JOURNAL OF MATHEMATICAL PSYCHOLOGY 39, 197–215 (1995)

Intersensory Facilitation of Reaction Time: Evaluation of Counter and Diffusion Coactivation Models

ADELE DIEDERICH

Universität Oldenburg

In a simple reaction time experiment a subject's task is to react to stimuli from three different modalities (visual, auditory, tactile) when any of the stimuli is presented alone, as a pair from two different modalities, or as a triple from all three modalities. The shorter reaction time to double stimuli compared to reaction time to single stimuli or the shorter reaction time to triple stimuli compared to reaction time to double and single stimuli is called "intersensory facilitation of reaction time." Two models were considered to explain this phenomenon: a counter model (Poisson superposition) and a diffusion model (Ornstein–Uhlenbeck process). Both models gave an excellent fit on the level of mean reaction times.

1. INTRODUCTION: INTERSENSORY INTERACTION AND THE FACILITATION OF REACTION TIME

Traditionally, perception and information processing have been studied in one single sensory modality, isolated from all remaining modalities. For example, in determining hearing thresholds the influence of other modalities such as vision or feeling is to be eliminated as much as possible. However, neuroanatomical and -physiological facts, psychological data, and, not least, common experience suggest that there is a more-or-less strong interaction among the modalities. For example, sometimes we close our eyes when listening to music so as not to be disturbed by visual influences. This interaction is sometimes called "intersensory interaction" (Welch & Warren, 1986) and different varieties of intersensory facilitation can be distinguished according its influence on perception. On the one hand, we can consider those intersensory interactions which cause an improvement or worsening of perception, e.g., increasing the threshold for visual or auditory stimuli, or faster processing of a stimulus in a given sensory modality. On the other hand, we can study those intersensory interactions which cause a qualitative change in perception, e.g., different perceptions of color when presenting high tones or low tones at the same time. Intersensory interaction is called "intersensory facilitation" or "facilitation," for short, if the interaction results in faster processing at some stages of processing, i.e., reducing reaction time, or if it improves perceptibility of stimuli, e.g. lowering of sensory thresholds.

Early reaction time (RT) studies in this area were conducted by Dunlap & Wells (1910), who conducted experiments with reactions to visual and auditory stimuli. Todd (1912) may have been the first to call the phenomenon of reducing reaction time "facilitation" when two or more stimuli are presented compared to the situation when only one stimulus is present. Performing an experiment where a light, a tone, and an electric shock were presented to the subject, he observed the following pattern of results which has been replicated many times (e.g., Bernstein, Clark, & Edelstein, 1969; Bernstein, Rose, & Ashe, 1970; Diederich, 1985, 1987; Diederich & Colonius, 1987; Gielen, Schmidt, & van den Heuvel, 1983; Hershenson, 1962; Miller, 1982; Morrell, 1967; Nickerson, 1973. All studies included only two different stimuli: light and tone): RT to double stimuli is shorter than RT to single stimuli and RT to triple stimuli is shorter than RT to double stimuli. Todd's explanation for these results is that "the more rapid responses to the simultaneous stimuli than to the single stimuli have been referred to in a previous division of this work as being 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." (Todd, 1912, p. 63).

To determine the amount of observable facilitation consider the following experimental design: In a simple reaction experiment (Donders, 1868) the subject has to respond to stimuli such as tone, light, or tone and light presented together or with a short delay between light and tone. The subject's only task is to press a button as soon as she notices any stimulus. The time from the onset of the stimulus (when pesenting two stimuli from the onset of the first stimulus) to the response of the subject is defined as reaction time. Typical results of this kind of experiment are as follows: with moderate intensities the RT to tone is shorter, on average, than the RT to light[1]; presenting light first and the

Correspondence and reprint requests should be addressed to Adele Diederich, Institut für Kognitionsforschung, Universität Oldenburg, P.O. Box 2503, D-26111 Oldenburg, Germany. E-mail: diederich@psychologie.uni-oldenburg.de.

[1] RT to sensory stimulation are reciprocally proportional to stimulus intensities up to an irreducible minimum (e.g., Woodworth & Schlosberg, 1954). Moderate intensities mean that the stimuli are well above perception threshold but not so intense that the RT is close to irreducible minimum.

0022-2496/95 $6.00

tone τ ms later the RT is, on average, shorter than the RT to tone plus the τ-ms delay. The decrease of RT when two or more stimuli are presented, compared to the RT when only one stimulus is presented, is here called "intersensory facilitation," i.e., facilitation is defined with respect to RT to single stimuli (reference point). Let $\overline{RT}_A$ and $\overline{RT}_V$ be the mean reaction time to auditory (A) and visual (V) stimuli, respectively, and $\overline{RT}_{VA}(\tau)$ denotes the reaction time when both signals are presented together, the auditory stimulus τ ms after the visual stimulus. The observed amount of intersensory facilitation is found to be

$$\text{observed facilitation} = \begin{cases} (\overline{RT}_A + \tau) - \overline{RT}_{VA}(\tau) \\ \qquad \text{for} \quad (\overline{RT}_A + \tau) < \overline{RT}_V \\ \overline{RT}_V - \overline{RT}_{VA}(\tau) \\ \qquad \text{for} \quad (\overline{RT}_A + \tau) > \overline{RT}_V \end{cases}$$

The difference in RT to light, tone, or tactile stimuli might be due not only to different modality dependend intensities (no cross modality matching is done in this kind of experiment) but also to different conduction times for the modalities. Hershenson (1962) argued that therefore intersensory facilitation occurs if neural events derived from the visual and auditory stimuli were somewhere contemporaneous. He claimed that "facilitation should be expected when the light and sound stimuli are offset in time by the amount approximately equal to the difference between the RTs to the single stimuli. The RT to this presentation should be faster than to either stimulus alone and the maximum effect should occur when Δt equals the difference between the RTs to the stimuli when presented singly." (Hershenson, 1962, p. 289.) That is, as Raab (1962) put it, "Physiological synchrony, rather than stimulus synchrony, is required." (Raab, 1962, p. 574.)

The primary purpose of this article is to present a new model of intersensory facilitation called the multichannel diffusion model. The remainder of the article is organized as follows. First we briefly review the earliest and simplest account of intersensory facilitation—the separate activation race models—and we summarize the basic evidence that rules out this class. Next we present two coactivation models of intersensory facilitation—the superposition counter model and the new multichannel diffusion model. Finally, these two competing models are compared by fitting both models to the results of a comprehensive experiment on intersensory facilitation. We conclude by pointing out some difficulties that these data pose for both coactivation models.

2. MODELING INTERSENSORY FACILITATION

Basically two different approaches have been suggested to explain the phenomenon of intersensory facilitation: *separate activation models* and *coactivation models* (Miller, 1982). Separate activation models assume that processing in different sensory channels is in parallel. Presenting multiple stimuli (e.g., light and tone, or light, tone, and vibration) produces separate activation in each sensory channel. In each channel activation builds to the level at which it can produce a response. The response is produced by the signal that is processed first. These models are also called "race models" since a race between several response activation processes takes place and the winner determines the response time. Assuming processing time to be a random variable, separate activation models predict faster average reaction to multiple stimuli compared to single stimuli since the expected value of the minimum of the processing times in different channels is smaller than or equal to the smaller expected value of the processing time in either channel (probability summation). Raab (1962) presented the first statistical model to explain the observed reaction time reduction. He assumed that processing time for visual and auditory stimuli are independent normally distributed random variables with the same variance differing only in the means. In experiments, however, it has been found that the amount of facilitation is larger than predicted by Raab's model. Gielen, *et al.* (1983) modified Raab's model by making less strict assumptions about the distribution of the processing time. But even then the model could not predict as much facilitation as observed in the data. Miller (1982) proposed an inequality for testing statistical facilitation models. The test is based on a prediction of the separate activation assumption for RT distributions where the RT to a double stimulus, say an auditory and a visual stimulus, is the minimum of the RT of the single stimuli, i.e.,

$$RT_{VA} = \min(RT_V, RT_A), \tag{1}$$

where RT_V and RT_A denote the reaction time random variable to a visual and an auditory stimulus, respectively, and RT_{VA} denotes the reaction time random variable when both signals are presented. Therefore

$$P(\min(RT_V, RT_A) \leqslant t) \leqslant P(RT_V \leqslant t) + P(RT_A \leqslant t). \tag{2}$$

As observed by Miller (1982) Eq. (2) is a prediction of separate activation models; the inequality puts an upper bound on the facilitation produced by double stimuli (cf. Colonius 1990; Ulrich & Giray, 1986). In Diederich (1992b) it is shown that for the triple stimulus case (e.g., auditory (A), visual (V), and tactile (T)), race models predict upper bounds for the triple stimulus condition of the type

$$P(\mathit{min}(RT_T, RT_A) \leqslant t) + P(\min(RT_V, RT_A) \leqslant t) - P(RT_A \leqslant t),$$

which are sharper than the bound analogous to Eq. (2), i.e.,

$$P(RT_V \leqslant t) + P(RT_A \leqslant t) + P(RT_T \leqslant t).$$

It is common pratice to divide observable reaction time into additive components, i.e., $RT = S + R$, where S is a random variable denoting the stimulus processing time, and R, another random variable, denotes all remaining processing times like stimulus encoding, stimulus preparation, motor delay, etc. Under this decomposition the hypothesis stated in Eq. (1) can be specified as follows:

$$RT_V = S_V + R, \qquad RT_A = S_A + R,$$

and

$$RT_{VA} = \min(S_V, S_A) + R.$$

Here, S_V and S_A are the modality specific stimulus processing times and the residual component, R, is assumed to be invariant across stimulus conditions. Note that under these assumptions, whether or not S and R are independent, the distribution inequality (Eq. 2) continue to hold (cf. Ulrich & Giray, 1986; and Colonius, 1986, for further results).

Diederich (1992a) reports experiments with stimuli from three different modalities, auditory, visual, and tactile. Assuming R to be a constant these empirical results, however, indicate serious violations of the separate activation assumption. In light of earlier results (Diederich & Colonius, 1987), supported by recent results by Giray & Ulrich (1994), indicating facilitation effects in the motor component, one might argue that the above assumptions on the relation between S and R are too simplistic. This issue will be taken up in the discussion at the end of the paper.

Coactivation models are an alternative to separate activation models to explain intersensory facilitation. They assume that interaction between different sensory modalities is possible. When presenting multiple stimuli activation of different channels can be combined to satisfy a single criterion for response initiation. Coactivation models predict faster average reaction time to multiple stimuli compared to single stimuli because the combined activation reaches this criterion faster. An early model belonging to this class of models is the so-called energy summation model. When presenting multiple stimuli, energy from these different sensory modalities is assumed to combine or to summate in some peripheral stages of processing. This leads to an earlier termination of the processing in these stages and therefore to shorter RT. Bernstein and his collaborators (Bernstein, 1970; Bernstein *et al.*, 1970) worked out an "energy integration model." The model hypothesizes that the intensities to visual and auditory stimuli integrate at some common point in the nervous system. Presenting two different stimuli is the same as increasing the intensity of either stimulus. An alternative approach to energy summation models is the preparation-enhancement model proposed by Nickerson (1973) and Bernstein, Chu, Briggs, & Schurman (1973). This model assumes that one stimulus plays an alerting role at many of the processing stages, so that the affected stages are terminated more quickly and the response comes earlier.

Whereas these models are mainly descriptive we will discuss two mathematical models that try to explain intersensory facilitation in the following section: a counter model and a diffusion model. Both models are information-accumulation models. These models assume that information about a presented signal accumulates steadily over time until some condition or criterion is satisfied, at which time a decision is made based on the information then available. It is assumed that the nervous system accumulates information about the stimulus all the time, the rate of accumulation being a monotonic function of signal intensity. Major distinctions among various types of information–accumulation models are whether they assume information is accumulated at discrete points in time or continuously over time and whether they assume increments of information are discrete or continuous.

2.1. The Superposition Model for Intersensory Facilitation

First, we consider counter models, models with discrete increments of information accumulated continuously over time.

Information about a signal is determined by the occurrence of events, whose only relevant property is their occurrence, but whose rate is affected by signal intensity. Since only the number of events $N(t)$ that have occurred by time t is relevant as information about a stimulus, these models are called counter models. Counter models have been used primarily in two-alternative choice reaction time tasks (see, e.g., Audley & Pike (1965), Green & Luce (1973, 1974), McGill (1963, 1967), Luce (1986) for more background, Luce & Green (1972), Pike (1966, 1968, 1971, 1973), Townsend and Ashby (1983) for a detailed description). These models assume the existence of a separate internal counter for each alternative. When a stimulus is presented the subject is assumed to accumulate information about a stimulus over time. The counter which first reaches a preset criterion initiates the response in favor of that particular stimulus. Thus, under this assumption counter models belong to the class of race models, since a race between counters takes place and the winner determines the response. But, as mentioned in the preceding section, all race models can be excluded to explain intersensory facilitation observed in our data. Thus, we assume that the counters combine somewhere to reach the preset criterion earlier. Schwarz (1989) proposed a Poisson superposition model for two processes. In the following this approach is extended to the three modality paradigm.

