Journal of Mathematical Psychology **45**, 421–451 (2001)
doi:10.1006/jmps.2000.1330, available online at http://www.idealibrary.com on IDEAL®

Unconditionally Selective Dependence of Random Variables on External Factors

Ehtibar N. Dzhafarov

Purdue University

What is the meaning of saying that random variables $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ (such as aptitude scores or hypothetical response time components), not necessarily stochastically independent, are selectively influenced respectively by subsets $\{\Gamma_1, ..., \Gamma_n\}$ of a factor set Φ upon which the joint distribution of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ is known to depend? One possible meaning of this statement, termed conditionally selective influence, is completely characterized in Dzhafarov (1999, *Journal of Mathematical Psychology*, **43**, 123–157). The present paper focuses on another meaning, termed unconditionally selective influence. It occurs when two requirements are met. First, for $i = 1, ..., n$, the factor subset Γ_i is the set of all factors that effectively change the marginal distribution of $\mathbf{X}_i$. Second, if $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ are transformed so that all marginal distributions become the same (e.g., standard uniform or standard normal), the transformed variables are representable as well-behaved functions of the corresponding factor subsets $\{\Gamma_1, ..., \Gamma_n\}$ and of some common set of sources of randomness whose distribution does not depend on any factors. Under the constraint that the factor subsets $\{\Gamma_1, ..., \Gamma_n\}$ are disjoint, this paper establishes the necessary and sufficient structure of the joint distribution of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ under which they are unconditionally selectively influenced by $\{\Gamma_1, ..., \Gamma_n\}$. The unconditionally selective influence has two desirable properties, uniqueness and nestedness: $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ cannot be influenced selectively by more than one partition $\{\Gamma_1, ..., \Gamma_n\}$ of the factor set Φ, and the components of any subvector of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ are selectively influenced by the components of the corresponding subpartition of $\{\Gamma_1, ..., \Gamma_n\}$.

The author is grateful to R. Thomas for helpful suggestions.

Address correspondence and reprint requests to Ehtibar N. Dzhafarov, Department of Psychological Sciences, Purdue University, 1364 Psychological Sciences Building, West Lafayette, IN 47907-1364.

0022-2496/01 $35.00

1. INTRODUCTION

1.1. Two Meanings of Selective Influence

Consider the following seemingly very simple problem. Let $\{\mathbf{X}_1, \mathbf{X}_2\}$ be two observable random variables, say, performance scores in two tests, known to have a bivariate normal distribution. Let γ_1 and γ_2 be two external factors (say, age and sex of the examinee), and the hypothesis be that γ_1 selectively influences $\mathbf{X}_1$, while γ_2 selectively influences $\mathbf{X}_2$. What is the meaning of this statement? If $\mathbf{X}_1$ and $\mathbf{X}_2$ are stochastically independent, the answer is trivial: the mean and the variance, μ_1, σ_1^2, of $\mathbf{X}_1$ change with γ_1 but not with γ_2, while μ_2 and σ_2^2 of $\mathbf{X}_2$ change with γ_2 but not with γ_1. If, however, $\mathbf{X}_1$ and $\mathbf{X}_2$ are known to be stochastically interdependent (which in this case means that the covariance between $\mathbf{X}_1$ and $\mathbf{X}_2$ is not always zero), the meaning of the selectiveness becomes less apparent. Should one say that the selectiveness in question takes place only when the correlation ρ does not depend on either γ_1 or γ_2 (but what should one make of the fact that the covariance then depends on both these factors)? Is it logically possible to speak of any form of selective influence if μ_1 and σ_1^2 are functions of γ_1 only, μ_2 and σ_2^2 are functions of γ_2 only, while ρ is a function of both γ_1 and γ_2?

Consider another example, leading to the same mathematical problem (bivariate-normal distribution that depends on two factors) but perhaps more appealing to those interested in mental architectures and response time decompositions (Dzhafarov, 1997). Let a simple response time $\mathbf{T}$, known to depend on two experimental factors γ_1 and γ_2 (say, stimulus intensity and response deadline), be assumed to be decomposable into a sum of two log-normally distributed components,[1]

$$\mathbf{T}(\gamma_1, \gamma_2) = \exp[\mathbf{X}_1(\gamma_1)] + \exp[\mathbf{X}_2(\gamma_2)],$$

where $\{\mathbf{X}_1, \mathbf{X}_2\}$ are bivariate-normally distributed and assumed to selectively depend on γ_1 and γ_2, respectively. Again, the meaning of this statement is clear when $\mathbf{X}_1$ and $\mathbf{X}_2$ are stochastically independent. It may be reasonable to assume, however, that $\mathbf{X}_1$ and $\mathbf{X}_2$ vary as a function of some state or process (arousal level, attention) that tends to prolong or shorten both of these variables simultaneously, introducing thereby a positive correlation between them. What then becomes of the hypothesis that $\{\mathbf{X}_1, \mathbf{X}_2\}$ are selectively influenced by γ_1 and γ_2, respectively?

Generalizing, what is the meaning in which one can say that stochastically interdependent random variables $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ (such as hypothetical response time components or observable aptitude scores) are selectively influenced by, respectively, subsets $\{\Gamma_1, ..., \Gamma_n\}$ of a factor set Φ upon which the joint distribution of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ is known to depend? Dzhafarov (1997, 1999) distinguishes two different meanings in which such selectiveness of influence can be understood.

[1] Response time components in this case can be thought of as durations of two consecutive processes, but this is only one possible interpretation. For a detailed discussion of response time components see Dzhafarov (1997) and Dzhafarov and Schweickert (1995).

One meaning, termed in Dzhafarov (1999) *conditionally selective influence*, is derived from the idea of indirect nonselectivity proposed in Townsend (1984) and studied in Townsend and Thomas (1994). Random variables $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ are conditionally selectively influenced by, respectively, $\{\Gamma_1, ..., \Gamma_n\}$ if, for any i, the conditional distribution of $\mathbf{X}_i$, given any fixed values for the rest of the variables, depends only on the factors comprising Γ_i. Dzhafarov (1999) provides a complete characterization (the necessary and sufficient structure of the joint density function) for this form of selective influence. When applied to the problem posed earlier, this characterization yields the following result. Bivariate normally distributed $\{\mathbf{X}_1, \mathbf{X}_2\}$ are conditionally selectively influenced by factors γ_1, γ_2, respectively, if and only if

(i) the correlation ρ can be presented as

$$\rho = kc_1(\gamma_1)\, c_2(\gamma_2),$$

where c_1 and c_2 are arbitrary functions and k is an arbitrary constant, except for the constraint $|\rho| < 1$;

(ii) the two variances are

$$\sigma_1^2 = \frac{c_1^2(\gamma_1)}{1-\rho^2}, \qquad \sigma_2^2 = \frac{c_2^2(\gamma_1)}{1-\rho^2};$$

(iii) if $k = 0$ (i.e., $\rho = 0$ identically), the means μ_1, μ_2 exclusively depend on γ_1, γ_2, respectively, but neither of them depends on either of the two factors if $k \neq 0$ (see Dzhafarov, 1999, pp. 140–141, for details).

The second meaning, which I term *unconditionally selective influence*, is the focus of the present study. This meaning, also discussed in Dzhafarov (1999), is derived from Dzhafarov (1992, 1997) and Dzhafarov and Schweickert (1995). According to the definition proposed in these papers, $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ are unconditionally selectively influenced by, respectively, $\{\Gamma_1, ..., \Gamma_n\}$ if these random variables can be presented as

$$\begin{bmatrix} \mathbf{X}_1 = X_1(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_1) \\ \cdots \\ \mathbf{X}_i = X_i(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_i) \\ \cdots \\ \mathbf{X}_n = X_n(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_n) \end{bmatrix}, \tag{1}$$

where $\{X_1, ..., X_n\}$ are some functions, while $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ are sources of randomness whose joint distribution does not depend on any factors (they can always be thought of as stochastically independent variables uniformly distributed between 0 and 1). The concept is schematically illustrated in Fig. 1: all stochasticity in $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ is due to factor-unrelated sources; stochastic interdependence follows from the fact that the same sources stochastically perturb different variables; and the factor subsets, $\{\Gamma_1, ..., \Gamma_n\}$, selectively modify the dependence of these variables on the random sources.

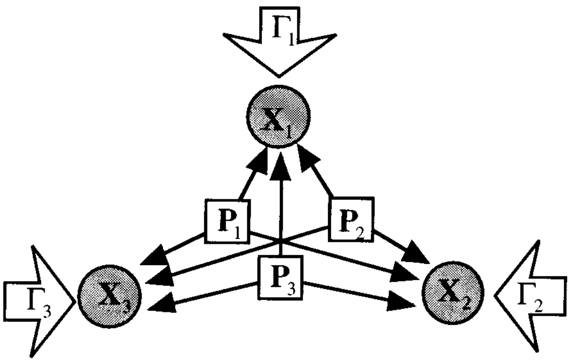

FIG. 1. Illustration to (1): $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$ selectively influenced by $\{\Gamma_1, \Gamma_2, \Gamma_3\}$. See text for details.

The unconditionally and conditionally selective forms of influence coincide when $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ are stochastically independent, but otherwise they generally exclude each other, if considered with respect to one and the same vector of factor subsets $\{\Gamma_1, ..., \Gamma_n\}$ (Dzhafarov, 1999).

The intuition underlying the representation (1) is clear and forms a departure point for the present development. It turns out, however, that to serve as a useful definition of unconditionally selective influence this representation should be amended along the following lines.

First, it is shown below that for one and the same distribution of random variables $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ (varying with Φ) there can be more than a single vector of factor subsets $\{\Gamma_1, ..., \Gamma_n\}$ which satisfies (1). To make the determination of $\{\Gamma_1, ..., \Gamma_n\}$ unique, which is an obvious desideratum, one needs one additional constraint, related to the notion of factor effectiveness (to be defined in the next subsection): each factor in Γ_i must effectively influence the marginal distribution of $\mathbf{X}_i$, $i = 1, ..., n$.

Second, certain constraints have to be imposed on the joint distribution for the random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$. In this paper I assume that the marginal distribution functions $\{F_1(x_1), ..., F_n(x_n)\}$ for this vector are continuous and that the random vector $\{\mathbf{U}_1, ..., \mathbf{U}_n\} = \{F_1(\mathbf{X}_1), ..., F_n(\mathbf{X}_n)\}$ (the copular base of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$, introduced in Subsection 1.2) possesses a continuous density on some n-dimensional convex area.

Third, to achieve a workable notion of unconditionally selective influence one has to assume that the functions $\{X_1, ..., X_n\}$ in (1) are sufficiently well-behaved. In the initial definition given in Section 2 the assumption is that $\{F_1 \circ X_1, ..., F_n \circ X_n\}$, where $F_i \circ X_i = F_i[X_i(\cdots)]$, form a homeomorphic transformation (i.e., a one-to-one transformation continuous together with its inverse).

At the end of Section 2 some of these requirements are relaxed to investigate the possibility that the number of the sources of randomness in (1) is greater or less than that of the random variables. In general, however, I do not set for myself the goal of making the technical *regularity constraints*, such as the continuity, convexity, and smoothness (existence of continuous density) just mentioned, as weak as possible. It appears that the numerous relaxations of these conditions which readily suggest themselves, while causing considerable technical difficulties, are unlikely to significantly broaden the scope of the conceivable applications of the theory.

There is one nontechnical aspect, however, in which the theory of unconditionally selective influence presented below is less general than that for the conditionally selective influence, as presented in Dzhafarov (1999): the theory below only applies to pairwise disjoint factor subsets $\{\Gamma_1, ..., \Gamma_n\}$,

$$i \neq j \Rightarrow \Gamma_i \cap \Gamma_j = \varnothing.$$

Since the union of all $\{\Gamma_1, ..., \Gamma_n\}$ is Φ (otherwise Φ can always be appropriately reduced), the vector $\{\Gamma_1, ..., \Gamma_n\}$ can be referred to as a *partition* of Φ, with the understanding that some of the subsets in this partition may be empty.

