

Possibly Dependent Probability Summation of Reaction Time

HANS COLONIUS

Purdue University

The redundant target effect is frequently observed in certain reaction time experiments. A probability summation mechanism is commonly invoked to explain the observed decrease of reaction time (RT). Here, probability summation models of RT are generalized by dropping the assumption of stochastic independence among processing times. An upper bound for the amount of statistical facilitation possible under negative dependence is derived for the bivariate case. Consequences for models of binocular summation and bisensory interaction are outlined and the main result is illustrated by a comparison of visual/auditory RT data with a simulated dependent probability summation mechanism. The impact of a base time component of reaction time in these simulations is characterized. Finally, it is shown that the detection paradigm can formally be treated under the reaction time paradigm. © 1990 Academic Press, Inc.

1. INTRODUCTION

Consider an experimental situation where the subject is required to discriminate between two visual targets, X and Y , say. For example, in a two-choice reaction time (RT) task the subject may be asked to respond with a left-hand button press if target X is detected and with a right-hand button press if Y is detected. On a given trial, only targets assigned to a single response are displayed. Now assume the number of simultaneously presented targets is varied, e.g., two X 's are presented within each trial. A *redundant targets effect* is obtained if there is a decrease of (average) RT with an increasing number of targets. The effect is pervasive; redundancy gains have been observed with visual as well as with bimodal (visual/auditory) targets in choice RT as well as in single RT situations (e.g., Hershenson, 1962; Miller, 1982; Gielen, Schmidt, & van den Heuvel, 1983).

The presence of a redundant targets effect suggests important conclusions about fundamental characteristics of the information processing mechanisms in a given experimental situation. For example, in the paradigm outlined above an assumption of serial processing of the redundant targets would obviously not be compatible with a decrease in mean RT (given the assumption that individual target processing is performed at the same average speed in both the redundant and the

Reprint requests should be addressed to Dr. Hans Colonius, Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907.

single target condition). Not only the existence but also the magnitude of the redundant target effect is of interest, however. To see this consider the following prominent explanation of the effect, which will be the main topic of this paper. Assume that the targets are being processed in parallel and that the processing time for each of the targets has a random duration. If these distributions overlap, then some of the 'longer' RTs to Target X will be replaced by 'shorter' RTs to Target Y . Thus, on the average, RT will be faster with multiple targets than with a single one. According to this *probability summation hypothesis*, this redundancy gain is a statistical effect of facilitation. The force of this explanation, however, hinges on whether a probability summation mechanism is able to predict the observed magnitudes of the effect. This then, as shown in the remainder of this introduction, depends on the specifics of the postulated probability summation mechanism.

In Section 2, two RT paradigms together with some empirical evidence for probability summation will be presented. In a subsequent theoretical section, the structure of bivariate dependence is explored. The concepts developed there are then employed in Section 4 to specify probability summation models in general and to characterize their RT predictions under different assumptions of dependence. Section 5 presents a small simulation study for demonstrational purposes. Moreover, effects of a base time (e.g. motor) component of RT are discussed. In a concluding section, it is shown that the detection paradigm can formally be treated under the RT paradigm and some open more general problems are touched upon.

The concept of probability summation has a long history in psychological theorizing (see, e.g., Pirenne, 1943; Brindley, 1963). The basic idea, in the context of a detection task, is as follows: there are n elementary detector mechanisms each reporting the presence of a stimulus with a certain probability p_i ($i = 1, \dots, n$), say; in the redundant target situation, each detector would be associated with one of the redundant targets. Assuming (1) that the subject reports detection if at least one detector is in the detection state, and (2) that all n detector mechanisms are stochastically independent, the probability of the subject reporting the presence of a stimulus amounts to

$$P_{1\dots n} = 1 - (1 - p_1)(1 - p_2) \cdots (1 - p_n). \quad (1)$$

The improvement of performance ensuing from having n detectors "at work" rather than one is articulated in the inequality following from (1),

$$P_{1\dots n} \geq \max_i p_i. \quad (2)$$

Moreover, $P_{1\dots n}$ is a nondecreasing function of n , the number of elementary detectors involved. Obviously, a number of other decision rules are conceivable and, perhaps, more realistic in certain detection situations; for example, one might require some fraction of the elementary detectors to report detection. The focus of this paper, however, is on modifying the other main assumption, viz. *stochastic independence*.

To appreciate the effect of dropping stochastic independence, consider the special case of $n = 2$. Let q_{ij} denote the probability that neither detector i nor detector j

is in the detection state. Under stochastic independence, we then have for the probability P_{ij} of detecting the stimulus with only detectors i and j involved

$$\begin{aligned} P_{ij} &= 1 - q_{ij} \\ &= 1 - (1 - p_i)(1 - p_j). \end{aligned} \quad (3)$$

Positive dependence between i and j can be defined as

$$q_{ij} > (1 - p_i)(1 - p_j); \quad (4)$$

this implies that P_{ij} is smaller than under stochastic independence. Analogously, negative dependence (reversing inequality (4)) implies that P_{ij} becomes larger.

In what follows, *simple reaction time* (*sRT*) will be considered as a measure of sensitivity rather than detection probability. In the last section, we show that the detection situation can be formally subsumed under the sRT situation. For RT experiments, the idea of probability summation typically takes the form of a statistical "race" between neural activities generated in different sensory channels. The "winning" channel will report activity to some central mechanism eliciting programming and execution of a button press response. In the next section, we sketch two empirical situations where probability summation mechanisms have been invoked to account for RT data. It will be argued that in order to improve assessment of the empirical validity of probability summation mechanisms, one should consider disposing of the commonly made a priori assumption of stochastic independence. To appreciate intuitively the effect of dropping independence, let us assume negative dependence between the channel processing times, random variables X and Y , say. Then large values of X tend to show up with small values of Y and vice versa. If the (marginal) distributions of X and Y are assumed invariant, increasing the degree of negative dependence between X and Y will force the mean of the minimum of X and Y to decrease, resulting in a speeding up to responses eventually beyond the minimum of the means of X and Y (see Proposition 4.5 below for a rigorous version of this argument). On the other hand, introducing positive dependence between X and Y has no facilitatory effect. In the case of very high positive correlation, for example, the mean of the minimum of X and Y cannot be inferior to the minimum of the means of X and Y .

While it seems that stochastic independence of channel processing times has been postulated mainly for reasons of analytic tractability, the most obvious theoretical justification for assuming negative dependent processing relies on an 'allocatable capacity' argument. If total processing capacity is constant for two channels across trials, then a trial-by-trial reallocation of capacity between two channels would produce negative dependence between the corresponding processing times. As we will see, such capacity need not be limited in the usual sense of degraded average performance on a single channel with increased load (see, e.g., Townsend & Ashby, 1983). Indeed, most of our later development will be devoted to unlimited capacity, taken in the sense that the marginal individual channel distribution is unchanged

whether or not the channel is active. Of course, a more specific justification for dependent processing is desirable and may in fact often be attainable by taking the particular experimental and theoretical setup into account.