We assume with Schwarz (1989) that the presentation of a stimulus induces a neural renewal counting process, i.e.,

$N \equiv \{N(t), t \geq 0\}$, starting at $t=0$.[2] A response will be executed as soon as a criterion c has been reached, i.e., as soon as c counts have been registered. Presenting a tactile stimulus, light, or tone induces an independent renewal counting process for each modality, i.e., $N_A \equiv \{N_A(t), t \geq 0\}$, $N_V \equiv \{N_V(t), t \geq 0\}$, and $N_T \equiv \{N_T(t), t \geq 0\}$ for the auditory, visual, and tactile stimuli, respectively. For the double-signal trials as well as for the triple signal trials, we define two new stochastic processes by $N_2 \equiv N_A + N_V$ and $N_3 \equiv N_A + N_V + N_T$, respectively.

The processes $N_2 = \{N_2(t), t \geq 0\}$ and $N_3 = \{N_3(t), t \geq 0\}$ are called the superposition of the processes N_A and N_V, and of N_A, N_V, and N_T, respectively. Let S_i, $i \geq 1$, denote the waiting time for the ith renewal; then it is obvious that a fixed number of counts c will, on average, be collected earlier in the superposition process $N_3(t)$ than in any single process. Evidently, the more processes that are running in parallel, the sooner a preset criterion, i.e., certain number of counts, is reached.

The theoretical count number c, or simply the *criterion c*, is assumed to be established by the subject for the number of units of information required for a response to be initiated. The criterion c can be influenced by experimental conditions; e.g., requiring high accuracy (no anticipation) might raise the criterion, whereas requiring high speed of responses might lower the criterion. The criterion is assumed to be a constant, however, over a given condition. For a discussion of all influencing variables for the criterion see Luce (1986).

2.1.1. *Mathematical Derivations for the Superposition Model*

The distribution function of the waiting time S_c for the cth count in the superposition of different processes can be expressed in terms of the waiting time distribution of each participating process. Let $P(S_c \leq t \mid k)$, $k =$ A, V, T, VA, TA, TV, TVA, be the distribution function of S_c under the auditory single-signal trials, the visual single-signal trials, the tactile single-signal trials, the double-signal trials, and the triple-signal trials, respectively. With $P(S_c \leq t) = P(N(t) \geq c)$ (see, e.g., Cinlar, 1975, p. 54) we get $P(S_c \leq t \mid \text{TVA})$ in terms of $P(S_c \leq t \mid \text{A})$, $P(S_c \leq t \mid \text{V})$, and $P(S_c \leq t \mid \text{T})$ as follows (for derivation see Appendix):

$$\begin{aligned} 1 - P(S_c \leq t \mid \text{TVA}) &= \sum_{n=0}^{c-1} P(N_3(t) = n) \\ &= \sum_{n=0}^{c-1} \sum_{i=0}^{n} \sum_{j=0}^{i} P(N_A(t) = j) \\ &\quad \times P(N_V(t) = i-j) P(N_T(t) = n-i) \\ &= \sum_{n=0}^{c-1} \sum_{i=0}^{n} \sum_{j=0}^{i} [P(S_j \leq t \mid A) - P(S_{j+1} \leq t \mid A)] \\ &\quad \times [P(S_{i-j} \leq t \mid \text{V}) - P(S_{i-j+1} \leq t \mid \text{V})] \\ &\quad \times [P(S_{n-i} \leq t \mid \text{T}) - P(S_{n-i+1} \leq t \mid \text{T})]. \end{aligned} \tag{3}$$

[2] This implies $N(0) = 0$. Strictly speaking this assumes that peripherical processing of the stimulus is concurrent with stimulus presentation.

In this way any distribution function for double or triple stimuli can be expressed in terms of the corresponding distribution functions of the single-signal trials. In our experiment the stimuli in the triple-signal trials were always presented in the order $s_T - s_V - s_A$ (denoting the tactile, visual, and auditory stimulus, respectively), since subjects tend to react fastest to tones and slowest to tactile stimuli. Letting τ_1 represent the time (in ms) between the presentation of s_T and s_V and τ_2 the time between s_T and s_A we get a process with counting function

$$\begin{gathered} N_{3,\tau_1,\tau_2}(t) = N_T(t) + N_V(t-\tau_1) + N_A(t-\tau_2), \\ \tau_2 \geq \tau_1 \geq 0, \qquad t \geq 0, \end{gathered} \tag{4}$$

where $N_T(s) = N_V(s) = N_A(s) = 0$, for $s < 0$. Figure 1 illustrates the superposition model for this situation.

In order to test the superposition model against empirical data, the interarrival time distribution must be specified. Although the above derivations hold for arbitrary interarrival time distributions, the simplest case is to assume exponentially distributed interarrival times yielding a Poisson counting process. It is adopted here as a starting point. With

$$\begin{aligned} P(N(t) \geq c) = P(S_c \leq t) &= \sum_{j=c}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!} \\ &= 1 - \sum_{j=0}^{c-1} e^{-\lambda t} \frac{(\lambda t)^j}{j!} \end{aligned} \tag{5}$$

we get

$$P_{3,\tau_1,\tau_2}(S_c \leq t)$$

$$= \begin{cases} 1 - \exp\{-\alpha_T t\} \sum_{j=0}^{c-1} \frac{(\alpha_T t)^j}{j!}, & 0 \leq t \leq \tau_1, \\ 1 - \exp\{-[\alpha_T \tau_1 + (\alpha_T + \alpha_V)(t - \tau_1)]\} \\ \quad \times \sum_{j=0}^{c-1} \frac{[\alpha_T \tau_1 + (\alpha_T + \alpha_V)(t - \tau_1)]^j}{j!}, & \tau_1 \leq t \leq \tau_2, \\ 1 - \exp\{-[(\alpha_T + \alpha_V + \alpha_A)\, t - \alpha_V \tau_1 - \alpha_A \tau_2]\} \\ \quad \times \sum_{j=0}^{c-1} \frac{[(\alpha_T + \alpha_V + \alpha_A)\, t - \alpha_V \tau_1 - \alpha_A \tau_2]^j}{j!}, & t \geq \tau_2, \end{cases} \tag{6}$$

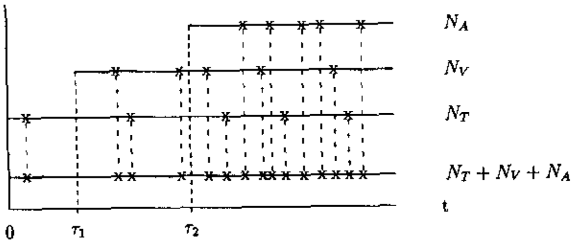

FIG. 1. The temporal course of (exemplary) events of the counting process $N_3(t)$ corresponding to Eq. (4).

where α_A, α_V, and α_T denote the intensity parameter of the auditory, visual, and tactile channel, respectively. Note that for $\tau_2 = \infty$, the distribution function reduces to the case when only two stimuli are presented. Setting $\tau_1 = \tau_2 = \infty$, the distribution function reduces to the case when only one stimulus is presented. Obviously the larger the critical count number c, the smaller the probability that the waiting time until the cth count occurs before time t.

2.1.2. *Expected Value and Variance of the Reaction Time for the Superposition Model*

The observed reaction time is assumed to be additively composed of the waiting time S_c plus all processes following (or preceding) it. The durations of these additional processes, which may include motor preparation and response execution components, are represented by a random variable, R, say. In general, S_c and R need not be stochastically independent, and for a test of the model at the level of the entire distribution functions, the bivariate distribution for (S_c, R) has to be specified. However, at the level of expectations, we have

$$E[\mathrm{RT}] = E[S_c + R] = E[S_c] + E[R]. \qquad (7)$$

If S_c and R are stochastically independent, the variance of the observable RT at interstimulus intervals (ISIs) of τ_1 and τ_2 ms will be

$$\mathrm{Var}[\mathrm{RT}] = \mathrm{Var}[S_c] + \mathrm{Var}[\mathrm{R}]. \qquad (8)$$

For the sake of parsimony, in the application below we assume $E(R)$ and $\mathrm{Var}(R)$ to be constant within an experimental condition, i.e., $E(R) = r$ and $\mathrm{Var}(R) = \sigma_r^2$. Equations (7) and (8) then simplify respectively to

$$E[\mathrm{RT}] = E[S_c] + r \qquad (9)$$

and

$$\mathrm{Var}[\mathrm{RT}] = \mathrm{Var}[S_c] + \sigma_r^2. \qquad (10)$$

The corresponding expected value $E_{\tau_1,\tau_2}[S_c]$ is (for derivation see Appendix)

$$\begin{aligned} E_{\tau_1,\tau_2}[S_c] = {} & \frac{c}{\alpha_T} - \exp\{-\alpha_T\tau_1\} \sum_{j=0}^{c-1} \frac{(\alpha_T\tau_1)^j (c-j)}{j!} \\ & \times \left[\frac{1}{\alpha_T} - \frac{1}{\alpha_T+\alpha_V}\right] \\ & - \exp\{-[\alpha_T\tau_2 + \alpha_V(\tau_2-\tau_1)]\} \\ & \times \sum_{j=0}^{c-1} \frac{[\alpha_T\tau_2+\alpha_V(\tau_2-\tau_1)]^j (c-j)}{j!} \\ & \times \left[\frac{1}{\alpha_T+\alpha_V} - \frac{1}{\alpha_T+\alpha_V+\alpha_A}\right]. \end{aligned} \qquad (11)$$

Setting $\tau_1 = \tau_2 = 0$ leads to a superposition model where all stimuli are presented at the same time. With $\alpha \equiv \alpha_T + \alpha_V + \alpha_A$, $E[S_c] = c/\alpha$ becomes the means of a gamma distribution with parameters c and α. From here it is easy to see that increasing c increases the mean response time; increasing α decreases the mean response time. The equation for the expected value for the situation when only two stimuli are presented can be found in Schwarz (1989), and the equation for the variance for two presented stimuli can be found in Diederich & Colonius (1991). The variance for the triple stimulus condition is

$$\mathrm{Var}[S_c] = E[S_c^2] - E[S_c]^2, \qquad (12)$$

where

$$\begin{aligned} E[S_c^2] = {} & \frac{c(c+1)}{\alpha_T^2} - 2\Bigg(\exp\{-\alpha_T\tau_1\} \sum_{j=0}^{c-1} \frac{[\alpha_T\tau_1)^j (c-j)}{j!} \\ & \times \left[\frac{1}{\alpha_T^2}\left(1+\alpha_T\tau_1+\frac{c-1-j}{2}\right) - \frac{1}{(\alpha_T+\alpha_V)^2} \right. \\ & \left. \times \left(1+(\alpha_T+\alpha_V)\,\tau_1+\frac{c-1-j}{2}\right)\right] \\ & - \exp\{-(\alpha_T\tau_2+\alpha_V(\tau_2-\tau-1))\} \\ & \times \sum_{j=1}^{c-1} \frac{\alpha_T\tau_2+\alpha_V(\tau_2-\tau_1)]^j (c-j)}{j!} \\ & \times \left[\frac{1}{(\alpha_T+\alpha_V)^2}\left(1+(\alpha_T+\alpha_V)\,\tau_2+\frac{c-1-j}{2}\right) \right. \\ & - \frac{1}{(\alpha_T+\alpha_V+\alpha_A)^2}\Bigg(1+(\alpha_T+\alpha_V+\alpha_A)\,\tau_2 \\ & \left. + \frac{c-1-j}{2}\Bigg)\right]\Bigg). \end{aligned} \qquad (13)$$

(For derivation see Appendix.)

2.2. The Multichannel Diffusion Model for Intersensory Facilitation

The counter model and, consequently, the superposition model are stochastic processes with continuous-time and discrete-state space. Considering a discrete-time process with continuous-state space leads to a class of models, so called general random walk models, that have functioned primarily as an alternative to counter models in choice tasks and choice reaction time (see, e.g., Ashby, 1983; Edwards, 1965; Heath, 1981; Laming, 1968; Link & Heath, 1975; Stone, 1960). For a detailed description of various types of random walk models for two-choice discrimination and response time data, see Townsend & Ashby (1983) and Luce (1986). (Note that in mathematical literature random walk models often exclusively refer to stochastic processes with discrete-time and discrete-state space). Recently developed alternatives to these models are diffusion models, stochastic processes with continuous-time and continuous-state space. Most research done on these models is also for choice reaction time. Ratcliff (1978) used a diffusion process in his theory of memory retrieval. Recently, Busemeyer, Townsend, and colleagues (Busemeyer & Townsend, 1992; Busemeyer & Goldstein, 1992; Busemeyer & Townsend, 1993) developed a dynamic stochastic model of decision making—decision field theory—which can be described by a diffusion process. Diffusion models also, though to a far less extent, entered simple reaction time research. Pacut (1980, 1982) applied a diffusion model to escape reaction and escape learning data; Pacut & Tych (1982) applied it to avoidance reaction latency and avoidance learning. Recently Smith (1990) proposed a diffusion model applied to simple reaction time. In the following section we develop a diffusion model called a "multichannel diffusion model" for intersensory facilitation.

