

# Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time

ADELE DIEDERICH

*International University Bremen, Bremen, Germany*

and

HANS COLONIUS

*Oldenburg University, Oldenburg, Germany*

Manual reaction times to visual, auditory, and tactile stimuli presented simultaneously, or with a delay, were measured to test for multisensory interaction effects in a simple detection task with redundant signals. Responses to trimodal stimulus combinations were faster than those to bimodal combinations, which in turn were faster than reactions to unimodal stimuli. Response enhancement increased with decreasing auditory and tactile stimulus intensity and was a U-shaped function of stimulus onset asynchrony. Distribution inequality tests indicated that the multisensory interaction effects were larger than predicted by separate activation models, including the difference between bimodal and trimodal response facilitation. The results are discussed with respect to previous findings in a focused attention task and are compared with multisensory integration rules observed in bimodal and trimodal superior colliculus neurons in the cat and monkey.

The question of how information from different sensory modalities is combined in the brain and how sensation and perception change as a function of the presence or absence of a particular stimulus modality has interested researchers for almost a century. Administering combinations of light, tone, and electric shock stimuli to participants, Todd (1912) observed a reduction of reaction time (RT) when bimodal stimuli were presented, as compared with unimodal ones. His explanation was that “the more rapid responses to the simultaneous stimuli than to the single stimuli . . . is . . . due to a virtual increase in the extensity or intensity of the stimulus. . . . The three simultaneous stimuli summate in excitatory effect and send their discharge down one common tract to the reaction finger” (p. 63). Many subsequent studies have replicated this *redundant signals effect* (RSE) for cross-modal stimuli (e.g., Corballis, 1998; Diederich & Colonius, 1987; Giray & Ulrich, 1993; Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994; Miller, 1982, 1986; Plat, Praamstra, & Horstink, 2000; Schwarz & Ischebeck, 1994), mostly with visual–auditory stimulus combinations, but also for unimodal stimulus combinations (Marzi et al., 1996; Murray, Foxe, Higgins, Javitt, & Schroeder, 2001). Recently, visual–tactile combinations have been studied as well (Amlôt, Walker, Driver, & Spence, 2003; Dieder-

ich, Colonius, Bockhorst, & Tabeling, 2003; Forster, Cavina-Pratesi, Aglioti, & Berlucchi, 2002).

In his pioneering study, Todd (1912) proposed a kind of “energy summation” mechanism to underlie the response speedup, but subsequently many alternative explanations have been advanced (for a recent review, see Diederich & Colonius, 2004). Raab (1962) suggested that presenting a multimodal stimulus complex produces parallel separate activations in different sensory channels that build to the level at which they can produce a response and that the response is triggered by the signal that reaches that level first. Assuming statistical variability in channel processing times this *race*, or *separate activation*, model predicts shorter average RTs to multimodal stimuli than to unimodal stimuli (*statistical facilitation*). For bimodal stimulus combinations, Miller (1982) developed a distribution inequality test of the separate activation model that has been shown to be violated in many subsequent RSE experiments (e.g., Mordkoff & Miller, 1993). An alternative to the separate activation concept is to assume some coactivation process in which the unimodal signals’ activations are first combined to jointly trigger a response (Grice, Canham, & Boroughs, 1984; Miller, 1982). Several coactivation models capture many aspects of observed RT data sets, at least at the level of means (Diederich, 1995; Schwarz, 1989, 1994).

However, the exact processing level at which coactivation might occur has yet to be determined. Whereas some studies have supported the idea that a substantial amount of the RSE may arise in the motor component (Diederich & Colonius, 1987; Giray & Ulrich, 1993),

---

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

more recent analyses of event-related potentials (ERPs; Mordkoff, Miller, & Roch, 1996; Murray et al., 2001; Schröger & Widmann, 1998), of latency changes in the activity of single area 4 neurons in the motor cortex of the monkey (Miller, Ulrich, & Lamarre, 2001), and of RTs with the stop-signal paradigm (Cavina-Pratesi, Bricolo, Prior, & Marzi, 2001) provide evidence against late motoric explanations of coactivation.

From neurophysiological single-cell recordings, several sites in the human brain have been identified as being involved in multisensory integration. Specifically, multisensory neurons in the superior colliculus (SC) of anesthetized cats (Meredith & Stein, 1986a, 1986b; Stein, Magalhães-Castro, & Kruger, 1976) and monkeys (Wallace, Wilkinson, & Stein, 1996) show an enhanced response to particular combinations of visual, auditory, and tactile stimuli paralleling the speedup observed in RT studies. Similar results have recently been obtained for recordings from unanesthetized cats by Wallace, Meredith, and Stein (1998) and from the awake behaving monkey by Bell, Corneil, Meredith, and Munoz (2001) and by Frens and Van Opstal (1998). Moreover, multisensory integration properties of most SC neurons, as well as observed orientation behavior, are mediated by influences from two cortical areas—the anterior ectosylvian sulcus (AES) and the rostral aspect of the lateral suprasylvian sulcus (rLS; Jiang, Jiang, & Stein, 2002; Jiang, Wallace, Jiang, Vaughan, & Stein, 2001).

The purpose of this study is to contribute to the strengthening of the tie between observable RT behavior and the underlying neural circuitry. Our point of departure is the little-noticed fact that, in the study by Todd (1912), trimodal stimulus combinations yielded RTs that not only were shorter than those to unimodal stimuli, but also were shorter than RTs to any bimodal stimulus combination. Although this trimodal facilitation effect may, at first sight, seem to be a rather trivial extension of the well-known bimodal effects, it will be shown here that a quantitative assessment of trimodal facilitation leads to some interesting new results and hypotheses. The first issue that arises is whether the effect of adding a third-modality stimulus can be predicted from the known bimodal effects on the basis of either statistical facilitation or coactivation. In particular, if trimodal facilitation is larger than predicted by statistical facilitation, this would suggest the presence of a specific trimodal interaction effect, in analogy to the occurrence of a trivariate interaction term in the analysis of variance (ANOVA) model. This test will be carried out employing a distribution inequality developed in Diederich (1992). A possible neural basis for such an interaction would be provided by the existence of *trimodal* multisensory neurons that are sensitive to visual, auditory, and somatosensory stimulation simultaneously. Although a strict categorization of multisensory neurons tends to be difficult, the proportion of trimodal neurons in the cat SC, for example, is estimated at 8% (see Stein & Meredith, 1993).<sup>1</sup>

The second issue concerns the effects on cross-modal RT from varying (1) the temporal alignment of stimuli from different modalities and (2) their intensity levels. For example, the speedup of responses to a visual–auditory stimulus tends to be maximal when the visual stimulus precedes the auditory by an interval that equals the difference in RT between unimodal responses to the visual and the auditory stimuli (Diederich & Colonius, 1987; Hershenson, 1962; Miller, 1986). Moreover, facilitation is sometimes found to be larger for low-intensity stimuli (Bernstein, Chu, & Briggs, 1973). This phenomenon, referred to as *inverse effectiveness*, is often seen as the major purpose of multisensory integration from a functional point of view (see Meredith & Stein, 1986b).

Except for Todd's (1912) study, this article seems to be the first to investigate to what degree these integration rules generalize to the interaction among three modalities. The results will be discussed with respect to their significance for the issue of separate activation versus coactivation. Given that some intriguing parallels between bimodal behavioral data and temporal integration rules at the level of individual multisensory neurons have been found (Meredith, Nemitz, & Stein, 1987), our findings in the trimodal setting will also suggest new hypotheses at the neuronal level.

## EXPERIMENT

In Todd's (1912) experiments, participants were instructed to respond to a specific target modality (visual, auditory, or somatosensory) when presented with a combined bimodal, or trimodal, stimulus set (*focused attention* paradigm). He did not find significant RT differences when the instruction was to respond to the stimulus set as a whole (*redundant target* paradigm). However, a host of subsequent studies revealed significant effects of instruction and context on cross-modal RTs (for an early review, see Nickerson, 1973). For example, the introduction of catch trials, in which the nontarget stimulus is occasionally presented alone and the participant must *not* respond if it is, usually inflates RTs somewhat because, presumably, the participant must not only detect the stimulus (set), but also discriminate and/or identify the stimulus modalities (see Gielen, Schmidt, & Van den Heuvel, 1983). Note that increased unimodal RTs may lead to larger estimates of the facilitation effect, since the latter is measured with respect to unimodal speed measures. Context effects have been observed, for example, by varying the composition of blocks of stimulus trials. When multimodal and unimodal stimulus conditions are mixed within a block, responses to the latter are usually somewhat delayed (Nickerson, 1973). In addition, factors directing attention to a specific modality or dividing attention among modalities have been shown to be involved in various context effects. In a speeded spatial discrimination task with visual, auditory, and tactile targets, Spence, Nicholls, and Driver (2001) observed

that participants responded more rapidly to tactile targets when attention was directed to the tactile modality in advance (by presenting the majority of targets in the tactile modality within a block) than when attention was divided between the modalities (by presenting equal numbers of targets in each modality). Besides this endogenous shift-of-attention effect, Spence et al. observed an exogenous attention shift: Participants responded more rapidly to tactile targets when they were preceded by another tactile target on the previous trial than when they were preceded by either a visual or an auditory target.

Since in this study we were primarily interested in the more low-level aspects of multisensory integration, the experimental design was set up in such a way that the effects of attention due to instruction or context were minimized. Therefore, a redundant target paradigm, rather than a focused attention instruction, was chosen, so that the participants were not required to discriminate between targets and nontargets but simply to respond to the detection of any stimulus combination. Moreover, stimulus modalities were blocked—that is, within a block of trials (20, 60, or 64 trials for uni-, bi-, and trimodal conditions, respectively), the stimuli never changed with respect to the modalities occurring, but only with respect to the stimulus onset asynchronies (SOAs). Thus, within a block, no division of attention among different modalities and no shift to a particular modality were required or encouraged. Consequently, we would expect RTs to be somewhat shorter than under mixed-block conditions and to obtain rather conservative estimates for facilitation from multisensory integration. If RT speedup is still larger than predicted from separate activation, this will be even more evidence for some contribution from a coactivation mechanism.

## Method

**Participants.** Four students (3 females) served as paid voluntary participants in the experiment. They gave informed consent prior to their inclusion in the study.

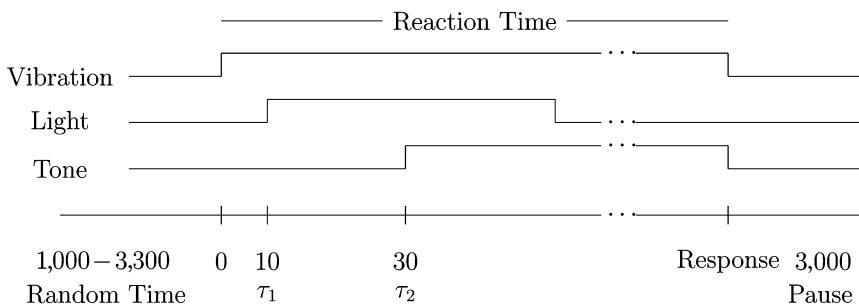
**Apparatus and Stimulus Presentation.** The visual stimulus consisted of a flash (250 lux) of 62.5-msec duration projected onto a screen (1 × 1.2 m) at a distance of 3.5 m. Sinusoidal tones of 1000 Hz and 70, 80, and 90 dB SPL steps served as the auditory

stimuli. They were presented diotically over closed headphones. The tactile stimuli were vibrations (sinusoidal voltage [27.7 Hz] at three levels—40, 145, and 448 mV) generated by an oscillation exciter (Mini-shaker, Type 4810, B & K). On this oscillation exciter, an aluminum plate (35 × 50 mm) was mounted on a threaded bolt. The aluminum plate served as a toe rest. For damping purposes, the oscillation exciter was encapsulated in a wooden box, which was mounted on a rubber foil and which had a little hole for the bolt with the aluminum plate. To suppress initial noise, the voltage for the vibration exciter was generated by a computer and started always at zero level. The computer signal was smoothed by a low-pass filter and was amplified according to the three intensities. The range of intensities included the maximum level below the participants' hearing threshold and the minimum excitation above the tactile threshold. For technical reasons, the flash always lasted 62.5 msec; the tone and the tactile stimulus were response terminated. The entire experiment (timing of the signals, RT measurement, randomization of trials, and calculation of the time between the attention signal and the first stimulus) was controlled by computer.