1.2. Preliminary Notions

We speak of a random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ and its joint distribution function as depending on Φ,

$$\begin{gathered}\{\mathbf{X}_1, ..., \mathbf{X}_n\} = \{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi) \\ \mathrm{Prob}[\mathbf{X}_1 \leqslant x_1, ..., \mathbf{X}_n \leqslant x_n] = F(x_1, ..., x_n; \Phi)\end{gathered} \tag{2}$$

even though it would be more precise to speak of a family of random vectors and their distribution functions indexed by all possible values of the factor set Φ. In all functions, such as $F(x_1, ..., x_n; \Phi)$, the factors separated from the arguments by a semicolon are treated as part of the functions' names rather than as their arguments. For example, instead of $F(x_1, ..., x_n; \Phi)$, one could write $F_{\Phi=\phi^1}(x_1, ..., x_n)$, $F_{\Phi=\phi^2}(x_1, ..., x_n)$, etc.

Marginal distribution functions are denoted as

$$\mathrm{Prob}[\mathbf{X}_i \leqslant x] = F_i(x; \Phi), \qquad i = 1, ..., n. \tag{3}$$

It is assumed that $F_i(x; \Phi)$ is continuous, for $i = 1, ..., n$. I call the vector of random variables

$$\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi) = \{F_1(\mathbf{X}_1; \Phi), ..., F_n(\mathbf{X}_n; \Phi)\} \tag{4}$$

the *copular base* of the random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$. Obviously, $\mathbf{U}_i$ ($i = 1, ..., n$) are standard uniformly distributed (i.e., uniformly distributed between 0 and 1): for any $0 \leqslant u \leqslant 1$,

$$\mathrm{Prob}[\mathbf{U}_i \leqslant u] = u, \qquad i = 1, ..., n. \tag{5}$$

The joint distribution function of $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$, however, generally depends on Φ:

$$\mathrm{Prob}[\mathbf{U}_1 \leqslant u_1, ..., \mathbf{U}_n \leqslant u_n] = C(u_1, ..., u_n; \Phi).$$

This joint distribution function is traditionally termed the *copula* of the random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ (see, e.g., Rüschendorf, Schweizer, & Taylor, 1996, and Nelsen, 1999; Colonius, 1990, was probably the first to use the notion of a copula in a psychological context). This is the term from which the "copular base" is derived. Obviously,

$$F(x_1, ..., x_n; \Phi) = C[F_1(x_1; \Phi), ..., F_n(x_n; \Phi)]. \tag{6}$$

The notion of a copula has been introduced in the theory of probability and statistics as a way of separating the pure stochastic interdependence of random variables from their marginal distributions. This intuition plays a prominent role in the present development. A random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is always viewed as being represented by (or decomposed into) its vectors of marginals $\{F_1(x_1; \Phi), ..., F_n(x_n; \Phi)\}$ and its copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$. Correspondingly, the unconditionally selective dependence of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ on a partition $\{\Gamma_1, ..., \Gamma_n\}$ of Φ is analyzed into two components: the selective dependence on $\{\Gamma_1, ..., \Gamma_n\}$ of the copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$, which constitutes the essence of the theory, and the far more trivial effect of $\{\Gamma_1, ..., \Gamma_n\}$ on the marginal distributions $\{F_1, ..., F_n\}$.

The copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ is assumed to possess a continuous joint density on some n-dimensional convex region of the standard n-dimensional unit cube, $[0, 1]^n$. Within this support, therefore, the joint distribution function $C(u_1, ..., u_n; \Phi)$ is continuous and strictly increasing in all arguments. The first-order conditional distribution functions for the copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ are defined as

$$\mathrm{Prob}\left[\mathbf{U}_i \leqslant u_i \,\middle|\, \mathop{\&}_{j \in \{1, ..., n\} - \{i\}} \mathbf{U}_j = u_j\right] = C_{i|}(u_1, ..., u_n; \Phi), \qquad i = 1, ..., n. \tag{7}$$

Within the support of $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ any conditional distribution function is continuous in all arguments and strictly increasing in the variable being conditioned (considered its main argument).

Each factor $\gamma \in \Phi$ attains its values in some set of possible values V_γ, containing at least two elements. A completely crossed factorial design is assumed, so that the possible values of the factor set Φ form the Cartesian product

$$V_\Phi = \mathop{\times}_{\gamma \in \Phi} V_\gamma.$$

A factor $\gamma \in \Gamma \subseteq \Phi$ is called effective with respect to a function $f(...; \Gamma)$ if this factor possesses at least two different values corresponding to two different values of f, at some fixed values of its arguments and the remaining factors in Γ. It is tacitly assumed throughout this paper that all factors in Φ are effective with respect to the joint distribution function $F(x_1, ..., x_n; \Phi)$ (otherwise Φ can always be rid of all ineffective factors).

1.3. *Overview of the Paper*

The aim of this paper is to develop a workable definition of the unconditionally selective influence and to ascertain the structure of the joint distribution for $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ that is necessary and sufficient for this random vector to be unconditionally selectively influenced by a given partition $\{\Gamma_1, ..., \Gamma_n\}$ of Φ.

Subsection 1.1 contains a discussion of the intuition underlying the notion of the unconditionally selective influence and the emendations that this intuition must undergo to evolve into a rigorous mathematical theory, as well as one important limitation (disjoint factor subsets) imposed on the generality of the theory. Subsection 1.2 introduces the preliminary notions and assumptions utilized in the theory to be presented: the decomposition of a random vector into its copular base and its vector of marginals (with certain regularity constraints imposed on both), and the notion of the factor effectiveness with respect to a function. I proceed now to overview the subsequent development, presented in Section 2.

The adjective *selective* (as in selective influence or selective attribution) hereafter means *unconditionally selective*, unless specified otherwise. I refer to vectors, such as $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ or $\{\Gamma_1, ..., \Gamma_n\}$, in both plural and singular grammatical forms: for example, the usage "$\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ are selectively attributable to $\{\Gamma_1, ..., \Gamma_n\}$" (meaning the random variables are) alternates with "$\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively attributable to $\{\Gamma_1, ..., \Gamma_n\}$" (meaning the random vector is).

1.3.1. *Basic Theory of Selective Attributability.* The intuition underlying (1) forms the basis for the notion of *selective attributability* whose introduction opens Section 2. In essence, a random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is considered to be selectively attributable to a partition $\{\Gamma_1, ..., \Gamma_n\}$ of Φ if a representation analogous to (1) holds for the copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ of this random vector (Definition 1, Subsection 2.1). The main results associated with this notion are as follows.

One can choose an arbitrary value Φ^0 of the factor set Φ as its anchoring value and take the random vector $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0)$ at this value to play the role of the sources of randomness, $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ in (1), for the copular bases $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ at all possible values of Φ (Lemma 2, Subsection 2.1).[2] Let the distribution of $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0)$ be known, as well as the distributions of the copular bases $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi_i^0)$, $i=1, ..., n$, taken at the values Φ_i^0 of the factor set Φ that are identical to the anchoring value Φ^0 in all components except for Γ_i. It turns out that then the fact of the selective attributability of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ to $\{\Gamma_1, ..., \Gamma_n\}$ allows one to determine the distribution of any other copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ essentially uniquely (leaving one a finite number of choices in the most general case). The form of this determination completely characterizes the notion of selective attributability and is derived in Theorems 2 and 3 (Subsection 2.5), based on Lemma 4 (Subsection 2.4). Recall that, given a vector of marginals $\{F_1, ..., F_n\}$, the distribution of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is uniquely determined by that of its copular base, $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$.

[2] The proof of all results stated in this paper, unless obvious, are given in the Appendix.

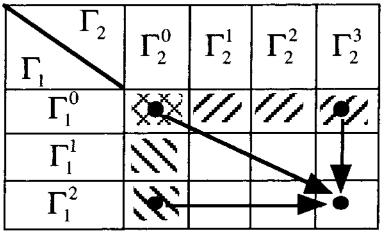

FIG. 2. Constraints imposed on copular bases $\{\mathbf{U}_1, \mathbf{U}_2\}$ by selective attribution of $\{\mathbf{X}_1, \mathbf{X}_2\}(\Gamma_1 \cup \Gamma_2)$ to $\{\Gamma_1, \Gamma_2\}$. See text for explanations.

Figure 2 provides an illustration for the case $n=2$, $V_{\Gamma_1}=\{\Gamma_1^0, \Gamma_1^1, \Gamma_1^2\}$, $V_{\Gamma_2}=\{\Gamma_2^0, \Gamma_2^1, \Gamma_2^2, \Gamma_2^3\}$ (i.e., Γ_1 and Γ_2 have, respectively, 3 and 4 possible values). Any particular pair of values, say, $\{\Gamma_1^0, \Gamma_2^0\}$, can be taken to serve as anchoring values for the factor set, and the distributions of $\{\mathbf{U}_1, \mathbf{U}_2\}(\Gamma_1 \cup \Gamma_2)$ corresponding to $\Gamma_1=\Gamma_1^0$ or $\Gamma_2=\Gamma_2^0$ (the shaded row and column in Fig. 2) are arbitrary (except for regularity constraints imposed on all copular bases). Given these distributions, however, the assumption that $\{\mathbf{X}_1, \mathbf{X}_2\}(\Gamma_1 \cup \Gamma_2)$ are selectively attributable to $\{\Gamma_1, \Gamma_2\}$, respectively, allows one to (essentially) uniquely determine the distributions corresponding to all other values of $\{\Gamma_1, \Gamma_2\}$. For example, the distribution of $\{\mathbf{U}_1, \mathbf{U}_2\}(\Gamma_1^2 \cup \Gamma_2^3)$ is determined by the distributions of $\{\mathbf{U}_1, \mathbf{U}_2\}(\Gamma_1^2 \cup \Gamma_2^0)$, $\{\mathbf{U}_1, \mathbf{U}_2\}(\Gamma_1^0 \cup \Gamma_2^3)$, and the anchoring distribution $\{\mathbf{U}_1, \mathbf{U}_2\}(\Gamma_1^0 \cup \Gamma_2^0)$.

The notion of selective attributability satisfies the following important requirement, termed *nestedness* (Subsection 2.6): if $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively attributable to $\{\Gamma_1, ..., \Gamma_n\}$, then any subvector of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$, say, $\{\mathbf{X}_1, \mathbf{X}_2\}(\Phi)$, is selectively attributable to the corresponding subpartition of $\{\Gamma_1, ..., \Gamma_n\}$, in this case $\{\Gamma_1, \Gamma_2\}$ (Theorem 5).

1.3.2. *Basic Theory of Selective Influence.* Selective attributability is a straightforward development of the naive intuition underlying the notion of selective influence, as proposed in Dzhafarov (1997, 1999). In particular, Lemma 1 (Subsection 2.1) establishes that selective attributability implies the representation (1), while Lemma 3 (ibid) shows that the sources of randomness, $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ in (1), can always be standardized as suggested in Dzhafarov (1997, 1999). (However, the identification of the randomness sources with the copular base at an anchoring value, mentioned above, turns out to be more useful than any such standardization.) The reason the notion of selective attributability has to be complemented by additional requirements has been mentioned in Subsection 1.1: one and the same random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ can generally be selectively attributed to different partitions of Φ. Example 1 (Subsection 2.2) shows that in some cases it can even be attributed selectively to *all* possible partitions of Φ. The theory turns therefore to another aspect of selectiveness in the dependence of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ on Φ, its manifestation on the level of marginal distributions. It is required that when $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively influenced by $\{\Gamma_1, ..., \Gamma_n\}$, any factor within Γ_i, $i=1, ..., n$, must be effective with respect to the marginal distribution $F_i(x_i)$. Selective dependence of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ on $\{\Gamma_1, ..., \Gamma_n\}$ is defined by combining this requirement with that of the selective attributability of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ to $\{\Gamma_1, ..., \Gamma_n\}$ (Definition 2, Subsection 2.3). Examples 1 and 2 (Subsections 2.2, 2.3) prove that these two requirements are logically independent, while Theorem 1 (Subsection 2.3) shows that with selective influence one does achieve the uniqueness that selective attributability lacks. The characterization of selective influence, that is, the necessary and sufficient structure that the joint distribution of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ has to have to be selectively influenced by $\{\Gamma_1, ..., \Gamma_n\}$, is trivially obtained from the characterization of selective attributability (Theorems 2 and 3, Subsection 2.5) by adding to it the requirement of the effectiveness of factors with respect to marginals (Theorem 4, ibid). Analogously, the nestedness of selective attributability (Theorem 5, Subsection 2.6) immediately implies the nestedness of selective influence (Theorem 6, ibid): if $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively influenced by $\{\Gamma_1, ..., \Gamma_n\}$, then any subvector of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively influenced by the corresponding subpartition of $\{\Gamma_1, ..., \Gamma_n\}$.