2. PROBABILITY SUMMATION MODELING OF REACTION TIME: TWO PARADIGMS

EXAMPLE 2.1: INTERSENSORY INTERACTION. A typical experimental setup is a bimodal detection task where auditory and visual stimuli are presented and the subject is instructed to respond with a button press as soon as a stimulus is perceived. On single-signal trials, only one signal is presented; on redundant-signal trials, both signals are presented either simultaneously or with a short interstimulus interval (ISI). The usual finding is that RT is shorter for redundant-signal trials. Raab (1962) proposed a probability summation mechanism ("statistical facilitation") where the two signals are processed in parallel in different sensory channels with RT being determined by whichever stimulus modality is registered first at some central mechanism activating the response process. Given some overlap of the distributions of finishing times for the two response-activation processes, responses to redundant signals will be faster, on average, than responses to either of the stimuli presented alone.

Raab's model fell slightly short of predicting all the facilitation reported in Hershenson (1962) and this was replicated in a number of later studies. Raab assumed the response-activation processes to be independent and normally distributed with equal variances. In a more recent study, Gielen *et al.* (1983) dropped these distributional assumptions by numerically estimating the redundant-signal distribution from the single-signal distributions. Nonetheless, they found mean RT in the redundant-signal situation to be 12–13 msec faster than predicted by their probability summation-type model ($p < .05$). As will become clear from our theoretical results below, it is conceivable that introducing a maximal negative correlation in Gielen *et al.*'s model could result in a better fit. On the other hand, in an interesting simulation study using the inversion method to generate negative dependence (see Section 5 below) Miller (1986) could still reject the probability summation model for his data.

EXAMPLE 2.2: BINOCULAR INTERACTION. The superiority of binocular over monocular viewing on a variety of visual tasks is well documented; see, e.g., the review papers by Blake and Fox (1973) and by Blake, Sloane, and Fox (1981). The phenomenon is often referred to as "binocular summation". While most studies deal with measuring binocular sensitivity vs. monocular sensitivity at threshold levels (e.g., Campbell & Green, 1965) or use matching procedures (e.g., Legge & Rubin, 1981), some studies have examined probability summation in binocular RT (Blake, Martens, & Di Gianfillipo, 1980; Blake, Martens, Garrett, & Westendorf, 1980;

Harwerth, Smith, & Levi, 1980). When simple RT to the onset of sinusoidal gratings is measured, binocular stimulation consistently yields shorter RTs regardless of grating contrast or spatial frequency. Binocular RT is about 10% faster than monocular RT. In analogy to the intersensory interaction studies, the issue is whether (or how much of) binocular summation could be due to a probability summation mechanism between the eyes. Again, the typical finding is that binocular probability summation falls short of predicting all of the facilitation reported for binocular RT. This conclusion, however, is based on assuming stochastically independent processing (see, e.g., Blake, Martens, Garrett, & Westendorf, 1980; Westendorf & Blake, 1988). On the other hand, even after negative dependence is introduced there will most likely remain a certain proportion of the binocular advantage that may be due to some genuine neural summation. This is suggested by studies involving conditions where neural interaction between the eyes is ruled out, such as stimulation of noncorresponding retinal areas, separation of the two monocular inputs over time, and stimulation with dissimilar monocular targets (e.g., Eriksen & Greenspon, 1968). Under conditions where only probability summation could be responsible for the binocular advantage, the improvement of binocular vs. monocular RT is less than observed under conditions allowing for neural summation to occur (see Blake, Martens, & Di Gianfillipo, 1980).

To sum up, in empirical RT studies of intersensory and of binocular interaction, the concept of probability summation is commonly invoked to account for part—though not all—of the RT facilitation observed with bimodal or binocular stimuli, respectively. Thus, in order to assess the true amount of neural summation showing up in these experiments, it seems essential to estimate the limits of performance set by probability summation alone.

3. THE STRUCTURE OF BIVARIATE DEPENDENCE

To prepare the definition of probability summation models given in the next section, we need some facts from probability theory characterizing the structure of bivariate distributions, which we collect in the lemmas below. However, readers more interested in the applications may skip this section.

For any cumulative distribution function (cdf) F , its inverse is defined by $F^{-1}(z) = \inf\{x: F(x) \geq z\}$.

LEMMA 3.1 (cf. Hoeffding, 1940; Fréchet, 1951; Whitt, 1976). *Let $\theta(F_X, F_Y)$ be the set of all bivariate cdf's H on R^2 having F_X and F_Y as marginal cdf's for random variables X and Y with finite positive variances. Then*

(a) *for any $H \in \theta(F_X, F_Y)$ and all $(x, y) \in R^2$,*

$$\max[F_X(x) + F_Y(y) - 1, 0] \leq H(x, y) \leq \min[F_X(x), F_Y(y)]; \quad (5)$$

(b) the bounds $H^-(x, y) = \max[F_X(x) + F_Y(y) - 1, 0]$ and $H^+(x, y) = \min[F_X(x), F_Y(y)]$ are elements of $\theta(F_X, F_Y)$ and achieve minimal and maximal correlation between X and Y , resp., among all bivariate distributions with these given marginals;

(c) let U be a uniform random variable; then

$$[F_X^{-1}(U), F_Y^{-1}(U)] \text{ has cdf } H^+$$

and

$$[F_X^{-1}(U), F_Y^{-1}(1 - U)] \text{ has cdf } H^-.$$

An outline of the proof of this lemma is given in the appendix. The left and right hand sides of inequality (5), H^- and H^+ , are known as the lower and upper Fréchet bounds. Both H^+ and H^- are singular bivariate distributions; H^+ assigns probability 1 to the set $\{(x, y): F_X(x) = F_Y(y)\}$ and H^- to the set $\{(x, y): F_X(x) + F_Y(y) = 1\}$. These extremal cdf's and the associated extreme correlations will be used to characterize dependent probability summation mechanisms. Part (c) of Lemma 3.1 is instrumental in simulating H^+ and H^- (see Section 5).

The sense in which the Fréchet bounds constitute dependence is directly opposite to that of stochastic independence: there exists a perfect monotone relation between the two random variables ('monotone dependence').

LEMMA 3.2 (Kimeldorf & Sampson, 1978, p. 898). *Let X, Y be random variables with continuous cdf's F_X and F_Y ; the joint cdf of (X, Y) is H^+ (H^-) if and only if there exists an increasing¹ (decreasing) function g for which $P[Y = g(X)] = 1$.*

The structure of a bivariate distribution can be decomposed into two logically independent parts, the marginals and the way the marginals are combined to generate the bivariate distribution. To be more precise, let $H(x, y)$ be a continuous cdf with marginal cdf's F_X and F_Y . The uniform representation U_H of H^2 as defined by Kimeldorf and Sampson (1975) then is

$$U_H(u, v) = H[F_X^{-1}(u), F_Y^{-1}(v)], \quad 0 \leq u, v \leq 1.$$

Observe that $U_H(u, v)$ is a cdf on the unit square with both marginals being uniform on $[0, 1]$. The class of all continuous cdf's can be decomposed into equivalence classes determined by the equivalence relation

$$H \sim K \quad \text{if and only if} \quad U_H = U_K.$$

¹ Note that by 'increasing' ('decreasing') is meant 'nondecreasing' ('nonincreasing') throughout this paper.