**Experimental procedure.** The participant was seated in a dark, sound-proofed room, 2.5 m away from the screen. Two response buttons mounted on a desk in front of him or her served as response device. The right foot (without shoe and socks) of the participant was put on the wooden box, and the big toe was put on the aluminum plate. To prevent it from getting out of place, the foot was strapped on. The participant was instructed to press the response buttons with the index finger as soon as any stimulus was detected (redundant target paradigm).<sup>2</sup>

The beginning of each trial was announced by an LED in front of the screen. A single stimulus or a stimulus combination was presented after a randomly generated time interval of 1,000 msec up to 3,300 msec. RT was recorded from onset of the first stimulus to keypressing for each hand separately. There was a 3-sec interval between the response and the signal for the beginning of the next trial. If a response occurred prior to stimulus onset, a rattle indicated this anticipatory response, and a repetition of the trial followed. Very few anticipatory responses actually occurred (about 0.002%).

Unimodal stimuli were visual stimuli (V), auditory stimuli with 70, 80, or 90 dB SPL (labeled  $A_{70}$ ,  $A_{80}$ , and  $A_{90}$ , respectively), and tactile stimuli with three different intensities (labeled  $T_1$ ,  $T_2$ ,  $T_3$ , where  $T_3$  was the most intense stimulus). Bimodal stimuli were combinations of any two stimuli from two different modalities. For example, the visual stimulus followed by a 70-dB tone  $\tau$  msec later is labeled  $V\tau A_{70}$ . The range of SOA for a particular stimulus combination was determined so that it included the range of RT differences to the respective unimodal stimuli, resulting in six different intervals in steps of 10 msec. Trimodal stimuli were combinations of three stimuli from different modalities—for example, a weak tactile stimulus followed by a visual stimulus  $\tau_1$  msec later and fol-



**Figure 1.** Time course for a trimodal stimulus with stimulus onset asynchrony (SOA)  $\tau_1$  between vibration and light and SOA  $\tau_2$  between vibration and tone.

lowed by a 70-dB tone  $\tau_2$  msec later, where  $\tau_2$  is always measured from the onset of the first stimulus (labeled  $T_1 \tau_1 V \tau_2 A_{70}$ ). Four different SOA values, in steps of 10 msec, were used for  $\tau_1$ , and four values were used for  $\tau_2$ . The time course for a trimodal stimulus trial with SOA  $\tau_1 = 10$  and  $\tau_2 = 30$  is depicted in Figure 1. To keep the number of conditions manageable, the presentation order of the modalities in the trimodal stimulus condition was fixed—that is, the tactile stimulus was always followed by the visual stimulus, followed by the auditory stimulus, and each participant was tested only on a (different) subset of the total number of conditions (see below). This presentation order in terms of increasing unimodal response speed was chosen in order to achieve a maximal temporal overlap of the modality-specific activation patterns. Table 1 summarizes all the experimental conditions.

The stimuli were presented in blocks of single-, double-, or triple-stimulus combinations. A block of single stimuli consisted of 20 trials for one modality. A block of double stimuli consisted of 60 trials for a particular stimulus combination. The SOAs between the stimuli were given in random order; each SOA was presented 10 times in a single block. A block of triple stimuli consisted of 64 trials for a particular stimulus combination. A specific SOA combination (between the first and the second and between the first and the third stimuli) was presented 4 times in each block. Each experimental session started with 10 triple-stimulus trials. This functioned as a warming-up period and was not used later in the data analysis; the intensities of the stimuli differed from those used in the experiment. Six to seven blocks were presented in a random balanced order in each session, including at least two single-, two double-, and two triple-stimulus blocks. After three experimental blocks, there was a rest period of about 10 min. Each session took between 60 and 75 min. By request, a participant was informed about his or her RT within a block. Including the training period (about 5 h), each participant was involved in the experiment for about 40 h. Each participant delivered 200 responses in each of 5 single-stimulus conditions; 100 responses in each of 48 double-stimulus conditions (8 different stimulus combinations and six different SOAs), and 100 responses in each of 48 triple-stimulus conditions (3 different stimulus combinations and 16 different SOA combinations). This resulted in a total of 10,600 responses collected from each participant.

**Results**

Since all the participants exhibited qualitatively similar results, across-subject data only are reported here.

**Mean unimodal RTs.** The mean response to the most intensive auditory stimulus was the fastest (132 msec); the mean response to the weakest tactile stimulus was the slowest (218 msec). Mean RTs and standard errors for the unimodal trials across all 4 participants are presented in Table 2. All means differed significantly ( $p < .01$ ).

**Mean bimodal and trimodal RTs.** Bimodal and trimodal stimuli produced shorter responses than did unimodal stimuli. Figure 2 shows mean RTs to bimodal stimuli (diamonds), as compared with unimodal stimuli (circles and squares), over SOA conditions. Standard errors are in the range of 0.6 to 1.9 msec, which are smaller than the markers and, therefore, omitted in the graphs. The longer unimodal mean RTs are presented by the horizontal line, and the shorter unimodal mean RTs plus the corresponding SOAs are represented by the linearly increasing line (note that the squares, for the longer RTs, and the circles, for the shorter RTs plus SOAs, are *not* observations but are included for a better comparison of these RTs with the bimodal stimulus means with respect to

**Table 1**  
**Experimental Conditions**

Unimodal Stimuli								
Tactile			Visual	Auditory				
$T_1$	$T_2$	$T_3$	V	$A_{70}$	$A_{80}$	$A_{90}$		
Bimodal Stimuli				SOA( $\tau$ )				
$V \tau A_{90}$			0	10	20	30	40	50
$V \tau A_{80}$			0	10	20	30	40	50
$V \tau A_{70}$			0	10	20	30	40	50
$T_3 \tau V$			0	10	20	30	40	50
$T_2 \tau V$			30	40	50	60	70	80
$T_1 \tau V$			30	40	50	60	70	80
$T_3 \tau A_{90}$			20	30	40	50	60	70
$T_3 \tau A_{80}$			10	20	30	40	50	60
$T_3 \tau A_{70}$			0	10	20	30	40	50
$T_2 \tau A_{90}$			30	40	50	60	70	80
$T_2 \tau A_{80}$			30	40	50	60	70	80
$T_1 \tau A_{90}$			50	60	70	80	90	100
$T_1 \tau A_{70}$			30	40	50	60	70	80
Trimodal Stimuli			SOA( $\tau_1$ )		SOA( $\tau_2$ )			
$T_3 \tau_1 V \tau_2 A_{80}$	0	10	20	30	30	40	50	60
$T_3 \tau_1 V \tau_2 A_{70}$	0	10	20	30	30	40	50	60
$T_2 \tau_1 V \tau_2 A_{90}$	10	20	30	40	40	50	60	70
$T_2 \tau_1 V \tau_2 A_{80}$	20	30	40	50	50	60	70	80
$T_1 \tau_1 V \tau_2 A_{90}$	30	40	50	60	60	70	80	90
$T_1 \tau_1 V \tau_2 A_{70}$	20	30	40	50	50	60	70	80

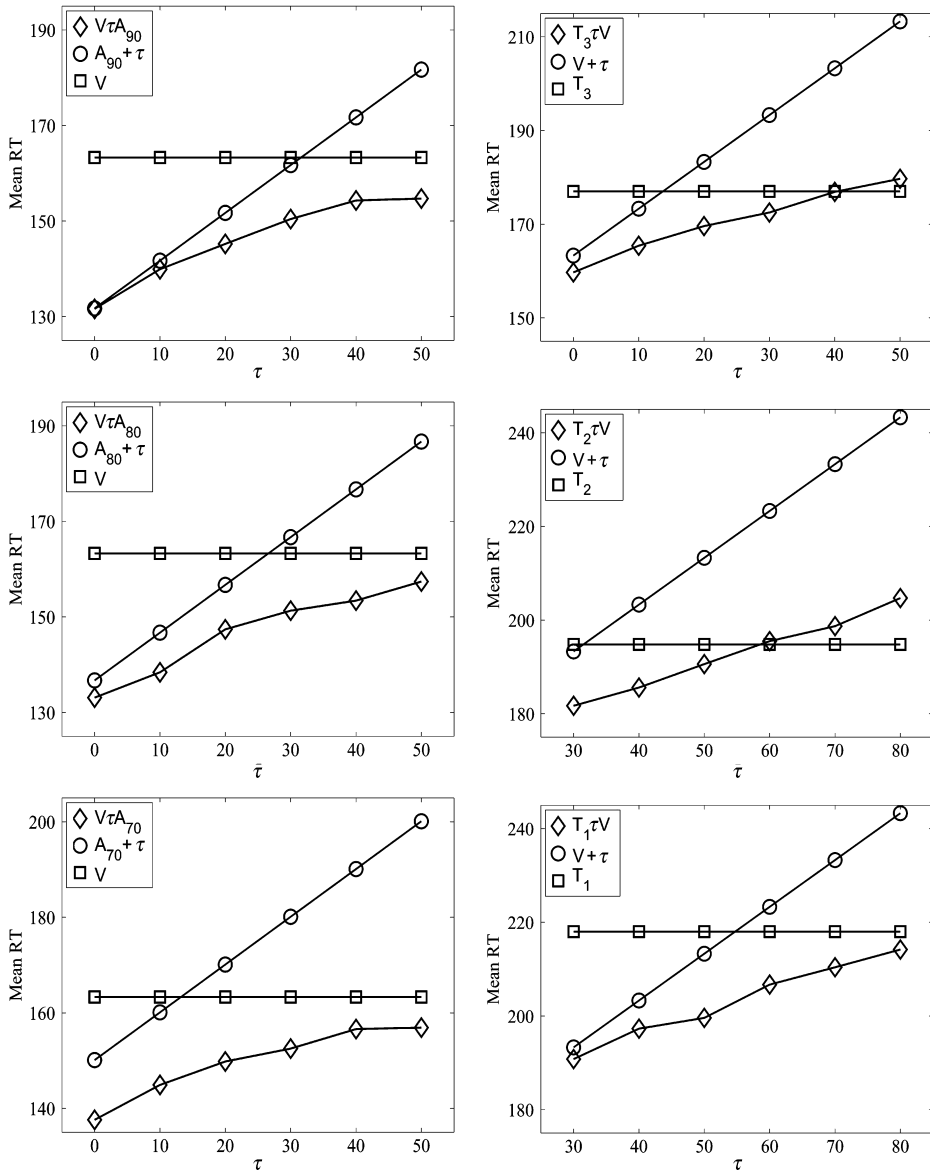
Note—Unimodal stimuli are labeled  $T_1$ ,  $T_2$ ,  $T_3$ , V,  $A_{70}$ ,  $A_{80}$ , and  $A_{90}$ , for the tactile stimuli with three intensities, the visual stimulus, and the auditory stimuli with 70, 80, or 90 db SPL, respectively. Bimodal stimuli are combinations of stimuli from different modalities, the second with a delayed onset time SOA( $\tau$ ). Trimodal stimuli include all three modalities, the second stimulus being delayed by  $\tau_1$  and the third delayed by  $\tau_2$ . Note that SOAs indicate the onset times with respect to the stimulus presented first.