1.3.3. *Generalized Theory of Selective Attributability and Influence.* The development presented in Subsections 2.1–2.6 is confined to the case when the number of the sources of randomness in (1) is the same as that of the random variables $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$. In Subsection 2.7 I consider the case when the sources of randomness $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ are allowed to be more numerous than $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$. It turns out (Theorem 7) that if the regularity (well-behavedness) constraints imposed on the mapping of $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ onto $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ in the case $m = n$ are suitably generalized to $m \geqslant n$, then the number of the sources of randomness can always be reduced to $m = n$. The case $m \geqslant n$ does not, therefore, lead to a more general theory. If, however, the sources of randomness $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ are allowed to be less numerous than $\{\mathbf{X}_1, ..., \mathbf{X}_n\}$ (Subsection 2.8), one obtains a genuine generalization of the theory that has important substantive applications: it is often reasonable to assume (or at least to consider the possibility) that some of the components of the random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ are uniquely determined by its other components, including the case when all these components are deterministic functions of each other (which happens when there is only one source of randomness, $m = 1$). The analysis presented in Subsection 2.8 is based on the notion of a *regular m-dimensionality* of the vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$, that generalizes the regularity constraints imposed on the joint distribution of this vector in the basic theory. With this notion, all definitions and results of the basic theory are generalized in a rather straightforward manner, including a complete characterization of selective influence, with the uniqueness and nestedness properties preserved. To facilitate comparisons, the formal statements presented in Subsection 2.8 (with the exception of Lemma 6) are labeled in the same way as their counterparts in the basic theory but with asterisks added to their numbers: thus Definition 1*, Lemma 4*, Theorem 3*, etc., include Definition 1, Lemma 4, Theorem 3, etc., as their respective special cases.

2. THEORY

2.1. *Selective Attribution*

I begin with a definition that is very close to that proposed in Dzhafarov (1997, 1999) for selective influence but turns out to be weaker than the notion of selective influence we arrive at in this paper.

Definition 1. A random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is said to be selectively attributable to a partition $\{\Gamma_1, ..., \Gamma_n\}$ of Φ if

$$\{F_1(x_1; \Phi), ..., F_n(x_n; \Phi)\} = \{F_1(x_1; \Gamma_1), ..., F_n(x_n; \Gamma_n)\} \tag{8}$$

and if there exist a random vector $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ (whose joint distribution does not depend on Φ) and a homeomorphic transformation $\{\xi_1, ..., \xi_n\}_\Phi$, such that the copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is representable as

$$\begin{bmatrix} \mathbf{U}_1 = \xi_1(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_1) \\ \cdots \\ \mathbf{U}_n = \xi_n(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_n) \end{bmatrix}. \tag{9}$$

The proof of the following lemma is obvious, on defining

$$F_i^{-1}(p; \Phi) = \max\{x : F_i(x; \Phi) \leqslant p\}, \qquad i = 1, ..., n. \tag{10}$$

Lemma 1. *If* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$, *then the representation* (1) *holds, with*

$$X_i(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_i) = F_i^{-1}[\xi_i(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_i)], \qquad i = 1, ..., n. \tag{11}$$

The transformation $\{X_1, ..., X_n\}$ *is homeomorphic if and only if the functions* $\{F_1(x_1; \Gamma_1), ..., F_n(x_n; \Gamma_n)\}$ *are strictly increasing* (*on arbitrary interval domains*).

The following simple observation plays a key role in the subsequent development. Since a homeomorphic relationship is transitive, one has

$$\begin{tikzcd}
\{\mathbf{P}_1, ..., \mathbf{P}_n\} \arrow[rr, leftrightarrow] \arrow[dr, leftrightarrow] & & \{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0) \arrow[dl, leftrightarrow] \\
& \{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi) &
\end{tikzcd}$$

for an arbitrary value Φ^0 of the factor set Φ. It immediately follows that we have

Lemma 2. *If* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$, *then* $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ *in the representation* (9) *can be chosen to be*

$$\{\mathbf{P}_1, ..., \mathbf{P}_n\} = \{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0), \tag{12}$$

where Φ^0 *is an arbitrary* (*anchoring*) *value of the factor set* Φ.

To bring Definition 1 even closer to the notion discussed in Dzhafarov (1997, 1999), it is useful to observe (even though this observation is not utilized in the subsequent development) that the sources of randomness $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ can always be standardized, in the following sense.

LEMMA 3. *If* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$, *then the random vector* $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ *in* (9) *can be chosen to consist of independent identically distributed random variables with any given distribution function strictly increasing on a given interval domain* (*in particular*, *uniform between* 0 *and* 1).

See the Appendix for the proof.

2.2. Factor Effectiveness with Respect to Marginals

Recall, from Subsection 1.2, that all factors in Φ are assumed to be effective with respect to the distribution function $F(x_1, ..., x_n; \Phi)$. It does not follow from (8) in Definition 1, however, that all factors in Γ_i are effective with respect to the marginal distribution functions $F_i(x_i; \Gamma_i)$, $i = 1, ..., n$. The following example demonstrates that it is even possible for $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ to be selectively attributable to $\{\Gamma_1, ..., \Gamma_n\}$ without any of the factors being effective with respect to any of the marginals.

EXAMPLE 1. Let $\Phi = \{\gamma_1, \gamma_2\}$ (two independent factors), and let $\{\mathbf{X}_1, \mathbf{X}_2\}$ be defined as

$$\begin{bmatrix} \mathbf{X}_1 = \begin{cases} \mathbf{P}_1 & \text{if} \quad \gamma_1 = \gamma_1^0 \\ \mathbf{P}_2 & \text{if} \quad \gamma_1 \neq \gamma_1^0 \end{cases} \\ \mathbf{X}_2 = \begin{cases} \mathbf{P}_1 & \text{if} \quad \gamma_2 = \gamma_2^0 \\ \mathbf{P}_2 & \text{if} \quad \gamma_2 \neq \gamma_2^0 \end{cases} \end{bmatrix}, \tag{13}$$

where $\{\mathbf{P}_1, \mathbf{P}_2\}$ are independent random variables uniformly distributed between 0 and 1 and $\{\gamma_1^0, \gamma_2^0\}$ are particular values of the two factors. The factors $\{\gamma_1, \gamma_2\}$ are clearly effective with respect to the joint distribution of $\{\mathbf{X}_1, \mathbf{X}_2\}$. Indeed, on observing that

$$\text{Prob}[\mathbf{X}_1 = \mathbf{X}_2] = \begin{cases} 1 & \text{if} \quad (\gamma_1 = \gamma_1^0 \,\&\, \gamma_2 = \gamma_2^0) \quad \text{or} \quad (\gamma_1 \neq \gamma_1^0 \,\&\, \gamma_2 \neq \gamma_2^0) \\ 0 & \text{if} \quad (\gamma_1 = \gamma_1^0 \,\&\, \gamma_2 \neq \gamma_2^0) \quad \text{or} \quad (\gamma_1 \neq \gamma_1^0 \,\&\, \gamma_2 = \gamma_2^0) \end{cases},$$

if follows that the distribution of $\{\mathbf{X}_1, \mathbf{X}_2\}$ changes whenever γ_1 changes from γ_1^0 to another value and whenever γ_2 changes from γ_2^0 to another value. It is also clear that $\{\mathbf{X}_1, \mathbf{X}_2\}$ is selectively attributable to the partition $\{\Gamma_1 = \{\gamma_1\}, \Gamma_2 = \{\gamma_2\}\}$. Indeed, since both $\mathbf{P}_1$ and $\mathbf{P}_2$ are standard uniformly distributed,

$$\{\mathbf{U}_1, \mathbf{U}_2\}(\gamma_1, \gamma_2) = \{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2) \tag{14}$$

and (13) holds for $\{\mathbf{U}_1, \mathbf{U}_2\}$. However, (14) also implies that neither of the two marginal distributions depends on either of the two subsets $\{\Gamma_1 = \{\gamma_1\}, \Gamma_2 = \{\gamma_2\}\}$.

We see that selective attribution does not imply effectiveness with respect to the marginals. An important related fact is that selective attribution generally is not unique: one and the same random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ can be selectively attributed to different partitions of Φ. This nonuniqueness is the reason why the notion of selective attribution (Definition 1) cannot by itself serve as a satisfactory depiction of one's intuitive picture of external factors selectively affecting random variables.

Example 1 (Continued). The representation (13) selectively attributes the random vector $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$ to the partition $\{\Gamma_1 = \{\gamma_1\}, \Gamma_2 = \{\gamma_2\}\}$. It is easy to verify, however, that the same random vector (i.e., with the same joint distribution function for all values of the two factors) can be represented as

$$\begin{bmatrix} \mathbf{X}_1 = \begin{cases} \mathbf{P}_1 & \text{if} \quad \gamma_2 = \gamma_2^0 \\ \mathbf{P}_2 & \text{if} \quad \gamma_2 \neq \gamma_2^0 \end{cases} \\ \mathbf{X}_2 = \begin{cases} \mathbf{P}_1 & \text{if} \quad \gamma_1 = \gamma_1^0 \\ \mathbf{P}_2 & \text{if} \quad \gamma_1 \neq \gamma_1^0 \end{cases} \end{bmatrix},$$

or it can be represented as

$$\begin{bmatrix} \mathbf{X}_1 = \mathbf{P}_1 \\ \mathbf{X}_2 = \begin{cases} \mathbf{P}_1 & \text{if} \quad (\gamma_1 = \gamma_1^0 \,\&\, \gamma_2 = \gamma_2^0) \quad \text{or} \quad (\gamma_1 \neq \gamma_1^0 \,\&\, \gamma_2 \neq \gamma_2^0) \\ \mathbf{P}_2 & \text{if otherwise} \end{cases} \end{bmatrix},$$

or it can be represented as

$$\begin{bmatrix} \mathbf{X}_1 = \begin{cases} \mathbf{P}_1 & \text{if} \quad (\gamma_1 = \gamma_1^0 \,\&\, \gamma_2 = \gamma_2^0) \quad \text{or} \quad (\gamma_1 \neq \gamma_1^0 \,\&\, \gamma_2 \neq \gamma_2^0) \\ \mathbf{P}_2 & \text{if otherwise} \end{cases} \\ \mathbf{X}_2 = \mathbf{P}_1 \end{bmatrix}.$$

Indeed, from all these representations, including (13), one can derive one and the same joint distribution function: for $0 < x_1, x_2 < 1$,

$$\begin{aligned} &F(x_1, x_2; \{\gamma_1, \gamma_2\}) \\ &\quad = \begin{cases} x_1 x_2 & \text{if} \quad (\gamma_1 = \gamma_1^0 \,\&\, \gamma_2 \neq \gamma_2^0) \quad \text{or} \quad (\gamma_1 \neq \gamma_1^0 \,\&\, \gamma_2 = \gamma_2^0) \\ \min\{x_1, x_2\} & \text{if} \quad (\gamma_1 = \gamma_1^0 \,\&\, \gamma_2 = \gamma_2^0) \quad \text{or} \quad (\gamma_1 \neq \gamma_1^0 \,\&\, \gamma_2 \neq \gamma_2^0) \end{cases}. \end{aligned}$$

Due to (14), the existence of these alternative representations means that $\{\mathbf{X}_1, \mathbf{X}_2\}\{\gamma_1, \gamma_2\}$ is selectively attributable to all four possible partitions of $\Phi = \{\gamma_1, \gamma_2\}$:

$$\{\Gamma_1 = \{\gamma_1\}, \Gamma_2 = \{\gamma_2\}\}, \quad \{\Gamma_1 = \{\gamma_2\}, \Gamma_2 = \{\gamma_1\}\},$$
$$\{\Gamma_1 = \emptyset, \Gamma_2 = \{\gamma_1, \gamma_2\}\}, \quad \{\Gamma_1 = \{\gamma_1, \gamma_2\}, \Gamma_2 = \emptyset\}.$$

2.3. *Selective Influence*

One can conjecture that the nonuniqueness of the selective attribution in Example 1 is a consequence of the fact that although both factors γ_1, γ_2 are effective with respect to the joint distribution of $\{\mathbf{X}_1, \mathbf{X}_2\}$, none of them is effective with respect to the distributions of $\mathbf{X}_1$ or $\mathbf{X}_2$ taken separately. If so, then a satisfactory definition of selective influence should involve both selective attributability to a partition of the factor set and the effectiveness of the factor subsets in the partition with respect to the marginal distributions.[3]

DEFINITION 2. A random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is said to be selectively influenced by a partition $\{\Gamma_1, ..., \Gamma_n\}$ of Φ if

(i) $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively attributable to $\{\Gamma_1, ..., \Gamma_n\}$;

(ii) for any $i=1, ..., n$, any factor $\gamma \in \Gamma$ is effective with respect to the marginal distribution $F_i(x_i; \Gamma_i)$ of $\mathbf{X}_i$.