² Continuity is assumed here to assure uniqueness of U_H . For a slightly more general definition ('copula') see Schweizer and Wolff (1981).

EXAMPLE 3.3: FARLIE–GUMBEL–MORGENSTERN (FGM) DISTRIBUTION. The bivariate cdf of the FGM distribution (see, e.g., Johnson & Kotz, 1972; for an application, see Colonius, 1986) with marginals F_X and F_Y is

$$H(x, y) = F_X(x) F_Y(y) \{1 + \alpha [1 - F_X(x)] [1 - F_Y(y)]\}, \tag{6}$$

with $\alpha \in [-1, 1]$; its corresponding uniform representation is

$$U_H(u, v) = uv [1 + \alpha(1 - u)(1 - v)] \tag{7}$$

with $0 \leq u, v \leq 1$ and $\alpha \in [-1, 1]$.

EXAMPLE 3.4: FRANK’S DISTRIBUTION. Frank (1979) introduced, and Genest (1987) further analyzed the bivariate cdf

$$H(x, y) = \log_\alpha \{1 + [(\alpha^{F(x)} - 1)(\alpha^{G(y)} - 1)/(\alpha - 1)]\} \tag{8}$$

with $\alpha \in (0, 1) \cup (1, \infty)$ and marginals $F(x)$ and $G(y)$; its uniform representation is

$$U_H(u, v) = \log_\alpha \{1 + [(\alpha^u - 1)(\alpha^v - 1)/(\alpha - 1)]\} \tag{9}$$

with $0 \leq u, v \leq 1$ and α as before.

The uniform representation of a bivariate distribution captures those properties of the joint distribution which are invariant under increasing transformations of the random variables. This is contained in the following lemma.

LEMMA 3.5 (Kimeldorf & Sampson, 1978, p. 900). *Let (X, Y) and (V, W) have continuous bivariate cdf’s H and K , respectively. Then $H \sim K$ if and only if there exist increasing functions g_1 and g_2 such that the joint cdf of $g_1(X)$ and $g_2(Y)$ is K .*

To summarize, any continuous bivariate distribution function can be decomposed into its ‘structure,’ defined by the equivalence class determined by its uniform representation, and its marginals. Conversely, given any uniform representation and a pair of continuous univariate distributions, there exists a unique bivariate distribution with these two components. According to Lemma 3.1, if the marginals are given, the structures achieving minimal and maximal correlation are the lower and the upper Fréchet bound, respectively, corresponding to the uniform representations $\max(u + v - 1, 0)$ and $\min(u, v)$.

On the other hand, for a given bivariate structure, one may also ask for the amount of dependence this structure allows to occur from choosing an appropriate pair of marginals. Kimeldorf and Sampson (1978) introduced a measure of dependence that is invariant under all order-preserving or order-reversing transformations of X and Y . Their *monotone correlation* ρ^* between two nondegenerate random variables X and Y is defined by

$$\rho^*(X, Y) = \sup \rho[f(X), g(Y)], \tag{10}$$

where the supremum is taken over all monotone functions f, g for which $0 < \text{Var}[f(X)] < \infty$ and $0 < \text{Var}[g(Y)] < \infty$ and with ρ the ordinary (Pearson product moment) correlation coefficient.

Obviously, $|\rho(X, Y)| \leq \rho^*(X, Y)$. It can be shown that $\rho^*(X, Y) = 0$ if and only if X and Y are stochastically independent, and that a perfect monotone relation between X and Y implies $\rho^*(X, Y) = 1$ (see Kimeldorf & Sampson, 1978). Moreover, it follows directly from Lemma 3.5 that

$$(X, Y) \sim (V, W) \quad \text{implies} \quad \rho^*(X, Y) = \rho^*(V, W).$$

Thus, for a given bivariate structure, ρ^* characterizes the amount of dependence. In particular, as the following lemma shows, the maximum achievable product moment correlation (from choosing an appropriate pair of marginals) is exactly the monotone correlation ρ^* .

LEMMA 3.7 (Kimeldorf and Sampson, 1978, Theorem 5). *Let (X, Y) have a continuous bivariate cdf H . Then*

$$\rho^*(X, Y) = \sup\{|\rho(V, W)| : (V, W) \sim (X, Y)\}.$$

The usefulness of ρ^* is limited, however, since there is no general procedure to compute it (see, however, Kimeldorf, May, & Sampson (1982) for the case of finite, discrete X and Y). Sometimes the possible range of dependence is severely restricted by the very structure of the bivariate distribution. For example, it follows from a result in Schucany, Parr, and Boyer (1978) (see also Huang & Kotz, 1984) that $|\rho(Y, Y)| \leq \alpha/3$ for the FGM distribution (Eq. 6) if X, Y have absolutely continuous marginals.

4. PROBABILITY SUMMATION MODELS OF RT

The aim of this section is to give a definition of probability summation of RT and to explore its properties. In general, we have to discriminate between two types of experimental conditions, the *single-channel condition* and the *multiple- or n -channel condition*, where $n \geq 2$. For example, the monocular viewing condition where only one eye is stimulated is a single-channel condition, while simultaneous visual and auditory stimulation in intersensory interaction experiments represents a 2-channel condition. In accordance with most of the RT literature (see, e.g., Luce, 1986), it is assumed here that observable RT can generally be decomposed into two additive random components, *processing time* and *base time*. In the single-channel-condition, processing time is the time from the presentation of the stimulus to the point where neural activity in the channel is first registered at some central mechanism. In the n -channel-condition, each channel is assigned some processing time starting from the presentation of the stimulus. Overall processing time then is assumed to be some nondecreasing function of these channel processing times. Probability summa-

tion models specify this function. The base time encompasses all other hypothetical components of the reaction time like motor programming and response execution. The results in the next two sections are limited to the $n=2$ case. In Section 6 specific problems for the $n > 2$ case are touched upon. Since there is no extra effort involved, however, the following definitions are written down for $n \geq 2$.