SOA). For example, consider the first panel with stimulus combination  $V \tau A_{90}$ . Mean RT to light (V) is 163 msec, and mean RT to a 90-dB tone ( $A_{90}$ ) is 132 msec. These longer RTs (163 msec) are indicated by the parallel line (squares). The shorter RTs to tone (132 msec) are presented by the circles. In the bimodal condition, they were presented  $\tau$  msec later than the light. To demonstrate the later onset, the SOA is simply added to the mean RT of 132 msec—that is, plus 10 msec, plus 20 msec, and so forth, resulting in the linearly increasing line. Obviously, observed RTs to bimodal stimuli are shorter than those to unimodal stimuli whenever the diamonds do not cross

**Table 2**  
**Means and Standard Errors of Reaction Times for the Unimodal Trials (Averaged Over Participants)**

Stimulus	Mean	Standard Error
$A_{90}$	132	0.6
$A_{80}$	137	0.9
$A_{70}$	150	1.1
V	163	0.7
$T_3$	177	0.8
$T_2$	195	1.2
$T_1$	218	1.2

Note—The means are based on 400 ( $A_{80}$ ,  $A_{70}$ ,  $T_2$ , and  $T_1$ ) and 800 ( $A_{90}$ , V, and  $T_3$ ) observations.



**Figure 2.** Observed mean response times (RTs, diamonds) to bimodal stimuli as a function of the stimulus onset asynchrony (SOA)  $\tau$ . Each panel represents different stimulus combinations. The horizontal line indicates the mean RT to the more slowly processed stimulus in each condition; the linearly increasing line refers to the SOA-shifted mean to the more quickly processed stimulus. Note that the squares (for the longer RTs) and the circles (for the shorter RTs plus SOAs) are *not* observations but are included for a better comparison of these RTs with the bimodal stimulus means with respect to SOA.

either of those two lines. This is, indeed, the case, except for some SOAs in the  $T_3V$  and  $T_2V$  conditions.

Mean bimodal RTs (diamonds) are shorter than mean unimodal RTs in almost all stimulus and SOA conditions. Exceptions are the mean RTs to tactile–visual stimuli for large SOAs. Furthermore, the amount of RT reduction depends on the stimulus combination and, within the stimulus condition, on SOA.

Figure 3 shows mean trimodal RTs, as compared with mean bimodal RTs, with SOAs ( $\tau_1$ ) between the first (tactile) and the second (visual) stimuli on the  $x$ -axis.

For all the SOA ( $\tau_1$ ) conditions, mean trimodal RTs are shorter than mean RTs to tactile–visual stimulation with the corresponding SOAs. Similarly, mean trimodal RTs are shorter than mean RTs to tactile–auditory stimuli for all the SOA ( $\tau_2$ ) conditions (between the first and

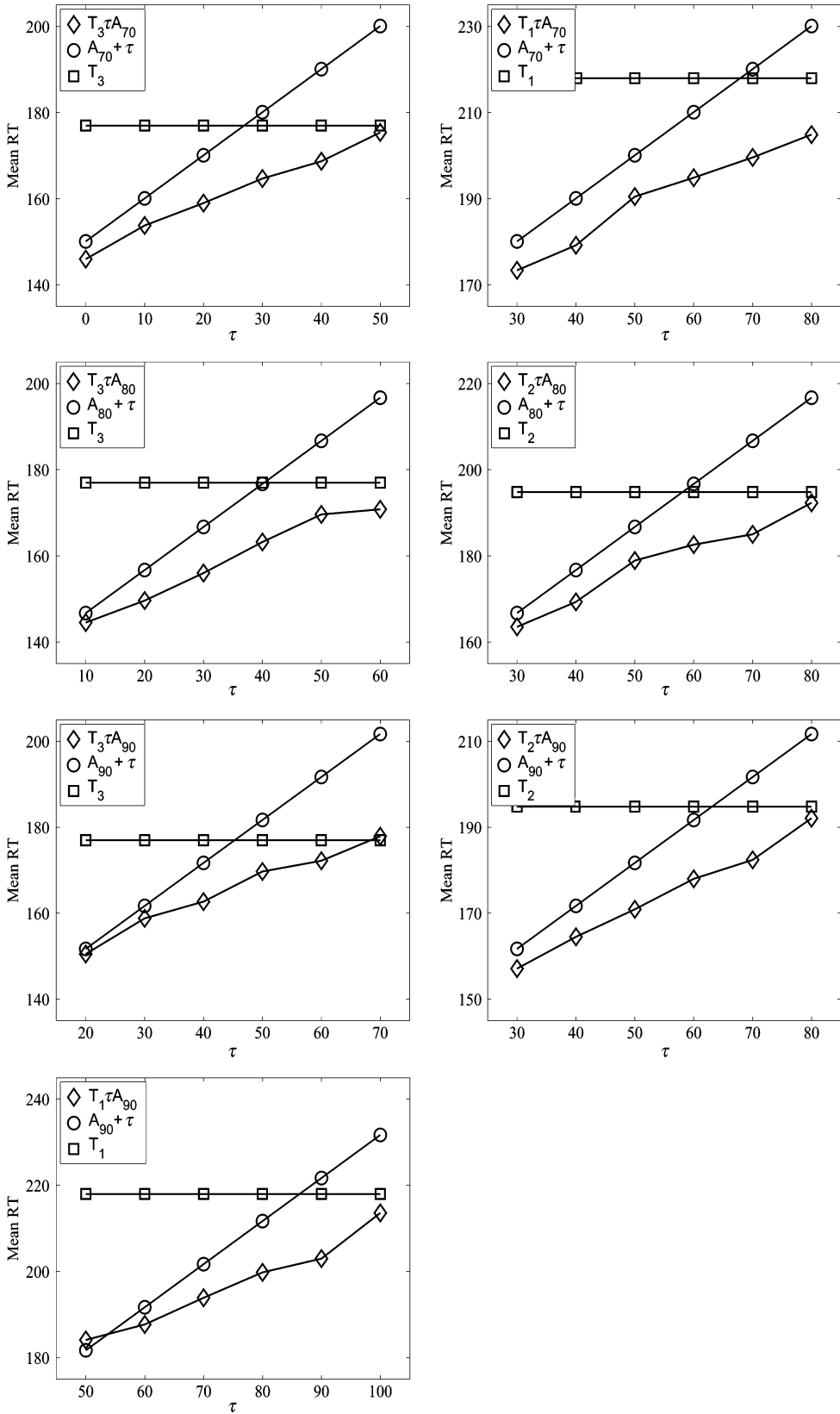
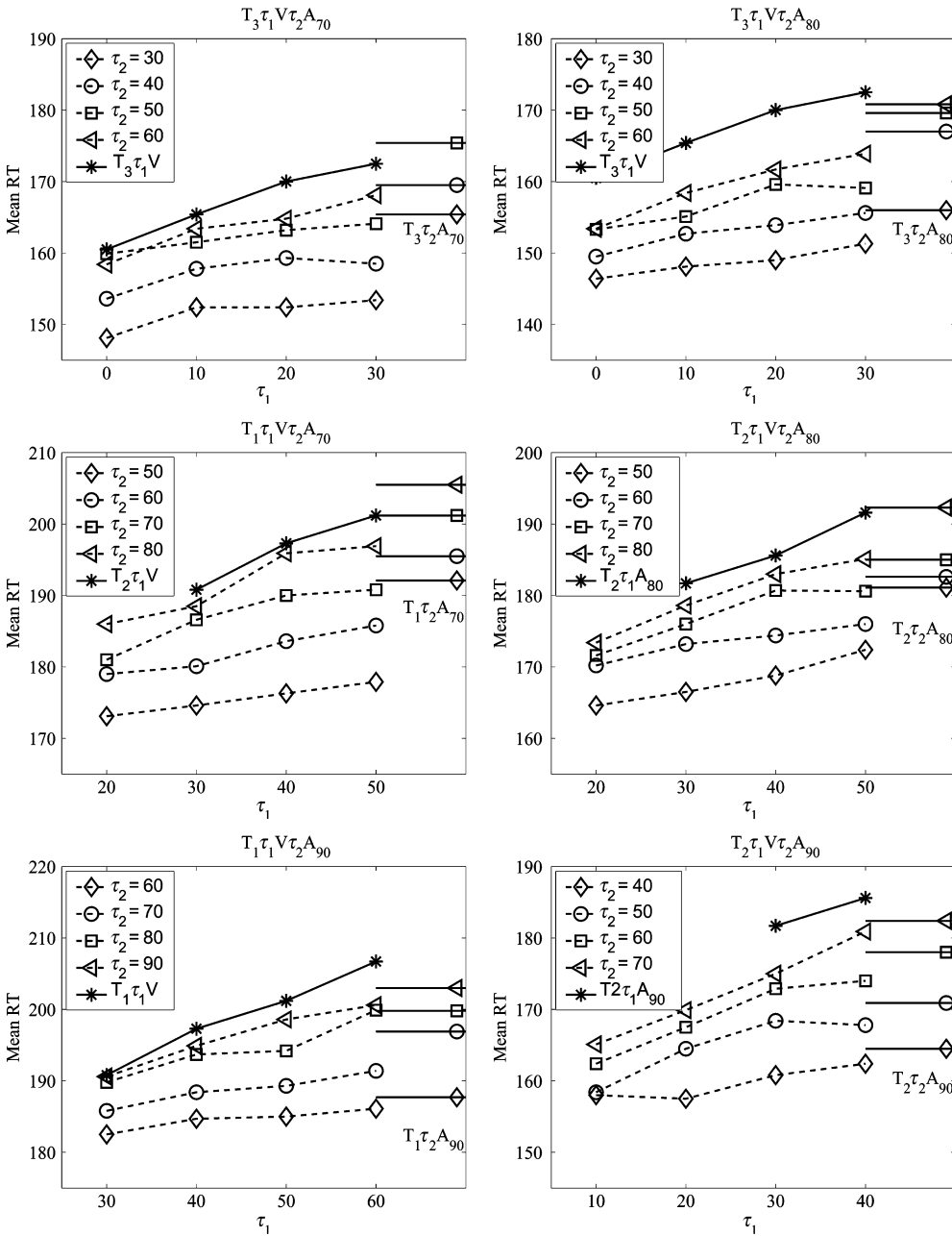


Figure 2 (Continued).



**Figure 3.** Observed mean response times (RTs) to trimodal stimuli as a function of stimulus onset asynchrony (SOA) between the first (tactile) and the second (visual) presented stimuli  $\tau_1$ . Symbols (diamonds, circles, squares, and triangles) refer to SOA ( $\tau_2$ ) conditions between the first (tactile) and the third (auditory) stimuli. The stars indicate the mean RTs to bimodal stimuli (tactile–visual) as a function of  $\tau_1$ . Mean RTs to tactile–auditory stimulation are presented at the far right of each panel (symbols with horizontal line). Different symbols refer to different SOA conditions for the bimodal stimuli. Note that these are the same SOAs as the second SOAs ( $\tau_2$ ) in the trimodal stimulus conditions. The horizontal lines are included only for a better comparison of the bimodal and the trimodal RTs. Consider, for example,  $T_130V60A_{70}$ . The RT is represented by the second circle on the dashed line and is compared with the RT in condition  $T_130V$  (first star of solid line) and with the RT to  $T_160A_{70}$  (circle with horizontal line on the far right). Each panel represents different stimulus intensity combinations.

the third stimuli). Standard errors are between 0.8 and 1.3 msec, again smaller than the markers. Note that RT to trimodal stimuli is an increasing function of both  $\tau_1$  and  $\tau_2$  for nearly all the stimulus conditions.