Assume that as soon as a signal is presented to the subject the nervous system stochastically accumulates very small amounts of activation continuously over time. Activation A grows nonlinearly towards an asymptote. After the signal is switched off, the amount of activation is assumed to decay to its starting level. This decay, which has been postulated in studies of neuronal activity dynamics (Ricciardi, 1977; Tuckwell, 1989), can be accounted for neither in the counter model nor in those random walk models developed in psychology. Further, it is assumed that activation never drops below its initial level. That is, the process is reflected on the lower boundary. Finally, a response is initiated as soon as a preset criterion level is reached. That is, the process is absorbed on the upper boundary. This process can be described by a diffusion process called the "Ornstein–Uhlenbeck process" with one reflecting and one absorbing boundary (see, e.g., Cox & Miller, 1965; Karlin & Taylor, 1981). Its drift, $\mu(x)$, is $\alpha - \beta x$ and its diffusion coefficient, $\sigma^2(x)$, is σ^2, where β and σ are arbitrary positive constants.

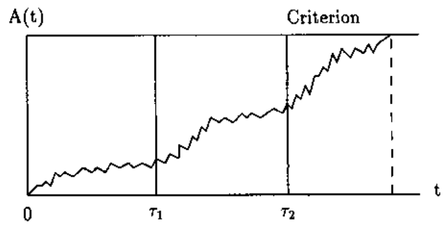

FIG. 2. A stochastic trajectory representing the information accumulation process when three different stimuli are presented, τ_1 and τ_2 ms apart.

α reflects the force directed towards the criterion, whereas β reflects a restoring force directed towards the origin, i.e., the initial level. $-\beta x$ is of a magnitude proportional to the distance to the origin, i.e., β represents the rate of growth towards an equilibrium when a signal is presented ($\alpha > 0$), and it reflects the rate of decay towards the initial level when the signal is switched off ($\alpha = 0$). Note that random walk models as referred to in mathematical literature are special cases of this type of model, i.e., the decay parameter β is zero. While the diffusion coefficient is independent of the state of information accumulation, the drift rate changes with the location x. In our experiment three different stimuli were presented, τ_1 and τ_2 ms apart and the information accumulation process for this situation can be illustrated as in Fig. 2.

The information accumulation process starts at time $t = 0$ with drift μ, say, and diffusion coefficient σ^2. At time $t = \tau_1$, when the second stimulus is presented, the process is located somewhere between the reflecting boundary and the absorbing boundary. The drift coefficient now changes since more information is available; i.e., $\mu(x)$ is assumed to become $(\alpha_T + \alpha_V) - \beta x$, with $\alpha_V > \alpha_T$. Finally, at time $t = \tau_2$, when the third stimulus is presented, the process is located somewhere between the reflecting boundary and the absorbing boundary.[3] The drift coefficient now is $\mu(x) = (\alpha_T + \alpha_V + \alpha_A) - \beta x$, with $\alpha_A > \alpha_V$. In analogy to the superposition model call the parameters $\alpha_T, \alpha_V, \alpha_A$, that represent the force directed toward the absorbing boundary (criterion) "intensity parameters." For simplicity the diffusion coefficient is assumed to be the same during any fixed stimulus period $(0 - \tau_1)$, $(\tau_1 - \tau_2)$, $(\tau_2 - \infty)$. Note that the diffusion process is thus time-homogeneous within each time interval $[0, \tau_1)$, $[\tau_1, \tau_2)$, and $[\tau_2, \infty)$ but non-homogeneous across the interval $[0, \infty)$.

Analogous to the counter model, the larger the criterion c assumed to be established by the subject, the larger is the number of steps, given a fixed step size, to reach the boundary; i.e., it takes more time until a response is initiated.

[3] This may be the typical case, but the model also allows for the possibility of absorption before τ_1 or τ_2.

2.3. Mathematical Derivations for the Diffusion Model

The time at which the stochastic process hits the upper boundary, i.e., when a response is initiated, is of special interest. We consider a single state and ask for the probability distribution of the time T at which $X(t)$ enters the state for the first time. In our case the state is an absorbing state and once the process has entered that state it will never leave it. Obviously T is a random variable taking values in the interval $[0, \infty)$ and it is called the "first passage time."[4] The backward or *Kolmogorov* equation is appropriate to calculate the probability distribution of the first passage time since our objective is to determine the first passage time distribution as a function of the initial state x_0 for a fixed final state (Cox & Miller, 1977, p. 216). In the case of an unrestricted process, i.e., the boundaries are $\pm\infty$ and starting at time $t_0 = 0$, the solution for the backward equation as a transition probability function for the Ornstein–Uhlenbeck process is a normal distribution with mean $x_0 \exp(-\beta t) + (\alpha/\beta)(1 - \exp(-\beta t))$ and variance $\sigma^2(1 - \exp(-2\beta t))/2\beta$. Unfortunately, no closed form solution is known for the case of an arbitrary boundary, e.g. one absorbing boundary and one reflecting boundary (see Ricciardi, 1977, p. 128; Bhattacharya & Waymire, 1990, p. 389). But there exist numerical methods based on limits of discrete-time Markov chains that provide practical approximations which will be used in the following (for a related approach see Busemeyer & Townsend, 1992). The discrete time Markov chain model is presented next.

Presenting a stimulus evokes an information accumulation process. This process is represented by a variable A, called the state of information accumulation or activation. The state of activation changes over time, depending on how much information is available, producing a new state of activation at each moment in time, a stochastic process denoted by $A(t)$. As soon as $A(t) \geqslant c$, a criterion set by the subject, the response is initiated. We assume that activation changes by very small amounts during small time intervals. Let $h, h > 0$, denote a very small time interval and $\Delta A(t+h) = A(t+h) - A(t)$ the change of activation during the interval h. $\Delta A(t+h)$ is a very small step equal to $\delta \equiv \sigma\sqrt{h}$. In letting $h \to 0$ the time interval and the step size as a function of h become infinitesimally small, and the continuous-time and Markov process can be approximated by a discrete-time Markov chain (see Bhattacharya & Waymire, 1990, pp. 386–388; Cox & Miller, 1977, pp. 213–215). The parameter δ is called the step size and the parameter σ^2 is called the diffusion coefficient. We assume a time-homogeneous Markov chain; i.e., the transition probabilities depend only on the difference $A(t+h) - A(t)$ instead of $A(t+h)$ and $A(t)$. Therefore the transition probabilities are determined as

$$\begin{aligned}\text{Prob}[A(t+h) - A(t) = -\delta \mid A(t) = x, A(t-h) = x_1, ...] \\ = \text{Prob}[\Delta A(t+h) = -\delta \mid A(t) = x] \\ \equiv p(x) \qquad (14)\end{aligned}$$

$$\begin{aligned}\text{Prob}[A(t+h) - A(t) = \delta \mid A(t) = x, A(t-h) = x_1, ...] \\ = \text{Prob}[\Delta A(t+h) = \delta \mid A(t) = x] \\ = 1 - p(x) \equiv q(x). \qquad (15)\end{aligned}$$

Note that the transition probability functions $p(x)$ and $q(x)$ depend on the state x but not on the time t. The drift coefficient, $\mu(x)$, and the diffusion coefficient, σ^2, for the discrete-time Markov chain can be defined by the first and second moments of $\Delta A(t)$, conditioned on the previous state of activation, as follows:

$$\mu(x) = \frac{E[\Delta A(t+h) \mid A(t) = x]}{h} = \frac{\delta[q(x) - p(x)]}{h} \qquad (16)$$

$$\sigma^2 = \frac{E[\Delta A(t+h)^2 \mid A(t) = x]}{h} = \frac{\delta^2[q(x) + p(x)]}{h}. \qquad (17)$$

The drift and diffusion coefficients are functions of the time unit h, which is taken to be a fixed constant. Its value is chosen a priori to be as close to zero as needed to make the discrete-time Markov chain model approximate the continuous-time Markov process as accurately as desired. The transition probabilities, $p(x)$ and $q(x)$, can be written as functions of the drift and diffusion coefficients, $\mu(x)$ and σ^2, by solving for $p(x)$ and $q(x)$ in Eqs. (16) and (17). Assuming $p(x) = 1 - q(x)$ and $\delta = \sigma\sqrt{h}$ yields

$$p(x) = \frac{1}{2}\left[1 - \frac{\mu(x)}{\sigma}\sqrt{h}\right], \qquad (18)$$

$$q(x) = \frac{1}{2}\left[1 + \frac{\mu(x)}{\sigma}\sqrt{h}\right]. \qquad (19)$$

In order to obtain $0 \leqslant p(x), q(x) \leqslant 1$, $\mu(x)/\sigma$ must be between $-1/\sqrt{h}$ and $1/\sqrt{h}$. The stochastic process obtained so far can be defined by a stochastic difference equation, i.e.,

$$A(t+h) = A(t) + \mu(A(t))\,h + \varepsilon(t)\sqrt{h}$$

or, writing $\Delta A(t+h)$ for $A(t+h) - A(t)$,

$$\Delta A(t+h) = \mu(A(t))\,h + \varepsilon(t)\sqrt{h}, \qquad (20)$$

where $\varepsilon(t)$, a random variable, is a stochastic disturbance term with mean 0 and variance σ^2 for small h (see Cox & Miller, 1977, p. 207). Since the drift coefficient of the

[4] An absorbing state is a recurrent state (Bhattacharya & Waymire, 1990, p. 419) that guarantees the finiteness (with probability 1) of the passage time (Bhattacharya & Waymire, 1990, p. 423).

Ornstein–Uhlenbeck process is $\mu(x) = \alpha - \beta x$ we get the following stochastic difference equation for changes in activation:

$$\Delta A(t+h) = (\alpha - \beta A(t))\, h + \varepsilon(t)\, \sqrt{h}. \tag{21}$$

To calculate the first passage time distribution and the expected value and the variance of the time the approximating process will need to hit the boundary, we have to determine the transition probability matrix for the Markov chain. This implies that we first have to define the state space of the process. Let S be a set of states of activation for the information accumulation process. The total number of states can be expressed as a function of the step size δ and the criterion c by setting $c = k\delta$. Consequently

$$S = \{0, \delta, 2\delta, ..., k\delta\}$$

and the cardinality of S is $k + 1 \equiv m$. Since the process is assumed never to skip states in its progress the transition matrix can be characterized as

$$p_{i,j} = \begin{cases} p((i-1)\,\delta), & j = i-1, \\ q((i-1)\,\delta), & j = i+1, \\ 0, & \text{otherwise,} \end{cases}$$

where p and q are defined earlier in Eq. (18) and Eq. (19). Note that the transition probabilities depend on the state in which the process is located.

