambda_V - \exp[-\lambda_V \tau][1/\lambda_V - 1/(\lambda_V + \lambda_A)] \\ &\quad + \mu - \pi * \delta, \end{aligned}$$

whereas for $\tau \leq 0$,

$$\begin{aligned} E[\text{RT}_{VA, \tau}] &= 1/\lambda_A + \tau - \exp[\lambda_A \tau][1/\lambda_A - 1/(\lambda_V + \lambda_A)] \\ &\quad + \mu - \pi * \delta. \end{aligned}$$

Similarly, for the V-version,

$$E[\text{RT}_{VA, \tau}] = 1/\lambda_V + \mu - \pi * \delta,$$

for all τ .

Variance and Covariance for Saccadic Response Time

Given the nonindependence between the first and the second stages of the model, the computation of the SRT variance is somewhat tedious (for details, see the Appendix). In particular, the covariance between the two stages has to be determined, and this turns out to give additional insight into the structure of the model.

Before considering a particular version of the model, the expression for the variance can be deduced as follows, making use of the conditional independence between W_1 and W_2 :

$$\begin{aligned} \text{Var}[W_1 + W_2] &= \text{Var}[W_1] + \text{Var}[W_2] + 2 \times \text{Cov}[W_1, W_2] \\ &= \text{Var}[W_1] + \pi \text{Var}[W_2 | I] \\ &\quad + (1 - \pi) \text{Var}[W_2 | I^c] \\ &\quad + \pi(1 - \pi) \{E[W_2 | I^c] - E[W_2 | I]\}^2 \\ &\quad + 2 \times \text{Cov}[W_1, W_2] = \text{Var}[W_1] \\ &\quad + \pi \text{Var}[W_2 | I] + (1 - \pi) \text{Var}[W_2 | I^c] \\ &\quad + \frac{\pi}{1 - \pi} \{E[W_2] - E[W_2 | I]\}^2 \\ &\quad + \frac{2\pi}{1 - \pi} \{E[W_1] - E[W_1 | I]\} \\ &\quad \times \{E[W_2] - E[W_2 | I]\}, \end{aligned}$$

where the last step follows from an expression for the covariance computed as

$$\begin{aligned} \text{Cov}[W_1, W_2] &= E[W_1 W_2] - E[W_1]E[W_2] \\ &= \pi E[W_1 W_2 | I] + (1 - \pi) E[W_1 W_2 | I^c] \\ &\quad - E[W_1]E[W_2] = \pi E[W_1 | I]E[W_2 | I] \\ &\quad + (1 - \pi) E[W_1 | I^c]E[W_2 | I^c] \\ &\quad - E[W_1] \{ \pi E[W_2 | I] + (1 - \pi) E[W_2 | I^c] \} \\ &= \pi \{E[W_1] - E[W_1 | I]\} \{E[W_2 | I^c] \\ &\quad - E[W_2 | I]\}. \end{aligned}$$

This covariance term includes the product of two differences of means: The second difference, $E[W_2 | I^c] - E[W_2 | I]$, is simply the amount of the IFE (δ), whereas the first difference, $E[W_1] - E[W_1 | I]$, measures the difference between the mean of the first stage and the mean of the first stage conditioned on the auditory peripheral's having won the race. It can be shown (see the Appendix) that for facilitatory neural summation ($\delta > 0$), the covariance in the V-version will always be negative (or zero), no matter what specific parameter values or SOAs are taken. This is clear intuitively: Assume that visual peripheral processing (V) takes a long time in a given trial; then, the auditory peripheral processing (A) has a high chance of winning the race. Consequently, there will be a speed-up in the second-stage processing. Since the first stage is determined by V , long first-stage processing will tend to go together with short second-stage processing, and vice versa. For the MIN-version of the model, the behavior of the covariance is more complex, because large V values do not imply long first-stage processing times [since here, $\min(V, A + \tau)$ determines the first stage]. Although the behavior also depends on the specific values of the distribution parameters, nonetheless, the following pattern emerges: For positive SOAs, the covariance is negative, whereas for negative SOAs, the covariance is positive.¹²

Since the covariance between the two stages is not an observable entity, these subtle differences between the two model versions cannot be exploited directly for testing between the MIN- and the V-versions. Nevertheless, the covariance is part of the (observable) SRT variance, to which we now return.

First note that in the variance expression above, owing to the mixture property of the model, $\text{Var}[W_2]$ is expressed as a weighted sum of the conditional variances $\text{Var}[W_2 | I]$ and $\text{Var}[W_2 | I^c]$ plus an *inflation term* that increases with the square of the amount of facilitation in the second stage (δ). This growth in variability, however, can be compensated for by a negative covariance term. Thus, given the different behaviors of the covariance term as a function of the SOA for the two versions, the fit of the variances (or, equivalently, the standard deviations) for various SOA values should be informative in evaluating the two different model versions. In order to include the variances in the model-fitting procedure, the explicit expressions for the variances for the exponential peripheral process-

ing times and the Gaussian central processing time have to be derived. These rather complex expressions must be derived separately for each model version and for positive and negative SOAs. Here is one example, the variance for the MIN-version in the case of positive SOAs:

$$\begin{aligned}
\text{Var}[W_1 + W_2] &= \text{Var}[W_1] + \pi \text{Var}[W_2 | I] \\
&+ (1 - \pi) \text{Var}[W_2 | I^c] \\
&+ \pi(1 - \pi) \{E[W_2 | I^c] \\
&- E[W_2 | I]\}^2 + 2\pi \{E[W_1] - E[W_1 | I]\} \\
&\times \{E[W_2 | I^c] - E[W_2 | I]\} \\
&= 2 \left\{ \frac{1}{\lambda_V^2} - \exp(-\lambda_V \tau) \left[\frac{1}{\lambda_V^2} - \frac{1}{(\lambda_V + \lambda_A)^2} \right. \right. \\
&\quad \left. \left. + \frac{\tau}{\lambda_V} - \frac{\tau}{\lambda_V + \lambda_A} \right] \right\} \\
&- \left\{ \frac{1}{\lambda_V} - \exp(-\lambda_V \tau) \left[\frac{1}{\lambda_V} - \frac{1}{\lambda_V + \lambda_A} \right] \right\}^2 \\
&+ \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \sigma(I)^2 \\
&+ \left[1 - \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \right] \sigma(I^c)^2 \\
&+ \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \left[1 - \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \right] \delta^2 \\
&+ 2 \left\{ \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \right\} \\
&\times \left\{ \left[\frac{1}{\lambda_V} - \frac{1}{\lambda_V + \lambda_A} \right] \right. \\
&\quad \left. [1 - \exp(-\lambda_V \tau)] - \tau \right\} \delta.
\end{aligned}$$

All formulas can be found in the Appendix.