DEFINITION 4.1. Let (X_1, \dots, X_n) be a continuous random vector, where X_i ($i = 1, \dots, n$) represents the channel processing time of channel i in an n -channel condition; let Z be a continuous random variable with cdf F_Z representing the overall time; an n -variate probability summation model is a pair (H, F_Z) , where H is an n -variate cdf of (X_1, \dots, X_n) such that

$$Z = \min(X_1, \dots, X_n). \quad (11)$$

Thus, according to probability summation models, overall processing time is determined by the channel processing time of the channel finishing first. Assuming the existence of a common underlying probability space for Z and B , where B is the corresponding base time component, observable RT in the n -channel condition is

$$\begin{aligned} \text{RT} &= Z + B \\ &= \min(X_1, \dots, X_n) + B. \end{aligned} \quad (12)$$

Commonly, stochastic independence between Z and B is surmised. In most empirical situations, including the two paradigms dealt with in Section 2, channel processing times are unobservable in the multiple-channel condition. Even if the contribution of the base time component is assumed negligible, the experimenter registers only the overall processing time, which—according to the probability summation model—is the minimum of the channel processing times. This severely limits identifiability of a given probability summation model (H, F_Z) : the n -variate cdf H cannot be determined without further assumptions. A popular assumption exploits the information given from the single-channel conditions. This assumption of *context independence*³ holds that the processing times in single-channel conditions that correspond to a multiple-channel condition have cdf's identical to the respective marginal cdf's in the multiple-channel condition. In binocular summation, for example, this amounts to saying that left and right eye processing times in both monocular conditions have cdf's identical to the left and right eye marginals in the binocular condition.

Even under context independence, however, the n -variate cdf H is still not uniquely determined without assuming, in addition, stochastic independence among the channel processing times. In fact, this is the starting point for considering possibly dependent probability summation models in this paper. If observed RT in

³ This notion corresponds to Ashby and Townsend's (1986) 'temporal separability' concept. It would typically be associated with unlimited, but perhaps allocatable, capacity, as noted earlier.

the multiple-channel condition is found to be faster than predicted from a stochastically independent probability summation model, introducing negative dependence may be an interesting alternative. The following fictitious example illustrates the effect.

EXAMPLE 4.2 (LOWER FRÉCHET BOUND OF A BIVARIATE EXPONENTIAL DISTRIBUTION). Let X, Y be exponentially distributed random variables with parameter 1, i.e.,

$$P(X \leq t) = P(Y \leq t) = 1 - e^{-t}.$$

Under stochastic independence, it can be shown that

$$E[\min(X, Y)] = 0.5.$$

To achieve maximum negative correlation we now take the lower Fréchet bound $H^-(x, y)$ as the bivariate cdf of (X, Y) (see Lemma 3.1):

$$\begin{aligned} H^-(x, y) &= \max[(1 - e^{-x}) + (1 - e^{-y}) - 1, 0] \\ &= \max[1 - e^{-x} - e^{-y}, 0]. \end{aligned}$$

To compute $E[\min(X, Y)]$, consider first

$$\begin{aligned} E[\max(X, Y)] &= \int_0^\infty [1 - H^-(t, t)] dt \\ &= \int_0^\infty \min(2e^{-t}, 1) dt \\ &= \ln 2 + 1, \end{aligned}$$

after some algebra. Obviously,

$$\begin{aligned} E[\min(X, Y)] &= E[X] + E[Y] - E[\max(X, Y)] \\ &= 1 + 1 - (\ln 2 + 1) \\ &\approx 0.31. \end{aligned}$$

In this case, by the introduction of negative dependence the expected value of $\min(X, Y)$ has been reduced by about 40%. The proposition below will show in which sense this example generalizes to arbitrary bivariate distributions. The following definition is useful in this context.

DEFINITION 4.3. Let (H, F_Z) be a bivariate probability summation model;

assume H has marginals F_X, F_Y ; then the *uniform representation of H* (cf. Section 3),

$$U_H(u, v) = H[F_X^{-1}(u), F_Y^{-1}(v)], \quad 0 \leq u, v \leq 1,$$

is called the *type of (H, F_Z)* .

It follows from the exposition in Section 3 that the dependence structure of a probability summation model is captured in its type. More precisely, the type of (H, F_Z) remains invariant under (possibly different) increasing transformations of X and Y and has a value of ρ^* , the monotone correlation, that is independent of the marginals. Given the marginals, a probability summation model is completely determined by its type. Moreover, the class of all bivariate probability summation models can be partitioned into equivalence classes defined by models of the same type.

EXAMPLE 4.4. The model

$$U_H(u, v) = \max[u + v - 1, 0], \quad 0 \leq u, v \leq 1,$$

is the type of all probability summation models having the lower Fréchet bound H^- as bivariate cdf. Since $\rho(U, V) = -1$, we have $\rho^*(U, V) = 1$.

PROPOSITION 4.5. *For given marginals F_X, F_Y , the bivariate probability summation model of the lower Fréchet bound type achieves minimum expected overall processing time; specifically,*

$$E[\min(X, Y)] = \int_0^{t^*} [1 - F_X(t) - F_Y(t)] dt, \quad (13)$$

where

$$t^* = \inf\{t : 1 - F_X(t) - F_Y(t) < 0\}. \quad (14)$$

Proof. For any probability summation model with bivariate cdf $H(x, y)$ and marginals F_X, F_Y , we have

$$P[\min(X, Y) \leq t] = F_X(t) + F_Y(t) - H(t, t) \quad \text{for all } t. \quad (15)$$

From Lemma 3.1, for all t

$$H^-(t, t) \leq H(t, t);$$

thus, inserting H^- for H in Eq. (15) maximizes the right hand side of Eq. (15) over all types of bivariate probability summation models. Considering the well-known identity

$$E[\min(X, Y)] = \int_0^\infty \{1 - P[\min(X, Y) \leq t]\} dt \quad (16)$$

settles the minimality assertion of the proposition. Inserting (15) into Eq. (16) with $H = H^-$ yields

$$\begin{aligned} E[\min(X, Y)] &= \int_0^\infty \{1 - [F_X(t) + F_Y(t) \\ &\quad - \max\{F_X(t) + F_Y(t) - 1, 0\}]\} dt \\ &= \int_0^\infty \max\{1 - F_X(t) - F_Y(t), 0\} dt \\ &= \int_0^{t^*} [1 - F_X(t) - F_Y(t)] dt \end{aligned}$$

with t^* defined as above.

Q.E.D.

Assuming context independence and neglecting the effect of the base time component, the right hand side of Eq. 13 may serve as a lower bound for mean RT under possibly dependent probability summation in the 2-channel condition. It can be evaluated by either simulation or numerical integration (see next section). From (13), we immediately have the upper bound

$$E[\min(X, Y)] \leq t^* \tag{17}$$

which need not be sharp, however (consider Example 4.2, where we have $t^* = \ln 2 \approx 0.69$).