Usually the transition matrix $\mathbf{T}$ is presented in its canonical form. Let the m-state Markov chain consist of r recurrent states and $(m-r)$ transient states. The transition matrix $\mathbf{T}$ can then be put in the form

$$\mathbf{T} = \left[\frac{\mathbf{0} \mid \mathbf{P}}{\mathbf{Q} \mid \mathbf{R}}\right],$$

where $\mathbf{P}$ is an $r \times r$ matrix with transition probabilities among the recurrent states for its elements, $\mathbf{Q}$ is an $(m-r) \times (m-r)$ substochastic matrix (with at least one row sum less than 1) with probabilities of transition only among the transient states for its elements, and $\mathbf{R}$ is an $(m-r) \times r$ matrix whose elements are the probabilities of the one-step transition from the $(m-r)$ transient states to the r recurrent states. In our case $\mathbf{P}$ is a 1×1 matrix containing the number 1 since we just have one absorbing boundary, the criterion c. $\mathbf{Q}$, an $(m-1) \times (m-1)$ matrix, contains the transition probabilities p_{ij} stated above except for the last row where $p_{1,2} = 1$, otherwise zero, since we have one reflecting boundary. $\mathbf{R}$, an $(m-1)$ vector, contains the transition probability from the transient to the absorbing state, with the first element being q_{m-1}, zero otherwise. The

	Z	1	2	3		⋯			m-2	m-1	m	
1	1	0	1	0	0	⋯	0	0	0	0	0	
2	0	p_2	0	q_2	0	⋯	0	0	0	0	0	
3	0	0	p_3	0	q_3	⋯	0	0	0	0	0	
		0	0	p_4	0	q_4	⋯					**R**
⋮						⋯						
						⋯	p_{m-3}	0	q_{m-3}	0	0	
m-2	0	0	0	0	0	⋯	0	p_{m-2}	0	q_{m-2}	0	
m-1	0	0	0	0	0	⋯	0	0	p_{m-1}	0	q_{m-1}	
m		0	0	0	0	⋯	0	0	0	0	1	**P**

FIG. 3. Transition matrix for the information accumulation process when one stimulus is presented.

starting position of the process is considered next. At time $t = 0$ the process is set in motion either by starting it at a fixed state i_0, called the *initial state*, or by randomly locating it in the state space according to a probability distribution $\mathbf{Z}$ on S, called the initial distribution. In the former case, $\mathbf{Z}$ is the distribution concentrated at the state i_0, i.e., $Z_j = 1$ if $j = i_0$, $Z_j = 0$ if $j \neq i_0$. In the latter case, the probability is z_i that at time zero the process will be found in state i, where $0 \leqslant z_i \leqslant 1$ and $\sum_i z_i = 1$. Thus the initial distribution $\mathbf{Z}$ is an $(m-r)$ vector containing the initial probability distribution over the transient states. We assume that the process starts at time $t = 0$ with probability 1, thus $\mathbf{Z}$ is an $(m-1)$ vector, with $z_1 = 1$, zero otherwise. Figure 3 shows the mapping of this process to a transition matrix.

To derive the equations for the mean reaction time we use standard methods developed in Markov chain theory (see, e.g., Bhat, 1984; Bhattacharya & Waymire, 1990).

Consider the probability of entering an absorbing state when the process leaves the class of transient states. It should be noted that if the process enters the absorbing state, it does not leave it. Let $f_{ij}^{(n)}$ be the probability that starting from transient state i, the process enters the absorbing state j in n steps. Noting that a transition from j to i is not possible, we can consider the number of steps as proportional to the time needed for a first passage transition from i to j. If we define T_{ij} to represent this random variable, we may write $f_{ij}^{(n)}$ as its distribution given by

$$P(T_{ij} = n) = f_{ij}^{(n)},$$

and let

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

be the probability to eventual passage to j. Let $\mathbf{F}^{(n)}$ be the matrix with elements $f_{ij}^{(n)}$ and $\mathbf{F}$ be the matrix with elements f_{ij}. Then

$$\mathbf{F}^n = \mathbf{Q}^{n-1}\mathbf{R}, \qquad n = 1, 2, ..., \infty,$$

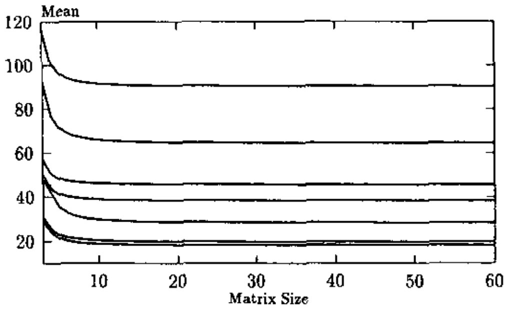

FIG. 4. The means of the discrete process as a function of the matrix size m.

and

$$\mathbf{F} = \sum_{n=1}^{\infty} \mathbf{F}^{(n)} = (\mathbf{I} - \mathbf{Q})^{-1}\, \mathbf{R},$$

where $\mathbf{Q}$ and $\mathbf{R}$ are defined earlier and $\mathbf{I}$ is the identity matrix (see, e.g., Bhat, 1984, p. 79). Considering that we start at a particular transient state, the probability distribution of the time to reach the criterion c starting at time $t = 0$, $T_{o,c} \equiv T$, equals

$$\text{Prob}[T = t] = \mathbf{Z}'\mathbf{Q}^{n-1}\mathbf{R}, \qquad n = 1, 2, ..., \infty,$$

where $t = (n+1)\,h$. The pth moment for the distribution of times to reach the boundary is therefore

$$E[T^p] = h^p\mathbf{Z}' \sum_n n^p \mathbf{Q}^{n-1}\mathbf{R}, \qquad n = 1, ..., \infty. \tag{22}$$

To see how well the continuous process can be mimicked by a discrete process the theoretical mean and the standard deviation are plotted as a function of the matrix size, m. Note that decreasing h, the time unit, corresponds to an increase in m, holding all other factors fixed. As can be seen from Figs. 4 and 5 the mean is almost constant already at a matrix size above 6 by 6. For the standard deviation a matrix size of 10 by 10 gives already a good approximation.

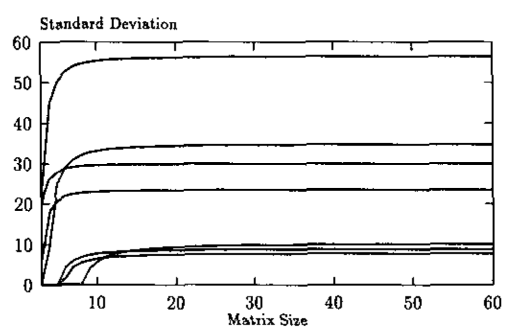

FIG. 5. The standard deviations of the discrete process as a function of the matrix size m.

TABLE 1

Parameters α, β, and c for the Ornstein–Uhlenbeck Process

α	0.2	0.2	0.2	0.2	0.5	0.5	0.5	0.5
β	0.01	0.001	0.01	0.001	0.01	0.001	0.01	0.001
c	15	15	10	10	15	15	10	10

Different curves reflect different α's, β's, and criteria c presented in Table 1 (α and β being parameters for the Ornstein–Uhlenbeck process). The first column of the table corresponds to the very upper lines in the figures, the second column to the second upper lines, etc.

In the experiment three different stimuli are presented, τ_1 and τ_2 ms apart. The first part of the process, i.e., when only one stimulus is presented, is exactly the same as described above. The process starts at z_1 at time $t = 0$ with probability 1. At time $t = \tau_1$, when the second stimulus is presented, the process may be somewhere between the absorbing and the reflecting boundary. The location of the process at that point in time is the starting position for the new process that starts with a different drift coefficient at time $t = \tau_1$. Clearly we now have a defective probability distribution over all transient states. At time $t = \tau_2$, when the third stimulus is presented, the process may be somewhere between the absorbing and the reflecting boundary. Again, a new process starts with a different drift coefficient at time $t = \tau_2$.

2.3.1. *Expected Value and Variance of the Reaction Time for the Multichannel Diffusion Model*

Like for the counter model, observed reaction time is assumed to be additively composed of the stimulus processing time, denoted by T, plus all processes following or preceding it (e.g., temporal uncertainty of stimulus onset), denoted by R. T and R are again assumed to be stochastically independent random variables. For the mean reaction time and for the variances respectively we have

$$E[\text{RT}] = E[T + R] = E[T] + E[R]. \tag{23}$$

$$\text{Var}[\text{RT}] = \text{Var}[T] + \text{Var}[R]. \tag{24}$$

Like for the counter model, if we assume the mean and the variance of R to be constant, i.e., $E(R) = r$ and $V(R) = \sigma_r^2$, Eq. (23) and Eq. (24) then simplify, respectively, to

$$E[\text{RT}] = E[T] + r$$

and

$$\text{Var}[\text{RT}] = \text{Var}[T] + \sigma_r^2.$$

The corresponding expected value $E[T]$ is

$$E[T]=h\left(\mathbf{Z}'\sum_{i=1}^{n_1} i\mathbf{Q}_1^{i-1}\mathbf{R}_1+\mathbf{Z}'\mathbf{Q}_1^{n_1}\sum_{i=n_1+1}^{n_2} i\mathbf{Q}_2^{(i-(n_1+1))}\mathbf{R}_2 + \mathbf{Z}'\mathbf{Q}_1^{n_1}\mathbf{Q}_2^{n_2-n_1}\sum_{i=n_2+1}^{\infty} i\mathbf{Q}_3^{(i-(n_2+1))}\mathbf{R}_3\right), \quad (25)$$

where $\mathbf{Q}_1, \mathbf{Q}_2$, and $\mathbf{Q}_3$ are the transition matrices with transient states for the first, second, and third part of the process, respectively; n_1 and n_2 are $\tau_1/h-1$ and $\tau_2/h-1$, respectively. $\mathbf{R}_1$, $\mathbf{R}_2$, and $\mathbf{R}_3$ are the vectors that contain the transition probability from the transient to the absorbing state for the first, second, and third part of the process. $\mathbf{Z}'\mathbf{Q}_1^{n_1}$ is the initial defective distribution for the second part and $\mathbf{Z}'\mathbf{Q}_1^{n_1}\mathbf{Q}_2^{n_2-n_1}$ is the initial defective distribution for the third part of the process. Setting the transition matrices $\mathbf{T}_1=\mathbf{T}_2=\mathbf{T}_3$ implies $\mathbf{Q}_1=\mathbf{Q}_2=\mathbf{Q}_3$ and $\mathbf{R}_1=\mathbf{R}_2=\mathbf{R}_3$. This leads to a diffusion process where only one stimulus is presented. The variance is

$$\mathrm{Var}[T]=E[T^2]-E[T]^2,$$

where

$$E[T^2]=h^2\left(\mathbf{Z}'\sum_{i=1}^{n_1} i^2\mathbf{Q}_1^{i-1}\mathbf{R}_1 + \mathbf{Z}'\mathbf{Q}_1^{n_1}\sum_{i=n_1+1}^{n_2} i^2\mathbf{Q}_2^{(i-(n_1+1))}\mathbf{R}_2 + \mathbf{Z}'\mathbf{Q}_1^{n_1}\mathbf{Q}_2^{n_2-n_1}\sum_{i=n_2+1}^{\infty} i^2\mathbf{Q}_3^{(i-(n_2+1))}\mathbf{R}_3\right). \quad (26)$$

For computational purposes the equations for the expected value and the second moment can be greatly simplified (for derivation see Appendix),

$$\begin{aligned} E[T]=h[&\mathbf{Z}'(\mathbf{I}-\mathbf{Q}_1)^{-1}((\mathbf{I}-\mathbf{Q}_1^{-1}(\mathbf{I}-\mathbf{Q}_1^{n_1})-n_1\mathbf{Q}_1^{n_1})\,\mathbf{R}_1 \\ &+\mathbf{Z}_2'(\mathbf{I}-\mathbf{Q}_2)^{-1}((\mathbf{I}-\mathbf{Q}_2)^{-1}(\mathbf{I}-\mathbf{Q}_2^{(n_2-n_1)}) \\ &+n_1\mathbf{I}-n_2\mathbf{Q}_2^{(n_2-n_1)})\,\mathbf{R}_2 \\ &+\mathbf{Z}_3'(\mathbf{I}-\mathbf{Q}_3)^{-1}((\mathbf{I}-\mathbf{Q}_3^{-1}+n_2\mathbf{I})\,\mathbf{R}_3] \end{aligned} \quad (27)$$

$$\begin{aligned} E[T^2]=h^2[&\mathbf{Z}'((2(\mathbf{I}-\mathbf{Q}_1)^{-3}-(\mathbf{I}-\mathbf{Q}_1)^{-2})(\mathbf{I}-\mathbf{Q}_1^{n_1}) \\ &-(2(\mathbf{I}-\mathbf{Q}_1)^{-2}+n_1(\mathbf{I}-\mathbf{Q}_1^{-1})\,n_1\mathbf{Q}_1^{n_1})\,\mathbf{R}_1 \\ &+\mathbf{Z}_2'((2(\mathbf{I}-\mathbf{Q}_2)^{-3}-(\mathbf{I}-\mathbf{Q}_2)^{-2})(\mathbf{I}-\mathbf{Q}_2^{(n_2-n_1)}) \\ &+(\mathbf{I}-\mathbf{Q}_2)^{-1}(2n_1(\mathbf{I}-\mathbf{Q}_2)^{-1}+n_1^2\mathbf{I}) \\ &-(2(\mathbf{I}-\mathbf{Q}_2)^{-2}+n_2(\mathbf{I}-\mathbf{Q}_2)^{-1})\,n_2\mathbf{Q}_2^{(n_2-n_1)})\,\mathbf{R}_2 \\ &+\mathbf{Z}_3'((2(\mathbf{I}-\mathbf{Q}_3)^{-3}-(\mathbf{I}-\mathbf{Q}_3)^{-2}) \\ &+(\mathbf{I}-\mathbf{Q}_3)^{-1}(2n_2(\mathbf{I}-\mathbf{Q}_3)^{-1}+n_2^2\mathbf{I}))\,\mathbf{R}_3], \end{aligned} \quad (28)$$

where $\mathbf{Z}_2'\equiv\mathbf{Z}'\mathbf{Q}_1^{n_1}$ and $\mathbf{Z}_3'\equiv\mathbf{Z}'\mathbf{Q}_1^{n_1}\mathbf{Q}_2^{(n_2-n_1)}$.