**Multisensory response enhancement.** A convenient way to assess the amount of intersensory facilitation or inhibition independently of the specific stimulus combination is to use a measure of response enhancement that

relates mean RT in the multimodal conditions to that in the unimodal condition. The following measure<sup>3</sup> quantifies the percentage of RT enhancement in analogy to a measure proposed for measuring multisensory enhancement in neural responses (Anastasio, Patton, & Belkacem-Boussaid, 2000; Colonius & Diederich, 2002; Meredith & Stein, 1986b).

$$MRE = \frac{\min(\overline{RT}_T, \overline{RT}_{V+\tau_1}, \overline{RT}_{A+\tau_2}) - \overline{RT}_{T\tau_1V\tau_2A}}{\min(\overline{RT}_T, \overline{RT}_{V+\tau_1}, \overline{RT}_{A+\tau_2})} \times 100,$$

where  $\overline{RT}_T$ ,  $\overline{RT}_V$ , and  $\overline{RT}_A$  refer to the observed mean RT to tactile, visual, or auditory unimodal stimuli, respectively.  $\tau_1$  and  $\tau_2$  are the SOAs between the tactile and the visual stimulus and between the tactile and the auditory stimulus in the trimodal stimulus condition.  $\overline{RT}_{T\tau_1V\tau_2A}$  is the mean RT to the corresponding trimodal stimulus combination. For example, an MRE value of 10 means that RT to the multimodal stimulus is reduced by 10%, as compared with the fastest mean response to any of the unimodal stimuli. Thus, the mean RT to the trimodal stimulus combination T30V50A, with  $\tau_1 = 30$  msec between the tactile and the visual stimulus and  $\tau_2 = 50$  msec between the tactile and the auditory stimulus, is related to the minimum RTs to T, V plus 30 msec, and A plus 50 msec. MRE for bimodal stimuli is defined accordingly.

Consider first the MRE to bimodal stimulus combinations. How does the RT facilitation depend on SOA? According to Raab (1962), in the separate activation model, maximum facilitation is expected when the SOA between the stimuli equals the time difference to the unimodal RTs (which he called “physiological synchronicity”), because this guarantees maximum overlap of the single-stimulus distributions. Thus, one would expect MRE to be the largest when the SOA is close to the difference in the mean RT of the unimodal stimuli. Table 3 presents the mean RT differences. Table 4 contains all MRE values as a function of bimodal stimulus condition (rows) and of SOA value (columns). Entries where the RT difference is closest to the SOA value are in boldface. As can be seen, boldface values are always maximal within a row, indicating that, indeed, the maximum amount of enhancement occurred for those bimodal stimulus combinations in which the SOA was closest to the difference in unimodal RTs.

The MREs for trimodal stimuli are larger than those for bimodal stimuli. Table 5 (MRE columns) shows the

**Table 3**  
Response Time (RT) Difference (in Milliseconds) Between Unimodal Stimuli

	V	A <sub>70</sub>	A <sub>80</sub>	A <sub>90</sub>
T <sub>1</sub>	55	68	81	—
T <sub>2</sub>	32	—	58	63
T <sub>3</sub>	14	27	40	45
V	—	13	26	31

Note—For example, the difference between the mean RTs to a tactile stimulus (T<sub>1</sub>) and a visual stimulus (V) is 55 msec.

**Table 4**  
Multisensory Response Enhancement (MRE) for All Bimodal Stimulus and Stimulus Onset Asynchrony (SOA) Conditions

Stimulus Condition	SOA (τ)										
	0	10	20	30	40	50	60	70	80	90	100
VτA <sub>90</sub>	<i>0</i>	<i>1.4</i>	<i>4.6</i>	<b>7.4</b>	<i>5.5</i>	<i>4.9</i>	—	—	—	—	—
VτA <sub>80</sub>	<i>2.9</i>	<i>6.1</i>	<i>6.4</i>	<b>7.4</b>	<i>6.1</i>	<i>3.7</i>	—	—	—	—	—
VτA <sub>70</sub>	<i>8.0</i>	<b>9.4</b>	<i>8.0</i>	<i>6.8</i>	<i>3.7</i>	<i>3.7</i>	—	—	—	—	—
T <sub>3</sub> τV	<i>1.8</i>	<b>4.6</b>	<i>4.0</i>	<i>2.3</i>	<i>0</i>	<i>-1.7</i>	—	—	—	—	—
T <sub>2</sub> τV	—	—	—	<b>5.7</b>	<i>4.6</i>	<i>2.1</i>	<i>0</i>	<i>-2.1</i>	<i>-5.1</i>	—	—
T <sub>1</sub> τV	—	—	—	<i>1.0</i>	<i>3.0</i>	<b>6.1</b>	<i>5.0</i>	<i>3.7</i>	<i>1.8</i>	—	—
T <sub>3</sub> τA <sub>90</sub>	—	—	<i>1.0</i>	<i>1.9</i>	<b>5.2</b>	<i>4.0</i>	<i>2.8</i>	<i>-1</i>	—	—	—
T <sub>3</sub> τA <sub>80</sub>	—	<i>2.0</i>	<i>4.5</i>	<i>6.6</i>	<b>7.9</b>	<i>4.0</i>	<i>3.4</i>	—	—	—	—
T <sub>3</sub> τA <sub>70</sub>	<i>2.7</i>	<i>3.8</i>	<i>6.5</i>	<b>6.8</b>	<i>4.5</i>	<i>1.1</i>	—	—	—	—	—
T <sub>2</sub> τA <sub>90</sub>	—	—	—	<i>3.1</i>	<i>4.7</i>	<i>6.0</i>	<b>7.3</b>	<i>6.7</i>	<i>1.5</i>	—	—
T <sub>2</sub> τA <sub>80</sub>	—	—	—	<i>1.8</i>	<i>4.5</i>	<i>4.3</i>	<b>6.2</b>	<i>5.1</i>	<i>1.5</i>	—	—
T <sub>1</sub> τA <sub>90</sub>	—	—	—	—	—	<i>-1</i>	<i>2.1</i>	<i>4.0</i>	<i>5.7</i>	<b>6.9</b>	<i>1.8</i>
T <sub>1</sub> τA <sub>70</sub>	—	—	—	<i>3.9</i>	<i>5.8</i>	<i>5.0</i>	<i>7.1</i>	<b>8.6</b>	<i>6.0</i>	—	—

Note—MREs in boldface indicate the largest expected enhancement according to the physiological synchronicity hypothesis. MREs in italics are based on nonsignificant mean differences between response times to bimodal stimuli and the smaller of the unimodal stimuli, including SOA adjustments.

MRE for all the trimodal stimulus and the SOA conditions. Moreover, MRE tends to be largest when both SOAs are close to the rounded difference in RT to unimodal stimuli. Boldface MRE values indicate the largest expected MRE increases according to the physiological synchronicity hypothesis. However, since not all SOA combinations correspond to unimodal RT differences, the results are not as clear cut as those for the bimodal stimuli.

Note, however, that the definition of MRE relates trimodal RT only to unimodal RT. In order to assess a specific effect of trimodal stimuli over and above the bimodal effects, we introduce a modified measure of multisensory response enhancement, MRE\*, relating trimodal RT to bimodal RTs from the corresponding bimodal stimulus conditions<sup>4</sup> as noted in Equation 1 at the bottom of the next page. For example, mean RT to stimulus T20V50A is related to RT to stimulus T20V, T50A, and to V30A +20 msec. Thus, MRE\* indicates the additional enhancement when a third stimulus is added to a bimodal stimulus condition. In other words, MRE\* measures a reduction in RT beyond the one measured by MRE. The results for 93 trimodal conditions<sup>5</sup> are shown in Table 5. Of these, 54 MRE\* values for the trimodal stimulus conditions are based on significant mean RT differences between trimodal RT and the minimum of the corresponding bimodal RTs (values based on nonsignificant differences have entries in italic).

With a few exceptions, the MRE\* values suggest that presenting a third stimulus caused a further enhancement of the responses.

**Effects of intensity.** It is well known that RT typically decreases with increasing stimulus intensity (Kohfeld, 1971; Piéron, 1914), and this is confirmed in our unimodal mean RTs for the different intensity levels for the auditory and the tactile stimuli (see Table 2). On the



**Table 5**  
**Multisensory Response Enhancement (MRE) for All**  
**Trimodal Stimulus and Stimulus Onset Asynchrony Conditions**

		$\tau_2$							
		30		40		50		60	
$\tau_1$	MRE	MRE*	MRE	MRE*	MRE	MRE*	MRE	MRE*	
$T_3\tau_1V\tau_2A_{80}$	0	10.4	3.3	9.2	3.3	8.6	5.1	7.4	5.6
	10	10.2	3.8	<b>11.6</b>	5.0	11.0	5.5	9.8	5.5
	20	8.4	1.9	12.4	4.9	9.6	5.9	10.2	6.5
	30	8.4	1.9	10.7	3.1	8.5	4.7	7.3	4.1
$T_3\tau_1V\tau_2A_{70}$	0	9.2	2.6	6.7	3.2	6.7	3.2	6.1	4.4
	10	<b>11.0</b>	3.8	8.7	2.5	8.1	3.6	8.1	3.6
	20	10.6	3.0	9.0	4.7	7.9	4.1	7.3	3.5
	30	10.7	4.2	7.9	3.6	6.8	4.6	5.1	2.9
$T_2\tau_1V\tau_2A_{90}$	10	8.1	1.3	9.2	4.3	6.9	2.4	6.4	—
	20	8.1	3.7	9.3	2.9	8.2	3.4	8.2	4.0
	30	5.8	1.2	7.7	1.8	<b>9.9</b>	2.8	10.4	4.9
	40	5.5	0.1	6.6	0.1	8.9	1.7	7.7	1.1
$T_2\tau_1V\tau_2A_{80}$	20	9.8	3.5	9.3	4.0	7.7	4.5	6.0	—
	30	11.9	4.0	<b>10.4</b>	4.4	9.8	4.4	8.8	4.4
	40	8.0	3.5	10.7	3.3	11.3	1.1	10.3	0.5
	50	7.5	3.4	9.1	2.2	11.6	1.1	13.1	3.1
$T_1\tau_1V\tau_2A_{90}$	30	5.2	-1.1	4.1	-0.6	4.1	0	3.6	2.6
	40	3.1	-0.5	6.9	1.1	6.9	2.6	5.9	3.1
	50	1.6	-1.1	4.5	0.1	8.5	3.0	7.0	1.0
	60	0.5	-1.6	4.0	0	6.1	0.1	7.8	0.1
$T_1\tau_1V\tau_2A_{70}$	20	5.5	0	4.9	1.7	3.8	0.6	3.3	—
	30	7.3	0.6	6.7	1.1	5.2	2.1	4.7	1.6
	40	11.3	2.7	8.4	2.1	7.4	2.1	6.4	3.6
	50	13.1	1.6	11.7	3.6	8.5	2.5	7.5	1.5

Note—MRE and MRE\* in italics are based on nonsignificant mean differences between the RTs to trimodal stimuli and to the smaller of the unimodal stimuli, including SOA adjustments.

other hand, RT facilitation, when measured with respect to unimodal RT (as in the MRE index), may not be a monotonically decreasing function of the intensity of the stimuli involved.