Proposition 4.5 gives a lower bound for the minimum achievable mean overall processing time under any probability summation model. Thus, in view of the discussion of intersensory and binocular interaction studies in Section 2, this bound can be used to assess the proportion of RT facilitation possibly due to a probability summation mechanism rather than some true neural interaction. On the other hand, if observed RT facilitation is less pronounced, a probability summation model of a moderate degree of negative dependence may be called for. In the following, we propose one way to order probability summation models by negative dependence and give some illustrative examples. The following definition of negative dependence was introduced by Lehmann (1966). A bivariate cdf H is *negatively quadrant dependent (NQD)* if

$$H(x, y) \leq F_X(x) F_Y(y) \quad \text{for all } (x, y) \in R^2, \tag{18}$$

where F_X, F_Y are the marginals of H . Using Hoeffding's Lemma (see Appendix) it is easy to show that NQD random variables have nonpositive product moment correlation coefficients. NQD distributions can be ordered as follows (Ebrahimi, 1982): bivariate cdf H_2 is *more negatively quadrant dependent than* H_1 ($H_2 \geq_{NQD} H_1$) if

$$H_2(x, y) \leq H_1(x, y) \quad \text{for all } (x, y) \in R^2, \tag{19}$$

where both H_1 and H_2 have marginals F_X and F_Y , i.e., $H_1, H_2 \in \theta(F_X, F_Y)$. The NQD ordering of distributions implies a corresponding ordering of correlations: let $\rho_1(X, Y)$, $\rho_2(X, Y)$ be the product moment correlation coefficients corresponding to H_1 and H_2 , respectively; then, again from Hoeffding's Lemma,

$$H_2 \geq_{NQD} H_1 \quad \text{implies} \quad \rho_2(X, Y) \leq \rho_1(X, Y) \leq 0. \quad (20)$$

The requirement of identical marginals is necessary here if the degree of negative dependence is to be assessed by rho, because rho is not invariant under monotone transformations of the scale. Note, moreover, that the above implication (20) is not reversible in general. The following definition seems natural in view of the above.

DEFINITION 4.6. Let $(H_1, F_Z^{(1)})$ and $(H_2, F_Z^{(2)})$ be probability summation models with $H_1, H_2 \in \theta(F_X, F_Y)$; then $(H_2, F_Z^{(2)})$ is more negative dependent than $(H_1, F_Z^{(1)})$ if $H_2 \geq_{NQD} H_1$.

This definition implies, in particular, that

$$\rho_2(X, Y) \leq \rho_1(X, Y).$$

Now consider two bivariate cdf's on the unit square with uniform marginals, U_H and U_K , representing two probability summation model types. If they can be ordered with respect to negative dependence, $U_H \geq_{NQD} U_K$, say, then we have an ordering of model types which is reflected in the corresponding rho values.

Comparing two probability summation models (with identical marginals) with respect to their degree of negative dependence is especially simple if their bivariate cdf's belong to the same one-parameter bivariate distribution family. The FGM distribution as well as Frank's distribution we considered in Section 3 belong to this category. For these distributions, the \geq_{NQD} ordering is isomorphic to the scalar ordering of the dependence parameter. In addition, if the product moment correlation coefficient is a monotone function of the dependence parameter, then implication (20) is in fact an equivalence. Let us consider a further example.

EXAMPLE 4.7: GUMBEL'S BIVARIATE EXPONENTIAL DISTRIBUTION. Gumbel (1960) has studied the bivariate cdf

$$H(x, y) = 1 - e^{-x} - e^{-y} + e^{-x-y-\alpha xy} \quad (x, y > 0; 0 \leq \alpha \leq 1).$$

It can be shown (see also Johnson & Kotz, 1972, p. 262) that $\rho(X, Y)$ is a decreasing function of α , reaching the minimum value $\rho(X, Y) = -0.40365\dots$ for $\alpha = 1$. X and Y are independent for $\alpha = 0$. Thus, the family of probability summation models corresponding to Gumbel's bivariate exponential distribution can be ordered with respect to negative dependence by using either the parameter values α or the \geq_{NQD} order relation on the bivariate cdf's.

We conclude this section by observing a simple but informative connection

between the Fréchet bounds and two well-known probability distribution inequalities for the overall processing time.

PROPOSITION 4.8. *Let (H^-, F_Z^-) and (H^+, F_Z^+) be bivariate probability summation models of the lower and, resp., upper Fréchet bound type with marginals F_X and F_Y ; then, for any bivariate probability summation model (H, F_Z) with the same marginals,*

$$F_Z^+(t) \leq F_Z(t) \leq F_Z^-(t) \quad (21)$$

for all t , where

$$F_Z^+(t) = \max[F_X(t), F_Y(t)]$$

and

$$F_Z^-(t) = \min[F_X(t) + F_Y(t), 1].$$

The left (resp. right) hand side of the above inequality follows easily by insertion of $H^+(t, t)$ (resp. $H^-(t, t)$) for $H(t, t)$ into Eq. (15). The upper bound in (21) was first used by Miller (1982) in testing the probability summation model. The lower bound in (21) was utilized by Grice, Canham, and Gwynne (1984) to evaluate distraction effects in redundant target trials (see also Ulrich & Giray, 1986). An analogous inequality is given in Ashby and Townsend (1986, Theorem 7) as a performance parity test of temporal separability (cf. Footnote 3).

5. A STUDY OF INTERSENSORY INTERACTION

The purpose of this section is to illustrate on a set of real data (a) the effect of dropping the stochastic independence assumption and (b) the computation of mean overall processing time according to Proposition 4.5. Moreover, eventual effects of the base time component will be discussed.

The data are taken from a study of intersensory interaction (see Example 2.1) presented in Diederich and Colonius (1987), where also the details of the experimental procedure can be found. Stimulus presentation in the redundant-signal condition consisted of a flash followed by a sinusoidal tone d msec later, with $d = 0, 10, \dots, 80$, the stimulus onset asynchrony (SOA) values. Subjects were instructed to press the response buttons with both index fingers as soon as they detected either signal (a double response simple RT task). In the single-signal conditions, only stimuli of one modality were presented. Each stimulus condition was presented in 10 blocks of 20 trials each in randomized order, amounting to 2200 responses from each of the subjects.