Data Fits for MIN-Version and V-Version

Parameters to be estimated are the intensities λ_V , λ_A for the peripheral processes V and A , as well as mean μ and standard deviations $\sigma(I)$ and $\sigma(I^c)$ for the central processing time W_2 , with and without interaction, respectively. In addition, visual–auditory interaction parameters (δ) have to be determined. There are different ways to proceed, depending on which spatial configurations are merged into one parameter value. If 1 parameter is estimated for each visual–auditory configuration, 20 parameters would have to be estimated. On the other hand, the data from Experiments 1 and 2 suggest a dependence of visual–auditory interaction on spatial dis-

tance. To reduce the number of parameters as much as possible, we assume a linear relationship between the horizontal distance and the strength of the visual–auditory interaction. Thus, only 2 more parameters have to be estimated. Obviously, assuming linearity is a simplification, but it resulted in interaction parameter estimates that were plausible, given the horizontal stimulus arrangement used here. Altogether, there were 8 parameters to fit 84 data points (means and standard deviations) for each subject.

All the parameters were estimated individually for each subject, using the Levenberg–Marquard algorithm (implemented in MATLAB 5.2) to minimize the (square root of the) sum-of-squares deviation between observed and expected mean SRTs (Dev. RT in Table 2) and standard deviations (Dev. SD in Table 2) over all experimental conditions, separately for the MIN-version and the V-version of the two-stage model. The intensity parameter estimates for all the subjects are presented in Table 2. To facilitate interpretation, the inverses of the intensity parameters are displayed, giving an estimate of the mean processing time for each of the peripheral processes.

The estimate for the auditory peripheral processing time is below 20 msec (except for A.M.) for the V-version but above 40 msec for the MIN-version. Thus, the estimates for the V-version are in the range of typical latencies of acoustically activated neurons in the cat SC (Meredith, Nemitz, & Stein, 1987). The estimated visual peripheral processing times range from about 20 msec to about 60 msec in both versions of the model—that is, some mean estimated visual processing times are clearly smaller than might be expected on the basis of physiological data (Meredith et al., 1987). In addition, sometimes the parameter values are not in the expected order. Specifically, in the MIN-version, estimated auditory peripheral processing times are longer than visual peripheral processing times for all the subjects. The V-version fares better in this respect. Moreover, peripheral processing time estimates for the more eccentric visual stimulus is longer than those for the less eccentric stimulus in both versions, except for subjects K.S. and R.V., who, however, show this reverse order in their unimodal visual RTs as well. In sum, the estimates for the peripheral processing times tend to give the V-version some advantage over the MIN-version. It should also be noted that the fit indices (Dev. RT and Dev. SD in Table 2) show a slight but consistent advantage for the V-version over the MIN-version.

Most estimates of the variance in the second processing stage are extremely small: $\sigma(I)$ is zero for the V-version for all the subjects and for 5 out of 7 subjects for the MIN-version. Differences between the two versions of the model can be found for $\sigma(I^c)$, with values close to zero for the MIN-version and values between 6.8 and 24.6 for the V-version. Parameter a , the slope of the fitted linear relation between visual–auditory interaction and horizontal distance is of the same order of magnitude for both

Table 2
Results of Fit to Two-Stage Model

Subject	$1/\lambda_{V25}$	$1/\lambda_{V15}$	$1/\lambda_A$	μ	$\sigma(I)$	$\sigma(I^c)$	a	Dev. RT	Dev. SD
MIN-Version									
A.M.	32	31	57	176	6.5	0	-0.19	4.5	3.9
B.G.	59	36	77	171	0	0	-0.47	8.6	10.9
K.S.	28	31	45	180	0	0	-0.46	5.3	5.4
M.I.	48	35	55	206	13.1	0	-0.47	7.9	7.0
M.R.	35	23	41	164	0	0	-0.36	5.4	5.1
O.V.	37	27	45	169	0	0	-0.19	7.7	6.4
R.V.	23	31	64	159	0	0	-0.35	5.2	4.8
V-Version									
A.M.	27	26	36	181	0	9.7	-0.16	4.5	3.1
B.G.	42	26	10	185	0	24.6	-0.26	7.5	10.6
K.S.	22	26	19	186	0	6.8	-0.37	4.7	4.4
M.I.	39	30	7	214	0	7.0	-0.32	6.1	6.1
M.R.	24	17	9	173	0	12.2	-0.23	4.9	4.1
O.V.	27	19	5	179	0	12.3	-0.11	6.1	5.7
R.V.	18	25	7	165	0	9.8	-0.21	3.3	4.3

Note— $1/\lambda_{V25}$ is the estimated processing time for visual peripheral processing, $1/\lambda_{V15}$ for auditory peripheral processing, and μ for central processing, for Experiment 2. Estimated standard deviations for central processing are $\sigma(I)$ in the case of intersensory interaction and $\sigma(I^c)$ if no interaction occurs. Since the interaction parameter δ is assumed to depend linearly on interstimulus distance, parameter a denotes the slope of this line for each subject. Dev. RT and Dev. SD refer to the (square root of the) sum-of-squares deviation between the observed and the expected mean saccadic reaction times and standard deviations, respectively.

versions. The negative a values indicate the decrease of the amount of interaction with increasing spatial distance, as was found in our data (Figure 5).

For the V-version, interaction parameter δ is positive for all the spatial distances and all the subjects, reflecting facilitation in the central processing stage (Figure 5). On the other hand, some small negative δ values (not larger than 5 msec in absolute value) are found for 3 subjects for the larger interstimulus distances in the MIN-version, indicating inhibition in the central stage. Inhibitory effects are in line with the results in Stein, Huneycott, and Meredith (1988), where spatial distances of 60° between auditory and visual stimuli produced response depression in multimodal neurons. Note, however, that this central inhibition—if it does exist in our data—is easily kept from showing up in the observed RTs because of a possible statistical facilitation effect in the first stage. In fact, given the intensity parameter estimates from the fitted MIN-version, those subjects with δ values suggesting inhibition have statistical facilitation values between 6 and 16 msec (for SOA = 0).