For bimodal stimuli, MRE decreases with increasing stimulus intensity, as is suggested by the *inverse effectiveness* hypothesis. For visual–auditory stimuli with a 70-dB tone, the maximum MRE is 9.4, as compared with 7.4 when the intensity of the tone is increased to 90 dB. Similarly, for tactile–visual stimuli, MRE is larger (6.1) with weak tactile stimuli ( $T_1$ ) than with strong tactile

stimuli ( $T_3$ ; MRE = 4.6). Finally, the most intense tactile–auditory combination ( $T_3A_{90}$ ) provides less enhancement (MRE = 5.2) than does the weakest ( $T_1A_{70}$ , MRE = 8.6).<sup>6</sup>

For trimodal stimuli, MRE decreases with increasing auditory intensity for  $T_2VA_{80}$  and  $T_2VA_{90}$  (MRE = 13.1 vs. MRE = 10.4), as for  $T_1VA_{70}$  and  $T_1VA_{90}$  (MRE = 13.1 vs. MRE = 8.5), but MRE increases with increasing auditory intensity for  $T_3VA_{70}$  and  $T_3VA_{80}$  (MRE = 11 vs. MRE = 12.4) and with increasing tactile intensity for  $T_1VA_{90}$  and  $T_2VA_{90}$  (MRE = 8.5 vs. MRE = 10.4).

### Discussion

Overall, a consistent pattern of effects on response speed emerges from our data. For unimodal stimuli, mean RTs to auditory stimuli were shorter than those to visual stimuli, which in turn were shorter than those to tactile RTs. Moreover, RT decreased with increasing auditory and tactile intensity. Mean response speed to bimodal and trimodal stimuli was faster than that to unimodal stimuli under almost all conditions. Specifically, bimodal RT increased monotonically, with SOA approaching unimodal RT when the second stimulus was presented late enough. Interestingly, trimodal RTs were still shorter than bimodal RTs, but increased monotonically with both SOA values.

In assessing the multisensory effects by relating bimodal and trimodal RTs to unimodal RTs, the MRE values displayed an inverted U-shaped function of SOA. Intriguingly, maximum MRE values were reached consistently for those SOA values that maximized the overlap of the unimodal RT distributions, supporting the *physiological synchronicity* hypothesis. Although this effect can be predicted by separate activation models (see Raab, 1962), it is also consistent with the general concept of coactivation: If stimuli from several modalities combine to produce a response, this will be most effective in the time range in which the activations from all modalities are most likely to be present (see Miller, 1986). Thus, a deeper analysis is required (see the next section). The majority of MRE\* values comparing bimodal with trimodal enhancement suggest a specific trimodal interaction effect, but again, whether this goes beyond what is expected from separate activation needs further scrutiny.

For bimodal stimuli, the intensity effects on response enhancement were unequivocal in exhibiting larger MRE values for lesser intensity levels (inverse effectiveness), whereas for trimodal stimuli the effect of intensity was less clear cut. A discussion of these results will be postponed to the final section.

### TESTING FOR SEPARATE ACTIVATION

It has become common usage to probe the status of the separate activation hypothesis by examining certain inequalities for the observable RT distribution functions

$$MRE^* = \frac{\min(\overline{RT}_{T_1V}, \overline{RT}_{T_2A}, \overline{RT}_{V(\tau_1-\tau_2)A} + \tau_1) - \overline{RT}_{T_1V\tau_2A}}{\min(\overline{RT}_{T_1V}, \overline{RT}_{T_2A}, \overline{RT}_{(\tau_2-\tau_1)A} + \tau_1)} \times 100 \quad (1)$$

(e.g., Forster et al., 2002; Miller, 1982, 1986; Townsend & Nozawa, 1995). Here, we go beyond this practice (1) in extending the test to three modalities and (2) in developing a specific method to test for a trimodal effect that is not predictable from the bimodal response enhancements.

Separate activation models assume the following: (1) Processing in different sensory channels occurs in parallel. (2) Presenting multimodal stimuli (e.g., light, tone, and vibration) produces separate activations in each of the corresponding channels and activation builds to the level at which it can produce a response. (3) The response is elicited by the signal that terminates processing first. Assuming statistical variability in the channel processing times, separate activation models predict shorter average RTs to multimodal stimuli than to unimodal stimuli, because the average of the winner's processing time is smaller than the average processing time in each single channel (statistical facilitation). For example, with **T**, **V**, and **A** denoting the random processing time for the tactile, visual, and auditory stimuli, respectively,

$$E[\min(\mathbf{T}, \mathbf{V}, \mathbf{A})] \leq \min[E(\mathbf{T}), E(\mathbf{V}), E(\mathbf{A})], \quad (2)$$

where *E* indicates the expected value of the random variables. If the RT reduction can be explained in terms of statistical facilitation, no assumption about an additional neural mechanism is needed. Miller (1982) suggested a distribution inequality test that since then has become the standard test for separate activation models:

$$\begin{aligned} P(RT_{VA} \leq t) &= P[\min(\mathbf{V}, \mathbf{A}) \leq t] \\ &= P(\mathbf{V} \leq t) + P(\mathbf{A} \leq t) - P(\mathbf{V} \leq t \ \& \ \mathbf{A} \leq t) \\ &\leq P(RT_V \leq t) + P(RT_A \leq t) \quad \text{for all } t, \end{aligned} \quad (3)$$

where  $RT_{VA}$ ,  $RT_V$ , and  $RT_A$  refer to the observable RTs with the respective bimodal and unimodal stimuli, as before. All separate activation models can be rejected if more facilitation is observed than what is consistent with this inequality. Note, however, that nonviolation of Inequality 3 does not provide unequivocal evidence in favor of separate activation models (Colonus, 1986, 1990, 1999; Diederich & Colonus, 1987; Miller, 1982; Ulrich & Giray, 1986). In most studies of bimodal interaction, Miller's test has provided evidence against statistical facilitation as the sole source of the response speedup (see Diederich & Colonus, 2004), and consequently, the upper bound of Inequality 3 has become a benchmark for assessing the amount of "true" facilitation (see Townsend & Nozawa, 1995).

Extending the test of separate activation models to three modalities, it should first be noted that the amount of statistical facilitation increases with the number of channels involved in the race. For example, in the case of the three variables above, it is obvious that  $\min(\mathbf{T}, \mathbf{V}, \mathbf{A}) \leq$

$\min(\mathbf{T}, \mathbf{V})$  is true, implying that  $E[\min(\mathbf{T}, \mathbf{V}, \mathbf{A})] \leq E[\min(\mathbf{T}, \mathbf{V})]$ , and the same holds with the pairs **T**, **A** and **V**, **A** instead of **T**, **V**. Thus, statistical facilitation with triple stimuli is at least as large as that with double stimuli.

On the level of the distributions, a straightforward extension of Inequality 3 to the trimodal case leads to the following test, where the SOAs between the first and the second stimuli ( $\tau_1$ ) and between the first and the third stimuli ( $\tau_2$ ) are also taken into account as shown in Equation 4 at the bottom of the page. This inequality puts an upper bound on the statistical facilitation produced by trimodal stimuli. Note that violations of Inequality 3 and Inequality 4, if any, are to be expected only for small values of *t*, because for large *t*, the right-hand side approaches 2 and 3, respectively, yielding the trivial inequalities  $1 < 1 + 1$  and  $1 < 1 + 1 + 1$ . Moreover, a null hypothesis in the form of inequalities involving the sum of two or more probabilities is not amenable to standard nonparametric testing procedures, and therefore, we settle for a descriptive evaluation of the evidence.

When at least three modalities are involved, there are various ways of obtaining further inequalities (Worsley, 1982). A generalized test of separate activation for multimodal stimuli including subsets of stimulus combinations was developed by Diederich (1992). For the special case of trimodal stimuli, three inequalities relating the observable RT distributions in uni-, bi-, and trimodal stimulus conditions can be derived. One of them is shown in Equation 5 at the top of the next page, and the other two follow by combining other pairs of modalities analogously.<sup>7</sup> Note that inequalities of this type relate the trimodal RT distribution to bimodal and unimodal distributions. Most important, these inequalities can be utilized to test whether response enhancement for trimodal stimuli is larger than that predicted from bimodal enhancement under the separate activation hypothesis.<sup>8</sup> In other words, a violation of this inequality indicates that there is a specifically trimodal enhancement effect over and above the bimodal enhancement that is not compatible with separate activation.

Applying Inequality 3 twice in the right-hand side of Inequality 5, it is easy to see that the right-hand side of Inequality 4 also constitutes an upper bound to the right-hand side of Inequality 5. Thus, we have the chain of inequalities shown in Equation 6 at the bottom of the next page. Note that violations of the second of the inequalities in Equation 6 would indicate that the summed bimodal interactions between touch–vision and vision–audition is larger than those predicted by separate activation.<sup>9</sup>

### Results

Figure 4 shows bimodal RT distributions (diamonds), together with the upper bound of Inequality 3 (circles) obtained from the sum of two unimodal RT distributions

$$\begin{aligned} P(RT_{T\tau_1V\tau_2A} \leq t) &= P[\min(\mathbf{T}, \mathbf{V} + \tau_1, \mathbf{A} + \tau_2) \leq t] \\ &\leq P(RT_T \leq t) + P(RT_V \leq t - \tau_1) + P(RT_A \leq t - \tau_2). \end{aligned} \quad (4)$$

$$\begin{aligned}
P(RT_{T\tau_1V\tau_2A} \leq t) &= P[\min(\mathbf{T}, \mathbf{V} + \tau_1, \mathbf{A} + \tau_2) \leq t] \\
&\leq P[\min(\mathbf{T}, \mathbf{V} + \tau_1) \leq t] + P[\min(\mathbf{V} + \tau_1, \mathbf{A} + \tau_2) \leq t] - P(\mathbf{V} + \tau_1 \leq t) \\
&= P(RT_{T\tau_1V} \leq t) + P[RT_{V(\tau_2-\tau_1)A} \leq t - \tau_1] - P(RT_V \leq t - \tau_1),
\end{aligned} \tag{5}$$

for various stimulus combinations and selected SOA conditions. All the distributions were obtained by Vincentizing across the 4 subjects.

The largest violations of the inequality occur for visual–auditory stimuli, and the smallest ones occur when  $T_3$ , the strongest tactile stimulus, is involved ( $T_3V$ ,  $T_3A_{70}$ ,  $T_3A_{90}$ ). This pattern is similar to the corresponding MRE values, although not so pronounced. Note, however, an important difference between the measures: MRE is a global measure of enhancement accumulated across the entire distribution (mean), whereas the inequality compares percentiles locally. MRE does not indicate whether the enhancement occurred by statistical or by “true” facilitation, whereas the inequality may rule out statistical facilitation as the sole source of the response speedup.

Figure 5 shows trimodal RT distributions (stars), together with the upper bound of Inequality 5 (diamonds) obtained from the sum of two unimodal distributions minus a unimodal distribution, and with the upper bound of Inequality 4 (circles) obtained from the sum of the three unimodal RT distributions, for selected SOA combinations with maximum MRE and MRE\* values in Tables 3 and 5, respectively.

Inequality 4 is clearly violated for all the conditions. A complete reversal of Inequality 6, indicating an effect of adding the third modality, is clearly apparent in a single condition ( $T_320V60A_{80}$ ) only. Conditions  $T_230V60A_{80}$ ,  $T_320V40A_{70}$ , and  $T_140V90A_{90}$  show but minimal violations of the inequality. Nevertheless, these conditions all exhibit maximum MRE\* values across their SOA conditions. Moreover, a lack of reversal should not be interpreted to indicate that the third modality had no effect: A ceiling effect, possibly caused by very strong bimodal interactions, may conceal it. It cannot be excluded, finally, that a more fine-grained net of SOA values would have generated more conclusive evidence of violation.

Overall, the distribution inequality tests presented strong evidence against the separate activation model. However, specific evidence that adding a third modality stimulus results in a response enhancement that is larger than that predicted by the bimodal enhancement effects was found in one condition only.