Let RT_V , RT_A , and $RT_{VA(d)}$ denote random variables corresponding to reaction times under the different stimulus conditions in obvious notation, where RT is

measured from the appearance of the (first) stimulus up to the button press. The amount of RT facilitation occurring in the redundant-signal conditions is defined by

$$\text{FAC}_d = \min[E(\text{RT}_V), E(\text{RT}_A + d)] - E(\text{RT}_{VA(d)}). \quad (22)$$

Ignoring effects of the base time component for a moment, the probability summation model stipulates

$$\text{RT}_{VA(d)} = \min(\text{RT}_V, \text{RT}_A + d). \quad (23)$$

The amount of RT facilitation is estimated from (22) by replacing the expected values by sample means ($N=200$). Observed mean RTs for the redundant-signal distributions are listed in column 2 of Table 1 below. Observed values of RT facilitation (in msec) are given in column 4. To estimate the amount of RT facilitation predicted from probability summation mechanisms, simulated RTs for the redundant-signal situations were obtained by sampling two RTs (one from each single-signal distribution), adjusting them appropriately for SOA, and selecting the minimum adjusted RT. Column 5 lists the RT facilitation obtained under independent sampling from the two single-signal distributions ($N=1000$ for all simulated values reported here). The values in parentheses in this column indicate the percentage of observed facilitation that can be accounted for by independent probability summation. Column 6 contains the RT facilitation that results from a probability summation model of the lower Fréchet bound type, where maximum negative correlation between the channel processing times is achieved. The mean RT under this model could have been estimated from a numerical computation of the integral in Proposition 4.5. Alternatively, we used a simulation by 'antithetic variates' (see, e.g., Hammersley and Handscomb, 1964; Miller, 1986). Essentially, the procedure is to construct pairs of RTs from the two single-signal distributions as follows: take the fastest (SOA adjusted) auditory RT and the slowest visual RT as the first pair, the next to the fastest auditory and the next to the slowest visual RT as the second pair, and so on. As a direct consequence of Lemma 3.1(b) and (c), this procedure does establish maximum negative correlation between the two channels (the rho values are the negative correlations so obtained for each subject). The parentheses in column 6 are again the percentage of observed facilitation that can be accounted for by a negatively dependent probability summation model.

Table 1 reveals a number of interesting regularities. The amount of RT facilitation observed strongly depends on the SOA value, achieving a maximum at about a value of d corresponding to the difference $E(\text{RT}_V) - E(\text{RT}_A)$. This is compatible with probability summation models since increasing the SOA results in shifting the auditory distribution towards the visual distribution. This in turn implies that, by virtue of the minimum rule, a larger proportion of the 'long' auditory processing times can be replaced by 'shorter' visual processing times. If d is increased further, the RT is mostly determined by the visual processing times alone. The exact SOA

TABLE 1
Observed Mean RT and Standard Deviation,
Observed and Simulated Facilitation in msec (Independence and Negative Dependence)

SOA	Obs. mean RT _{VA}	Obs. stand. dev.	Facilitation		
			Obs.	Indep.	Neg. Dep.
Subject 1 (mean RT _A = 126; mean RT _V = 160; N = 200; rho = -.865)					
0	127	8.3	-1	1	2
10	135	8.8	1	0	2 (200)
20	141	7.6	5	2 (40)	3 (60)
30	149	7.9	7	5 (71)	7 (100)
40	151	10.6	9	4 (44)	6 (67)
50	154	14.2	6	2 (33)	3 (50)
60	155	11.1	5	1 (20)	2 (40)
70	160	14.0	0	1	1
80	160	15.0	0	0	1
Subject 2 (mean RT _A = 124; mean RT _V = 162; N = 200; rho = -.880)					
0	120	7.5	4	2 (50)	2 (50)
10	131	6.9	3	1 (33)	2 (67)
20	139	6.9	5	2 (40)	3 (60)
30	145	8.1	9	4 (44)	6 (67)
40	150	8.7	12	7 (58)	9 (75)
50	152	11.6	10	4 (40)	5 (50)
60	155	11.9	7	2 (29)	3 (43)
70	161	13.5	1	0	2
80	163	15.0	-1	1	1
Subject 3 (mean RT _A = 124; mean RT _V = 153; N = 200; rho = -.642)					
0	122	10.8	2	2 (100)	4 (200)
10	131	11.7	3	3 (100)	6 (200)
20	135	11.5	9	5 (56)	9 (100)
30	142	10.1	11	9 (82)	12 (109)
40	143	9.6	10	5 (50)	8 (80)
50	147	12.8	6	4 (67)	6 (100)
60	151	11.7	2	2 (100)	4 (200)
70	154	16.5	-1	1	3)
80	153	17.1	0	2	2

Note. The parentheses in columns 5 and 6 denote the percentage of observed facilitation reached by the simulation. The rho values refer to the maximum possible negative correlation achieved by the method of antithetic variates.

value of maximum statistical facilitation, however, will depend on the specifics of the respective shapes of the auditory and visual distributions. On the other hand, it is obvious from the percentage values in column 5 that only part of the amount of observed facilitation can be accounted for by an independent probability summation model. Invoking a negatively dependent probability summation mechanism improves the prediction considerably, as can be seen from the parentheses in column 6. For Subjects 1 and 2, however, these values indicate that the proportion of observed facilitation that can be accounted for by a negatively dependent probability summation process reaches 100% only in two of the cases. For Subject 3, the facilitation predicted by maximum negative dependence exceeds the facilitation observed for five out of nine SOA values. Thus, a bivariate dependent distribution with a more moderate degree of dependence is called for in this case. One such distribution, the Farlie–Gumbel–Morgenstern distribution (cf. Example 3.3), was proposed earlier (Colonius, 1986) but, obviously, there are other viable candidates. We have not yet made an attempt to find the best fitting distribution for this subject (for Subjects 1 and 2, another source of RT facilitation has to be identified, anyhow).

Moreover, the pattern of rho values is somewhat puzzling. Remember that, for each subject, rho is the maximum negative correlation between the channel processing times the marginal distributions allow to occur. For Subjects 1 and 2, the rho values are $-.865$ and $-.880$, respectively, while for Subject 3, rho is only $-.642$. This is at odds with the fact that negative channel dependence suffices to generate the amount of facilitation observed for Subject 3 but not for Subjects 1 and 2. It is conceivable, however, that the possibly achievable degree of negative dependence varies for different SOA values due to a violation of the assumption of context independence.⁴ To summarize these data indicate that a probability summation mechanism—*independent* or *negatively dependent*—may be only part of a model accounting for the observed facilitation in this bimodal detection task. Nonetheless, as this exemplary study illustrates, just how big that part is depends very much on whether independent or dependent probability summation is taken into consideration. Whether, and under what conditions, the marginal distribution functions remain invariant under different SOA values and/or under different redundancy conditions is an open issue. Since the marginal distributions are not directly observable empirically (only their minima are, according to probability summation), there is no direct way of testing context independence. Any evidence relating to it would have to follow from particular features of the experimental situation or from the postulated processing mechanisms. Intuitively, a violation of context independence

⁴ The point is subtle. Given context independence, a variation of negative dependence over SOA values should show up in the correlation parameter of a bivariate distribution fitted to the data. However, this would not explain the failure of dependent probability summation to account for the entire amount of observed facilitation. On the other hand, if context independence is violated, the degree of negative correlation achievable might so vary from one SOA value to the next as to prevent statistical facilitation from occurring sufficiently often.

may appear more plausible in a system with stochastically dependent processing than in an independent one. It should be emphasized, however, that there is no cogent statistical argument for this to be the case. For example, assume in the visual/auditory interaction paradigm that the visual channel processing times correlate differently with the auditory channel processing times depending on the given SOA value. Nonetheless, the marginal visual processing time distribution may well remain the same under all SOA values.