With this definition one achieves the desired uniqueness.

THEOREM 1. *If a random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively influenced by a partition $\{\Gamma_1, ..., \Gamma_n\}$ of Φ, then it is not selectively influenced by any other partition of Φ.*

This uniqueness theorem follows from Definition 2 immediately upon observing that a partition $\{\Gamma_1, ..., \Gamma_n\}$ can only selectively influence $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ if Γ_i is precisely the set of factors that are effective with respect to the distribution function of $\mathbf{X}_i$, $i=1, ..., n$.

The requirements (i) and (ii) in Definition 2 are logically independent. Example 1 above illustrates the fact that (i) can be satisfied without (ii). To demonstrate the reverse, consider

EXAMPLE 2. Let $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$ be defined by

$$\begin{bmatrix} \mathbf{X}_1 = \gamma_1 \times \begin{cases} \mathbf{P}_1 & \text{if} \quad \gamma_1 = \gamma_2 \\ \mathbf{P}_2 & \text{if} \quad \gamma_1 \neq \gamma_2 \end{cases} \\ \mathbf{X}_2 = \gamma_2 \mathbf{P}_2 \end{bmatrix}, \tag{15}$$

$\{\mathbf{P}_1, \mathbf{P}_2\}$ being, again, independent random variables uniformly distributed between 0 and 1 and $\{\gamma_1, \gamma_2\}$ being arbitrary real numbers. Clearly, both factors are effective with respect to the joint distribution function. It is also clear that $\Gamma_1 = \{\gamma_1\}$ is the precise subset of factors effective with respect to (the marginal distribution of) $\mathbf{X}_1$, while $\Gamma_2 = \{\gamma_2\}$ is the precise subset of factors effective with respect to $\mathbf{X}_2$. The copular base of $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$ is

$$\begin{bmatrix} \mathbf{U}_1 = \begin{cases} \mathbf{P}_1 & \text{if} \quad \gamma_1 = \gamma_2 \\ \mathbf{P}_2 & \text{if} \quad \gamma_1 \neq \gamma_2 \end{cases} \\ \mathbf{U}_2 = \mathbf{P}_2 \end{bmatrix}. \tag{16}$$

[3] Townsend and Schweickert (1989) introduce the notion of *marginal selectivity* that corresponds to (8) in this paper. It is possible that they implicity assume that Γ_i in $F_i(x_i; \Gamma_i)$, $i=1, ..., n$, cannot contain ineffective factors, in which case their marginal selectivity also includes the present notion of effectiveness with respect to marginals.

It is easy to realize that $\{\mathbf{U}_1, \mathbf{U}_2\}(\gamma_1, \gamma_2)$ cannot be selectively attributed to the partition $\{\{\gamma_1\}, \{\gamma_2\}\}$. Indeed, otherwise the functions v_1, v_2 in any representation

$$\begin{bmatrix} \mathbf{U}_1 = v_1(\mathbf{R}_1, \mathbf{R}_2; \{\gamma_1\}) \\ \mathbf{U}_2 = v_2(\mathbf{R}_1, \mathbf{R}_2; \{\gamma_2\}) \end{bmatrix},$$

where $\mathbf{R}_1, \mathbf{R}_2$ are some random variables, would have to be different when $\gamma_1 = \gamma_2$ but identical when $\gamma_1 \neq \gamma_2$. This is impossible because v_1 cannot depend on γ_2, or v_2 on γ_1.

2.4. *Choice of Sources of Randomness*

The main advantage gained by the assumption that the transformation $\{\xi_1, ..., \xi_n\}_\Phi$ in (9) is a homeomorphism consists in the applicability of Lemma 2: the vector $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ of the sources of randomness can be identified with $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ itself, taken at a particular, anchoring, value Φ^0 of the factor set Φ. It is shown in the next subsection that with this choice of $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ the transformation $\{\xi_1, ..., \xi_n\}_\Phi$ lends itself to a complete characterization in terms of the conditional distribution functions for $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$, taken at certain special values of Φ. The following two properties of the functions $\{\xi_1, ..., \xi_n\}_\Phi$ help one in achieving this goal.

LEMMA 4. *Let $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ be selectively attributable to $\{\Gamma_1, ..., \Gamma_n\}$, and let $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ be chosen according to* (12),

$$\{\mathbf{P}_1, ..., \mathbf{P}_n\} = \{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0).$$

Let $\{\Gamma_1, ..., \Gamma_n\} = \{\Gamma_1^0, ..., \Gamma_n^0\}$ for the anchoring value Φ^0. Then, within the support of $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$, for any $i = 1, ..., n$ and any value of Γ_i in (9),

(i) $\xi_i(p_1, ..., p_n; \Gamma_i^0) = p_i$;

(ii) $\xi_i(p_1, ..., p_n; \Gamma_i)$ *is strictly monotonic* (*increasing or decreasing*) *in* p_i.

See the Appendix for the proof.

2.5. *Characterization of Selective Influence*

The theorems below play a central role in this paper. They express the functions $\{\xi_1, ..., \xi_n\}_\Phi$ in (9) through the conditional distribution functions taken at special values of Φ,

$$C_{i|}(p_1, ..., p_n; \Phi^0) \qquad \text{and} \qquad C_{i|}(p_1, ..., p_n; \Phi_i^0), \qquad i = 1, ..., n,$$

where Φ_i^0 is defined as follows: given a partition $\{\Gamma_1, ..., \Gamma_n\}$ of Φ and a choice of anchoring values $\{\Gamma_1^0, ..., \Gamma_n^0\}$, Φ_i^0 denotes the factor set with freely varying Γ_i but with $\Gamma_j = \Gamma_j^0$ for all $j \neq i$ (the shaded cells in Fig. 2). Recall that, due to the convexity of the support for the continuous density of the copular base, any of these special conditionals is continuously increasing in its main argument (on some interval) and can therefore be inverted. Thus

$$u = C_{i|}(p_1, ..., p_{i-1}, p, p_{i+1}, ..., p_n; \Phi_i^0) \Leftrightarrow p = C_{i|}^{-1}(p_1, ..., p_{i-1}, u, p_{i+1}, ..., p_n; \Phi_i^0).$$

I also make use of the structural symbol $\updownarrow$ preceding a variable belonging to $[0, 1]$. The meaning of $\updownarrow a$ is "either a or $1-a$."

Theorem 2. *If* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$, *and if in the representation*

$$\begin{bmatrix} \mathbf{U}_1 = \xi_1(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_1) \\ \cdots \\ \mathbf{U}_n = \xi_n(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma) \end{bmatrix}$$

the vector $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ *is chosen according to* (12),

$$\{\mathbf{P}_1, ..., \mathbf{P}_n\} = \{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0),$$

then, for $i = 1, ..., n$,

$$\xi_i(p_1, ..., p_n; \Gamma_i) = C_{i|}^{-1}[p_1, ..., p_{i-1}, \updownarrow C_{i|}(p_1, ..., p_n; \Phi^0), p_{i+1}, ..., p_n; \Phi_i^0], \tag{17}$$

where the specification of $\updownarrow C_{i|}$ *(i.e., the choice between* $C_{i|}$ *or* $1 - C_{i|}$*) may only depend on* Γ_i.

See the Appendix for the proof.

Stated directly in terms of random variables, Theorem 2 says that the only possible representation (9) is

$$\begin{bmatrix} \mathbf{U}_1 = C_{1|}^{-1}[\updownarrow C_{1|}(\mathbf{P}_1, ..., \mathbf{P}_n; \Phi^0), \mathbf{P}_2, ..., \mathbf{P}_n; \Phi_1^0] \\ \cdots \\ \mathbf{U}_i = C_{i|}^{-1}[\mathbf{P}_1, ..., \mathbf{P}_{i-1}, \updownarrow C_{i|}(\mathbf{P}_1, ..., \mathbf{P}_n; \Phi^0), \mathbf{P}_{i+1}, ..., \mathbf{P}_n; \Phi_i^0] \\ \cdots \\ \mathbf{U}_n = C_{n|}^{-1}[\mathbf{P}_1, ..., \mathbf{P}_{n-1}, \updownarrow C_{n|}(\mathbf{P}_1, ..., \mathbf{P}_n; \Phi^0); \Phi_n^0] \end{bmatrix}, \tag{18}$$

Clearly, the representation (18) is not only necessary but is also sufficient for selective attributability. It is also clear that once the functions $\{\xi_1, ..., \xi_n\}_\Phi$ are specified, the density function $\psi(u_1, ..., u_n; \Phi)$ for $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$, at arbitrary Φ, can be found from the density function $\psi(u_1, ..., u_n; \Phi^0)$ for $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0)$, taken at the anchoring value of Φ. This leads one to the following characterization of selective attribution.

THEOREM 3. *$\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively attributable to $\{\Gamma_1, ..., \Gamma_n\}$ if and only if its copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ is related to $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0)$ (taken at an anchoring value Φ^0 of Φ) by means of* (18). *In this case*

$$\psi[\xi_1(p_1, ..., p_n; \Gamma_1), ..., \xi_n(p_1, ..., p_n; \Gamma_n); \Phi]\, |d\xi(p_1, ..., p_n; \Gamma_1)\ldots d\xi_n(p_1, ..., p_n; \Gamma_n)| = \psi(p_1, ..., p_n; \Phi^0)\, |dp_1, ..., dp_n|, \tag{19}$$

where $\{\xi_1, ..., \xi_n\}_\Phi$ are defined by (17). *In particular, if the density of $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ is continuously differentiable, then $\{\xi_1, ..., \xi_n\}_\Phi$ is a diffeomorphism* (*i.e., it is continuously differentiable with a nonvanishing Jacobian*), *and the criterion* (19) *can be written as*

$$\psi[\xi_1(p_1, ..., p_n; \Gamma_1), ..., \xi_n(p_1, ..., p_n; \Gamma_n); \Phi] = \psi(p_1, ..., p_n; \Phi^0) \left|\frac{d\xi_1(p_1, ..., p_n; \Gamma_1)\ldots d\xi_n(p_1, ..., p_n; \Gamma_n)}{dp_1\ldots dp_n}\right|^{-1} \tag{20}$$

The proof of this statement is trivial: (19) merely asserts that probability elements of $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ and $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0)$ at corresponding points are the same, and (20) restates this by making use of the existence of the nonzero Jacobian

$$\left|\frac{d\xi_1(p_1, ..., p_n; \Gamma_1)\ldots d\xi_n(p_1, ..., p_n; \Gamma_n)}{dp_1\ldots dp_n}\right|.$$

To complete the analysis, it remains to observe that Definition 2 together with Theorems 2 and 3 immediately lead to the following characterization of selective influence.