## GENERAL DISCUSSION

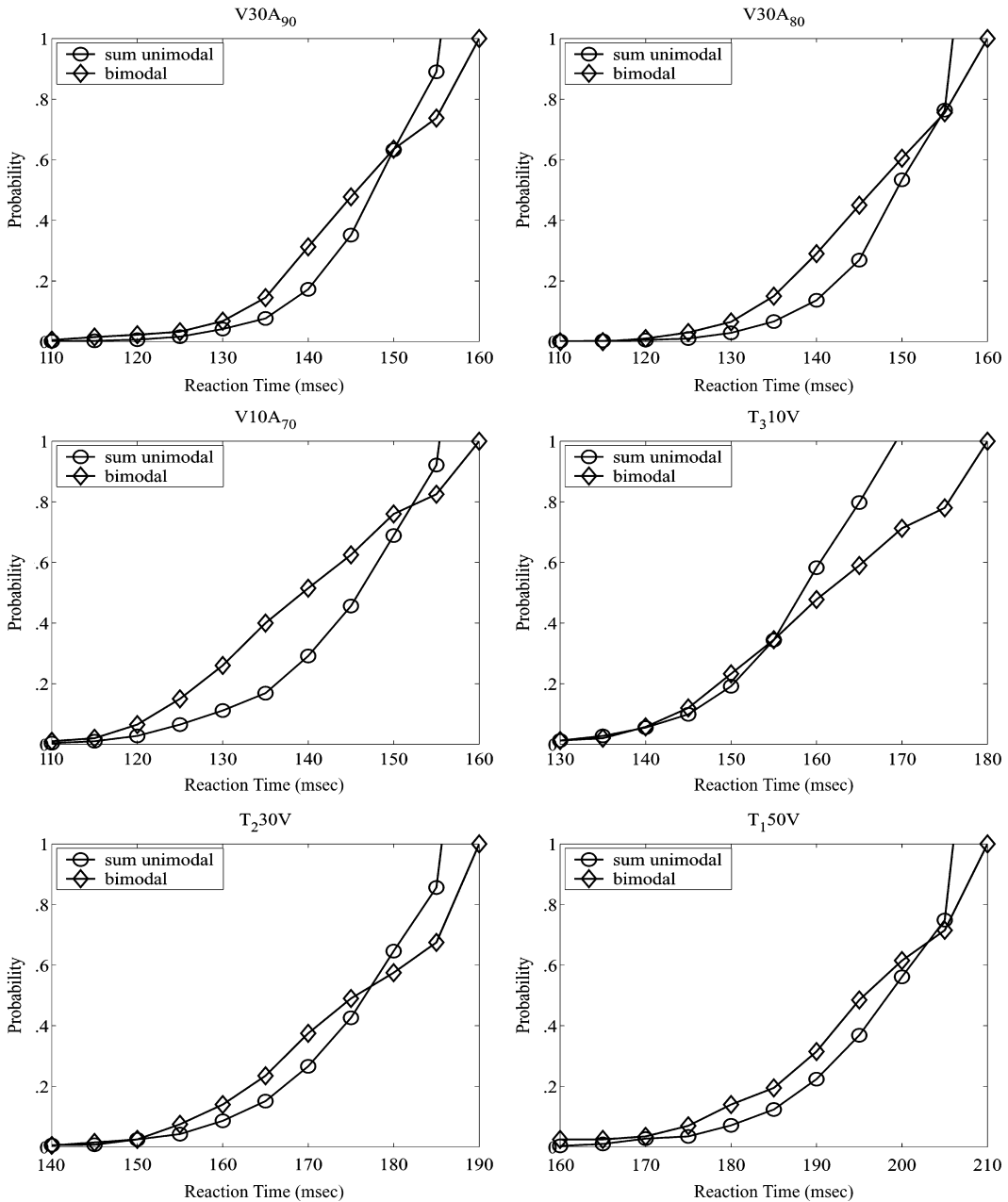
Our main experimental finding is a trimodal multisensory enhancement effect on RT. There is some indi-

cation that the effect existed over and above the bimodal enhancement effects. Second, varying the temporal alignment of the stimuli resulted in a rather regular pattern of SOA effects on response enhancement, demonstrating the importance of *physiological synchronicity* for maximum enhancement to show up. Third, decreasing intensity of the auditory or the tactile stimulus increased multisensory enhancement in the bimodal combinations, supporting the *inverse effectiveness* hypothesis.

We assessed the size of the enhancement effects by a set of distribution inequality tests leading to the unambiguous result that the observed RT enhancement was larger than that predicted by separate activation models. These tests also suggested that the increase of enhancement from bimodal to trimodal responses was larger than that predicted from the bimodal effects on the basis of separate activation. However, further empirical studies to probe this hypothesis are called for. Our results favor an explanation of the multisensory enhancement effects in terms of a coactivation mechanism that combines activations from the different modalities to jointly trigger a response. As was outlined in the introduction, this model captures some important features of many data sets quantitatively, but we are only beginning to unravel which neural substrates in the human brain are involved in coactivation. If our finding of a specific trimodal enhancement effect is borne out by further empirical evidence, this would hint at a possible contribution from multisensory neurons that are sensitive to visual, auditory, and somatosensory stimulation simultaneously. Although this type of neuron has been identified in the cat SC, it has been classified as *trimodal* only because it is responsive to different pairs of bimodal stimulation—that is, to both visual–auditory and auditory–tactile stimulus combinations—but no simultaneous trimodal stimulations have been performed, to our knowledge (see Stein & Meredith, 1993). On the other hand, a trimodal–bimodal enhancement difference in RT suggests the existence of trimodal neurons that should show an increased response activity for simultaneous trimodal stimulation, notwithstanding the practical difficulties of such an experimental setup.

It must be noted, however, that the separate activation concept (race model) cannot be dismissed entirely, for several reasons. First, responses faster than those predicted by the race model may indicate only that, in addition to *statistical facilitation*, other mechanisms for re-

$$\begin{aligned}
P(RT_{T\tau_1V\tau_2A} \leq t) &\leq P(RT_{T\tau_1V} \leq t) + P[RT_{V(\tau_2-\tau_1)A} \leq t - \tau_1] - P(RT_V \leq t - \tau_1) \\
&\leq P(RT_T \leq t) + P(RT_V \leq t - \tau_1) + P(RT_A \leq t - \tau_2)
\end{aligned} \tag{6}$$



**Figure 4.** Testing separate activation with Inequality 3 for all the bimodal stimulus conditions for selected stimulus onset asynchronies. Diamonds refer to the cumulative distribution to bimodal stimuli, and circles refer to the sum of two cumulative distributions to unimodal stimuli (upper bound in Inequality 3).

sponse speedup may be effective, including interchannel crosstalk from interstimulus contingencies (see Mordkoff & Yantis, 1993) and warning signal or attention effects (see Spence, 2002); for other recent quantitative proposals for the RSE that partially retain the race concept, see Miller and Ulrich (2003), Colonius and Arndt (2001), and Colonius and Diederich (2004). Second, the behavior of multisensory neurons is complex, defying any simple extrapolation from changes in spike rate ac-

tivity to RT changes. Consider, for example, a bimodal (visual–auditory) stimulus condition. A faster reaction to a bimodal stimulus cannot be attributed solely to the *existence* of a bimodal neuron, because a bimodal neuron is typically responsive to unimodal input as well. Thus, an RT speedup must be a function of an *enhanced* response of the bimodal neurons to the visual–auditory stimulus. The amount of the latter, however, critically depends on external variables, such as the spatiotemporal

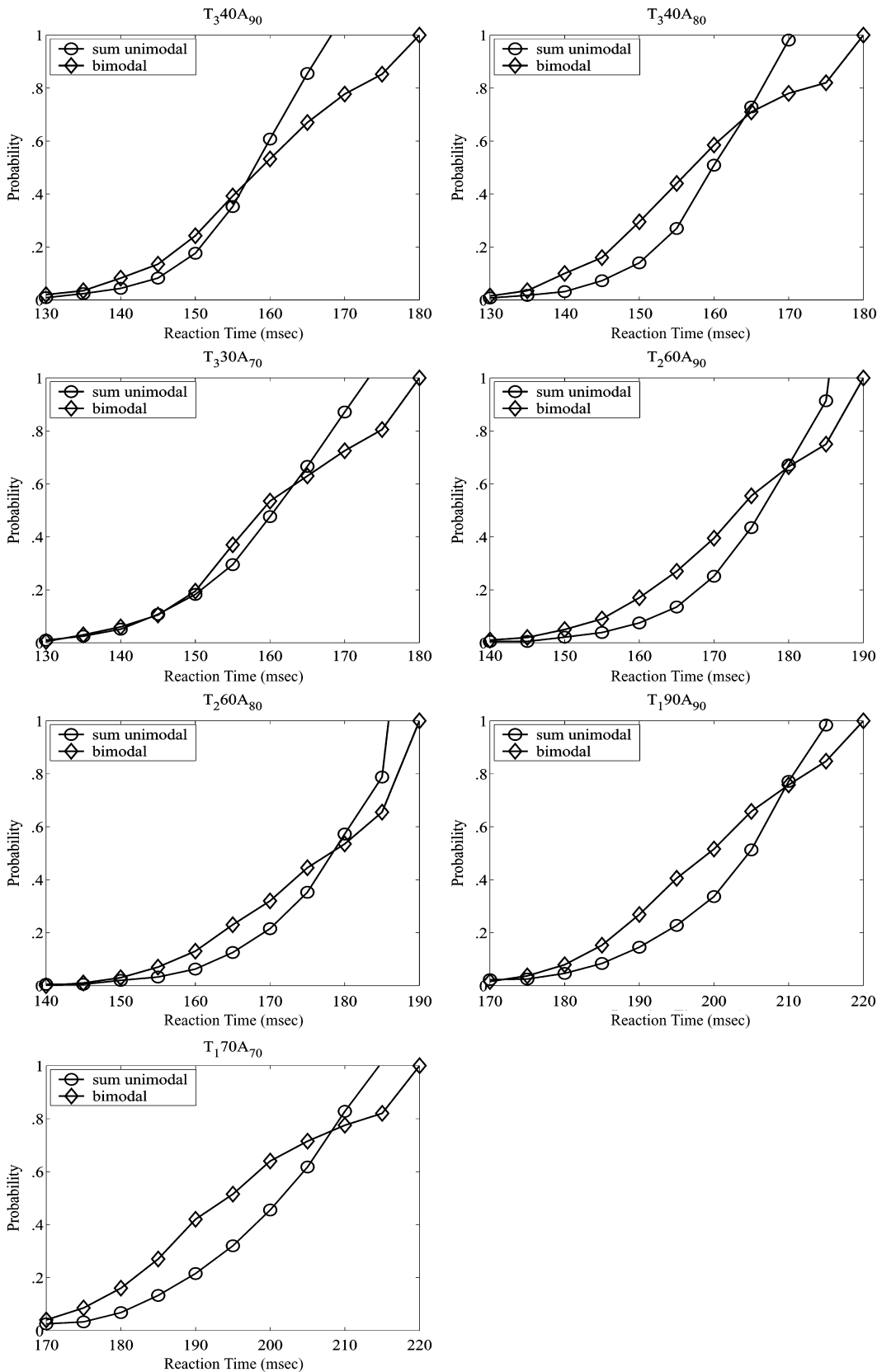
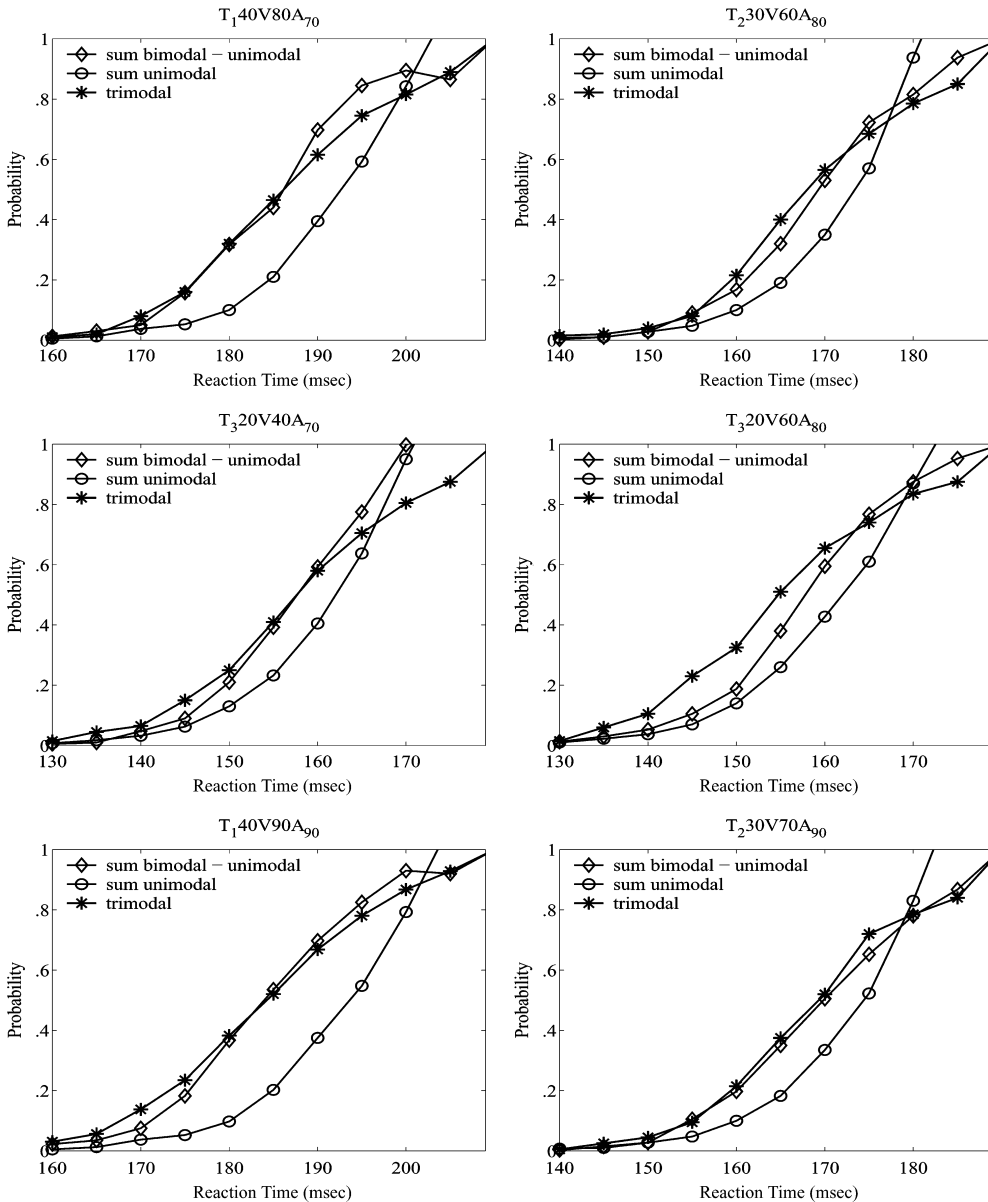