Figure 6 displays the observed and the predicted mean SRTs and standard deviations as functions of the SOA over all visual-auditory configurations, including purely visual SRTs separately for 1 subject, the other subjects showing very similar fits. Although a goodness-of-fit test would indicate significant violations,¹³ both model versions show a qualitatively satisfactory fit for the SOA functions for many of the configurations. In particular, the acceptable fit of the standard deviations over different SOA values for some of the configurations is quite remarkable. In general, RT standard deviations (equiva-

lently, variances) are much more difficult to predict than mean values, because they are more sensitive to outliers and to processes not captured within the model (e.g., motor variability).

Discussion of the Ex-Gaussian Two-Stage Model

Although the overall performance of the ex-Gaussian two-stage model in predicting means and standard deviations of the SRTs has been encouraging, there are some obvious shortcomings that call for further investigation. As was observed above, some of the parameter estimates for some of the subjects are not in a plausible range of values. Another peculiar feature of the parameter estimates are the zero values for the standard deviation of the second-stage duration for nearly all the subjects [with the exception of the $\sigma(I^c)$ values in the V-version]. Given the inherent variability of the duration of neural processes, a constant value of the central processing time is quite unlikely. Rather, this points to a possible artifact in the parameter estimation procedure.

A more conclusive assessment of the ex-Gaussian two-stage model could be achieved along the following lines. First, the properties of the parameter estimation procedure should be explored in more detail. It is desirable, in particular, to have a measure of the stability of the parameter estimates. How much variability in the estimates is to be expected? In a large-scale simulation study, many sets of data could be generated from the model with known parameter values, and for each data set the parameters could be estimated, yielding a distribution of the parameter estimates around each (known) true parameter value. Similarly, to assess the overall chances of distin-

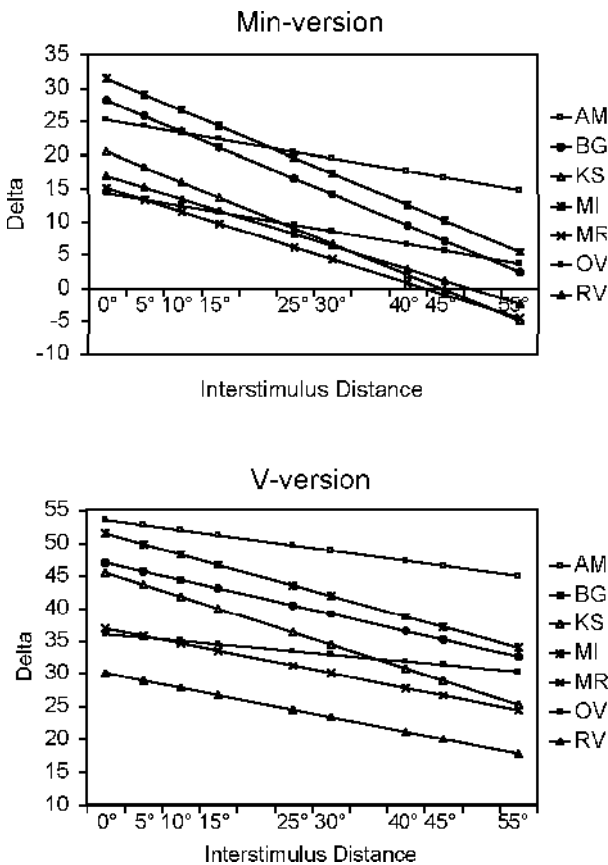


Figure 5. Estimated linear relationship between δ values as a function of interstimulus distance for MIN-version (upper panel) and V-version (lower panel) of the model for all the subjects in Experiment 2. (Note the different scales on the ordinates.)

guishing between the two model versions, one could simulate data from one model version and try to fit them with the other model version, and vice versa.¹⁴

Finally, a more direct test of the distributional assumptions (exponential and Gaussian) is desirable. For this, however, a larger sample size for any given experimental condition is needed. The primary goal of the experiments in this paper was to observe the behavior of the intersensory facilitation effect over a large number of spatial and temporal conditions. The next step will be to obtain a large number of observations for a few selected conditions and to fit the model to the empirical distributions.

DISCUSSION AND CONCLUSION

The experimental results of this paper provide clear evidence for the existence of an intersensory facilitation effect on SRT for visual targets with auditory accessory stimulation. Specifically, it was demonstrated that the speed-up of RTs is a decreasing function of the horizontal distance between the visual target and the auditory distractor (Experiment 1). Moreover, the facilitatory effect

of the auditory nontarget is the larger the earlier the auditory stimulus is presented, at least in the range of SOAs employed in Experiment 2 (−30 to 120 msec). Our observations are in accordance with earlier studies (e.g., Corneil & Munoz, 1996; Frens et al., 1995; Hughes et al., 1998) and extend the effect to a situation in which the auditory stimuli are presented via a virtual auditory setup.

Although multiple visual stimuli typically have dramatic effects both on saccade amplitude and on SRT, depending on the specific stimulus arrangement (see Findlay & Walker, 1999), the effect of auditory accessory stimuli is basically limited to SRT. This finding is in line with the hypothesis of a separate programming of saccadic latency (*when*-pathway) and amplitude (*where*-pathway) by Becker and Jürgens (1979). The increasing facilitatory effect of the auditory stimulus when presented earlier, relative to the visual target, suggests the existence of a nonspecific warning effect of the auditory accessory, in analogy to the preparation enhancement concept of Nickerson (1973). Although this would predict a general speed-up of SRT in the presence of an auditory distractor, it does not suffice to explain the specific dependence of the IFE on the spatiotemporal arrangement between visual and auditory signals. Rather, this indicates the existence of a neural integration stage at which the saccadic response is modulated by a salience map of multimodal activity determined by the spatiotemporal stimulus configuration. As has been demonstrated by a large number of neurophysiological studies at the cellular level, multimodal cells in the deep layers of the SC respond to multimodal stimulation in accordance with the relative spatial and temporal alignment of the stimuli (see Stein & Meredith, 1993).