THEOREM 4. *$\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively influenced by $\{\Gamma_1, ..., \Gamma_n\}$ if and only if*

(i) *for any $i = 1, ..., n$, any factor $\gamma \in \Gamma_i$ is effective with respect to the marginal distribution $F_i(x_i; \Gamma_i)$ of $\mathbf{X}_i$;*

(ii) *the copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ is related to $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi^0)$ (taken at an anchoring value Φ^0 of Φ) by means of* (18).

2.6. Nestedness of Selective Influence

The uniqueness of the partition $\{\Gamma_1, ..., \Gamma_n\}$ that may selectively influence a random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ (Theorem 1), though critical, is not by itself sufficient to ensure that the notion of selective influence is well-constructed. The second obvious desideratum for a well-constructed definition is that if $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively influenced by $\{\Gamma_1, ..., \Gamma_n\}$, then any subvector of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ must be selectively influenced by the corresponding subpartition of $\{\Gamma_1, ..., \Gamma_n\}$. This is indeed the case.

THEOREM 5. *If a random vector* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to a partition* $\{\Gamma_1, ..., \Gamma_n\}$ *of* Φ, *then, for any subset* $\{i_1, ..., i_k\}$ *of* $\{1, ..., n\}$ $(1 \leqslant k \leqslant n)$, *the subvector* $\{\mathbf{X}_{i_1}, ..., \mathbf{X}_{i_k}\}$ *only depends on* $\Phi' = \Gamma_{i_1} \cup \cdots \cup \Gamma_{i_k}$, *and it is selectively attributable to the partition* $\{\Gamma_{i_1}, ..., \Gamma_{i_k}\}$ *of* Φ'.

See the Appendix for the proof.

Since the property of the effectiveness of Γ_i with respect to the marginal distribution $F_i(x_i; \Gamma_i)$ is unrelated to other factor subsets or other marginals, one immediately obtains

THEOREM 6. *If a random vector* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively influenced by a partition* $\{\Gamma_1, ..., \Gamma_n\}$ *of* Φ, *then, for any subset* $\{i_1, ..., i_k\}$ *of* $\{1, ..., n\}$ $(1 \leqslant k \leqslant n)$, *the subvector* $\{\mathbf{X}_{i_1}, ..., \mathbf{X}_{i_k}\}$ *only depends on* $\Phi' = \Gamma_{i_1} \cup \cdots \cup \Gamma_{i_k}$, *and it is selectively influenced by the partition* $\{\Gamma_{i_1}, ..., \Gamma_{i_k}\}$ *of* Φ'.

2.7. *Excessive Sources of Randomness*

The requirement that $\{\xi_1, ..., \xi_n\}_\Phi$ in Definition 1 be a homeomorphism implies that the number of the sources of randomness $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ coincides with that of the random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$. One may wonder whether the notion of selective attributability can be generalized in an interesting way if one replaces (9) with

$$\begin{bmatrix} \mathbf{U}_1 = \xi_1(\mathbf{P}_1, ..., \mathbf{P}_m; \Gamma_1) \\ \cdots \\ \mathbf{U}_n = \xi_n(\mathbf{P}_1, ..., \mathbf{P}_m; \Gamma_n) \end{bmatrix}, \tag{21}$$

where m in $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ is allowed to be greater than n or less than n (Fig. 3). The answer to this question turns out to be different for the two possibilities, $m \geqslant n$ and $m \leqslant n$. Both these inequalities are made to be nonstrict in order to emphasize that well-constructed generalizations in both cases should reduce to Definition 1 when $m = n$. This implies, in particular, that the mapping $\{\xi_1, ..., \xi_n\}_\Phi$ should be sufficiently well-behaved in the general case to become homeomorphic at $m = n$.

I consider the $m \geqslant n$ version first. A natural well-behavedness constraint for the mapping $\{\xi_1, ..., \xi_n\}_\Phi$ in this situation consists in requiring that this mapping be *regularly continuous*, in the following sense: it is to be continuous and there are to

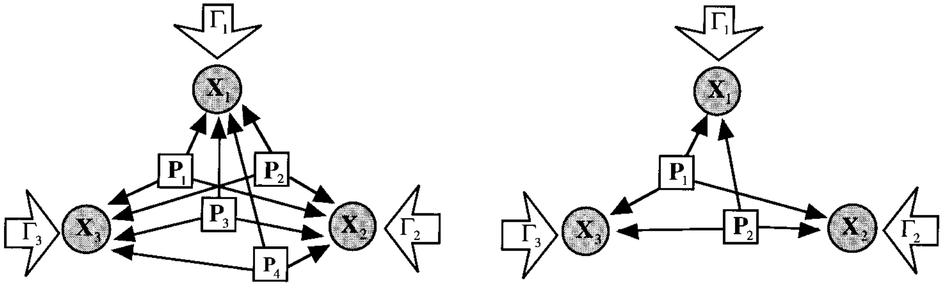

FIG. 3. Excessive (left) and defective (right) sources of randomness. See text for details.

be $m-n$ components of $\{p_1, ..., p_m\}$ (with no loss of generality, let them be $\{p_{n+1}, ..., p_m\}$) such that, for any fixed values of $\{p_{n+1}, ..., p_m\}$,

(i) the relationship between the remaining n components $\{p_1, ..., p_n\}$ and $\{u_1 ..., u_n\}(\Phi)$ is homeomorphic;

(ii) the random vector $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(p_{n+1}, ..., p_m, \Phi)$ has a continuous density function on a convex n-dimensional region (of the n-dimensional standard unit cube).

The key to dealing with regularly continuous mappings is provided by the following lemma, whose proof is obvious.

LEMMA 5. *If the copular base* $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ *of* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is representable according to* (21) *with* $m \geqslant n$ *and with regularly continuous* $\{\xi_1, ..., \xi_n\}_\Phi$, *then the mapping*

$$\begin{bmatrix} \mathbf{U}_1 = \xi_1(\mathbf{P}_1, ..., \mathbf{P}_m; \Gamma_1) \\ \cdots \\ \mathbf{U}_n = \xi_n(\mathbf{P}_1, ..., \mathbf{P}_m; \Gamma_n) \\ \mathbf{U}_{n+1} = \mathbf{P}_{n+1} \\ \cdots \\ \mathbf{U}_m = \mathbf{P}_m \end{bmatrix} \tag{22}$$

is a homeomorphism.

In other words, given a regularly continuous mapping from $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ to $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$, one gets a homeomorphism between $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ and the extended random vector $\{\mathbf{U}_1, ..., \mathbf{U}_n, \mathbf{U}_{n+1}, ..., \mathbf{U}_m\}(\Phi)$, where the subvector $\{\mathbf{U}_{n+1}, ..., \mathbf{U}_m\}$ is identical to $\{\mathbf{P}_{n+1}, ..., \mathbf{P}_m\}$ at any value of Φ. Using this lemma, one can simply repeat the steps of the previous development to obtain the following results. First, $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ can be chosen to coincide with $\{\mathbf{U}_1, ..., \mathbf{U}_n, \mathbf{U}_{n+1}, ..., \mathbf{U}_m\}(\Phi^0)$, taken at an anchoring value of Φ. Second, Lemma 4 holds for any of the functions $\{\xi_1, ..., \xi_n\}_\Phi$: in particular, $\xi_i(p_1, ..., p_m; \Gamma_i)$ is strictly monotonic in p_i, $i = 1, ..., n$. Finally, Theorems 2 and 4 also apply to these functions, and by considering $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ as a subvector of $\{\mathbf{U}_1, ..., \mathbf{U}_n, \mathbf{U}_{n+1}, ..., \mathbf{U}_m\}(\Phi)$ one concludes that (22) implies the representability

$$\begin{bmatrix} \mathbf{U}_1 = \zeta_1(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_1) \\ \cdots \\ \mathbf{U}_n = \zeta_n(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_n) \end{bmatrix}, \tag{23}$$

where $\{\zeta_1, ..., \zeta_n\}_\Phi$ is a homeomorphism. This proves

THEOREM 7. *If the copular base* $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ *of* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is representable according to* (21) *with* $m \geqslant n$ *and with regularly continuous* $\{\xi_1, ..., \xi_n\}_\Phi$, *then* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$ *in the sense of Definition* 1.

If, in addition, any factor $\gamma \in \Gamma_i$ *is effective with respect to the marginal distribution* $F_i(x_i; \Gamma_i)$ *of* $\mathbf{X}_i$, $i = 1, ..., n$, *then* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively influenced by* $\{\Gamma_1, ..., \Gamma_n\}$ *in the sense of Definition* 2.

In other words, the $m \geqslant n$ version of (21) is immediately reducible to $m = n$: excessive sources of randomness can always be eliminated.

2.8. Defective Sources of Randomness

It might appear that the $m \leqslant n$ version of (21) too can be immediately reduced to $m = n$, by adding to the list of the sources of randomness $n - m$ dummy variables. The theory as presented so far, however, will not be applicable then, because $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ representable by means of a homeomorphism (21) with $m \leqslant n$ cannot have a density on an n-dimensional region.

As an example of the situation in question, Dzhafarov and Schweickert (1995), Cortese and Dzhafarov (1996), and Dzhafarov and Cortese (1996) analyzed the possibility that the random variables $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ are *perfectly positively stochastically interdependent*,

$$\begin{bmatrix} \mathbf{X}_1 = X_1(\mathbf{P}; \Gamma_1) \\ \cdots \\ \mathbf{X}_n = X_n(\mathbf{P}; \Gamma_n) \end{bmatrix}, \tag{24}$$

where all functions are increasing in the value of $\mathbf{P}$, a single common source of randomness. In terms of the present paper this means

$$\begin{bmatrix} \mathbf{U}_1 = \mathbf{P} \\ \cdots \\ \mathbf{U}_n = \mathbf{P} \end{bmatrix}, \tag{25}$$

that is, the copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ is simply $\{\mathbf{P}, ..., \mathbf{P}\}$, for any value of Φ. Obviously, the joint density of $\{\mathbf{P}, ..., \mathbf{P}\}$ is not a function on an n-dimensional region.

Motivated by this example, let us say that a random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is *regularly m-dimensional* $(1 \leqslant m \leqslant n)$ if

(i) all its marginal distribution functions are continuous;

(ii) there is at least one m-component subvector of its copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ that possesses a continuous density;

(iii) the support for the density of any such a subvector is an m-dimensional convex region (of the m-dimensional standard unit cube);

(iv) for any such a subvector, the remaining components of $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ are continuously differentiable functions of this subvector (depending on the factor set Φ).

A subvector just defined is called a *regular copular subbase* of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$. The copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ as defined in Subsection 1.2 is the only regular copular subbase when $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is regularly n-dimensional (i.e., when $m = n$).

In the general case, the random variables can always be so arranged that $\{\mathbf{U}_1, ..., \mathbf{U}_m\}(\Phi)$ is a regular copular subbase, while

$$\mathbf{U}_i = \omega_i(\mathbf{U}_1, ..., \mathbf{U}_m; \Phi), \qquad i = m+1, ..., n, \tag{26}$$

are continuously differentiable functions. Clearly, the dimensionality m is determined uniquely, while the choice of a regular copular subbase need not be unique. In particular, it is possible that any m-component subvector of $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ can be chosen to serve as $\{\mathbf{U}_1, ..., \mathbf{U}_m\}(\Phi)$ above. Whatever the case, if a subvector of $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ is a regular copular subbase at some value of the factor set Φ, then, according to the definition, it has to be a regular copular subbase for all other values of Φ. One is now naturally led to the following generalization of Definition 1.

