

Measuring channel dependence in separate activation models

HANS COLONIUS

Universität Oldenburg, Oldenburg, Federal Republic of Germany

In the bimodal detection task, the observer must respond as soon as a signal is presented in either of two modalities (e.g., a tone or a flash). A typical finding is a facilitation of reaction time for redundant-signal trials, that is, when both signals are presented simultaneously. Separate-activation models assume that the two signals are processed simultaneously within different channels and that a response is initiated as soon as an activation level is exceeded in either channel. This paper examines a number of approaches to formulating a model of the stochastic mechanisms involved. Conditions allowing inference of the correlational structure of the channel processes are discussed, and new models are proposed. Moreover, distribution inequalities for testing of the models under weakened independence assumptions are derived.

This paper addresses a problem in the reaction time (RT) analysis of a bimodal detection task. In this experimental situation, the observer must respond as soon as a signal is presented in either of two modalities (e.g., a tone or a flash). On single-signal trials, only one signal is presented; on redundant-signal trials, both signals are presented simultaneously or with a short interstimulus interval (ISI) delay. The typical finding is that RT is shorter for redundant-signal trials. One type of model that has been advanced for the redundant-signals effect is the separate-activation model (cf. Miller, 1982). It is assumed that the two signals are processed simultaneously within different channels and that each channel produces a separate activation. The response is initiated as soon as an activation level is exceeded in either channel. Responses to redundant signals are especially fast because they are produced by the faster of two processes with randomly varying durations. Raab (1962) summarized these "race models" by introducing the term "statistical facilitation."

Recently, a number of papers concerned with the stochastic modeling of the separate-activation mechanism have been published. Specifically, it has been asked whether the detection times of the two channels are independent random variables, and, if independence is rejected, whether they are positively or negatively correlated (Grice, Canham, & Borroughs, 1984; Ulrich & Giray, 1986; van der Heijden, Schreuder, Maris, & Neerinx, 1984). In this paper, we critically review some of these results and propose a number of alternative approaches.

Let RT_x (RT_y) denote the reaction time to stimulus S_x (S_y) in the single-stimulus condition, and RT_{xy} the reaction time to the redundant signal, that is, when both sig-

nals are presented. The corresponding cumulative distribution functions are

$$P(RT_x \leq t) = F_x(t),$$

$$P(RT_y \leq t) = F_y(t),$$

$$P(RT_{xy} \leq t) = F_{xy}(t).$$

In testing the separate activation model, it is commonly assumed that the marginal distributions in the redundant-signal condition can be equated with the RT distributions in the single-signal conditions, implying

$$RT_{xy} = \min(RT_x, RT_y).$$

The plausibility of this empirically untestable assumption rests on the details of the experimental setup and will be taken for granted here. From elementary probability, we have for any t

$$F_{xy}(t) = F_x(t) + F_y(t) - P[\max(RT_x, RT_y) \leq t]. \quad (1)$$

Omitting the unobservable negative term on the right-hand side of Equation 1, Miller (1982) obtained

$$F_{xy}(t) \leq F_x(t) + F_y(t). \quad (2)$$

Violations of this inequality allow rejection of the entire class of separate activation models. It has been observed, however, that the upper bound in Equation 2 is not very sharp. Let us rewrite Equation 1 in the following way:

$$\begin{aligned} F_{xy}(t) &= F_x(t) + F_y(t) - F_x(t)F_y(t) \\ &\quad - \{P[\max(RT_x, RT_y) \leq t] - F_x(t)F_y(t)\} \\ &= F_x(t) + F_y(t) - F_x(t)F_y(t) - D(t), \end{aligned} \quad (3)$$

where $D(t)$ stands for the expression in braces.

If the processing times are assumed to be independent, we have $D(t) = 0$ and Equation 3 can, in principle, be used to check this independent channels model. Furthermore, a number of authors (e.g., Grice et al., 1984; van

I am grateful to Adele Diederich for helpful comments and to Stan Gielen for providing part of his data. The author's address is: Institut für Kognitionsforschung, Universität Oldenburg, P.O. Box 2503, D2900 Oldenburg, Federal Republic of Germany.

der Heijden et al., 1984) have taken positive (negative) values of $D(t)$ as evidence for positive (negative) correlation between the processing times. Unfortunately, this last line of reasoning is not valid in general. Only if certain specific distributional assumptions not directly testable from the available data are introduced can the direction and/or size of the correlation be inferred. The problem resides in the fact that in order to assess the dependence of two random variables, knowledge of their entire bivariate distribution is needed. In the empirical context of separate-activation models, however, only the marginals and the distribution of the minimum of the processing times are available at best.