Analyzing the RT difference distributions (left hand RT minus right hand RT) Diederich and Colonius (1987) found some evidence for an additional facilitatory effect in the base time (motor) components of the RTs. However, the size of this effect most likely is too small to account for all of the facilitation left. Moreover, the subsequent analysis in this section of the influence of the base time suggests that the amount of facilitation obtained in our simulations is not a conservative estimate of the actual facilitation occurring under a probability summation mechanism. To allow for an effect of the base time component let us assume an additive decomposition of RT as introduced in Section 4. Specifically, let us write

$$RT_V = V + B_V, \quad RT_A = A + B_A, \quad RT_{VA} = \min(V, A) + B_{VA},$$

where V, A denote the channel processing times and B_V, B_A, B_{VA} refer to the random base time components, resp. The following proposition implies that the amount of facilitation obtainable from a (possibly dependent) probability summation model with nonconstant base time may be smaller than what our simulations indicated.⁵

PROPOSITION 5.1. *Let V, A, B_V, B_A and B_{VA} be nonnegative random variables on some probability space such that B_V and B_A are stochastically independent and B_V, B_A, B_{VA} have the same marginal cdf; then for all $t \in R^+$*

$$P[\min(V + B_V, A + B_A) \leq t] \geq P[\min(V, A) + B_{VA} \leq t]. \quad (24)$$

Proof. For convenience, we assume the existence of a bivariate density $g_{VA}(v, a)$ for (V, A) ; it is easy to generalize the argument. Take some fixed (v, a) and consider

$$\begin{aligned} &P[\min(V + B_V, A + B_A) \leq t \mid (V, A) = (v, a)] \\ &= P[\min(v + B_V, a + B_A) \leq t] \\ &= 1 - [1 - F(t - v)][1 - F(t - a)] \\ &= F(t - v) + F(t - a) - F(t - v)F(t - a), \end{aligned} \quad (25)$$

⁵ I am grateful to Wolfgang Schwarz (Marburg) for suggesting this inequality.

where F is the common marginal cdf of B_V, B_A, B_{VA} . Similarly,

$$\begin{aligned}
 P[\min(V, A) + B_{VA} \leq t \mid (V, A) = (v, a)] \\
 &= P[\min(v, a) + B_{VA} \leq t] \\
 &= F(t - \min(v, a)) \\
 &= \begin{cases} F(t - a) & \text{if } v \leq a \\ F(t - v) & \text{if } a < v. \end{cases} \tag{26}
 \end{aligned}$$

Obviously, (25) is greater or equal to (26). Multiplying both sides by $g_{VA}(v, a)$ and integrating over $R^+ \times R^+$ yields the result. Q.E.D.

The left hand side of inequality (24) corresponds to the probability summation mechanism underlying the simulations, while the right hand side represents the probability summation mechanism of the model to be tested. Since stochastic ordering of distributions implies an ordering of expected values, we have

$$E[\min(V + B_V, A + B_A)] \leq E[\min(V, A) + B_{VA}].$$

This, in turn, implies that the amount of facilitation observed in the simulations may overestimate the facilitation actually occurring.

6. CONCLUDING REMARKS

It should be mentioned that our results generalize, at least to some degree, to the case of more than two channels. The main problem in dealing with this is to specify the multivariate dependence structure. The basic intuition that increasing negative dependence results in reducing the mean of the minimum time has to be reconciled with the elementary fact that three random variables cannot be negatively correlated in pairs to some arbitrary degree. There are various concepts of multivariate positive and negative dependence in the literature (see, e.g., Joag-Dev & Proschan, 1983) that become relevant in this context.

In the introductory section we discussed the effect of dropping stochastic independence in the context of a probability summation model in a detection task with n elementary detector mechanisms involved. As long as the number of false alarms is small enough to be ignored (a questionable assumption, of course) the definition of probability summation models in Section 4 formally encompasses the detection situation. To see this, define for each elementary detector i ($i = 1, \dots, n$) a binary random variable

$$D_i = \begin{cases} 1 & \text{if } i \text{ is in the detection state} \\ 0 & \text{if otherwise.} \end{cases}$$

Then $p_i = P[D_i = 1]$, while the probability P_{ij} of detecting the stimulus with only detectors i and j involved is

$$P_{ij} = P[D_i = 1 \vee D_j = 1].$$

In tight analogy to the RT situation facilitation is defined by the gain in detection probability that occurs from having two detectors 'at work' rather than one,

$$\text{FAC}_{\text{detection}} = P_{ij} - \max(p_i, p_j).$$

To model the 'race' we have to associate with each detector i a nonnegative random variable

$$D'_i = 1 - D_i.$$

Facilitation as defined for RT then is (see Section 5)

$$\begin{aligned} & \min[E(D'_i), E(D'_j)] - E[\min(D'_i, D'_j)] \\ &= \min[E(1 - D_i), E(1 - D_j)] - E[\min(1 - D_i, 1 - D_j)] \\ &= \min(1 - p_i, 1 - p_j) - P[D_i = 0 \wedge D_j = 0] \\ &= 1 - \max(p_i, p_j) - (1 - P_{ij}) \\ &= \text{FAC}_{\text{detection}}. \end{aligned}$$

Thus both definitions coincide.

The main finding of this paper, though, relates to the analysis of reaction time in the context of the redundant target paradigm. It has been shown that replacing the assumption of stochastic independence by negative dependence in the probability summation hypothesis allows a larger part of the redundancy gain to be attributed to statistical facilitation rather than to some neural coactivation. Consequently, whenever a realistic assessment of the role of statistical facilitation for a redundant target effect is to be made, the lower bound for the expected processing time given in Proposition 4.5 should be taken into consideration.

APPENDIX: OUTLINE OF PROOF OF LEMMA 3.1

(a) For any $(x, y) \in R^2$,

$$H(x, y) = P(X \leq x, Y \leq y) \leq P(X \leq x) = F_X(x).$$

Thus, by symmetry,

$$H(x, y) \leq \min[F_X(x), F_Y(y)],$$

which settles the right hand side of (5). Since

$$P(X > x, Y > y) = 1 - F_X(x) - F_Y(y) + H(x, y) \geq 0,$$

we have

$$H(x, y) \geq F_X(x) + F_Y(y) - 1$$

in addition to $H(x, y) \geq 0$, so

$$H(x, y) \geq \max\{F_X(x) + F_Y(y) - 1, 0\},$$

the left hand side of (5).