Definition 1*. A regularly m-dimensional ($1 \leqslant m \leqslant n$) random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is said to be selectively attributable to a partition $\{\Gamma_1, ..., \Gamma_n\}$ of Φ if

$$\{F_1(x_1; \Phi), ..., F_n(x_n; \Phi)\} = \{F_1(x_1; \Gamma_1), ..., F_n(x_n; \Gamma_n)\}$$

and if there exist a random vector $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ (whose joint distribution does not depend on Φ) and a homeomorphism $\{\xi_1, ..., \xi_n\}_\Phi$, such that the copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is representable according to (21),

$$\begin{bmatrix} \mathbf{U}_1 = \xi_1(\mathbf{P}_1, ..., \mathbf{P}_m; \Gamma_1) \\ \cdots \\ \mathbf{U}_n = \xi_n(\mathbf{P}_1, ..., \mathbf{P}_m; \Gamma_n) \end{bmatrix}.$$

Since any regular copular subbase of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is homeomorphically related to the copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$, one obtains

Lemma 6. *Definition* 1* *implies that* $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ *is homeomorphically related to any regular copular subbase of* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$.

Assuming, as above, that $\{\mathbf{U}_1, ..., \mathbf{U}_m\}(\Phi)$ is a regular copular subbase, one obtains the following obvious generalizations of Lemmas 2 and 3.

Lemma 2*. *If a regularly m-dimensional* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$, *then* $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ *in the representation* (21) *can be chosen to be*

$$\{\mathbf{P}_1, ..., \mathbf{P}_m\} = \{\mathbf{U}_1, ..., \mathbf{U}_m\}(\Phi^0), \tag{27}$$

where Φ^0 *is an arbitrary* (*anchoring*) *value of the factor set* Φ.

Lemma 3*. *If a regularly m-dimensional* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$, *then the random vector* $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ *in* (21) *can be chosen to consist of independent identically distributed random variables with any given distribution function strictly increasing on a given interval domain* (*in particular, uniform between* 0 *and* 1).

With the generalized meaning for selective attributability, Definition 2 for selective influence remains unchanged. It is easy to see that the same is true for Theorem 1 asserting the uniqueness of the partition $\{\Gamma_1, ..., \Gamma_n\}$ that may selectively influence a given (regularly m-dimensional) $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$. The following generalized versions of Lemma 4 and Theorem 2 are also obvious.

LEMMA 4*. *Let a regularly m-dimensional* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *be selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$, *and let* $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ *be chosen according to* (27). *Let* $\{\Gamma_1, ..., \Gamma_n\} = \{\Gamma_1^0, ..., \Gamma_n^0\}$ *for the anchoring value* Φ^0. *Then, within the support of* $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$, *for any* $i = 1, ..., m$ *and any value of* Γ_i *in* (21),

(i) $\xi_i(p_1, ..., p_m; \Gamma_i^0) = p_i$;

(ii) $\xi_i(p_1, ..., p_m; \Gamma_i)$ *is strictly monotonic* (*increasing or decreasing*) *in* p_i.

THEOREM 2*. *If a regularly m-dimensional* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$ *and if in the representation* (21) *the vector* $\{\mathbf{P}_1, ..., \mathbf{P}_m\}$ *is chosen according to* (27),

$$\{\mathbf{P}_1, ..., \mathbf{P}_m\} = \{\mathbf{U}_1, ..., \mathbf{U}_m\}(\Phi^0),$$

then, for $i = 1, ..., m$,

$$\xi_i(p_1, ..., p_m; \Gamma_i) = C_{i|}^{-1}[p_1, ..., p_{i-1}, \updownarrow C_{i|}(p_1, ..., p_m; \Phi^0), p_{i+1}, ..., p_m; \Phi_i^0], \tag{28}$$

where the specification $\updownarrow C_{i|}$ (*i.e., the choice between* $C_{i|}$ *or* $1 - C_{i|}$) *may only depend on* Γ_i.

Unlike its prototype, this theorem only establishes the form of the functions $\{\xi_1, ..., \xi_m\}_\Phi$. It has, therefore, to be complemented by a description of the remaining functions $\{\xi_{m+1}, ..., \xi_n\}_\Phi$. As does (28), this description also refers to the distributions of the copular bases at special values of the factor set, Φ_i^0, $i = m+1, ..., n$.

THEOREM 2* (Continued). *For* $i = m+1, ..., n$,

$$\xi_i(p_1, ..., p_m; \Gamma_i) = \omega_i[p_1, ..., p_m; \Phi_i^0], \tag{29}$$

where the functions $\{\omega_{m+1}, ..., \omega_n\}_\Phi$ *are defined by* (26).

The proof of this statement is obvious. Stated directly in terms of random variables, Theorem 2* says that if a regularly m-dimensional $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is selectively attributable to $\{\Gamma_1, ..., \Gamma_n\}$, then it can be represented as

$$\begin{bmatrix} \mathbf{U}_1 = C_{1|}^{-1}[\updownarrow C_{1|}(\mathbf{P}_1, ..., \mathbf{P}_m; \Phi^0), \mathbf{P}_2, ..., \mathbf{P}_m; \Phi_1^0] \\ \cdots \\ \mathbf{U}_i = C_{i|}^{-1}[\mathbf{P}_1, ..., \mathbf{P}_{i-1}, \updownarrow C_{i|}(\mathbf{P}_1, ..., \mathbf{P}_m; \Phi^0), \mathbf{P}_{i+1}, ..., \mathbf{P}_m; \Phi_i^0] \\ \cdots \\ \mathbf{U}_m = c_{m|}^{-1}[\mathbf{P}_1, ..., \mathbf{P}_{m-1}, \updownarrow C_{m|}(\mathbf{P}_1, ..., \mathbf{P}_m; \Phi^0); \Phi_m^0] \\ \mathbf{U}_{m+1} = \omega_{m+1}[\mathbf{P}_1, ..., \mathbf{P}_m; \Phi_{m+1}^0] \\ \cdots \\ \mathbf{U}_n = \omega_n[\mathbf{P}_1, ..., \mathbf{P}_m; \Phi_n^0] \end{bmatrix}. \tag{30}$$

A complete characterization of selective attributability (and selective influence) is given by the following straightforward generalization of Theorem 3.

THEOREM 3*. *A regularly m-dimensional* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively attributable to* $\{\Gamma_1, ..., \Gamma_n\}$ *if and only if its copular base* $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ *is related to* $\{\mathbf{U}_1, ..., \mathbf{U}_m\}(\Phi^0)$ (*taken at an anchoring value* Φ^0 *of* Φ) *by means of* (30). *In this case* (*denoting density functions by* ψ),

$$\psi[\xi_1(p_1, ..., p_m; \Gamma_1), ..., \xi_n(p_1, ..., p_m; \Gamma_n); \Phi] \\ \times |d\xi_1(p_1, ..., p_m; \Gamma_1)...d\xi_m(p_1, ..., p_m; \Gamma_m)| \\ = \psi(p_1, ..., p_m; \Phi^0)\, |dp_1...dp_m|, \tag{31}$$

where $\{\xi_1, ..., \xi_n\}_\Phi$ *are defined by* (28) *and* (29). *In particular, if the* (*m-dimensional*) *density of* $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ *is continuously differentiable, then* $\{\xi_1, ..., \xi_n\}_\Phi$ *is a diffeomorphism, and the criterion* (31) *can be written as*

$$\psi[\xi_1(p_1, ..., p_m; \Gamma_1), ..., \xi_m(p_1, ..., p_m; \Gamma_n); \Phi] \\ = \psi(p_1, ..., p_m; \Phi^0) \left|\frac{d\xi_1(p_1, ..., p_m; \Gamma_1)...d\xi_m(p_1, ..., p_m; \Gamma_m)}{dp_1...dp_m}\right|^{-1}. \tag{32}$$

If, in addition, any factor $\gamma \in \Gamma_i$ *is effective with respect to the marginal distribution* $F_i(x_i; \Gamma_i)$ *of* $\mathbf{X}_i$, $i = 1, ..., n$, *then* $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ *is selectively influenced by* $\{\Gamma_1, ..., \Gamma_n\}$ *in the sense of Definition* 2 (*based on Definition* 1*).

Note that the value of (32) does not depend on $\{\Gamma_{m+1}, ..., \Gamma_n\}$, but this subpartition, together with $\{\Gamma_1, ..., \Gamma_m\}$, determines the point $\{u_1, ..., u_n\}$ to which this value is assigned. Note also that the representations (30) and (31)–(32) can be constructed in as many different ways as there are regular copular subbases in $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$.

To complete the analysis, it only remains to verify that the generalized notion of selective attributability satisfies Theorem 5: a subvector of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ selectively attributable to (or influenced by) a partition $\{\Gamma_1, ..., \Gamma_n\}$ is selectively attributable to (influenced by) the corresponding subpartition (a detailed formulation coincides with that of Theorem 5 and is therefore omitted). The proof of this statement (Theorem 5 under Definition 1*) is given in the Appendix.

2.9. Transformed Copular Bases and Distribution Functions

I conclude this section with a simple technical observation that proves to be useful in applications. The extraction from a random vector $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ of its copular base $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ plays an important role in the present theory. It is often more convenient, however, to use instead of $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ a *transformed copular base* $\{T(\mathbf{U}_1), ..., T(\mathbf{U}_n)\}(\Phi)$, with T being some strictly monotonic function. Clearly, this substitution (an example of which is given in the next section) should not change the theory in any nontrivial way. Returning, for simplicity, to the case of regularly n-dimensional random vectors and presenting the principal formula of the theory, (17), as

$$C_{i|}[p_1, ..., p_{i-1}, \xi_i(p_1, ..., p_n; \Gamma_i), p_{i+1}, ..., p_n; \Phi_i^0] = \updownarrow C_{i|}(p_1, ..., p_n; \Phi^0),$$
$$i = 1, ..., n, \qquad (33)$$

one can always rewrite it as

$$G_{i|}[t_1, ..., t_{i-1}, \tau_i(t_1, ..., t_n; \Gamma_i), t_{i+1}, ..., t_n; \Phi_i^0] = \updownarrow G_{i|}(t_1, ..., t_n; \Phi^0),$$
$$i = 1, ..., n, \qquad (34)$$

where

$$G_{i|}(t_1, ..., t_n; \Phi) = C_{i|}[T^{-1}(t_1), ..., T^{-1}(t_n); \Phi],$$
$$\tau_i(t_1, ..., t_n; \Gamma_i) = \xi_i[T^{-1}(t_1), ..., T^{-1}(t_n); \Gamma_i].$$

It is also easy to see that the validity of (34) does not change if both its sides are subjected to any strictly monotonic transformation, R, that may or may not coincide with T. It is especially convenient to choose R so that

$$R(1-u) = -R(u).$$

Examples are $R(u) = 2u - 1$ or $R(u) = Z^{-1}(u)$, where Z is the standard normal integral. Under such a transformation, (34) acquires a more conventional form

$$H_{i|}[t_1, ..., t_{i-1}, \tau_i(t_1, ..., t_n; \Gamma_i), t_{i+1}, ..., t_n; \Phi_i^0]$$
$$= \pm H_{i|}(t_1, ..., t_n; \Phi^0), \qquad i = 1, ..., n, \qquad (35)$$

where $H_{i|}$ stands for the composition $R \circ G_{i|}$, and the choice of the sign only depends on Γ_i.