The above is nicely illustrated by the following representation of the covariance of two arbitrary random variables due to Hoeffding (see Lehmann, 1966):

$$\text{cov}(RT_x, RT_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(s,t) ds dt, \tag{4}$$

where

$$D(s,t) = P(RT_x \leq s, RT_y \leq t) - P(RT_x \leq s)P(RT_y \leq t).$$

Obviously, $D(s,t) \geq 0$ for all s,t implies nonnegative covariance. However, positive covariance does not ensure that $D(s,t)$ will be positive for all s,t . In particular, positivity of $D(t,t) = D(t)$, the term that occurs in Equation 3, is neither necessary nor sufficient for positive covariance (or correlation).

To illustrate this argument, it may be useful to distinguish a notion of "global" dependence from one of "local" dependence. If the function $D(s,t)$ is positive at some point (s,t) , we have *local* positive dependence at that point. On the other hand, covariance (or correlation) is an overall *global* measure of dependence, since it is obtained by integrating over all points (s,t) , as demonstrated by Equation 4. The point then is that local positive dependence everywhere, that is, at all points (s,t) , implies global positive dependence but not vice versa. The following example shows, more specifically, that global positive dependence does not even imply local positive dependence in terms of $D(t)$ for some t .

Example:

Let random variables X and Y take on the values 1, 2, 3 with probabilities given in Table 1. Thus, we have $\text{cov}(X,Y) = 0.44$, while $D(1) = -0.01$, $D(2) = -0.04$, and $D(3) = 0.0$.

BOUNDS FOR $\text{corr}(RT_x, RT_y)$

As pointed out above, assessing the dependence of two random variables without introducing any specific assumptions requires knowledge of their entire bivariate distribution function. It is interesting to note, however, that nonetheless some bounds on $\text{corr}(RT_x, RT_y)$ can be given (cf. Whitt, 1976):

Table 1
An Example of Global Positive But Local Nonpositive Dependence

	3	0.1	0.1	0.6
X	2	0.0	0.0	0.1
	1	0.0	0.0	0.1
		1	2	3
			Y	

Lemma:

Let random variables RT_x, RT_y have a bivariate distribution,

$$G_{xy}(s,t) = P(RT_x \leq s, RT_y \leq t), \tag{5}$$

with marginal distributions $F_x(s)$ and $F_y(t)$; then the correlation between RT_x and RT_y , $\text{corr}(RT_x, RT_y)$, is bounded, that is,

$$\text{corr}^- \leq \text{corr}(RT_x, RT_y) \leq \text{corr}^+;$$

here, corr^- (corr^+) is the smallest (largest) possible correlation for two random variables having a bivariate distribution with marginals identical to F_x and F_y . (An outline of the proof is contained in Appendix A.)

Estimation of corr^- and corr^+ is demonstrated below. The usefulness of these bounds is, of course, not guaranteed a priori. Consider, for example, the case $\text{corr}^- = -1$ and $\text{corr}^+ = 1$. Nonetheless, if one of the bounds is equal to (or near) zero, for instance, the sign of $\text{corr}(RT_x, RT_y)$ is determined. The numerical estimates of the bounds depend on the empirical distributions \hat{F}_x and \hat{F}_y from the single-signal situations. (Details are provided in Appendix A.)

For illustrative purposes, we use part of the data of two experiments reported in Gielen, Schmidt, and van den Heuvel (1983). The data of Experiment 1 consist of three blocks of 25 responses of 1 subject to a visual, an auditory, and a combined signal, respectively. In the redundant-signal situation, the auditory stimulus was presented 46.5 msec after the visual stimulus. The data of Experiment 2 consist of three blocks of 25 responses of 1 subject to a visual, a kinesthetic (a torque to the responding limb), and a combined signal, respectively. In the redundant-signal situation, the torque stimulus was presented 100 msec after the visual stimulus. The estimates for the bounds of the correlation between the processing times in the redundant signal situation are as given in Table 2. Admittedly, these bounds are of limited value in assessing the actual degree of dependence between the processing times. The width of the interval may in part be due to an insufficient approximation of the Hoeffding integrals for computing the covariances with too small a number of data points available. In any case, they may serve as a baseline in the following sense: any separate activation model for the redundant-signal situation consistent with the observed marginal distributions

Table 2
Maximal and Minimal Correlations for Fixed Marginals
in Gielen et al.'s Experiments

	corr ⁻	corr ⁺
Experiment 1	-.799	.955
Experiment 2	-.860	.830

predicts a correlation lying in the range of corr⁻ to corr⁺. Some more specific models will be discussed in the next section. The reader should note that in estimating these bounds, no data from the redundant-signal trials are being used. If they are, the lower bound can presumably be sharpened. However, this, as well as diagnostic properties on the marginals implying closer bounds for the correlation, deserves further study.