Using spectral analytic methods the matrix computation can be greatly simplified. A tridiagonal transition matrix $\mathbf{Q}$ is similar to a symmetric tridiagonal matrix, i.e., there exists a real valued matrix $\mathbf{D}$ such that $\mathbf{DQD}^{-1}=\mathbf{S}$, where $\mathbf{S}$ is a symmetric tridiagonal matrix (see Horn & Johnson, 1990, p. 174). This fact guarantees that $\mathbf{Q}$ has $m-1$ linearly independent eigenvectors and $m-1$ real eigenvectors (Horn & Johnson, 1990). According to the *Frobenius–Perron* theorem (Bhattacharya & Waymire, 1990, Chap. 3), all of the eigenvalues in $\mathbf{Q}$ are less than 1 in magnitude. Therefore, we factor $\mathbf{Q}$ as

$$\begin{aligned} \mathbf{Q}&=\mathbf{P}\,\Lambda\mathbf{P}^{-1} \\ \mathbf{Q}^2&=(\mathbf{P}\,\Lambda\mathbf{P}^{-1})(\mathbf{P}\,\Lambda\mathbf{P}^{-1})=\mathbf{P}\,\Lambda^2\mathbf{P}^{-1} \\ &\vdots \\ \mathbf{Q}^n&=\mathbf{P}\,\Lambda^n\mathbf{P}^{-1}, \end{aligned}$$

where $\mathbf{P}$ is the matrix of linear independent eigenvectors, and Λ is the diagonal matrix of real valued eigenvalues less than one in magnitude for the matrix $\mathbf{Q}$.

3. AN EXPERIMENTAL TEST OF INTERSENSORY FACILITATION

As the preceding sections have shown, both the counter and the diffusion model are mathematically quite distinct. The superposition model assumes discrete increments of information accumulated continuously over time. The increments occur according to a Poisson distribution, the time elapsing between two pieces of information is exponentially distributed. On the other side the diffusion model assumes continuous accrual of information accumulated continuously over time having a Gaussian transition probability function for the unrestricted case. The consideration of a special diffusion process, the Ornstein–Uhlenbeck process, was motivated by studies of neuronal activity dynamics postulating that activity grows nonlinearly toward an asymptote and that the amount of activation decays to its starting level when information is no longer available.

The experiment, described in detail below, was primarily designed to test the presence of intersensory facilitation. Moreover, it was intended to test whether and to what degree both models are able to describe the intersensory facilitation effect. However, it was not meant to differentiate between the models.

Subjects. Four right-handed subjects (PM, KK, GH, BO), three female and one male, associated with the psychology department of Oldenburg University, participated in the experiment. They received required subject hours and/or 10 DM per hour for their participation.

Apparatus. The visual stimulus was a flash (250 Lux) of $\frac{1}{16}$ s duration generated by a strobe (SUNPAK auto 36 DX, Thyristor) and projected onto a projection screen (REVUE LUXUS, 1 m × 1.2 m) at a distance of 3.5 m. The strobe was mounted in a separate room behind a small hole in the wall so that the click of the release was not audible. The auditory stimulus was a sinusoidal tone of 1000 Hz which could be varied in 10 dB SPL steps. It was presented diotically by closed headphones (Philipp HP 8665/00). To generate the tactile stimulus an oscillation exciter (Brüel & Kjær, Type 4809) was used. On this oscillation exciter an aluminum plate (35 mm × 50 mm) was mounted on a threaded bolt. The aluminum plate served as 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 small hole for the bolt with the aluminum plate. The oscillation exciter was controlled by sinusoidal voltage (27.7 Hz) at three levels. The intensities and the frequencies of the tactile stimuli were determined by a pretest for getting a maximal level below the subjects' hearing threshold and the minimal excitation above the tactile threshold. To suppress initial noise the voltage for the vibration exciter was generated by a computer (type PDP 11/34) and started always at zero. The computer signal was smoothed by a low pass filter and amplified according to three intensities. The tone and the tactile stimulus were response terminated to avoid switch-off effects. The whole experiment (timing of the signals, reaction time measurement, randomization of trials, calculation of the time between the attention signal and the first stimulus) was controlled by the computer.

Procedure. The subject was seated in a dark, sound-absorbing room, 2.5 m away from the screen. Two response buttons mounted on a desk in front of her/him served as response device. The right foot (without shoe and sock) of the subject was put on a wooden box and the big toe was put on an aluminum plate. To prevent it from slipping out of place the foot was strapped on. The subject was instructed to press the response buttons with the right and the left index finger at the same time as soon as she/he noticed any of the stimuli. The beginning of each trial was signaled to the subject by a small warning lamp in front of the screen. After a random foreperiod a single stimulus or a stimulus combination was presented. The single stimuli were visual stimuli (Condition V), auditory stimuli with 70, 80, or 90 dB SPL (Condition A70, A80, or A90, respectively), and tactile stimuli with three different intensities (Condition T1, T2, T3, where T3 was the most intense stimulus). Due to individual differences in reaction times subjects PM and KK were presented with conditions V, A80, A90, T2, and T3; subjects GH and BO were presented with conditions V, A70, A90, T1, and T3 to have the reaction time in a certain range, i.e., to have the reaction times to the different single stimulus conditions not too close together and not too far apart. The double stimuli were combinations of any two stimuli of two different modalities. For example, the visual stimulus was followed by a 70-dB tone τ ms later (Condition VA70(τ)). The range of the interstimulus intervals was determined for each subject and for each stimulus combination individually. The difference in reaction time between the reaction times of two different single stimulus conditions after the training session was the mean around which the interstimulus intervals were established. Maximal intersensory facilitation can be expected when the interstimulus interval between two signals equals this difference, i.e., when physiological synchrony, rather than stimulus synchrony, is achieved (see Hershenson, 1962, Nickerson, 1973). The range enclosed six different intervals in steps of 10 ms. The triple stimuli were combinations of three stimuli of three different modalities, e.g., a weak tactile stimulus followed by the visual stimulus τ_1 later and followed by a 70-dB tone τ_2 ms later (Condition T1VA70(τ_1, τ_2)). For the triple stimulus condition, four different interstimulus intervals in steps of 10 ms were determined between the first and the second presented stimulus as well as between the second and the third stimulus. All stimulus combinations presented to the subjects can be seen in Table 2.

TABLE 2

Stimulus Combinations for All Subjects

Subject	Single	Double		Triple	
PM,KK	V	VA80(τ)	τ= 0,...,50	T2VA80(τ_1, τ_2)	τ_1= 20,...,50
	A80	VA90(τ)	τ= 0,...,50		τ_2= 50,...,80
	A90	T2V(τ)	τ= 30,...,80	T3VA80(τ_1,τ_2)	τ_1= 0,...,30
	T2	T3V(τ)	τ= 0,...,50		τ_2= 30,...,60
	T3	T2A80(τ)	τ= 30,...,80	T2VA90(τ_1,τ_2)	τ_1= 10,...,40
		T2A90(τ)	τ= 30,...,80		τ_2= 40,...,70
		T3A80(τ)	τ= 10,...,60		
		T3A90(τ)	τ= 20,...,70		
GH	V	VA70(τ)	τ= 0,...,50	T1VA70(τ_1,τ_2)	τ_1= 20,...,50
	A70	VA90(τ)	τ= 0,...,50		τ_2= 50,...,80
	A90	T1V(τ)	τ= 30,...,80	T3VA70(τ_1,τ_2)	τ_1= 0,...,30
	T1	T3V(τ)	τ= 0,...,50		τ_2= 30,...,60
	T3	T1A70(τ)	τ= 30,...,80	T1VA90(τ_1,τ_2)	τ_1= 40,...,70
		T1A90(τ)	τ= 50,...,100		τ_2= 70,...,100
		T3A70(τ)	τ= 0,...,50		
		T3A90(τ)	τ= 20,...,70		
BO	V	VA70(τ)	τ= 0,...,50	T1VA70(τ_1,τ_2)	τ_1= 20,...,50
	A70	VA90(τ)	τ= 0,...,50		τ_2= 50,...,80
	A90	T1V(τ)	τ= 30,...,80	T3VA70(τ_1,τ_2)	τ_1= 0,...,30
	T1	T3V(τ)	τ= 0,...,50		τ_2= 30,...,60
	T3	T1A70(τ)	τ= 30,...,80	T1VA90(τ_1,τ_2)	τ_1= 30,...,60
		T1A90(τ)	τ= 30,50,...,90		τ_2= 60,...,90
		T3A70(τ)	τ= 0,...,50		
		T3A90(τ)	τ= 10,...,60		

Note. τ increases in 10-ms intervals. For example, τ = 0, ..., 50 would be interpreted as 0, 10, 20, 30, 40, 50.

The subject was instructed to press the response buttons with the index fingers as fast as possible upon detecting either signal (no catch trials). The reaction time was recorded from stimulus onset to keypress for each hand separately. After the last response there was a 3-s pause before the beginning of the next trial was signaled. The first response in each trial terminated the auditory or the tactile signal. The purpose of the random foreperiod was to avoid anticipatory responses. The random foreperiod was never less than 1 s. In every trial, an exponentially distributed random duration was added to a 1-s base time, i.e., $1000 + 1000(-ln(x))$, $x \in [0.1, 1]$, where x is a uniformly distributed random number. The maximum duration of the foreperiod was therefore about 3.3 s. 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 1 in 500 trials). A block of single stimuli consisted of 20 trials for one modality (no mixture of stimuli for different modalities). A block of double stimuli consisted of 60 trials for a particular stimulus combination. The interstimulus intervals between the stimuli were in random order, and each ISI was presented ten times in a single block. A block of triple stimuli consisted of 64 trials for a particular stimulus combination. The ISIs between the stimuli were presented randomly, and a certain ISI combination (between the first and the second and the first and the third) was presented four times in each block. A single experimental session started with ten triple stimuli trials. It functioned as a warming-up period and was not used later in the data analysis. The intensities of the stimuli were different from those used in the experiment. In a random balanced order two or three single stimulus blocks, two or three double stimuli blocks, and two or three triple stimuli blocks were presented, but at most seven blocks were presented per session. After the training block and three experimental blocks there was a rest period of at most 10 min (the subject decided the actual length). After three further blocks there was again a rest period of maximum 10 min. Each session took about 60–75 min. By request the subject was informed about her/his reaction times. Including the training period (about 5 h) each subject was involved in the experiment for about 40 h. For each subject a sample of 200 responses was taken in each of the five single stimulus conditions; a sample of 100 responses was taken in each of the 48 double stimulus conditions (8 different stimulus combinations and 6 different ISIs), and a sample of 100 responses was taken in each of the 48 triple-stimulus conditions (3 different stimulus combinations and 16 different ISI combinations). In total, each subject's responses numbered 10,600.

4. PARAMETER ESTIMATION

The following strategy was used to test the coactivation models. First, we estimated all of the crucial intensity parameters from the mean response times to the triple-stimulus conditions. Then we used these estimates to calculate predictions for the means and variances for the double-stimulus conditions. The triple-stimulus conditions were used to estimate the intensity parameters because these conditions provided a more comprehensive database than the double-stimulus conditions. Although it would, in principle, be possible to predict and test the entire RT distribution functions for these models, the number of observations per condition (100) was too small to apply efficient nonparametric estimation procedures (e.g., spline estimation). For the superposition model five parameters were estimated from the means for a triple stimulus combination, e.g., T1VA70: the residual constant r, the critical count number c, and the intensity parameters α_{A70}, α_V, and α_{T1} for the auditory, visual, and tactile stimulus, respectively. For one triple stimulus condition we have 16 different interstimulus interval combinations and thus at least 16 mean reaction times.

For the diffusion model, seven parameters for each triple stimulus combination need to be determined: the intensity parameters α_A, α_V, and α_T for the auditory, visual, and tactile stimulus, respectively; the parameter β, the force directed toward the origin; the residual constant r; the diffusion coefficient σ^2; and the boundary c in terms of the size of the transition matrix, i.e., m. The time constant h is not estimated from the data but is chosen as close to zero as needed to approximate a continuous time process. For our data we decided h to be 1 ms. Considering Eq. (18) and Eq. (19) for $p(x)$ and $q(x)$ it is obvious that nothing is lost in setting $\sigma = 1$. Since the step size δ was defined as $\delta = \sigma\sqrt{h}$, we have $\delta = 1$. The criterion c was expressed in terms of the step size, i.e., $c = k\delta$, and the size of the transition matrix m was defined as $m = k + 1$. Thus the matrix size is $c + 1$. β was thought to be a function of the intensity parameters and the matrix size, i.e., $\beta = \gamma(1 + \alpha)/m$, α being the corresponding intensity parameter or one of its combinations and γ a number between 0 and 1 to be estimated. A technical reason for this was to constrain the drift rates $\mu(x)$ to take on values between 0 and 1, since the drift rates functioned as probabilities in the transition matrix. From a psychological point of view it is reasonable because with increasing intensity of a stimulus, the force directed to the origin should become larger. Thus, there remain six parameters to be estimated from the means: $\alpha_A, \alpha_V, \alpha_T, r, m, \gamma$.