At the behavioral level, the two-stage model proposed in this paper explicates the temporal aspects of the generation of saccades to visual targets in the presence of auditory distractors (*when*-pathway). As was described in detail above, in the MIN-version of the model, the auditory accessory acts both as a nonspecific warning signal for the onset of the visual target stimulus and as a component in the bimodal convergence taking place in the neural integration stage, whereas in the V-version of the model, the role of the auditory stimulus in the first stage is limited to modulating the probability that intersensory interaction effects will occur in the second stage.¹⁵ Most important, both model versions achieved a rather encouraging fit for both the means of the SRTs and their variance. Although the fit of the ex-Gaussian variant of the two-stage model hinted at a possible advantage of the V-version over the MIN-version, it is too early at this point to draw a definite conclusion.

There are a number of steps to be taken in order to further probe the two-stage model. In order to test the merits of the two different versions, a larger range of SOAs is necessary. In particular, the predictions of the two versions will differ more with the auditory accessory being presented even earlier than the 30 msec used here. In order to test the basic two-stage assumption, a variation

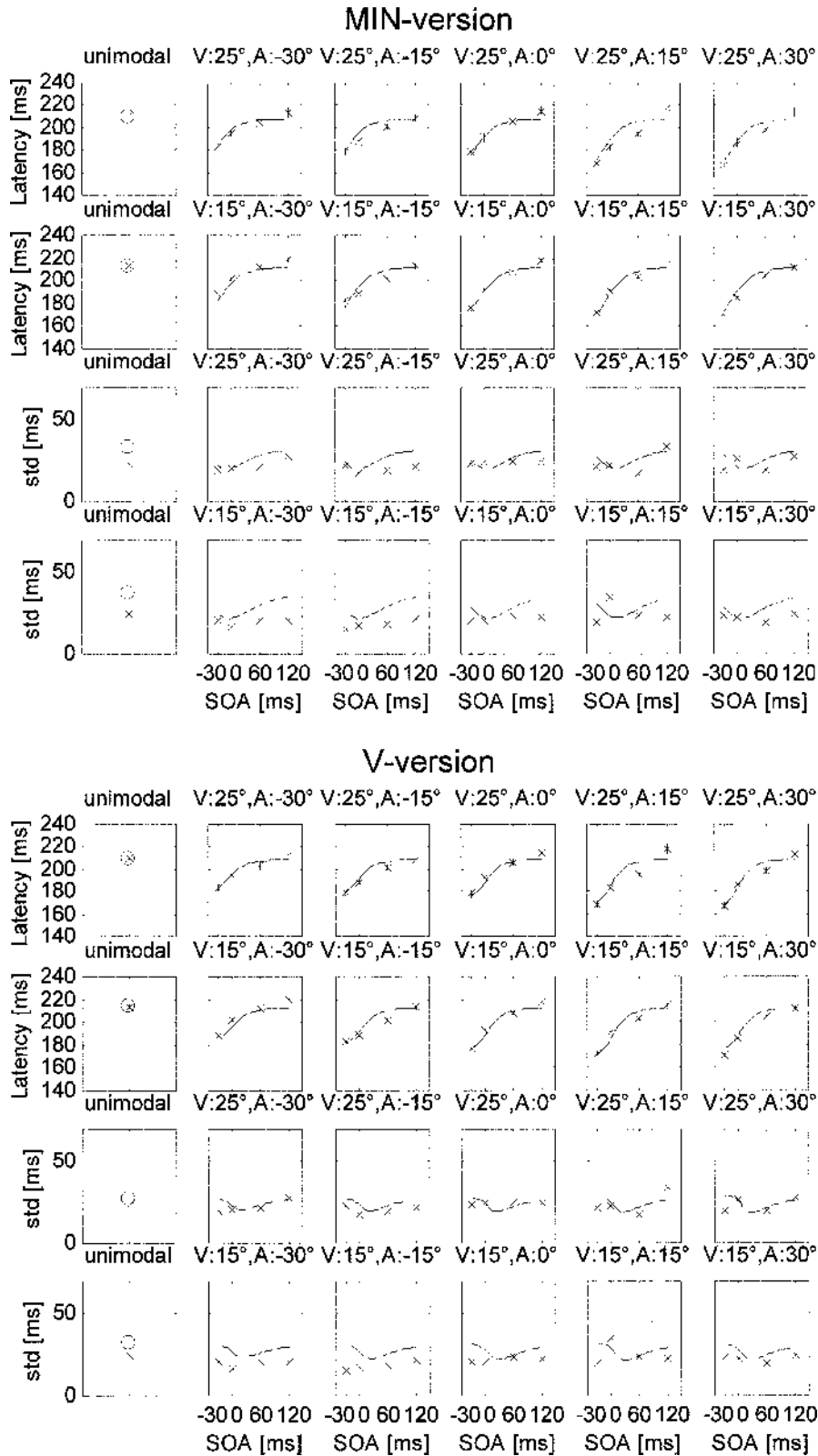


Figure 6. Individual data fits for MIN-version (upper panel) and V-version (lower panel) of the model for 1 subject (K.S.). Empirical latency mean values are represented by x; error bars (for means) indicate standard errors. In the unimodal condition, \circ represents the predicted mean values. The eccentricities of the visual and auditory stimuli are given above each panel, with negative values indicating a contralateral position of the auditory stimulus with respect to the visual stimulus.

of the stimulus intensities is called for. For example, increasing the intensity of the auditory accessory should increase its chances of winning the race in the first stage, and specific predictions could then be tested. As has been suggested by one of the reviewers of this paper, another possibility is to interchange the role of the visual and the auditory stimuli—that is, to have the auditory as the imperative stimulus and the visual as the accessory stimulus. It should be noted, however, that the two modalities per se are not symmetric and, thus, not necessarily interchangeable with respect to the task of saccade generation.