MODELING DEPENDENCE

It is obvious from the last section that some assumptions on the bivariate distribution are needed to derive more specific information on the degree of dependence between the processing times.

In an attempt to measure the strength of channel dependence, Grice et al. (1984) proposed computation of the tetrachoric correlation coefficient. The proportions of responding and not responding to the single signals provide estimates of the marginals of a fourfold table; the proportion of not responding in the redundant-signal situation fills one cell of the table. Since the table has only one degree of freedom, the entire table is thus determined. The tetrachoric correlation coefficient can be estimated at any quantile of the redundant-signal distribution. However, the validity of the whole approach rests on the assumption of bivariate normality. If the computations yield inconsistent estimates (as found, e.g., by van der Heijden et al., 1984), it remains unclear whether this is evidence against the model of separate activation or merely against the assumption of bivariate normality.

Let us outline another approach of modeling dependence, one which may be called "semiparametric." The idea here is to construct a bivariate distribution from specified marginals. An example is given by Huang and Kotz (1984) in their discussion of the bivariate Farlie-Gumbel-Morgenstern (FGM) system of distributions with bivariate cumulative distribution

$$G_{xy}(s,t) = F_x(s)F_y(t)[1 + \alpha\bar{F}_x(s)\bar{F}_y(t)], \tag{6}$$

where $\bar{F} = 1 - F$, α a real number. There are other systems of distributions available and some remarks on choosing an appropriate form are made below. This semiparametric approach is attractive for a number of reasons. Note that for these models the bivariate distribution is characterized by the univariate marginal distributions. Thus, the assumptions these models make only amount to specifying the way the bivariate distribution is constructed from the marginals. It is not necessary to postulate a special distributional form of the marginals in order to test the model. On the other hand, if for some reason

the experimenter prefers to assume a parametric form for the marginals (e.g., normal distributions), he or she may still stay within the semiparametric approach by just putting these marginals into the model equation. Moreover, models like the FGM system (Equation 6) have a simple analytical form; their dependence structure is entirely captured by just one parameter (for several empirical applications, see, e.g., Conway, 1983).

Before the semiparametric approach is illustrated, some comments on testing of the separate activation model in general are in order. The first step in testing the model is to use Equation 2 (or a sharpened version of Equation 2 to be presented later). If no rejection is possible, some more specific model could be proposed and tested. However, as pointed out earlier, apart from the marginal distributions, only the minimum processing times are observable and, hence, amenable to testing. Consequently, evidence in favor of any specific model is somewhat limited. In our view, this supports the adequacy of an approach such as the semiparametric one with rather large model subclasses.

Apart from the FGM system of distributions, essentially two other one-parameter families that possess various different properties are known (the interested reader should consult Johnson & Kotz, 1972). For the present, we adopt the FGM system here for reasons of simplicity and because it contains the case of independent processing. It is known, however, that this model allows for only rather weak degrees of dependence. If stronger dependence is suspected, then either such models as the iterated FGM systems, which introduce one or more additional parameters, or the other one-parameter systems are possible alternatives (cf. Huang & Kotz, 1984).

For the dependence parameter α in the FGM model, there is a simple approximation to its maximum likelihood estimate that depends on the marginals only (see Appendix B for details). Moreover, the Pearson product-moment correlation coefficient is given by

$$\text{corr}(RT_x, RT_y) = \frac{\alpha * \left\{ \int_{-\infty}^{\infty} [1 - F_x(s)] ds \right\} * \left\{ \int_{-\infty}^{\infty} [1 - F_y(t)] dt \right\}}{[SD(X)SD(Y)]} \tag{7}$$

In particular, if both marginals are normal, we have

$$\text{corr}(RT_x, RT_y) = \alpha/\pi. \tag{8}$$

For the data from Gielen et al.'s (1983) Experiment 2, computation of the correlation via Equation 7 yields .09. However, this estimate should be taken with a grain of salt, since it is based on only three blocks of 25 data points. If the marginals are assumed to be normal, the corresponding estimate via Equation 8 is .012. For van der Heijden et al.'s (1984, Table 1, p. 581) data, the correlation estimate via Equation 8 is .481.