For each subject, we have three different triple stimulus combinations (e.g., T1VA70, T3VA70, T1VA90); thus we have five different intensity parameters: α_V for light, α_{A70} for the 70-dB tone, α_{A90} for the 90-dB tone, α_{T1} for the weak tactile stimulus, α_{T3} for the strong tactile stimulus. To obtain a single common estimate of each intensity parameter, we estimated all of the parameters simultaneously across all three triple stimulus conditions. The criterion c was assumed to be constant across all three conditions, because the same qualitative stimuli always appeared and in the same order. Thus, for all three triple stimulus conditions (48 means), we estimated nine parameters from the data for the superposition model: the intensity parameters $\alpha_V, \alpha_{A70}, \alpha_{A90}, \alpha_{T1}$, and α_{T3}; three residual parameters r_1, r_2, and r_3; and one critical count number c. For the multichannel diffusion model we estimated ten parameters from the means: the intensity parameters $\alpha_V, \alpha_{A70}, \alpha_{A90}, \alpha_{T1}$, and α_{T3}; three residual parameters r_1, r_2, and r_3; the matrix size m; and γ.

These parameters were estimated by minimizing the function

$$SSE = \sum_{j=1}^{n} \sum_{i=1}^{m} (M_{ij} - (E_{ij}[T] + r_j))^2 \tag{29}$$

using the optimization program GAUSS.OPTMUM. M_{ij} represents the observed mean RT for the ith stimulus interval combination (τ_1, τ_2) in the jth triple stimulus combination; $E_{ij}[T]$ represents the predicted stimulus processing time for the ith interstimulus interval combination in the jth triple stimulus combination; r_j represents the residual time in the jth triple stimulus combination. The variances were not included in the estimation procedure because they were considered too noisy.[5]

The predictions for the double stimulus conditions were then generated as follows. The intensity parameters estimated from the triple stimulus conditions were used to compute the predictions for the double stimulus conditions. In addition, for the diffusion model, the value of the parameter gamma (used to determine the restoring force) estimated from the triple stimulus conditions was used again with the double stimulus conditions. However, the residual time, r, and the criterion, c, for the counter and the diffusion models may have changed across blocks of trials with different qualitative stimulus combinations. Therefore, one criterion, c, and one residual, r, were re-estimated for each of double stimulus condition (6 means). With these estimated parameters we also predicted the standard deviations of the RT for all triple stimulus and double stimulus conditions.

5. RESULTS

Before describing the main results, we note that the differences between the RT of the right hand and the RT of the left hand were negligible. In the remaining we consider only the RT of the left hand (chosen by chance). First consider the results for the single stimulus conditions.

Consistent with previous research, the mean response times observed in the present experiment (Table 3) generally imply the following ordering of stimulus conditions for subjects BO and GH:

$$T1 > T3 > V > A70 > A90.$$

[5] When the variances were included the objective function grew dramatically up to the 10-fold, 20-fold, and more.

TABLE 3

Mean Response Time for the One Stimulus Conditions for All Subjects

	BO	KK	GH	PM
T1	216.74	-	220.39	-
T2	-	202.48	-	187.98
T3	174.84	183.75	176.70	172.82
V	165.45	166.42	164.80	156.60
A70	152.44	-	148.56	-
A80	-	143.02	-	130.31
A90	140.37	130.13	132.93	123.51

From this order we expect the relation between the intensity parameters to be

$$\alpha_{T1} < \alpha_{T3} < \alpha_V < \alpha_{A70} < \alpha_{A90} \tag{30}$$

since the subjects generally responded fastest to the 90-dB tone, second fastest to the 70-dB tone, etc., and slowest to the weakest tactile stimulus. While there are some individual speed differences, the pattern is about the same for the subject group PM and KK as for the subject group GH and BO, which in each case had different stimulus conditions.

For the multichannel diffusion model the expected order of the intensity parameters (estimated from the triple stimulus conditions) was observed for all subjects (see Table 4).

TABLE 4

Estimated Parameters of the Triple Stimulus Condition for the Multichannel Diffusion Model

	BO	KK	GH	PM
m	13	16	13	12
α_{T1}	0.0929	-	0.0698	-
α_{T2}	-	0.0965	-	0.0723
α_{T3}	0.2258	0.1726	0.1266	0.1759
α_V	0.2946	0.2866	0.3151	0.2496
α_{A70}	0.3704	-	0.5494	-
α_{A80}	-	0.4850	-	0.7356
α_{A90}	0.5655	0.7692	0.8695	0.8384
β	0.1512	0.1523	0.2544	0.1603
r_1	130.56	116.70	126.04	119.59
r_2	127.18	116.33	118.65	118.00
r_3	129.39	118.38	120.59	117.28

TABLE 5

Estimated Parameters of the Triple Stimulus Condition for the Superposition Model

	BO	KK	GH	PM
c	2	2	7	4
α_{T1}	0.0200	-	0.0486	-
α_{T2}	-	0.0134	-	0.0470
α_{T3}	0.0440	0.0311	0.0646	0.0775
α_V	0.0601	0.0380	0.0982	0.0775
α_{A70}	0.2073	-	0.1235	-
α_{A80}	-	0.1719	-	2.3217
α_{A90}	0.8030	0.4361	0.1235	2.3217
r_1	137.18	122.58	118.70	122.72
r_2	133.78	125.36	111.25	121.10
r_3	136.28	124.52	118.68	129.47

However, for the superposition model, the expected order of the intensity parameters (also estimated from the triple stimulus conditions) was not observed for two subjects, GH and PM. For them an appropriate restriction in the program led to the parameters shown in Table 5. Thus, for subject GH the intensity parameters for the 70-dB tone and the 90-dB tone are identical. For subject PM the intensity parameters for the strong tactile stimulus and the light are identical as are the parameters for the tones of different intensities.

For both models the estimated residual times r_1, r_2, and r_3 are relatively invariant within subjects and they are large compared to the overall observed mean RTs. For both models the fits of the means for the triple stimulus condition are very good for all subjects. The predictions of the standard deviations for the triple stimulus conditions are poor for all subjects under most conditions. This is supported by looking at the standard error $d(s)$ of the sample standard deviation (Cramér, 1974)

$$d(s) = \frac{\sqrt{(1/n) \sum [x_i - \bar{x}]^4 - s^4}}{2s\sqrt{n}},$$

where x_i denotes the ith observation, $\bar{x}$ the sample mean, s the sample standard deviation, and n the sample size, $n = 100$. $d(s)$ was never larger than 3. The predictions were either too small (sometimes smaller than 1, the smallest observed standard deviation being about 10) or too large (about 37, the largest observed standard deviation being less than 29). For each subject the fit was somewhat better in 3 to 5 out of 12 stimulus conditions, but the best fitting conditions varied for each subject. For our parameter estimates, increasing, decreasing, or constant standard deviations as a function of the first interstimulus interval could be found. The patterns of the predicted standard deviations are very similar in both models. For some triple stimulus conditions the predicted standard deviations are almost identical for both models, for other conditions the predictions differ only by a constant. Table 6 depicts some of these results. The whole set of results can be found in Diederich (1992a).

TABLE 6

Subject GH: Observed and Predicted Mean Reaction Times and Standard Deviations for Triple Condition T1VA70 (Weak Tactile Stimulus, T1, Followed by a Visual Stimulus, V, Followed by a 70-dB Tone, A70)

	Observed	Prediction	Observed	Prediction
(τ_1, τ_2)	Mean RT	Superposition / Multi channel	SD	Superposition / Multi channel
(20,50)	171.85	173.60 / 173.40	13.20	11.37 / 11.51
(30,50)	177.22	177.91 / 177.84	13.82	10.50 / 10.02
(40,50)	179.14	181.85 / 181.33	15.48	10.04 / 9.46
(50,50)	185.60	185.52 / 183.95	19.42	9.99 / 9.85
(20,60)	175.28	175.29 / 176.07	19.16	13.12 / 14.39
(30,60)	181.74	181.30 / 181.53	17.84	11.92 / 13.00
(40,60)	185.84	185.82 / 186.10	13.47	10.93 / 12.12
(50,60)	188.39	189.89 / 189.23	14.76	10.39 / 12.07
(20,70)	176.14	177.98 / 177.80	14.76	14.85 / 16.78
(30,70)	183.44	183.64 / 183.99	16.86	13.72 / 15.72
(40,70)	190.02	188.86 / 189.56	20.10	12.54 / 14.96
(50,70)	197.95	193.58 / 193.65	21.50	11.55 / 14.75
(20,80)	177.53	178.98 / 178.89	16.24	16.18 / 18.64
(30,80)	186.71	185.05 / 185.57	17.74	15.35 / 17.95
(40,80)	191.95	190.87 / 191.85	14.25	14.34 / 17.50
(50,80)	198.47	196.28 / 196.79	15.10	13.29 / 17.43

Note. The first and the second stimulus are τ_1 ms apart, the first and the third stimulus are τ_2 ms apart.

TABLE 7

Subject GH: Observed and Predicted Mean Reaction Times and Standard Deviations for Double Stimulus Condition VA70 (visual Stimulus, V, Followed by a 70-dB Tone, A70, τ ms Apart)

	Observed	Prediction		Observed	Prediction	
τ	Mean RT	Superposition	Multi channel	SD	Superposition	Multi channel
0	138.21	136.64	137.65	25.37	9.02	3.47
10	141.51	142.19	142.94	16.71	9.06	4.34
20	149.02	147.38	148.28	18.24	9.65	6.88
30	150.83	151.65	151.70	14.30	11.26	10.06
40	154.76	154.72	153.75	15.57	13.46	12.77
50	154.96	156.71	154.97	15.50	15.56	14.84

TABLE 8

Subject GH: Observed and Predicted Mean Reaction Times and Standard Deviations for Double Condition T1A90 (Weak Tactile Stimulus, T1, Followed by a 90dB Tone, A90, τ ms Apart)

	Observed	Prediction		Observed	Prediction	
τ	Mean RT	Superposition	Multi channel	SD	Superposition	Multi channel
50	181.99	180.97	180.48	13.82	15.54	11.21
60	187.46	188.00	188.01	12.62	15.85	14.54
70	196.02	194.87	195.07	23.82	16.48	19.16
80	200.91	201.49	201.67	19.10	17.54	21.96
90	206.75	207.77	207.85	17.09	19.10	25.87
100	213.58	213.62	213.63	22.61	22.91	29.84

TABLE 9

Subject GH: Observed and Predicted Mean Reaction Times and Standard Deviations for Double Stimulus Condition T3V (Strong Tactile Stimulus, T1, Followed by a Visual, V, τ ms Apart)

	Observed	Prediction		Observed	Prediction	
τ	Mean RT	Superposition	Multi channel	SD	Superposition	Multi channel
0	160.73	160.11	160.42	21.04	10.64	12.12
10	166.56	166.09	165.06	18.85	10.71	12.47
20	172.77	171.65	171.00	16.69	11.42	13.73
30	174.04	176.35	176.16	17.26	13.09	16.07
40	179.45	180.01	180.41	19.05	15.39	18.97
50	183.38	182.70	183.86	21.47	17.80	22.02

For all double stimulus conditions the fit for the means was almost perfect for all subjects for both models. The predictions for the standard deviation for the double stimulus conditions was poor for all subjects under almost all conditions. Tables 7, 8, and 9 depict some of these results. (See Diederich (1992a) for detailed results) Comparing both models by the sum of squared errors for the means gave a slight advantage to the multichannel diffusion model (see Table 10). Note that the number of parameters estimated for the double stimulus condition is the same for both models.