Figure 4 (Continued).



**Figure 5.** Testing separate activation with Inequality 4 and Inequality 5 for all trimodal stimulus conditions for selected stimulus onset asynchronies. Stars refer to the cumulative distribution to trimodal stimuli, diamonds refer to the sum of two cumulative distributions to bimodal stimuli minus a unimodal cumulative distribution (upper bound in Inequality 5), and circles refer to the sum of three cumulative distributions to unimodal stimuli (upper bound in Inequality 4).

stimulus configuration and stimulus intensities (see below). Third, it is well known that individual neurons can differ widely in their spatiotemporal response properties, and observable behavior must be seen as the result of a large number of neurons contributing to the response from different sites of the brain.

Temporal integration rules similar to those shown in our data have also been observed for responses at the level of individual multisensory neurons. Bimodal cells in the cat SC show maximal levels of response enhance-

ment when the peak discharge periods evoked by each modality overlap in time (Meredith et al., 1987). This overlap is not correlated with simultaneous stimulus presentation, but with time differences in stimulus presentation that match their latency differences, again similar to the findings reported here. Moreover, Meredith and Stein (1986b) observed that for bimodal neurons in the SC of the cat, response enhancement was largest for visual-auditory stimulus combinations. The mean number of impulses increased by up to 1,207% over that

elicited by the most effective unimodal stimulus—namely, the visual stimulus. Somatosensory–auditory and somatosensory–visual stimuli produced far less response enhancement (79% and 75%, respectively), paralleling our behavioral findings in which enhancement tends to be largest for visual–auditory stimuli and smallest for visual–tactile stimuli.

Although these correspondences between the behavioral and the neurophysiological data are striking, there are also some discrepancies that need explanation. Even taking into account that the quantitative measures of multisensory enhancement for RT and for average neural firing rates are not identical, the size of the RT effects appears, in general, to be less dramatic than that of the neural enhancement effects. It is possible that the RT effects measured here do underestimate the true size of the effect. As was mentioned above, due to the blocking of different stimulus conditions, our measures of RT enhancement effects must be considered to be conservative. In addition, it is well known that MRE decreases with increasing spatial disparity between the stimuli from different modalities (Colonus & Arndt, 2001; Frens, Van Opstal, & Van der Willigen, 1995; Harrington & Peck, 1998; Hughes, Nelson, & Aronchick, 1998; Spence & Driver, 1997a, 1997b). Due to limitations of our experimental setup, the stimuli from different modalities could not be presented at the same spatial location, possibly contributing to an underestimation of the true effect as well. It should be noted, however, that this spatial factor is not likely to undermine our general conclusions, since the spatial configuration has been kept constant across all conditions, and presenting all the stimuli from the same spatial position probably would have resulted in a general elevation of the enhancement effects.

Another discrepant finding is the occurrence of a massive response depression (as compared with unimodal stimulation) in multisensory SC neurons under significant spatial disparity (e.g., Kadunce, Vaughan, Wallace, Benedek, & Stein, 1997; Meredith & Stein, 1986a), whereas a similar lengthening of human RT to disparate spatial stimulation has not generally been observed (e.g., Frens et al., 1995). Note that in our experiment, auditory stimuli were presented over headphones with identical input to either ear, so visual–auditory distance was not well defined, whereas visual–tactile distance was significant, because the tactile stimulus was applied to the toe. Interestingly, the only cases of significant response depression occurred for bimodal tactile–visual stimulus combinations (see Table 4). Systematic variations of spatial stimulus conditions have been investigated mainly with saccadic eye movements under the focused attention paradigm, and the lack of RT depression has usually been attributed to a general warning signal effect of the nontarget stimulus, to the effects of expecting a specific modality, or to other strategies (e.g., Corneil & Munoz, 1996; Frens et al., 1995). However, recently, Amlôt et al. (2003) observed increased saccadic RT when visual distractors were presented opposite to somatosensory targets.

Response depression in the orientation behavior of the cat under disparate cross-modal stimulation has been demonstrated by Stein, Meredith, Huneycutt, and McDade (1989). Lately, it has been shown that response depression in multisensory neurons is not a function of spatial disparity per se. Two cross-modal (visual–auditory) stimuli, whether or not they were widely disparate, produced the same response enhancement as long as they were located within the visual–auditory receptive field overlap of the neuron (Kadunce, Vaughan, Wallace, & Stein, 2001). Given that behavior must be seen as the result of a large number of neurons with different receptive field properties, any simple extrapolation from these findings to RTs seems futile, however. We conclude that more systematic variation of spatial stimulus configurations in a redundant signals task is needed to identify the conditions for response depression to occur in RTs.

Our finding that response enhancement increases for low-intensity bimodal stimulus combinations (inverse effectiveness) stands out, relative to the findings in most previous studies. Although this effect is considered an important, ubiquitous property of multisensory neurons (see Stein & Meredith, 1993), human studies have not demonstrated the expected increase of response enhancement when stimulus intensity is lowered (Arndt & Colonus, 2003; Frens et al., 1995; Hughes et al., 1994). From an ecological point of view, inverse effectiveness would be most important for an animal in a potentially hostile environment: Near-threshold stimuli in several modalities would combine to create an above-threshold warning signal. This may be one reason why it has not shown up in most behavioral studies with humans. Corneil, Van Wanrooij, Munoz, and Van Opstal (2002) suggested some further reasons. First, in many studies, the stimuli may not have been close enough to threshold. They were not close to threshold in this study, either, which may explain why we did find a consistent, but not a dramatic, effect of inverse effectiveness. Second, most previous studies were of the focused attention type, so that participants had to discriminate between target and nontarget modalities, possibly prolonging the RTs. The study by Corneil et al. provides some support for an explanation along this line. In a redundant signals saccadic response task, they presented unimodal and cross-modal visual–auditory targets at various spatial positions embedded in an audiovisual background noise. Varying auditory intensity, they found multisensory response enhancement to decrease with an increasing signal-to-noise ratio, as predicted by inverse effectiveness. More studies with near-threshold stimulus combinations are needed to determine whether inverse effectiveness at the level of the neural data can be obtained in behavioral settings.

To conclude, despite many caveats and discrepant findings, it seems to us that the study of multisensory integration rules in RT can benefit from relating it to findings at the level of individual multisensory neurons or assemblies of such neurons. Further modeling at the behavioral level should also take into account a number of recent devel-

opments in those areas. For example, besides the common excitatory–excitatory type of multisensory convergence, there is evidence for an excitatory–inhibitory type of neural circuit by which inputs from one modality inhibit those from the other (see Meredith, 2002). Moreover, recent functional magnetic resonance imaging (fMRI) and ERP studies suggest that multisensory processing does not take place entirely in feedforward convergent pathways but that it can also modulate early cortical unisensory processing (see Laurienti et al., 2002, and Macaluso, Frith, & Driver, 2000, for fMRI; see McDonald, Teder-Sälejärvi, Di Russo, & Hillyard, 2003, and Molholm et al., 2002, for ERP).