REFERENCES

- BECKER, W. (1991). Saccades. In R. H. S. Carpenter (Ed.), *Vision and visual dysfunction: Vol. 8. Eye movements* (pp. 95-137). London: Macmillan.
- BECKER, W., & JÜRGENS, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, **19**, 967-983.
- BERNSTEIN, I. H., & EDELSTEIN, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology*, **87**, 241-247.
- BERNSTEIN, I. H., ROSE, R., & ASHE, V. M. (1970). Energy integration in intersensory facilitation. *Journal of Experimental Psychology*, **86**, 196-203.
- BERTELSON, P., & TISSEYRE, E. (1969). The time course of preparation: Confirmatory results with visual and auditory warning signals. *Acta Psychologica*, **30**, 145-154.
- CARPENTER, R. H. S., & WILLIAMS, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*, **377**, 59-62.
- COLONIUS, H. (1990). Possibly dependent probability summation of reaction times. *Journal of Mathematical Psychology*, **24**, 253-275.
- COLONIUS, H., & ARNDT, P. A. (1999). The effect of auditory distractors on saccades towards visual targets. *Behavioral & Brain Sciences*, **22**, 677-678.
- COLONIUS, H., & TOWNSEND, J. T. (1997). Activation-state representation of models for the redundant signals effect. In A. A. J. Marley (Ed.), *Choice, decision, and measurement: Essays in honor of R. Duncan Luce* (pp. 245-254). Mahwah, NJ: Erlbaum.
- 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.
- DIEDERICH, A. (1992). Probability inequalities for testing separate activation models of divided attention. *Perception & Psychophysics*, **52**, 714-716.
- 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. (1987). Intersensory facilitation in the motor component? *Psychological Research*, **49**, 23-29.
- DZHAFAROV, E. N., & SCHWEICKERT, R. (1995). Decompositions of response times: An almost general theory. *Journal of Mathematical Psychology*, **39**, 285-314.
- ENGELKEN, E. D., & STEVENS, K. W. (1989). Saccadic eye movements in response to visual, auditory, and bisensory stimuli. *Aviation, Space, & Environmental Medicine*, **60**, 762-768.
- FINDLAY, J. M. (1982). Global visual processing for saccadic eye movements. *Vision Research*, **22**, 1033-1045.
- FINDLAY, J. M. (1983). Visual information processing for saccadic eye movements. In A. Hein & M. Jeannerod (Eds.), *Spatially oriented behavior* (pp. 281-303). New York: Springer-Verlag.
- FINDLAY, J. M., & WALKER, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral & Brain Sciences*, **22**, 661-674.
- FRENS, M. A., & VAN OPSTAL, A. J. (1998). Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Research Bulletin*, **46**, 211-224.
- FRENS, M. A., VAN OPSTAL, A. J., & VAN DER WILLIGEN, R. F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception & Psychophysics*, **57**, 802-816.
- GIELEN, S. C. A. M., SCHMIDT, R. A., & VAN DEN HEUVEL, P. J. M. (1983). On the nature of intersensory facilitation of reaction time. *Perception & Psychophysics*, **34**, 161-168.
- GIRAY, M., & ULRICH, R. (1993). Motor coactivation revealed by response force in divided and focused attention. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 1278-1291.
- HACKLEY, S. A., & VALLE-INCLÁN, E. (1998). Automatic alerting does not speed late motoric processes in a reaction time task. *Nature*, **391**, 786-788.
- HACKLEY, S. A., & VALLE-INCLÁN, E. (1999). Accessory stimulus effects on response selection: Does arousal speed decision making? *Journal of Cognitive Neuroscience*, **11**, 321-329.
- HANES, D. P., & SCHALL, J. D. (1996). Neural control of voluntary movement initiation. *Science*, **274**, 427-430.
- HARRINGTON, L. K., & PECK, C. K. (1998). Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Experimental Brain Research*, **122**, 247-252.
- HEUERMAN, H., & COLONIUS, H. (1999). Localization experiments with saccadic responses in virtual auditory environments. In T. Dau, V. Hohmann, & B. Kollmeier (Eds.), *Psychophysics, physiology and models of hearing* (pp. 89-92). Singapore: World Scientific Publishing.
- 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 & Performance*, **20**, 131-153.
- KOPECZ, K. (1995). Saccadic reaction times in gap/overlap paradigms: A model based on integration of intentional and visual information on neural, dynamic fields. *Vision Research*, **35**, 2911-2925.
- LUECK, C. J., CRAWFORD, T. J., SAVAGE, C. J., & KENNARD, C. (1990). Auditory-visual interaction in the generation of saccades in man. *Experimental Brain Research*, **82**, 149-157.
- MEREDITH, M. A., NEMITZ, J. W., & STEIN, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons: I. Temporal factors. *Journal of Neuroscience*, **10**, 3215-3229.
- MEREDITH, M. A., & STEIN, B. E. (1986). Visual, auditory and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, **56**, 640-662.
- MIDDLEBROOKS, J. C., & GREEN, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, **42**, 135-159.
- MILLER, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, **14**, 247-279.
- MILLER, J. O. (1986). Time course of coactivation in bimodal divided attention. *Perception & Psychophysics*, **40**, 331-343.
- MUNOZ, D. P., & WURTZ, R. H. (1993a). Fixation cells in monkey superior colliculus: I. Characteristics of cell discharge. *Journal of Neurophysiology*, **70**, 559-575.
- MUNOZ, D. P., & WURTZ, R. H. (1993b). Fixation cells in monkey superior colliculus: II. Reversible activation and deactivation. *Journal of Neurophysiology*, **70**, 576-589.
- NICKERSON, R. S. (1973). Intersensory facilitation of reaction time: Energy summation or preparation enhancement. *Psychological Review*, **80**, 489-509.
- NOZAWA, G., REUTER-LORENZ, P. A., & HUGHES, H. C. (1994). Parallel and serial processes in the human oculomotor system: Bimodal integration and express saccades. *Biological Cybernetics*, **72**, 19-34.
- PICK, H. L., JR., & WARREN, D. H. (1969). Sensory conflict in judgments of spatial direction. *Perception & Psychophysics*, **6**, 203-205.
- POSNER, M. I., NISSEN, M. J., & KLEIN, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, **83**, 157-171.
- RAAB, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, **24**, 574-590.