TABLE 10

Sum of Squared Errors of the Means for the Poisson Superposition Model (SP) and the Multichannel Diffusion Model (DI)

	BO		KK		GH		PM	
	SSE_{SP}	SSE_{DI}	SSE_{SP}	SSE_{DI}	SSE_{SP}	SSE_{DI}	SSE_{SP}	SSE_{DI}
VA70	24.0	8.0	–	–	9.6	4.7	–	–
VA80	–	–	21.8	4.9	–	–	17.5	11.4
VA90	26.0	13.0	48.7	11.9	45.0	31.2	10.8	14.0
T1V	41.3	22.3	–	–	19.1	20.0	–	–
T2V	–	–	32.6	2.9	–	–	20.0	18.1
T3V	17.1	9.6	14.8	9.5	7.9	11.1	4.9	6.0
T1A70	16.5	8.6	–	–	29.9	21.6	–	–
T1A90	65.3	26.8	–	–	4.0	5.3	–	–
T2A80	–	–	41.0	12.4	–	–	16.7	16.2
T2A90	–	–	19.4	11.2	–	–	47.7	26.8
T3A70	5.4	17.0	–	–	17:5	17.5	–	–
T3A80	–	–	15.6	11.6	–	–	18.8	7.3
T3A90	41.5	14.2	12.2	10.4	5.5	5.4	24.7	14.8

Including the deviation from the empirical variances in the estimation procedure did not improve the overall results. The variances were still far from an optimal fit and the fit for the mean worsened considerably.

6. DISCUSSION AND SUMMARY

An experiment was conducted investigating intersensory facilitation of reaction times with stimuli of three different modalities (light, tone in two different intensities, vibration in two different intensities). Any combination of two different modalities and some combinations of three different modalities were presented at different interstimulus intervals. Three different models to explain the observed facilitation of reaction time were considered: a separate activation model, a superposition model, and a diffusion model, the last two belonging to the general class of coactivation models. Using Boole's inequality as a test for separate activation models Diederich (1992a, b) was able to show that these models cannot predict as much facilitation as observed in the data. This result is consistent with earlier work by Diederich (1985), Diederich & Colonius (1987), Gielen *et al.* (1983), Miller (1982), and others. Note that the test applies to the observable reaction time distribution. Assuming the observable reaction time is divided into additive components of processing time and residual time, the test applies only if the distributions of the processing time, S, and the residual time, R, are assumed to be invariant across stimulus conditions, i.e., the test assumes context independence. This assumption on the relation between S and R might be too simplistic and was dropped in the case of considering coactivation models. To be more fair to separate activation models specific assumptions about the variation of the residual time across various stimulus conditions have to be established requiring further research.

Both coactivation models make identical predictions about the ordering of the intensity parameters. For the diffusion model the expected ordering was observed for all four subjects whereas for the superposition model the relation was observed for two subjects. The fit for both the superposition model and the diffusion model was very good with respect to the means in the triple stimulus condition for all subjects. The intensity parameters estimated from triple stimulus conditions were then used to predict the means for the double stimulus conditions. The fit was very good for the superposition model and almost perfect for the diffusion model measured against the objective functions stated in Eq. (29). The fit for the standard deviations was poor for all subjects for both the triple stimulus conditions and the double stimulus conditions under both coactivation models. These results suggest that coactivation type models are able to account for intersensory facilitation at the level of the mean reaction times.

Our results, however, also indicate that the variability in the data has not been captured sufficiently by the coactivation models studied here. One reason for the poor fit of the standard deviation might be that the residual time is not independent of the processing time. We assumed that the observed reaction time consists of two additive variables: the stimulus processing time and a residual time in which all remaining processes were lumped together. For simplicity we chose the mean and the variance of residual time to be constant within a given condition. But since the residual time is much larger than the observed reaction time it is plausible that the residual time contains processes that are not independent from the processing time but contribute an important part to the overall variance. However, simply adding a constant to the predicted variance does not solve the problem because the difference between the observed and predicted variance is not constant across ISIs within a given stimulus configuration. The pattern of the deviations of the variance predictions is complex but there is no systematic pattern observable across subjects and/or experimental conditions. Thus, the next step in developing theory would be to give up the independence between the stimulus processing time and the residual time or, more generally, to make an assumption about the bivariate distribution of S and R.

APPENDIX

Derivation of Eq. (3)

Let $N_3(t) \equiv N_A(t) + N_V(t) + N_T(t)$ and let $N_2(t) \equiv N_A(t) + N_V(t)$. Then

$$P(N_2(t) = i) = \sum_{j=0}^{i} P(N_A(t) = j)\, P(N_V(t) = i - j)$$

and

$$\begin{aligned} P(N_3(t) = n) &= P(N_A(t) + N_V(t) + N_T(t) = n) \\ &= \sum_{i=0}^{n} P(N_2(t) = i)\, P(N_T(t) = n - i) \\ &= \sum_{i=0}^{n} \sum_{j=0}^{i} P(N_A(t) = j)\, P(N_V(t) = i - j) \\ &\quad \times P(N_T(t) = n - i) \end{aligned}$$

and therefore

$$\begin{aligned} 1 - P(S_c \leqslant t \mid \text{TVA}) &= \sum_{n=0}^{c-1} P(N_3(t) = n) \\ &= \sum_{n=0}^{c-1} \sum_{i=0}^{n} \sum_{j=0}^{i} P(N_A(t) = j) \\ &\quad \times P(N_V(t) = i - j)\, P(N_T(t) = n - i). \end{aligned}$$

With

$$P(S_i \leqslant t \mid k) - P(S_{i+1} \leqslant t \mid k) = P(N_k(t) = i)$$

it follows

$$\begin{aligned}1 - P(S_c \leqslant t) = \sum_{n=0}^{c-1} \sum_{i=0}^{n} \sum_{j=0}^{i} &[P(S_j \leqslant t \mid \mathrm{A}) - P(S_{j+1} \leqslant t \mid \mathrm{A})] \\ &\times [P(S_{i-j} \leqslant t \mid \mathrm{V}) - P(S_{i-j+1} \leqslant t \mid \mathrm{V})] \\ &\times [P(S_{n-i} \leqslant t \mid \mathrm{T}) - P(S_{n-i+1} \leqslant t \mid \mathrm{T})]\end{aligned}$$

Derivation of Eq. (11)

For positive random variables,

$$E(X) = \int_0^\infty (1 - F(x))\, dx.$$

Derivation in three parts:

(1) $0 \leqslant t \leqslant \tau_1$.

$$\begin{aligned}E_1 &= \int_0^{\tau_1} \exp\{\alpha_\mathrm{T} t\} \sum_{j=0}^{c-1} \frac{(\alpha_\mathrm{T} t)^j}{j!}\, dt \\ &= -\frac{\exp\{-\alpha_\mathrm{T} t\}}{\alpha_\mathrm{T}} \sum_{j=0}^{c-1} \frac{(\alpha_\mathrm{T} t)^j}{j!}\bigg|_0^{\tau_1} + \int_0^{\tau_1} \frac{\exp\{\alpha_\mathrm{T} t\}}{\alpha_\mathrm{T}} \sum_{j=1}^{c-1} \frac{(\alpha_\mathrm{T} t)^{j-1} \alpha_\mathrm{T}}{(j-1)!}\, dt \\ &= -\frac{\exp\{-\alpha_\mathrm{T} t\}}{\alpha_\mathrm{T}} \sum_{j=0}^{c-1} \frac{(\alpha_\mathrm{T} t)^j}{j!}\bigg|_0^{\tau_1} - \frac{\exp\{-\alpha_\mathrm{T} t\}}{\alpha_\mathrm{T}} \sum_{j=0}^{c-2} \frac{(\alpha_\mathrm{T} t)^j}{j!}\bigg|_0^{\tau_1} \\ &\quad \cdots - \frac{\exp\{-\alpha_\mathrm{T} t\}}{\alpha_\mathrm{T}} \sum_{j=0}^{1} \frac{(\alpha_\mathrm{T} t)^j}{j!}\bigg|_0^{\tau_1} - \frac{\exp\{-\alpha_\mathrm{T} t\}}{\alpha_\mathrm{T}} \\ &= \frac{c}{\alpha_\mathrm{T}} - \frac{\exp\{-\alpha_\mathrm{T} \tau_1\}}{\alpha_\mathrm{T}} \sum_{j=0}^{c-1} \frac{(\alpha_\mathrm{T} \tau_1)^j}{j!} (c-j).\end{aligned}$$

(2) $\tau_1 \leqslant t \leqslant \tau_2$

$$\begin{aligned}E_2 &= \int_{\tau_1}^{\tau_2} \exp\{-((\alpha_\mathrm{T} + \alpha_\mathrm{V}) t - \alpha_\mathrm{V} \tau_1)\} \times \sum_{j=0}^{c-1} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V}) t - \alpha_\mathrm{V} \tau_1)^j}{j!}\, dt \\ &= -\frac{\exp\{-(\alpha_\mathrm{T} + \alpha_\mathrm{V}) t + \alpha_\mathrm{V} \tau_1\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V}} \times \sum_{j=0}^{c-1} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V}) t - \alpha_\mathrm{V} \tau_1)^j}{j!}\bigg|_{\tau_1}^{\tau_2} \\ &\quad + \int_{\tau_1}^{\tau_2} \frac{\exp\{-(\alpha_\mathrm{T} + \alpha_\mathrm{V}) t + \alpha_\mathrm{V} \tau_1\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V}} \times \sum_{j=1}^{c-1} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V}) t - \alpha_\mathrm{V} \tau_1)^{j-1} (\alpha_\mathrm{T} + \alpha_\mathrm{V})}{(j-1)!}\, dt \\ &= -\frac{\exp\{-(\alpha_\mathrm{T} + \alpha_\mathrm{V}) t + \alpha_\mathrm{V} \tau_1\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V}} \times \sum_{j=0}^{c-1} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V}) t - \alpha_\mathrm{V} \tau_1)^j}{j!}\bigg|_{\tau_1}^{\tau_2} \\ &\quad - \frac{\exp\{-(\alpha_\mathrm{T} + \alpha_\mathrm{V}) t + \alpha_\mathrm{V} \tau_1\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V}} \times \sum_{j=0}^{c-2} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V}) t - \alpha_\mathrm{V} \tau_1)^j}{j!}\bigg|_{\tau_1}^{\tau_2} \\ &\quad \vdots \\ &= -\frac{\exp\{-(\alpha_\mathrm{T} + \alpha_\mathrm{V}) \tau_2 + \alpha_\mathrm{V} \tau_1\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V}} \times \sum_{j=0}^{c-1} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V}) \tau_2 - \alpha_\mathrm{V} \tau_1)^j}{j!} (c-j) \\ &\quad + \frac{\exp\{-(\alpha_\mathrm{T} + \alpha_\mathrm{V}) \tau_1 + \alpha_\mathrm{V} \tau_1\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V}} \times \sum_{j=0}^{c-1} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V}) \tau_1 - \alpha_\mathrm{V} \tau_1)^j}{j!} (c-j) \\ &= \frac{\exp\{-\alpha_\mathrm{T} \tau_1\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V}} \sum_{j=0}^{c-1} \frac{(\alpha_\mathrm{T} \tau_1)^j}{j!} (c-j) - \frac{\exp\{-((\alpha_\mathrm{T} + \alpha_\mathrm{V}) \tau_2 - \alpha_\mathrm{V} \tau_1)\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V}} \times \sum_{j=0}^{c-1} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V}) \tau_2 - \alpha_\mathrm{V} \tau_1)^j}{j!} (c-j).\end{aligned}$$

(3) $t \geqslant \tau_2$.

$$\begin{aligned}E_3 &= \int_{\tau_2}^{\infty} \exp\{-((\alpha_\mathrm{T} + \alpha_\mathrm{V} + \alpha_\mathrm{A}) t - \alpha_\mathrm{V} \tau_1 - \alpha_\mathrm{A} \tau_2)\} \times \sum_{j=0}^{c-1} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V} + \alpha_\mathrm{A}) t - \alpha_\mathrm{V} \tau_1 - \alpha_\mathrm{A} \tau_2)^j}{j!}\, dt \\ &= -\frac{\exp\{-((\alpha_\mathrm{T} + \alpha_\mathrm{V} + \alpha_\mathrm{A}) t - \alpha_\mathrm{V} \tau_1 - \alpha_\mathrm{A} \tau_2)\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V} + \alpha_\mathrm{A}} \times \sum_{j=0}^{c-1} \frac{((\alpha_\mathrm{T} + \alpha_\mathrm{V} + \alpha_\mathrm{A}) t - \alpha_\mathrm{V} \tau_1 - \alpha_\mathrm{A} \tau_2)^j}{j!}\bigg|_{\tau_2}^{\infty} (c-j) \\ &= \frac{\exp\{-(\alpha_\mathrm{T} \tau_2 + \alpha_\mathrm{V} (\tau_2 - \tau_1))\}}{\alpha_\mathrm{T} + \alpha_\mathrm{V} + \alpha_\mathrm{A}} \times \sum_{j=0}^{c-1} \frac{(\alpha_\mathrm{T} \tau_2 + \alpha_\mathrm{V} (\tau_2 - \tau_1))^j}{j!} (c-j).\end{aligned}$$

Adding E_1, E_2, and E_3 yields Eq. (11).

Derivation of Eq. (14)

For positive random variables,

$$E[X^2] = 2\int_0^\infty x(1-F(x))\,dx.$$

Using the same procedure as for Eq. (11) yields Eq. (14).