A test of the FGM model could be carried through by a chi-square test on the data observed in the redundant-signal situation. If no parametric assumption on the marginals is made, the values of the observed marginals go-

ing into the prediction will have to be counted as parameters that reduce the number of degrees of freedom in the chi-square. Obviously, this suggests a split-half technique; that is, half of the data should be used to estimate the marginals and the dependence parameter, while the other half could be taken for the comparison with the predictions. If a parametric assumption on the marginals is made, only the number of estimated parameters of these functions would have to be taken into account for computing the number of degrees of freedom for the test. In any case, our present data base is too small to allow a reasonable check of the model.

AN INEQUALITY TESTING FOR POSITIVE (NEGATIVE) DEPENDENCE

Let us return to the issue of testing the separate activation model before subscribing to a specific model. In the proposition given below, we state a simple inequality strictly sharper than the one given by Miller (1982) (cf. Equation 2), allowing for the rejection of a number of models. Two random variables, such that $D(s,t) \geq 0$ (resp. ≤ 0) for all s,t , are called *positively (resp. negatively) quadrant dependent*, PQD (NQD) for short (cf. Lehmann, 1966). The class of PQD (NQD) random variables is a large one, but it is entirely contained in the class of all positively (negatively) correlated random variables; that is, PQD (NQD) implies positive (negative) correlation. Two random variables are independent if and only if they are both PQD and NQD.

Proposition 1:

If RT_x, RT_y are PQD random variables, then

$$F_{xy}(t) \leq F_x(t) + F_y(t) - F_x(t)F_y(t). \quad (9)$$

Moreover, for NQD random variables, the above inequality is reversed.

This proposition is an immediate consequence of Equality 3 above. The inequality is put to use as follows. Suppose, for a given set of data some specific model, such as the bivariate normal or the FGM, is to be tested. Then Inequality 9 should be checked first. The reason is that these specific models (and a number of others) imply either Inequality 9 or its reverse, depending on whether the processing times are positively or negatively correlated. For example, van der Heijden et al. (1984) actually tested Inequality 9 (and its reverse), and found that it was clearly violated (it seems, however, they were not aware of the consequence, namely rejection of the bivariate normal). The role this proposition plays is, loosely speaking, somewhat destructive: Violation of the inequality (or its reverse) rules out a number of models for separate activation; nonviolation, however, does not allow any inference on the direction of the correlation, as we demonstrated in the first section of this paper.

Finally, we turn to a generalization of the above proposition. The observable reaction time is typically conceived of as the sum of the detection time and a common variable base time, B , summarizing all stages that follow stimulus detection (e.g., motor components, cf. Ulrich & Giray, 1986). We thus have

$$RT_x = X+B, RT_y = Y+B, RT_{xy} = \min(X,Y)+B.$$

The following proposition amounts to saying that the above test extends to the case of an additive decomposition. It generalizes a corollary in Ulrich and Giray (1986, p. 249) by assuming X and Y to be PQD instead of independent.

Proposition 2:

Suppose X and Y are independent of B ; if X and Y are PQD random variables, then so are $RT_x = X+B$ and $RT_y = Y+B$ and the bound in Proposition 1 applies.

For a proof, the reader is referred to Tong (1980, Theorem 5.1.4).

REFERENCES

- CONWAY, D. (1983). Farlie-Gumbel-Morgenstern distributions. *Encyclopedia of Statistical Sciences*, 3, 28-31.
- GIELEN, S. C. A. M., SCHMIDT, R. A., & VAN DEN HEUVEL, P. J. M. (1983). On the nature of intersensory facilitation of reaction time. *Perception & Psychophysics*, 34, 161-168.
- GRICE, G. R., CANHAM, L., & BOROUGHS, J. M. (1984). Combination rule for redundant information in reaction time tasks with divided attention. *Perception & Psychophysics*, 35, 451-463.
- HUANG, J. S., & KOTZ, S. (1984). Correlation structure in iterated Farlie-Gumbel-Morgenstern distributions. *Biometrika*, 71, 633-636.
- JOHNSON, N. L., & KOTZ, S. (1972). *Distributions in statistics: Continuous multivariate distributions*. New York: Wiley.
- LEHMAN, E. L. (1966). Some concepts of dependence. *Annals of Mathematical Statistics*, 39, 1137-1153.
- MILLER, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247-279.
- RAAB, D. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, 24, 574-590.
- TONG, Y. L. (1980). *Probability inequalities in multivariate distributions*. New York: Academic Press.
- ULRICH, R., & GIRAY, M. (1986). Separate-activation models with variable base-times: Testability and checking cross-channel dependency. *Perception & Psychophysics*, 39, 248-254.
- VAN DER HEIJDEN, A. H. C., SCHREUDER, R., MARIS, L., & NEERINCKX, M. (1984). Some evidence for correlated separate activation in a simple letter-detection task. *Perception & Psychophysics*, 36, 577-585.
- WHITT, W. (1976). Bivariate distributions with given marginals. *The Annals of Statistics*, 4, 1280-1289.