3. CONCLUSION

I forgo summarizing the theory, as this is adequately done in the abstract and in Subsection 1.3. Instead, this concluding section provides an illustration of how the theory applies to the problem with which this paper opens: What is the meaning of saying that $\mathbf{X}_1$ is selectively influenced b γ_1 and $\mathbf{X}_2$ is selectively influenced by γ_2, when $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$ is bivariate-normally distributed with parameters, $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$? Using the terminology of Section 2, the question is about $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$ being selectively influenced by the partition $\{\Gamma_1 = \{\gamma_1\}, \Gamma_2 = \{\gamma_2\}\}$ of the factor set $\Phi = \{\gamma_1, \gamma_2\}$. Clearly, all the regularity conditions stipulated in this paper are satisfied: the marginal distribution functions are continuous and the copular base of $\{\mathbf{X}_1, \mathbf{X}_2\}\{\gamma_1, \gamma_2)$ has a continuous density on a convex two-dimensional region (in this case, the entire standard unit square). According to Definition 2, we begin by requiring that the marginal distributions of $\mathbf{X}_1$ and $\mathbf{X}_2$ be effectively influenced by γ_1 and γ_2, respectively. This means that γ_1 is a nondummy argument in at least one of the two equations

$$\mu_1 = \mu_1(\gamma_1), \qquad \sigma_1^2 = \sigma_1^2(\gamma_1), \tag{36}$$

while γ_2 is a nondummy argument in at least one of the two equations

$$\mu_2 = \mu_2(\gamma_2), \qquad \sigma_2^2 = \sigma_2^2(\gamma_2). \tag{37}$$

We proceed now to establishing the conditions under which $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$ is selectively attributable to $\{\Gamma_1 = \{\gamma_1\}, \Gamma_2 = \{\gamma_2\}\}$. Obviously, these conditions should only relate to the value of the correlation ρ at different values of the factors. It is convenient here to invoke the consideration presented in Subsection 2.9 and to replace the copular base $\{\mathbf{U}_1, \mathbf{U}_2\}(\gamma_1, \gamma_2)$ with its transformation

$$\{\mathbf{T}_1, \mathbf{T}_2\}(\gamma_1, \gamma_2) = \{Z^{-1}(\mathbf{U}_1), Z^{-1}(\mathbf{U}_2)\}(\gamma_1, \gamma_2),$$

where Z is the standard normal integral. This is equivalent to the conventional transformation of $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$ into

$$\{\mathbf{T}_1, \mathbf{T}_2\}(\gamma_1, \gamma_2) = \left\{\frac{\mathbf{X}_1 - \mu_1(\gamma_1)}{\sigma_1(\gamma_1)}, \frac{\mathbf{X}_2 - \mu_2(\gamma_2)}{\sigma_2(\gamma_1)}\right\}(\gamma_1, \gamma_2).$$

The random vector $\{\mathbf{T}_1, \mathbf{T}_2\}(\gamma_1, \gamma_2)$ is bivariate-normally distributed with standard normal marginals and the correlation $\rho = \rho(\gamma_1, \gamma_2)$. Using the fact that the conditional distribution of $\mathbf{T}_1$ given $\mathbf{T}_2 = t_2$ is normal with the mean ρt_2 and variance $1 - \rho^2$ (and analogously for $\mathbf{T}_2$ given $\mathbf{T}_1 = t_1$), one can write (35) with $R \equiv Z^{-1}$ in the form

$$\frac{\tau_1(t_1, t_2; \gamma_1) - \rho(\gamma_1, \gamma_2^0)\, p_2}{\sqrt{1 - \rho^2(\gamma_1, \gamma_2^0)}} = \pm \frac{p_1 - \rho(\gamma_1^0, \gamma_2^0)\, p_2}{\sqrt{1 - \rho^2(\gamma_1^0, \gamma_2^0)}},$$

$$\frac{\tau_2(t_1, t_2; \gamma_2) - \rho(\gamma_1^0, \gamma_2)\, p_1}{\sqrt{1 - \rho^2(\gamma_1^0, \gamma_2)}} = \pm \frac{p_2 - \rho(\gamma_1^0, \gamma_2^0)\, p_1}{\sqrt{1 - \rho^2(\gamma_1^0, \gamma_2^0)}},$$

where the choice of the sign in the first equation only depends on the value of γ_1, and that in the second depends only on γ_2.

This leads to the following version of selective attribution:

$$\mathbf{T}_1 = \pm \mathbf{P}_1 \frac{\sqrt{1 - \rho^2(\gamma_1, \gamma_2^0)}}{\sqrt{1 - \rho^2(\gamma_1^0, \gamma_2^0)}} + \mathbf{P}_2 \left[\rho(\gamma_1, \gamma_2^0) \mp \rho(\gamma_1^0, \gamma^0{}_2) \frac{\sqrt{1 - \rho^2(\gamma_1, \gamma_2^0)}}{\sqrt{1 - \rho^2(\gamma_1^0, \gamma_2^0)}} \right]$$

$$\mathbf{T}_2 = \pm \mathbf{P}_2 \frac{\sqrt{1 - \rho^2(\gamma_1^0, \gamma_2)}}{\sqrt{1 - \rho^2(\gamma_1^0, \gamma_2^0)}} + \mathbf{P}_1 \left[\rho(\gamma_1^0, \gamma_2) \mp \rho(\gamma_1^0, \gamma^0{}_2) \frac{\sqrt{1 - \rho^2(\gamma_1^0, \gamma_2)}}{\sqrt{1 - \rho^2(\gamma_1^0, \gamma_2^0)}} \right].$$

Since $\{\mathbf{T}_1, \mathbf{T}_2\}(\gamma_1, \gamma_2)$ is uniquely characterized by $\rho(\gamma_1, \gamma_2)$, it remains to compute this correlation from the above representation:

$$\rho(\gamma_1, \gamma_2) = \mathrm{E}(\mathbf{T}_1 \mathbf{T}_2].$$

If one denotes

$$\rho(\gamma_1, \gamma_2^0) = \cos \phi_1(\gamma_1), \quad \rho(\gamma_1^0, \gamma_2) = \cos \phi_2(\gamma_2),$$

and

$$\rho(\gamma_1^0, \gamma_2^0) = \cos \phi_1(\gamma_1^0) = \cos \phi_2(\gamma_2^0) = \cos \phi,$$

the computation yields a remarkable result,

$$\rho(\gamma_1, \gamma_2) = \cos[\phi \pm \phi_1(\gamma_1) \pm \phi_2(\gamma_2)], \tag{38}$$

where the choice of the signs at $\phi_1(\gamma_1)$ and $\phi_2(\gamma_2)$ may only depend on the corresponding factors. If the factors (γ_1, γ_2) are continuous variables, then it is reasonable to require, in addition, that $\rho(\gamma_1, \gamma_2)$ be a continuous function. One can easily verify, by considering the convergence $(\gamma_1, \gamma_2) \to (\gamma_1^0, \gamma_2^0)$, that under this requirement the only possible combinations of the signs at $\phi_1(\gamma_1)$ and $\phi_2(\gamma_2)$ are

$$\rho(\gamma_1, \gamma_2) = \begin{cases} \cos[\phi + \phi_1(\gamma_1) - \phi_2(\gamma_2)] \\ \qquad\qquad \text{or} \\ \cos[\phi - \phi_1(\gamma_1) + \phi_2(\gamma_2)] \end{cases}. \tag{39}$$

Together with (36) and (37), the formulas (38) and (39) establish a complete characterization of the selective influence effected by $\{\Gamma_1 = \{\gamma_1\}, \Gamma_2 = \{\gamma_2\}\}$ on a bivariate normally distributed $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$. If the distribution of $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$ is known on a sample level only, and if the sample statistically supports the assumption that the distribution of the random vector is bivariate normal, then the hypothesis of selective influence can be tested by comparing the fit to the data of the unconstrained bivariate normal distribution with that of the bivariate normal distribution whose parameters satisfy (36)–(37) and (38) or (39).

The solution has a surprisingly nice mathematical form, and although it satisfies all the intuitive requirements underlying the theory of selective influence, it would be difficult to foresee this solution based on one's intuition alone.

It is instructive to contrast this solution with that obtained when the same situation is being approached from the point of view of *conditionally* selective influence (see Subsection 1.1). Not only is the solution here less elegant mathematically, it also lends itself to a much less satisfactory interpretation. This is especially apparent when one considers the first of the two examples of Subsection 1.1, when $\{\mathbf{X}_1, \mathbf{X}_2\}(\gamma_1, \gamma_2)$ are two directly observable aptitude scores. Clearly, in this example nothing prevents one from focusing only on one of the two scores, say, $\mathbf{X}_1$, or from not conducting the second test altogether: this should not affect the distribution of

$\mathbf{X}_1$ if the two tests, as normally happens, are conducted separately. When taken separately, however, the distribution of $\mathbf{X}_1$ has glaringly counterintuitive properties. To begin with, the variance

$$\sigma_1^2 = \frac{c_1^2(\gamma_1)}{1 - kc_1^2(\gamma_1)\, c_2^2(\gamma_2)}$$

depends on both γ_1 and γ_2, which means that $\mathbf{X}_1$ is not selectively influenced by γ_1. Using the language of this paper, conditionally selective influence does not have the property of nestedness. Another counterintuitive feature is that the way the mean value μ_1 of $\mathbf{X}_1$ depends on the two factors is different depending on the correlation ρ between $\mathbf{X}_1$ and $\mathbf{X}_2$: if ρ is identically zero, then μ_1 may depend on γ_1, but it may not depend on either of the two factors if ρ can attain nonzero values. Put differently, if μ_1 is observed to depend on γ_1, then $\mathbf{X}_1$ cannot be conditionally selectively influenced by γ_1 if taken in combination with any score with which it may be correlated.

Returning to the unconditionally selective influence, two issues should be mentioned among those that remain to be investigated. First, it is not clear at present to what extent the pairwise disjointness of the factor subsets $\{\Gamma_1, ..., \Gamma_n\}$ is a critical limitation of the theory. Second, perhaps more importantly, it is not clear at present whether the theory of selective influence can be constructed without the requirement of the factor effectiveness with respect to marginals. It is shown in this paper that this requirements is sufficient for achieving the uniqueness of selective attribution, but there is no proof that this is also necessary. It seems to me that the theory would be more satisfying if the uniqueness could be ensured by means of a constraint formulated entirely in terms of copular bases (with the understanding, of course, that no $\mathbf{X}_i$ is allowed to depend on any factors outside its factor subset Γ_i, $i = 1, ..., n$). Then the notion of selective influence would be well-defined even when the marginal distributions do not vary at all.

APPENDIX: PROOFS

Proof of Lemma 3. Since $\{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi)$ has a continuous density with an n-dimensional convex support, by Lemma 2 $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ in (9) can always be chosen to have the same property. With this choice, the chained conditional

$$G_{i|}(p_1, ..., p_{i-1}, p_i) = \mathrm{Prob}[\mathbf{P}_i \leqslant p_i \mid \mathbf{P}_1 = p_1 \,\&\, \ldots \,\&\, \mathbf{P}_{i-1} = p_{i-1}]$$

is continuous in all arguments and increasing in its main argument (i.e., the variable being conditioned). The random vector $\{\mathbf{R}_1, ..., \mathbf{R}_n\}$, defined by

$$\begin{bmatrix} \mathbf{R}_1 = G_1(\mathbf{P}_1) \\ \mathbf{R}_2 = G_{2|}(\mathbf{P}_1, \mathbf{P}_2) \\ \cdots \\ \mathbf{R}_n = G_{n|}(\mathbf{P}_1, ..., \mathbf{P}_{n-1}, \mathbf{P}_n) \end{bmatrix},$$

consists of independent random variables uniformly distributed between 0 and 1, and $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ can be presented as

$$\begin{bmatrix} \mathbf{P}_1 = G_1^{-1}(\mathbf{R}_1) \\ \mathbf{P}_2 = G_{2|}^{-1}(\mathbf{P}_1, \mathbf{R}_2) \\ \cdots \\ \mathbf{P}_n = G_{n|}^{-1}(\mathbf{P}_1, ..., \mathbf{P}_{n-1}, \mathbf{R}_n) \end{bmatrix},$$

where the inverses of conditional distribution functions are taken with respect to the main arguments. Obviously, the relationship between $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ and $\{\mathbf{R}_1, ..., \mathbf{R}_n\}$ is homeomorphic. By a chain of nested substitutions,

$$\begin{bmatrix} \mathbf{P}_1 = G_1^{-1}(\mathbf{R}_1) = \theta_1(\mathbf{R}_1) \\ \mathbf{P}_2 = G_{2|}^{-1}(\theta_1(\mathbf{R}_1), \mathbf{R}_2) = \theta_2(\mathbf{R}_1, \mathbf{R}_2) \\ \cdots \\ \mathbf{P}_n = G_{n|}^{-1}(\theta_1(\mathbf{R}_1), ..., \theta_{n-1}(\mathbf{R}_1, ..., \mathbf{R}_{n-1}), \mathbf{R}_n) = \theta_n(\mathbf{R}_1, ..., \mathbf{R}_n) \end{bmatrix}.$$

By choosing any continuous and increasing on some interval domain distribution function $H(q)$, one can replace $\{\mathbf{R}_1, ..., \mathbf{R}_n\}$ in the representation above with $\{H(\mathbf{Q}_1), ..., H(\mathbf{Q}_n)\}$, where $\mathbf{Q}_i$ is distributed according to $H(q)$, $i = 1, ..., n$. Substituting these expressions for $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$ in (9) and renaming $\{\mathbf{Q}_1, ..., \mathbf{Q}_n\}$ into $\{\mathbf{P}_1, ..., \mathbf{P}_n\}$, we obtain the statement of the lemma.