- ROSS, S. M. (1983). *Stochastic processes*. New York: Wiley.
- SANDERS, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, **53**, 61-97.
- SCHMIDT, R. A., GIELEN, S. C. A. M., & VAN DEN HEUVEL, P. J. M. (1984). The locus of intersensory facilitation of reaction time. *Acta Psychologica*, **57**, 147-164.
- SIMON, J. R., & CRAFT, J. L. (1970). Effects of an irrelevant auditory stimulus on visual choice reaction time. *Journal of Experimental Psychology*, **86**, 272-274.
- STEIN, B. E., HUNEYCOTT, W. S., & MEREDITH, M. A. (1988). Neurons and behavior: The same rules of multisensory integration apply. *Brain Research*, **448**, 355-358.
- STEIN, B. E., & MEREDITH, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
- STERNBERG, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II* (pp. 276-315). Amsterdam: North-Holland.
- TODD, J. W. (1912). Reaction to multiple stimuli. In R. S. Woodworth (Ed.), *Archives of Psychology: No. 25. Columbia Contributions to Philosophy and Psychology* (Vol. XXI, No. 8). New York: Science Press.
- TOWNSEND, J. T., & ASHBY, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.
- VAN OPSTAL, A. J., & VAN GISBERGEN, J. A. M. (1989). A nonlinear model for collicular spatial interactions underlying the metrical properties of electrically elicited saccades. *Biological Cybernetics*, **60**, 171-183.
- VAN ZANDT, T., COLONIUS, H., & PROCTOR, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, **7**, 208-256.
- VAN ZANDT, T., & RATCLIFF, R. (1995). Statistical mimicking of reaction time data: Single-process models, parameter variability, and mixtures. *Psychonomic Bulletin & Review*, **2**, 20-54.
- 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.
- WELCH, R. B., & WARREN, D. H. (1986). Intersensory interactions. In K. R. Boff, L. Kaufmann, J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception*. New York: Wiley.
- WENZEL, E. M., ARRUDA, M., KISTLER, D. J., & WIGHTMAN, F. L. (1993). Localization using nonindividualized head-related transfer function. *Journal of the Acoustical Society of America*, **94**, 111-123.
- WURTZ, R. H., BASSO, M. A., PARÉ, M., & SOMMER, M. A. (2000). The superior colliculus and the cognitive control of movement. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 573-588). Cambridge, MA: MIT Press.
- YANTIS, S., MEYER, D. E., & SMITH, J. E. K. (1991). Analysis of multinomial mixture distributions: New tests for stochastic models of cognition and action. *Psychological Bulletin*, **110**, 350-374.
- YAO, L., & PECK, C. K. (1997). Saccadic eye movements to visual and auditory targets. *Experimental Brain Research*, **115**, 25-34.

NOTES

1. One could argue that Gielen et al.'s (1983) bimodal condition was not a proper focused attention task, but this is inconsequential for our discussion.

2. This happens when a negative growth rate is sampled and the threshold cannot be reached. Although the probability for this event can be made arbitrarily small by shifting the Gaussian distribution to the right, the occurrence of infinitely long latencies can only be avoided by choosing a distribution function with no probability mass on the negative real line.

3. Strictly speaking, this can be assured only for the noncoincident visual-auditory configurations—that is, for 18 out of 20 configurations.

4. It is conceivable that there is partial temporal overlap between the stages, but introducing this generalization should only be considered if it is suggested by empirical evidence.

5. This cannot happen for distributions with a nonfinite domain, such as exponential, normal, or gamma distributions.

6. In catch trials, only the accessory stimulus is presented, and subjects are instructed to withhold their responses.

7. We refrain from giving a formal definition of such concepts as direct selective or nonselective influence here; see, however, Townsend and Ashby (1983) and Dzhabarov and Schweickert (1995).

8. In particular, a simplifying assumption here is that changing the position of the auditory stimulus does not modify its effect as a warning signal in the first processing stage.

9. By *increasing/decreasing* is meant *non-decreasing/non-increasing*, respectively.

10. For details, see Yantis et al. (1991).

11. Of course, there are many other distributions possessing this property as well.

12. The intuition here is that when $A + \tau < V$ for (relatively large) positive τ values, the minimum tends to be large, whereas for (relatively large) negative τ values, the minimum tends to be small. Thus, in the former case, long first-stage processing goes together with short second-stage processing, whereas in the latter case, short first-stage processing goes together with short second-stage processing.

13. Note that the data points do not constitute an independent sample, making a statistical test dubious.

14. This strategy has proven quite successful in a recent study comparing diffusion and race models for RTs (Van Zandt, Colonus, & Proctor, 2000).

15. At the neurophysiological level, the distinction between the model versions could have the following interpretation: The warning effect of the auditory nontarget in the MIN-Version is construed as an inhibitory effect on the fixation neurons located in the SC and/or on the omnipause neurons in the brain stem leading to a faster build-up of the saccade-generating neuronal activity, whereas in the V-version such a direct effect on the dynamics of saccade would not be possible.

APPENDIX
Variance and Covariance for the Ex-Gaussian Two-Stage Model

First, the general expressions for the covariance and the variance without any distributional assumptions holding both for the MIN-version and the V-version of the model are presented (see the main text). Defining $I = \{A + \tau < V\}$ and $\pi = P[I]$,

$$\begin{aligned} \text{Cov}[W_1 W_2] &= E[W_1 W_2] - E[W_1]E[W_2] = \pi E[W_1 W_2 | I] + (1 - \pi)E[W_1 W_2 | I^c] - E[W_1]E[W_2] \\ &= \pi E[W_1 | I]E[W_2 | I] + (1 - \pi)E[W_1 | I^c]E[W_2 | I^c] - E[W_1]\{ \pi E[W_2 | I] + (1 - \pi)E[W_2 | I^c] \} \\ &= \pi \{ E[W_1] - E[W_1 | I] \} \{ E[W_2 | I^c] - E[W_2 | I] \} \end{aligned}$$

$$\text{Note: } E[W_i] - E[W_i | I] = (1 - \pi) \{ E[W_i | I^c] - E[W_i | I] \} \quad i = 1, 2$$

$$\begin{aligned} \text{Var}[W_1 + W_2] &= \text{Var}[W_1] + \text{Var}[W_2] + 2 \times \text{Cov}[W_1, W_2] = \text{Var}[W_1] + \pi \text{Var}[W_2 | I] + (1 - \pi) \text{Var}[W_2 | I^c] \\ &\quad + \pi(1 - \pi) \{ E[W_2 | I^c] - E[W_2 | I] \}^2 + 2 \times \text{Cov}[W_1, W_2] = \text{Var}[W_1] + \pi \text{Var}[W_2 | I] \\ &\quad + (1 - \pi) \text{Var}[W_2 | I^c] + \frac{\pi}{1 - \pi} \{ E[W_2] - E[W_2 | I] \}^2 + \frac{2\pi}{1 - \pi} \{ E[W_1] - E[W_1 | I] \} \\ &\quad \times \{ E[W_2] - E[W_2 | I] \} \end{aligned}$$