## REFERENCES

- AMLÖT, R., WALKER, R., DRIVER, J., & SPENCE, C. (2003). Multimodal visual–somatosensory integration in saccade generation. *Neuropsychologia*, **41**, 1-15.
- ANASTASIO, T. J., PATTON, P. E., & BELKACEM-BOUSSAID, K. (2000). Using Bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Computation*, **12**, 1165-1187.
- ARNDT, A., & COLONIUS, H. (2003). Two separate stages in crossmodal saccadic integration: Evidence from varying intensity of an auditory accessory stimulus. *Experimental Brain Research*, **150**, 417-426.
- BELL, A. H., CORNEIL, B. D., MEREDITH, M. A., & MUNOZ, D. P. (2001). The influence of stimulus properties on multisensory processing in the awake primate superior colliculus. *Canadian Journal of Experimental Psychology*, **55**, 123-132.
- BERNSTEIN, I. H., CHU, P. K., & BRIGGS, P. (1973). Stimulus intensity and foreperiod effects in intersensory facilitation. *Quarterly Journal of Experimental Psychology*, **25**, 171-181.
- CAVINA-PRATESI, C., BRICOLO, E., PRIOR, M., & MARZI, C. A. (2001). Redundancy gain in the stop-signal paradigm: Implications for the locus of coactivation in simple reaction time. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 932-941.
- COLONIUS, H. (1986). Measuring channel dependence in separate activation models. *Perception & Psychophysics*, **40**, 251-255.
- COLONIUS, H. (1990). Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology*, **34**, 253-275.
- COLONIUS, H. (1999). A theorem on parallel processing models with a generalized stopping rule. *Mathematical Social Sciences*, **38**, 247-258.
- COLONIUS, H., & ARNDT, P. (2001). A two-stage model for visual–auditory interaction in saccadic latencies. *Perception & Psychophysics*, **63**, 126-147.
- COLONIUS, H., & DIEDERICH, A. (2002). A maximum-likelihood approach to modeling multisensory enhancement. In T. G. Dieterich, S. Becker, & Z. Ghahramani (Eds.), *Advances in neural information processing systems 14* (pp. 181-187). Cambridge, MA: MIT Press.
- COLONIUS, H., & DIEDERICH, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *Journal of Cognitive Neuroscience*, **16**, 1000-1009.
- CORBALLIS, M. C. (1998). Interhemispheric neural summation in absence of the corpus callosum. *Brain*, **121**, 1795-1807.
- CORNEIL, B. D., & MUNOZ, D. P. (1996). The influence of auditory and visual distractors on human orienting gaze shifts. *Journal of Neuroscience*, **16**, 8193-8207.
- CORNEIL, B. D., VAN WANROOIJ, M., MUNOZ, D. P., & VAN OPSTAL, A. J. (2002). Auditory–visual interactions subserving goal-directed saccades in a complex scene. *Journal of Neurophysiology*, **88**, 438-454.
- 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.
- DIEDERICH, A., & COLONIUS, H. (2004). Modeling the time course of multisensory interaction in manual and saccadic responses. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes* (pp. 395-408). Cambridge, MA: MIT Press.
- DIEDERICH, A., COLONIUS, H., BOCKHORST, D., & TABELING, S. (2003). Visual–tactile spatial interaction in saccade generation. *Experimental Brain Research*, **148**, 328-337.
- FORSTER, B., CAVINA-PRATESI, C., AGLIOTI, S. M., & BERLUCCHI, G. (2002). Redundant target effect and intersensory facilitation from visual–tactile interactions in simple reaction time. *Experimental Brain Research*, **143**, 480-487.
- 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.
- GRICE, G. R., CANHAM, L., & BOROUGHS, J. M. (1984). Combination rule for redundant information in reaction time tasks with divided attention. *Perception & Psychophysics*, **35**, 451-463.
- HARRINGTON, L. K., & PECK, C. K. (1998). Spatial disparity affects visual–auditory interactions in human sensorimotor processing. *Experimental Brain Research*, **122**, 247-252.
- HERSHENSON, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, **63**, 289-293.
- HUGHES, H. C., NELSON, M. D., & ARONCHICK, D. M. (1998). Spatial characteristics of visual–auditory summation in human saccades. *Vision Research*, **38**, 3955-3963.
- HUGHES, H. C., REÜTER-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.
- JIANG, W., JIANG, H., & STEIN, B. E. (2002). Two corticotectal areas facilitate orientation behavior. *Journal of Cognitive Neuroscience*, **14**, 1240-1255.
- JIANG, W., WALLACE, M. T., JIANG, H., VAUGHAN, J. W., & STEIN, B. E. (2001). Two cortical areas mediate multisensory integration in superior colliculus neurons. *Journal of Neurophysiology*, **85**, 506-522.
- KADUNCE, D. C., VAUGHAN, J. W., WALLACE, M. T., BENEDEK, G., & STEIN, B. E. (1997). Mechanisms of within- and cross-modality suppression in the superior colliculus. *Journal of Neurophysiology*, **78**, 2834-2847.
- KADUNCE, D. C., VAUGHAN, J. W., WALLACE, M. T., & STEIN, B. E. (2001). The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Experimental Brain Research*, **139**, 303-310.
- KOHFELD, D. L. (1971). Simple reaction time as a function of stimulus intensity in decibels of light and sound. *Journal of Experimental Psychology*, **88**, 251-257.
- LAURIENZI, P. J., BURDETTE, J. H., WALLACE, M. T., YEN, Y., FIELD, A. S., & STEIN, B. E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *Journal of Cognitive Neuroscience*, **14**, 420-429.
- MACALUSO, E., FRITH, C. D., & DRIVER, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, **289**, 1206-1208.
- MARZI, C. A., SMANIA, N., MARTINI, M. C., GAMBINA, G., TOMELLERI, G., PALAMARA, A., ALESSANDRINI, F., & PRIOR, M. (1996). Implicit redundant-targets effect in visual extinction. *Neuropsychologia*, **34**, 9-22.
- MCDONALD, J. J., TEDER-SÄLEJÄRVI, W. A., DI RUSSO, F., & HILLYARD, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, **15**, 10-19.
- MEREDITH, M. A. (2002). On the neural basis for multisensory convergence: A brief overview. *Cognitive Brain Research*, **14**, 31-40.
- MEREDITH, M. A., NEMITZ, J. W., & STEIN, B. E. (1987). Determinants



- of multisensory integration in superior colliculus neurons: I. Temporal factors. *Journal of Neuroscience*, **7**, 3215-3229.
- MEREDITH, M. A., & STEIN, B. E. (1986a). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research*, **365**, 350-354.
- MEREDITH, M. A., & STEIN, B. E. (1986b). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, **56**, 640-662.
- MILLER, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, **14**, 247-279.
- MILLER, J. O. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics*, **40**, 331-343.
- MILLER, J. O., & ULRICH, R. (2003). Simple reaction time and statistical facilitation: A parallel grains model. *Cognitive Psychology*, **46**, 101-151.
- MILLER, J. O., ULRICH, R., & LAMARRE, Y. (2001). Locus of the redundant-signals effect in bimodal divided attention: A neurophysiological analysis. *Perception & Psychophysics*, **63**, 555-562.
- MOLHOLM, S., RITTER, W., MURRAY, M. M., JAVITT, D. C., SCHROEDER, C. E., & FOXE, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, **14**, 115-128.
- MORDKOFF, J. T., & MILLER, J. O. (1993). Redundancy gains and coactivation with two different targets: The problem of target preferences and the effects of display frequency. *Perception & Psychophysics*, **53**, 527-535.
- MORDKOFF, J. T., MILLER, J. O., & ROCH, A. C. (1996). Absence of coactivation in the motor component: Evidence from psychophysiological measures of target detection. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 25-41.
- MORDKOFF, J. T., & YANTIS, S. (1993). Dividing attention between color and shape: Evidence of coactivation. *Perception & Psychophysics*, **53**, 357-366.
- MURRAY, M. M., FOXE, J. J., HIGGINS, B. A., JAVITT, D. C., & SCHROEDER, C. E. (2001). Visuo-spatial neural response interactions in early cortical processing during a simple reaction time task: A high-density electrical mapping study. *Neuropsychologia*, **39**, 828-844.
- NICKERSON, R. S. (1973). Intersensory facilitation of reaction time: Energy summation or preparation enhancement. *Psychological Review*, **80**, 489-509.
- PIÉRON, H. (1914). Recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices. *L'Année Psychologique*, **20**, 17-96.
- PLAT, F. M., PRAAMSTRA, P., & HORSTINK, M. W. (2000). Redundant-signals effects on reaction time, response force, and movement-related potentials in Parkinson's disease. *Experimental Brain Research*, **130**, 533-539.
- RAAB, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, **24**, 574-590.
- SCHRÖGER, E., & WIDMANN, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology*, **35**, 755-759.
- SCHWARZ, W. (1989). A new model to explain the redundant-signals effect. *Perception & Psychophysics*, **46**, 498-500.
- SCHWARZ, W. (1994). Diffusion, superposition, and the redundant-targets effect. *Journal of Mathematical Psychology*, **38**, 504-520.
- SCHWARZ, W., & ISCHEBECK, A. (1994). Coactivation and statistical facilitation in the detection of lines. *Perception*, **23**, 157-168.
- SPENCE, C. (2002). Multisensory attention and tactile information-processing. *Behavioural Brain Research*, **135**, 57-64.
- SPENCE, C., & DRIVER, J. (1997a). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, **59**, 1-22.
- SPENCE, C., & DRIVER, J. (1997b). On measuring selective attention to an expected sensory modality. *Perception & Psychophysics*, **59**, 389-403.
- SPENCE, C., NICHOLLS, M. E. R., & DRIVER, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, **63**, 330-336.
- STEIN, B. E., MAGALHÃES-CASTRO, B., & KRUGER, L. (1976). Relationship between visual and tactile representations in cat superior colliculus. *Journal of Neurophysiology*, **39**, 401-419.
- STEIN, B. E., & MEREDITH, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
- STEIN, B. E., MEREDITH, M. A., HUNEYCUTT, W. S., & MCDADE, L. (1989). Behavioral indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience*, **1**, 12-24.
- TODD, J. W. (1912). *Reaction to multiple stimuli* (Archives of Psychology, No. 25). New York: Science Press.
- TOWNSEND, J. T., & NOZAWA, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, **39**, 321-359.
- ULRICH, R., & GIRAY, M. (1986). Separate-activation models with variable base times: Testability and checking of cross-channel dependency. *Perception & Psychophysics*, **39**, 248-254.
- WALLACE, M. T., MEREDITH, M. A., & STEIN, B. E. (1998). Multisensory integration in the superior colliculus of the alert cat. *Journal of Neurophysiology*, **80**, 1006-1010.
- WALLACE, M. T., WILKINSON, L. K., & STEIN, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology*, **76**, 1246-1266.
- WORSLEY, K. J. (1982). An improved Bonferroni inequality and applications. *Biometrika*, **69**, 297-302.

## NOTES

1. First recordings from trimodal neurons date back to the 1970s, but few have been examined for multisensory integration. Those that have been examined were studied with two sensory stimuli (not three) and showed the same responses found in bimodal neurons, though their receptive fields were larger (Barry Stein, personal communication, July 2002).

2. Actually, the participant was instructed to respond with the right and the left index fingers simultaneously, to test for motor effects not related to this study. Only right-hand responses were evaluated here, since no significant difference between the hands was found.

3. It differs slightly from the more common measure of simply subtracting bimodal from unimodal RT (see Miller, 1986).

4. In simplified notation, writing in analogy to an ANOVA scheme, for example,

$$RT_T = RT_{TVA} + (RT_T - RT_{TA}) + (RT_{TA} - RT_{TVA}).$$

MRE\* relates to the expression in the last parenthesis, whereas MRE relates to the expression in the parenthesis in

$$RT_T \equiv RT_{TVA} + (RT_T - RT_{TVA}).$$

5. Out of a total of 96 conditions, for 3 conditions the corresponding bimodal conditions had not been presented (see Table 1), so MRE\* values could not be computed.

6. These conclusions are based on significantly different RT means ( $p \leq .01$ ), since no test statistic for the MRE differences is known to us.

7. Moreover, taking at each  $t$  the minimum of all three upper bounds may lead to a sharper upper bound.

8. To substantiate this claim, let us introduce the shorthand notation  $P(T \leq t) \equiv P(T)$ , and so on, neglecting all SOA and  $t$  values. Then  $P(\min(T, V + \tau_1, A + \tau_2) \leq t) \equiv P(T \cup V \cup A)$ . Note that Inequality 5 corresponds to  $P(T \cup V \cup A) \leq P(T \cup V) + P(V \cup A) - P(V)$ . Expanding both sides and canceling terms, this inequality is equivalent to  $P(T \cap V \cap A) \leq P(V \cap A)$ . Thus, a reversal of this inequality would mean that adding a third modality (T) increases the probability  $P(V \cap A) \equiv P(V \leq t \cap A \leq t)$ , in contradiction to the separate activation assumption.

9. From Inequality 3,  $P(T \cup V) \leq P(T) + P(V)$  and  $P(V \cup A) \leq P(V) + P(A)$ . Adding both inequalities, subtracting  $P(V)$  on both sides, and using Equation 5 yields the second inequality in Equation 6, as claimed.