APPENDIX A

Outline of Proof for the Lemma

For any bivariate distribution, G , with marginals F_x and F_y , the following bounds hold:

$$G_{xy}^- = \max(F_x + F_y - 1, 0) \leq G_{xy} \leq \min(F_x, F_y) = G_{xy}^+. \quad (A1)$$

G_{xy}^- and G_{xy}^+ constitute bivariate distributions with marginals identical to those of G_{xy} . Putting

$$D^+(s,t) = G_{xy}^+(s,t) - F_x(s)F_y(t)$$

$$D^-(s,t) = G_{xy}^-(s,t) - F_x(s)F_y(t)$$

and inserting D^+ and D^- into Hoeffding's representation of the covariance (Equation 4) yields immediately the covariances cov^+ and cov^- of the random variables corresponding to G_{xy}^+ and G_{xy}^- , respectively. Since G_{xy} , G_{xy}^+ , and G_{xy}^- have the same marginals, their corresponding standard deviations are identical, too. Thus,

$$\text{corr}^+ = \text{cov}^+ / [SD(X) * SD(Y)]$$

and

$$\text{corr}^- = \text{cov}^- / [SD(X) * SD(Y)].$$

Applying Inequality A1 then implies the lemma.

Estimating the Bounds corr^+ and corr^-

First, the empirical (cumulative) distribution functions \hat{F}_x and \hat{F}_y are used to approximate G_{xy}^+ and G_{xy}^- :

$$\hat{G}_{xy}^+ = \max(\hat{F}_x + \hat{F}_y - 1, 0) \text{ and } \hat{G}_{xy}^- = \min(\hat{F}_x, \hat{F}_y).$$

This yields corresponding estimates

$$\hat{D}^+ = \hat{G}_{xy}^+ - \hat{F}_x \hat{F}_y \text{ and } \hat{D}^- = \hat{G}_{xy}^- - \hat{F}_x \hat{F}_y.$$

Then the Hoeffding integrals for cov^+ and cov^- are approximated by the sums generated by replacing D^+ and D^- by \hat{D}^+ and \hat{D}^- , respectively. For Experiment 1 of Gielen et al. (1983), these estimates are about

$$\text{cov}^+ = 190 \text{ and } \text{cov}^- = -159.$$

Dividing by the observed standard deviations, $SD(X) = 17.6$ and $SD(Y) = 11.3$ (with X denoting visual, Y auditory stimuli), then delivers the estimates of corr^+ and corr^- reported in the first

row of Table 2. For Experiment 2, the corresponding estimates are

$$\hat{\text{cov}}^+ = 186.4 \text{ and } \hat{\text{cov}}^- = -193.4$$

$$SD(X) = 23 \text{ and } SD(Y) = 9.76$$

(with X denoting visual, Y torque stimuli), leading to the estimates reported in the second row of Table 2.

APPENDIX B

The approximate maximum likelihood estimate for α is known to be (cf. Conway, 1983)

$$\hat{\alpha} \approx \sum_{i=1}^n [1 - 2 * \hat{F}_x(s_i)] * [1 - 2 * \hat{F}_y(t_i)] / N, \tag{B1}$$

where

$$N = \sum_{i=1}^n [1 - 2 * \hat{F}_x(s_i)]^2 * [1 - 2 * \hat{F}_y(t_i)]^2,$$

given the observations (s_i, t_i) $i = 1, \dots, n$. We computed $\hat{\alpha}$ by taking the observations (t_i) in the redundant signal situation and inserting the corresponding values of the empirical distribution functions $\hat{F}_x(t_i)$ $\hat{F}_y(t_i)$ into Approximation B1. For Gielen et al.'s (1983) Experiment 2 data, this yields $\hat{\alpha} = 0.039$. To evaluate Equation 7, the integrals have to be approximated by the corresponding sums over the empirical survivor functions $(1 - F_x)$ and $(1 - F_y)$, yielding 17.12, and 31.64, respectively. The standard deviations are $SD(X) = 9.76$ and $SD(Y) = 23.00$.

(Manuscript received March 17, 1986;
revision accepted for publication August 1, 1986.)