A.1 Variance for MIN-Version

A.1.1 $\tau \geq 0$:

$$P[I] = \pi = \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A}$$

$$E[W_1] = E[\min(V, A + \tau)] = \frac{1}{\lambda_V} - \exp(-\lambda_V \tau) \left[\frac{1}{\lambda_V} - \frac{1}{\lambda_V + \lambda_A} \right]$$

$$E[W_1 | I] = E[\min(V, A + \tau) | A + \tau < V] = \tau + \frac{1}{\lambda_V + \lambda_A}$$

$$E[W_1] - E[W_1 | I] = \left[\frac{1}{\lambda_V} - \frac{1}{\lambda_V + \lambda_A} \right] [1 - \exp(-\lambda_V \tau)] - \tau$$

$$\begin{aligned} \text{Var}[W_1] &= \text{Var}[\min(V, A + \tau)] = 2 \left\{ \frac{1}{\lambda_V^2} - \exp(-\lambda_V \tau) \left[\frac{1}{\lambda_V^2} - \frac{1}{(\lambda_V + \lambda_A)^2} + \frac{\tau}{\lambda_V} - \frac{\tau}{\lambda_V + \lambda_A} \right] \right\} \\ &\quad - \left\{ \frac{1}{\lambda_V} - \exp(-\lambda_V \tau) \left[\frac{1}{\lambda_V} - \frac{1}{\lambda_V + \lambda_A} \right] \right\}^2 \end{aligned}$$

$$\begin{aligned} \text{Var}[W_1 + W_2] &= \text{Var}[W_1] + \pi \text{Var}[W_2 | I] + (1 - \pi) \text{Var}[W_2 | I^c] + \pi(1 - \pi) \{ E[W_2 | I^c] - E[W_2 | I] \}^2 \\ &\quad + 2\pi \{ E[W_1] - E[W_1 | I] \} \times \{ E[W_2 | I^c] - E[W_2 | I] \} \\ &= 2 \left\{ \frac{1}{\lambda_V^2} - \exp(-\lambda_V \tau) \left[\frac{1}{\lambda_V^2} - \frac{1}{(\lambda_V + \lambda_A)^2} + \frac{\tau}{\lambda_V} - \frac{\tau}{\lambda_V + \lambda_A} \right] \right\} \\ &\quad - \left\{ \frac{1}{\lambda_V} - \exp(-\lambda_V \tau) \left[\frac{1}{\lambda_V} - \frac{1}{\lambda_V + \lambda_A} \right] \right\}^2 \\ &\quad + \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \sigma(I)^2 + \left[1 - \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \right] \sigma(I^c)^2 \\ &\quad + \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \left[1 - \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \right] \delta^2 \\ &\quad + 2 \left\{ \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \right\} \left\{ \left[\frac{1}{\lambda_V} - \frac{1}{\lambda_V + \lambda_A} \right] [1 - \exp(-\lambda_V \tau)] - \tau \right\} \delta \end{aligned}$$

APPENDIX (Continued)

A.1.2 $\tau \leq 0$:

$$P[I] = \pi = 1 - \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} = \frac{\lambda_V + \lambda_A - \exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A}$$

$$E[W_1] = E[\min(V, A + \tau)] = \frac{1}{\lambda_A} + \tau - \exp(\lambda_A \tau) \left[\frac{1}{\lambda_A} - \frac{1}{\lambda_V + \lambda_A} \right]$$

$$E[W_1 | I] = E[\min(V, A + \tau) | A + \tau < V] = \frac{2\lambda_A + \lambda_V}{\lambda_A(\lambda_V + \lambda_A)} + \frac{\tau(\lambda_V + \lambda_A) - 1}{\lambda_A + \lambda_V [1 - \exp(\lambda_A \tau)]}$$

$$E[W_1] - E[W_1 | I] = \frac{1}{\lambda_A} - \frac{2\lambda_A + \lambda_V}{\lambda_A(\lambda_V + \lambda_A)} - \exp(\lambda_A \tau) \left[\frac{1}{\lambda_A} - \frac{1}{\lambda_V + \lambda_A} \right] + \tau - \frac{\tau(\lambda_V + \lambda_A) - 1}{\lambda_A + \lambda_V [1 - \exp(\lambda_A \tau)]}$$

$$\text{Var}[W_1] = \text{Var}[\min(V, A + \tau)] = \frac{2[1 - \exp(\lambda_A \tau) + \lambda_A \tau] + (\lambda_A \tau)^2}{\lambda_A^2} + \frac{2\exp(\lambda_A \tau)}{(\lambda_V + \lambda_A)^2}$$

$$- \left\{ \frac{1}{\lambda_A} + \tau - \exp(\lambda_A \tau) \left[\frac{1}{\lambda_A} - \frac{1}{\lambda_V + \lambda_A} \right] \right\}^2$$