Derivation of Eq. (27)

$$\mathbf{S}_1 \equiv \sum_{i=k}^{N} \mathbf{Q}^{i-k} = \mathbf{Q}^0 + \mathbf{Q}^1 + \mathbf{Q}^2 + \cdots + \mathbf{Q}^{N-k}$$

$$\mathbf{QS}_1 = \sum_{i=k}^{N} \mathbf{Q}^{i-k+1} = \mathbf{Q}^1 + \mathbf{Q}^2 + \cdots + \mathbf{Q}^{N-k} + \mathbf{Q}^{N-k+1}$$

$$\mathbf{S}_1 - \mathbf{QS}_1 = (\mathbf{I}-\mathbf{Q}) - \mathbf{S}_1 = \mathbf{I} - \mathbf{Q}^{N-k+1}$$

$$\mathbf{S}_1 = (\mathbf{I}-\mathbf{Q})^{-1}(\mathbf{I}-\mathbf{Q}^{N-k+1}).$$

$$\mathbf{S}_2 \equiv \sum_{i=k}^{N} i\mathbf{Q}^{i-k} = k\mathbf{Q}^0 + (k+1)\,\mathbf{Q}^1 + \cdots + (N-1)\,\mathbf{Q}^{N-1-k} + N\mathbf{Q}^{N-k}$$

$$\mathbf{QS}_2 = k^1 + (k+1)\,\mathbf{Q}^2 + \cdots + (N-1)\,\mathbf{Q}^{N-k} + N\mathbf{Q}^{N-k+1}$$

$$\mathbf{S}_2 - \mathbf{QS}_2 = k\mathbf{I} + \mathbf{Q}^1 + \mathbf{Q}^2 + \cdots + \mathbf{Q}^{N-k} - N\mathbf{Q}^{N-k+1}$$

$$(\mathbf{I}-\mathbf{Q})\,\mathbf{S}_2 = (\mathbf{I}-\mathbf{Q}^{N-k+1}) - \mathbf{I} + k\mathbf{I} - N\mathbf{Q}^{N-k+1}$$

$$\mathbf{S}_2 = (\mathbf{I}-\mathbf{Q})^{-1}\,[(\mathbf{I}-\mathbf{Q})^{-1}(\mathbf{I}-\mathbf{Q}^{N-k+1}) + (k-1)\,\mathbf{I} - N\mathbf{Q}^{N-k+1}].$$

Setting $k=1$, $N=n_1$, $\mathbf{Q}=\mathbf{Q}_1$, multiplying this left by $h\mathbf{Z}'$ and right by $\mathbf{R}_1$ leads to the first part of the sum in Eq. (27). Setting $k=n_1+1$, $N=n_2$, $\mathbf{Q}=\mathbf{Q}_2$, multiplying this left by $h\mathbf{Z}'_2$ and right by $\mathbf{R}_2$ leads to the second part of the sum in Eq. (27). Setting $k=n_2+1$, $N=\infty$, $\mathbf{Q}=\mathbf{Q}_3$, multiplying this by $h\mathbf{Z}'_3$, and right by $\mathbf{R}_3$ leads to the third part of the sum in Eq. (27).

Derivation of Eq. (28)

$$\mathbf{S}_3 \equiv \sum_{i=k}^{N} i^2\mathbf{Q}^{i-k} = k^2\mathbf{Q}^0 + (k+1)^2\,\mathbf{Q}^1 + (k+2)^2\,\mathbf{Q}^2 + \cdots + N^2\mathbf{Q}^{N-k}$$

$$\mathbf{QS}_3 = \sum_{i=k}^{N} i^2\mathbf{Q}^{i-k+1} = k^2\mathbf{Q}^1 + (k+1)^2\,\mathbf{Q}^2 + \cdots + (N-1)^2\,\mathbf{Q}^{N-k} + N^2\mathbf{Q}^{N-k+1}$$

$$\mathbf{S}_3 - \mathbf{QS}_3 = k^2\mathbf{Q}^0 + [(k+1)^2 - k^2]\,\mathbf{Q} + \cdots + [N^2-(N-1)^2]\,\mathbf{Q}^{N-k} - N^2\mathbf{Q}^{N-k+1}$$

$$(\mathbf{I}-\mathbf{Q})\mathbf{S}_3 = \sum_{i=k}^{N}[i^2-(i-1)^2]\,\mathbf{Q}^{i-k} + (k-1)^2\,\mathbf{I} - N^2\mathbf{Q}^{N-k+1}$$

$$= \sum_{i=k}^{N}(2i-1)\,\mathbf{Q}^{i-k} + (k-1)^2\,\mathbf{I} - N^2\mathbf{Q}^{N-k+1}$$

$$= 2\sum_{i=k}^{N} i\mathbf{Q}^{i-k} - \sum_{i=k}^{N}\mathbf{Q}^{i-k} + (k-1)^2\,\mathbf{I} - N^2\mathbf{Q}^{N-k+1}$$

$$\mathbf{S}_3 = 2(\mathbf{I}-\mathbf{Q})^{-2}\,[(\mathbf{I}-\mathbf{Q})^{-1}(\mathbf{I}-\mathbf{Q}^{N-k+1}) + (k-1)\,\mathbf{I} - N\mathbf{Q}^{N-k+1}] - (\mathbf{I}-\mathbf{Q})^{-1}\,[(\mathbf{I}-\mathbf{Q})^{-1}(\mathbf{I}-\mathbf{Q}^{N-k+1}) - (k-1)^2\,\mathbf{I} + N^2\mathbf{Q}^{N-k+1}.$$

REFERENCES

Ashby, F. G. (1983). A biased random walk model for two choice reaction times. *Journal of Mathematical Psychology*, **27**, 277–297.

Audley, R. J., & Pike, A. R. (1965). Some alternative stochastic models of choice. *British Journal of Mathematical and Statistical Psychology*, **18**, 207–255.

Bernstein, I. H., Clark, M. H., & Edelstein, B.H. (1969). Effects of an auditory signal on visual reaction time. *Journal of Experimental Psychology*, **80**, 567–569.

Bernstein, I. H., Rose, R., & Ashe, V. M. (1970). Energy integration in intersensory facilitation. *Journal of Experimental Psychology*, **86**, 196–203.

Bernstein, I. H., Chu, Ph. K., Briggs, P., & Schurman, D. L. (1973). Stimulus intensity and foreperiod effects in intersensory facilitation. *Quartely Journal of Experimental Psychology*, **25**, 171–181.

Bhat, U. N. (1984). *Elements of applied stochastic processes*. New York: Wiley.

Bhattacharya, R. N., & Waymire, E. C. (1990). *Stochastic processes with applications*. New York: Wiley.

Busemeyer, J. R., & Goldstein, W. M. (1992). Linking together different measures of preference: A dynamic model of matching derived from decision field theory. *Organizational Behavior and Human Decision Processes*, **52**, 370–396.

Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences*, **23**, 155–182.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic–cognitive approach to decision making. *Psychological Review*, **100**, 432–459.

Cinlar, E. (1975). *Introduction to stochastic processes*. New York: Prentice–Hall.

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.

Cox, D. R., & Miller, H. D. (1977). *The theory of stochastic processes*. London: Chapman & Hall.

Cramér, H. (1974). Mathematical methods of statistics. Princeton: Princeton Univ. Press.

Diederich, A. (1985). *Eine experimentelle und theoretische Analyse der intersensorischen Bahnung*. Unpublished Diplomarbeit.

Diederich, A. (1987). Measuring facilitation in the motor component. In E. E. Roskam, & R. Suck (Eds.), *Progress in Mathematical Psychology*, 1 (pp. 209–217). Amsterdam: North-Holland.

Diederich, A. (1992a). *Intersensory facilitation: Race, superposition, and diffusion models for reaction time to multiple stimuli* (Dissertation). Frankfurt: Verlag Peter Lang.

Diederich, A. (1992b). Probability inequalities for testing separate activation models of divided attention. *Perception & Psychophysics*, **52**(6), 714–716.

Diederich, A., & Colonius, H. (1987). Intersensory facilitation in the motor component? A reaction time analysis. *Psychological Research*, **49**, 23–29.

Diederich, A., & Colonius, H. (1991). A further test of the superposition model for redundant signals effect in bimodal detection. *Perception & Psychophysics*, **50**, 83–85.

Donders, F. C. (1868). Die Schnelligkeit psychischer Prozesse. *Archiv für Anatomie und Physiologie*, 657–681.

Dunlap, K., & Wells, G. E. (1910). Some experiments with reactions to visual and auditory stimuli, *Psychological Review*, **17**, 319–335.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology*, **2**, 312–329.

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**(2), 161–168.

Giray, M., & Ulrich, R. (1994). Motor coactivation revealed by response force in devided and focused attention. *Journal of Experimental Psychology: Human Perception and Performance*, **19**(6), 1278–1291.

Green, D. M., & Luce, R. D. (1973). Speed-accuracy trade off in auditory detection. In S. Kornblum (Ed.). *Attention and performance*, IV (pp. 547–569). New York: Academic Press.

Green, D. M., & Luce, R. D. (1974). Timing and counting mechanisms in auditory discrimination and reaction time. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.). *Contemporary developments in mathematical psychology* (Vol. II, pp. 372–415). San Francisco: Freeman.

Heath, R.A. (1981). A tandem random walk model for psychological discrimination. *British Journal of Mathematical and Statistical Psychology*, **34**, 76–92.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, **63**, 289–293.

Horn, R. A., & Johnson, Ch. R. (1990). *Matrix analysis*. Cambridge: Cambridge Univ. Press.

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London: Academic Press.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, **40**, 77–105.

Luce, R. D. (1986). *Response times*. New York: Oxford Univ. Press.

Luce, R. D., & Green, D. M. (1972). A neural timing theory for response times and the psychophysics of intensity. *Psychological Review*, **79**, 14–57.

McGill, W. J. (1963). Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds). *Handbook of mathematical psychology* (Vol I, pp. 309–360). New York: Wiley.

McGill, W. J. (1967). Neural counting mechanisms and energy detection in audition. *Journal of Mathematical Psychology*, **4**, 351–376.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, **14**, 247–279.

Morrell, L. K. (1967). Intersensory facilitation of reaction time. *Psychonomic Science*, **8**, 77–78.

Nickerson, R. S. (1973). Intersensory facilitation of reaction time: Energy summation or preparation enhancement? *Psychological Review*, **80**, 489–509.

Pacut, A. (1980). Mathematical modelling of reaction latency: The structure of the models and its motivation. *Acta Neurobiologiae Experimentalis*, **40**, 199–215.

Pacut, A. (1982). Mathematical modelling of reaction latency. Part II. A model of escape reaction and escape learning. *Acta Neurobiologiae Experimentalis*, **42**, 379–395.

Pacut, A., & Tych, W. (1982). Mathematical modelling of reaction latency. Part III. A model of avoidance reaction latency and avoidance learning. *Acta Neurobiologiae Experimentalis*, **42**, 397–420.

Pike, A. R. (1966). Stochastic models of choice behavior: Response probabilities and latencies of finite Markov chain systems. *British Journal of Mathematical and Statistical Psychology*, **19**, 15–32.

Pike, A. R. (1968). Latency and relative frequency of response in psychophysical discrimination. *British Journal of Mathematical and Statistical Psychology*, **21**, 161–182.

Pike, A. R. (1971). The latencies of correct and incorrect responses in discrimination tasks: Their interpretation in terms of a model based on simple counting. *Perception & Psychophysics*, **9**, 455–460.

Pike. A. R. (1973). Response latency models for signal detection. *Psychological Review*, **80**, 53–68.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transaction of the New York Academy of Sciences*, **24**, 574–590.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59–108.

Ricciardi, L. M. (1977). *Lecture notes in biomathematics*. Berlin: Springer-Verlag.

Schwarz, W. (1989). A new model to explain the redundant-signals effect. *Perception & Psychophysics*, **46**(5), 498–500.

Smith, P. L. (1990). *Attention, redundancy, and perceptual integration*. Paper presented at the 23rd Annual Mathematical Psychology Meetings, August, University of Toronto, Canada.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, **25**, 251–260.

Todd, J. W. (1912). Reaction to the multiple stimuli. *Archives of Psychology*, **25**, 1–65.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge Univ. Press.

Tuckwell, H. C. (1989). *Stochastic processes in neurosciences*. Philadelphia: SIAM.

Ulrich, R., & Giray, M. (1986). Separate-activation models with variable base time: Testabolity and checking of cross-channel dependency. *Perception & Psychophysics*, **36**, 545–558.

Welch, R. B., & Warren, D. V. (1986). Intersensory interaction. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.). *Handbook of perception and human performance* (Vol. I, pp. 25-1–25-35). New York: Wiley.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology*. New York: Holt.

Received: May 17, 1993