(b) Mardia (1970, p. 31) verifies sufficient conditions for H^+ and H^- to belong to $\theta(F_X, F_Y)$. The extreme correlation property is an immediate consequence of Hoeffding's Lemma (see Lehmann, 1966, p. 1139): for any random vector (X, Y) with cdf H and marginals F_X, F_Y , the covariance $\text{cov}(X, Y)$ can be represented as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [H(x, y) - F_X(x) F_Y(y)] dx dy.$$

(c) Since $F[F^{-1}(U)] \geq U$, we have

$$\begin{aligned} P[F_X^{-1}(U) \leq x, F_Y^{-1}(U) \leq y] &= P[U \leq F_X(x), U \leq F_Y(y)] \\ &= \min[F_X(x), F_Y(y)] \\ &= H^+(x, y); \end{aligned}$$

moreover,

$$\begin{aligned} P[F_X^{-1}(U) \leq x, F_Y^{-1}(1 - U) \leq y] &= P[U \leq F_X(x), 1 - U \leq F_Y(y)] \\ &= P[1 - F_Y(y) \leq U \leq F_X(x)] \\ &= \max[F_X(x) + F_Y(y) - 1, 0]. \end{aligned}$$

ACKNOWLEDGMENTS

This work was supported by Grants Co 94/I + II from Deutsche Forschungsgemeinschaft and by a grant from Stiftung Volkswagenwerk to the author. Helpful comments by two anonymous reviewers and, in particular, by the Editor, Jim Townsend, are gratefully acknowledged.

REFERENCES

ASHBY, F. G., & TOWNSEND, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, **93**, 154-179.

- BLAKE, R., & FOX, R. (1973). The psychophysical inquiry into binocular summation. *Perception and Psychophysics*, **14**, 161–185.
- BLAKE, R., MARTENS, W., & DI GIANFILLIP, A. (1980). Reaction time as a measure of binocular interaction in human vision. *Investigative Ophthalmology and Visual Science*, **19**, 930–941.
- BLAKE, R., MARTENS, W., GARRETT, A., & WESTENDORF, D. (1980). Estimating probability summation for binocular reaction time data. *Perception and Psychophysics*, **27**, 375–378.
- BLAKE, R., SLOANE, M., & FOX, R. (1981). Further developments in binocular summation. *Perception and Psychophysics*, **30**, 266–276.
- BRINDLEY, G. S. (1963). The relation of frequency of detection to intensity of stimulus for a system of many independent detectors each of which is stimulated by a m -quantum coincidence. *Journal of Physiology*, **169**, 412–415.
- CAMPBELL, F. W. & GREEN, D. G. (1965). Monocular versus binocular visual acuity. *Nature*, **208**, 191–192.
- COLONIUS, H. (1986). Measuring channel dependence in separate activation models. *Perception and Psychophysics*, **40**, 251–255.
- DIEDERICH, A. & COLONIUS, H. (1987). Intersensory facilitation in the motor component? A reaction time analysis. *Psychological Research*, **49**, 23–29.
- EBRAHIMI, N. (1982). The ordering of negative dependence. *Communications in Statistics-Theory and Methods*, **11**, 2389–2399.
- ERIKSEN, C. W. & GREENSPON, T. S. (1968). Binocular summation over time in the perception of form at brief durations. *Journal of Experimental Psychology*, **76**, 331–336.
- FRANK, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Mathematicae*, **19**, 194–226.
- FRÉCHET, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Annales de l'Université de Lyon, Section A, Série 3*, **14**, 53–77.
- GENEST, C. (1987). Frank's family of bivariate distributions. *Biometrika*, **74**, 549–555.
- 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 and Psychophysics*, **34**, 161–168.
- GRICE, G. R., CANHAM, L., & GWYNNE, J. W. (1984). Absence of a redundant-signals effect in a reaction time task with divided attention. *Perception and Psychophysics*, **36**, 565–570.
- GUMBEL, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, **59**, 698–707.
- HAMMERSLEY, J. M., & HANDSCOMB, D. C. (1964). *Monte Carlo methods*. London: Methuen.
- HARWERTH, R. S., SMITH, E. L., & LEVI, D. M. (1980). Supra-threshold binocular interactions for grating patterns. *Perception and Psychophysics*, **27**, 43–50.
- HERSHENSON, M. (1962). Reaction time as a measure of inter-sensory facilitation. *Journal of Experimental Psychology*, **63**, 289–293.
- HOEFFDING, W. (1940). Massstabinvariante Korrelationstheorie. *Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin*, **5**, 179–233.
- HUANG, J. S., & KOTZ, S. (1984). Correlation structure in iterated Farlie-Gumbel-Moregenstern distributions. *Biometrika*, **71**, 633–636.
- JOAG-DEV, K. & PROSCHAN, F. (1983). Negative association of random variables, with applications. *Annals of Statistics*, **11**, 286–295.
- JOHNSON, N. L., & KOTZ, S. (1972). *Distributions in Statistics: Continuous multivariate distributions*. New York: Wiley.
- KIMELDORF, G., MAY, J. H., & SAMPSON, A. R. (1982). Concordant and discordant monotone correlations and their evaluation by nonlinear optimization. In S. H. Zanakis & J. S. Rustagi (Eds.), *TIMS/Studies in the Management Sciences*, **19**, 117–130. Amsterdam: North-Holland.
- KIMELDORF, G., & SAMPSON, A. (1975). Uniform representations of bivariate distributions. *Communications in Statistics*, **4**, 617–627.
- KIMELDORF, G., & SAMPSON, A. (1978). Monotone dependence. *Annals of Statistics*, **6**, 895–903.
- LEGGE, G. E., & RUBIN, G. S. (1981). Binocular interaction in suprathreshold contrast interaction. *Perception and Psychophysics*, **30**, 49–61.

- LEHMANN, E. L. (1966). Some concepts of dependence. *Annals of Mathematical Statistics*, **39**, 1137-1153.
- LUCE, R. D. (1986). *Response times. Their role in inferring elementary mental organization*. New York: Oxford Univ. Press.
- MARDIA, K. V. (1970). *Families of bivariate distributions*. London: Griffin.
- MILLER, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, **14**, 247-279.
- MILLER, J. O. (1986). Timecourse of coactivation in bimodal divided attention. *Perception and Psychophysics*, **40**, 331-343.
- PIRENNE, M. H. (1943). Binocular and unocular thresholds for vision. *Nature*, **153**, 698-699.
- RAAB, D. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, **24**, 574-590.
- SCHUCANY, W. R., PARR, W. C., & BOYER, J. E. (1978). Correlation structure in Farlie-Gumbel-Morgenstern distributions. *Biometrika*, **65**, 650-653.
- SCHWEIZER, B., & WOLFF, E. F. (1981). On nonparametric measures of dependence for random variables. *Annals of Statistics*, **9**, 879-885.
- TOWNSEND, J. T., & ASHBY, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge Univ. Press.
- ULRICH, R. & GIRAY, M. (1986). Separate-activation models with variable base times: Testability and checking of cross-channel dependency. *Perception & Psychophysics*, **39**, 248-254.
- WESTENDORF, D., & BLAKE, R. (1988). Binocular reaction times to contrast increments. *Vision Research*, **28**, 355-359.
- WHITT, W. (1976). Bivariate distributions with given marginals. *Annals of Statistics*, **4**, 1280-1289.

RECEIVED: August 8, 1988