$$\begin{aligned} \text{Var}[W_1 + W_2] &= \text{Var}[W_1] + \pi \text{Var}[W_2 | I] + (1 - \pi) \text{Var}[W_2 | I^c] + \pi(1 - \pi) \{E[W_2 | I^c] - E[W_2 | I]\}^2 \\ &\quad + 2\pi \{E[W_1] - E[W_1 | I]\} \times \{E[W_2 | I^c] - E[W_2 | I]\} \\ &= \frac{2[1 - \exp(\lambda_A \tau) + \lambda_A \tau] + (\lambda_A \tau)^2}{\lambda_A^2} + \frac{2\exp(\lambda_A \tau)}{(\lambda_V + \lambda_A)^2} \left\{ \frac{1}{\lambda_A} + \tau - \exp(\lambda_A \tau) \left[\frac{1}{\lambda_A} - \frac{1}{\lambda_V + \lambda_A} \right] \right\}^2 \\ &\quad + \left(1 - \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \right) \sigma(I)^2 + \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \sigma(I^c)^2 + \left(1 - \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \right) \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \delta^2 \\ &\quad + 2 \left\{ 1 - \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \right\} \left\{ \frac{1}{\lambda_A} - \frac{2\lambda_A + \lambda_V}{\lambda_A(\lambda_V + \lambda_A)} - \exp(\lambda_A \tau) \left[\frac{1}{\lambda_A} - \frac{1}{\lambda_V + \lambda_A} \right] \right. \\ &\quad \left. + \tau - \frac{\tau(\lambda_V + \lambda_A) - 1}{\lambda_A + \lambda_V [1 - \exp(\lambda_A \tau)]} \right\} \delta \end{aligned}$$

A.2 Variance for V-Version

A.2.1 $\tau \geq 0$:

$$P[I] = \pi = \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A}$$

$$E[W_1] = E[V] = \frac{1}{\lambda_V}$$

$$E[W_1 | I] = E[V | A + \tau < V] = \frac{\lambda_A(1 + \lambda_V \tau) + \lambda_V(2 + \lambda_V \tau)}{\lambda_V(\lambda_A + \lambda_V)}$$

$$E[W_1] - E[W_1 | I] = -\frac{1 + \tau(\lambda_V + \lambda_A)}{\lambda_V + \lambda_A}$$

$$\begin{aligned} \text{Var}[W_1 + W_2] &= \text{Var}[W_1] + \pi \text{Var}[W_2 | I] + (1 - \pi) \text{Var}[W_2 | I^c] + \pi(1 - \pi) \{E[W_2 | I^c] - E[W_2 | I]\}^2 \\ &\quad + 2\pi \{E[W_1] - E[W_1 | I]\} \times \{E[W_2 | I^c] - E[W_2 | I]\} \\ &= \frac{1}{\lambda_V^2} + \exp(-\lambda_V \tau) \frac{\lambda_A}{\lambda_V + \lambda_A} \sigma(I)^2 + \left[1 - \exp(-\lambda_V \tau) \frac{\lambda_A}{\lambda_V + \lambda_A} \right] \sigma(I^c)^2 \\ &\quad + \exp(-\lambda_V \tau) \frac{\lambda_A}{\lambda_V + \lambda_A} \left[1 - \exp(-\lambda_V \tau) \frac{\lambda_A}{\lambda_V + \lambda_A} \right] \delta^2 - \frac{2\delta \exp(-\lambda_V \tau) \lambda_A [1 + \tau(\lambda_A + \lambda_V)]}{(\lambda_V + \lambda_A)^2} \end{aligned}$$

APPENDIX (Continued)

A.2.2 $\tau \leq 0$:

$$P[I] = \pi = 1 - \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} = \frac{\lambda_V + \lambda_A - \exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A}$$

$$E[W_1] = E[V] = \frac{1}{\lambda_V}$$

$$E[W_1 | I] = E[V | A + \tau < V] = \frac{(\lambda_A + \lambda_V) \left(\frac{1}{\lambda_V} - \frac{\lambda_V \exp(\lambda_A \tau)}{(\lambda_A + \lambda_V)^2} \right)}{\lambda_A + \lambda_V - \exp(\lambda_A \tau) \lambda_V}$$

$$E[W_1] - E[W_1 | I] = - \frac{\exp(\lambda_A \tau) \lambda_A}{(\lambda_A + \lambda_V) [\lambda_A + \lambda_V - \exp(\lambda_A \tau) \lambda_V]}$$

$$\begin{aligned} \text{Var}[W_1 + W_2] &= \text{Var}[W_1] + \pi \text{Var}[W_2 | I] + (1 - \pi) \text{Var}[W_2 | I^c] + \pi(1 - \pi) \{E[W_2 | I^c] - E[W_2 | I]\}^2 \\ &\quad + 2\pi \{E[W_1] - E[W_1 | I]\} \times \{E[W_2 | I^c] - E[W_2 | I]\} \\ &= \frac{1}{\lambda_V^2} + \left[1 - \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \right] \sigma(I)^2 + \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \sigma(I^c)^2 \\ &\quad + \left[1 - \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \right] \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \delta^2 - \frac{2\delta \exp(\lambda_A \tau) \lambda_A}{(\lambda_V + \lambda_A)^2} \end{aligned}$$

A.3 Covariance for MIN-Version

A.3.1 $\tau \geq 0$:

$$\text{Cov}[W_1, W_2] = \left\{ \frac{\exp(-\lambda_V \tau) \lambda_A}{\lambda_V + \lambda_A} \right\} \left\{ \left[\frac{1}{\lambda_V} - \frac{1}{\lambda_V + \lambda_A} \right] [1 - \exp(-\lambda_V \tau)] - \tau \right\} \delta$$

A.3.2 $\tau \leq 0$:

$$\text{Cov}[W_1, W_2] = \left\{ 1 - \frac{\exp(\lambda_A \tau) \lambda_V}{\lambda_V + \lambda_A} \right\} \left\{ \frac{1}{\lambda_A} - \frac{2\lambda_A + \lambda_V}{\lambda_A(\lambda_V + \lambda_A)} - \exp(\lambda_A \tau) \left[\frac{1}{\lambda_A} - \frac{1}{\lambda_V + \lambda_A} \right] + \tau - \frac{\tau(\lambda_V + \lambda_A) - 1}{\lambda_A + \lambda_V [1 - \exp(\lambda_A \tau)]} \right\} \delta$$

A.4 Covariance for V-Version

A.4.1 $\tau \geq 0$:

$$\text{Cov}[W_1, W_2] = \frac{\delta \exp(-\lambda_V \tau) \lambda_A [1 + \tau(\lambda_A + \lambda_V)]}{(\lambda_V + \lambda_A)^2}$$

A.4.2 $\tau \leq 0$:

$$\text{Cov}[W_1, W_2] = \frac{\delta \exp(\lambda_A \tau) \lambda_A}{(\lambda_V + \lambda_A)^2}$$