Proof of Lemma 4. At $\Phi = \Phi^0 = \Gamma_1^0 \cup \cdots \cup \Gamma_n^0$ the statement (i) is true for any i, by construction. Since $\xi_i(p_1, ..., p_n; \Gamma_i^0)$ does not depend on factors outside Γ_i, the statement must hold irrespective of the values of the other factor subsets.

Putting now $\Gamma_j = \Gamma_j^0$ for all $j \neq i$ and allowing Γ_i to vary freely, we have

$$\begin{bmatrix} u_1 = \xi_1(p_1, ..., p_n; \Gamma_1^0) = p_1 \\ \cdots \\ u_i = \xi_n(p_1, ..., p_i, ..., p_n; \Gamma_i) \\ \cdots \\ u_n = \xi_n(p_1, ..., p_n; \Gamma_n^0) = p_n \end{bmatrix}.$$

Since $\{\xi_1, ..., \xi_n\}_\Phi$ is a homeomorphism for any Φ, the relationship between u_i and p_i must be continuous and one-to-one. By agreement, the domain of $\{p_1, ..., p_n\}$ is a convex region of the standard unit cube, because of which the domain of p_i, for fixed values of other arguments, is a certain interval within $[0, 1]$. Then $\xi_n(p_1, ..., p_i, ..., p_n; \Gamma_i)$ must be either strictly increasing or strictly decreasing in p_i.

Proof of Theorem 2. By Lemma 4, at $\Phi = \Phi_i^0$ we have

$$\begin{bmatrix} \mathbf{U}_1 = \xi_1(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_1^0) = \mathbf{P}_1 \\ \cdots \\ \mathbf{U}_i = \xi_i(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_i) \\ \cdots \\ \mathbf{U}_n = \xi_n(\mathbf{P}_1, ..., \mathbf{P}_n; \Gamma_n^0) = \mathbf{P}_n \end{bmatrix},$$

where $\xi_i(p_1, ..., p_n; \Gamma_i)$ is either increasing or decreasing in p_i. It the former case

$$\begin{aligned} &C_{i|}[p_1, ..., p_{i-1}, \xi_i(p_1, ..., p_n; \Gamma_i), p_{i+1}, ..., p_n; \Phi_i^0] \\ &\quad = \text{Prob}\left[\mathbf{U}_i \leqslant \xi_i(p_1, ..., p_n; \Gamma_i) \,\middle|\, \mathop{\&}_{j \in \{1, ..., n\} - \{i\}} \mathbf{U}_j = p_j; \Phi_i^0 \right] \\ &\quad = \text{Prob}\left[\mathbf{P}_i \leqslant p_i \,\middle|\, \mathop{\&}_{j \in \{1, ..., n\} - \{i\}} \mathbf{P}_j = p_j; \Phi_i^0 \right] \\ &\quad = C_{i|}(p_1, ..., p_{i-1}, p_i, p_{i+1}, ..., p_n; \Phi^0), \end{aligned}$$

while for the decreasing relationship

$$\begin{aligned} &C_{i|}[p_1, ..., p_{i-1}, \xi_i(p_1, ..., p_n; \Gamma_i), p_{i+1}, ..., p_n; \Phi_i^0] \\ &\quad = \text{Prob}\left[\mathbf{U}_i \leqslant \xi_i(p_1, ..., p_n; \Gamma_i) \,\middle|\, \mathop{\&}_{j \in \{1, ..., n\} - \{i\}} \mathbf{U}_j = p_j; \Phi_i^0 \right] \\ &\quad = \text{Prob}\left[\mathbf{P}_i \geqslant p_i \,\middle|\, \mathop{\&}_{j \in \{1, ..., n\} - \{i\}} \mathbf{P}_j = p_j; \Phi_i^0 \right] \\ &\quad = 1 - C_{i|}(p_1, ..., p_{i-1}, p_i, p_{i+1}, ..., p_n; \Phi^0). \end{aligned}$$

Since the conditional distribution functions are continuous in all arguments, the choice between these two possibilities cannot depend on the values of $\{p_1, ..., p_n\}$, but it may depend on Φ_i^0 whose only varying part is Γ_i. The statement of the theorem now follows immediately.

Proof of Theorem 5. It is sufficient to prove the theorem for the subvector $\{\mathbf{X}_2, ..., \mathbf{X}_n\}$. Rewrite (18) as

$$\begin{bmatrix} C_{1|}[\mathbf{U}_1, \mathbf{P}_2, ..., \mathbf{P}_n; \Phi_1^0] = \updownarrow C_{1|}(\mathbf{P}_1, ..., \mathbf{P}_n; \Phi^0) \\ \cdots \\ C_{i|}[\mathbf{P}_1, ..., \mathbf{P}_{i-1}, \mathbf{U}_i, \mathbf{P}_{i+1}, ..., \mathbf{P}_n; \Phi_i^0] = \updownarrow C_{i|}(\mathbf{P}_1, ..., \mathbf{P}_n; \Phi^0) \\ \cdots \\ C_{n|}[\mathbf{P}_1, ..., \mathbf{P}_{n-1}, \mathbf{U}_n; \Phi_n^0] = \updownarrow C_{n|}(\mathbf{P}_1, ..., \mathbf{P}_n; \Phi^0) \end{bmatrix}.$$

Excluding the first equation and putting $\mathbf{P}_1 = p$, we have

$$\left[\begin{array}{c} C_{2|}[p, \mathbf{U}_2, ..., \mathbf{P}_n \colon \Phi_1^0] = \updownarrow C_{1|}(p, \mathbf{P}_2, ..., \mathbf{P}_n; \Phi^0) \\ \cdots \\ C_{i|}[p, \mathbf{P}_2, ..., \mathbf{P}_{i-1}, \mathbf{U}_i, \mathbf{P}_{i+1}, ..., \mathbf{P}_n; \Phi_i^0] = \updownarrow C_{i|}(p, \mathbf{P}_2, ..., \mathbf{P}_n; \Phi^0) \\ \cdots \\ C_{n|}[p, \mathbf{P}_2, ..., \mathbf{P}_{n-1}, \mathbf{U}_n; \Phi_n^0] = \updownarrow C_{n|}(p, \mathbf{P}_2, ..., \mathbf{P}_n; \Phi^0) \end{array}\right].$$

It follows that

$$\int_0^1 C_{i|}[p, \mathbf{P}_2, ..., \mathbf{P}_{i-1}, \mathbf{U}_i, \mathbf{P}_{i+1}, ..., \mathbf{P}_n; \Phi_i^0]\, dp = \updownarrow \int_0^1 C_{i|}(p, \mathbf{P}_2, ..., \mathbf{P}_n; \Phi^0)\, dp,$$
$$i = 2, ..., n. \qquad (*)$$

On observing that $\mathbf{P}_1$ is distributed uniformly between 0 and 1, and that, due to Lemma 4, $\{\mathbf{P}_1, ..., \mathbf{P}_{i-1}, \mathbf{U}_i, \mathbf{P}_{i+1}, ..., \mathbf{P}_n\} = \{\mathbf{U}_1, ..., \mathbf{U}_n\}(\Phi_i^0)$, we have

$$\begin{aligned} &\int_0^1 C_{i|}[p, p_2, ..., p_{i-1}, u_i, p_{i+1}, ..., p_n; \Phi_i^0]\, dp \\ &\quad = \mathrm{Prob}\left[\mathbf{U}_i \leqslant u_i \,\middle|\, \mathop{\&}_{j \in \{2, ..., n\} - \{i\}} \mathbf{P}_j = p_j\right] \\ &\quad = \bar{C}_{i|}(p_2, ..., p_{i-1}, u_i, p_{i+1}, ..., p_n; \Phi_i^0), \qquad i = 2, ..., n. \end{aligned}$$

Analogously,

$$\begin{aligned} &\int_0^1 C_{i|}(p, p_2, ..., p_n; \Phi^0)\, dp \\ &\quad = \mathrm{Prob}\left[\mathbf{P}_i \leqslant p_i \mid \mathop{\&}_{j \in \{2, ..., n\} - \{i\}} \mathbf{P}_j = p_j\right] \\ &\quad = \bar{C}_{i|}(p_2, ..., p_n; \Phi^0), \qquad i = 2, ..., n. \end{aligned}$$

Equation $(*)$ therefore is equivalent to

$$\bar{C}_{i|}(\mathbf{P}_2, ..., \mathbf{P}_{i-1}, \mathbf{U}_i, \mathbf{P}_{i+1}, ..., \mathbf{P}_n; \Phi_i^0) = \updownarrow \bar{C}_{i|}(\mathbf{P}_2, ..., \mathbf{P}_n; \Phi^0), \qquad i = 2, ..., n.$$

Solving with respect to $\mathbf{U}_i$, we get the representation

$$\begin{aligned} \mathbf{U}_i &= \bar{C}_{i|}^{-1}[\mathbf{P}_2, ..., \mathbf{P}_{i-1}, \updownarrow \bar{C}_{i|}(\mathbf{P}_2, ..., \mathbf{P}_n; \Phi^0), \mathbf{P}_{i+1}, ..., \mathbf{P}_n; \Phi_i^0] \\ &= \zeta_i(\mathbf{P}_2, ..., \mathbf{P}_n; \Gamma_i), \qquad i = 2, ..., n. \end{aligned}$$

This means that $\{\mathbf{U}_2, ..., \mathbf{U}_n\}$ is selectively attributable to $\{\Gamma_2, ..., \Gamma_n\}$. Obviously, $\{\Gamma_2, ..., \Gamma_n\}$ is a partition of $\Phi' = \Phi - \Gamma_1$ and $\{\mathbf{U}_2, ..., \mathbf{U}_n\} = \{\mathbf{U}_2, ..., \mathbf{U}_n\}(\Phi')$.

Proof of Theorem 5 *under Definition* 1*. Recall that $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ is so arranged that $\{\mathbf{U}_1, ..., \mathbf{U}_m\}(\Phi)$ is one of its regular copular subbases. It is sufficient to prove the theorem for the subvector obtained from $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ by removing one of its components, $\mathbf{X}_i$. The statement of the theorem is trivially true if $\mathbf{U}_i \in \{\mathbf{U}_{m+1}, ..., \mathbf{U}_n\}$. Assume therefore that $\mathbf{U}_i \in \{\mathbf{U}_1, ..., \mathbf{U}_m\}$, say, $\mathbf{U}_i = \mathbf{U}_1$. If at least one regular copular subbase of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ does not contain $\mathbf{U}_1$, then the vector $\{\mathbf{X}_2, ..., \mathbf{X}_n\}(\Phi')$, $\Phi' = \Phi - \Gamma_1$, is regularly m-dimensional. In this case we simply delete the first equation from (30) to obtain

$$\begin{bmatrix} \mathbf{U}_2 = C_{2|}^{-1}[\mathbf{P}_1, \updownarrow C_{2|}(\mathbf{P}_1, ..., \mathbf{P}_m; \Phi^0), ..., \mathbf{P}_m; \Phi_1^0] \\ \cdots \\ \mathbf{U}_m = C_{m|}^{-1}[\mathbf{P}_1, ..., \mathbf{P}_{m-1}, \updownarrow C_{m|}(\mathbf{P}_1, ..., \mathbf{P}_m; \Phi^0); \Phi_m^0 \\ \mathbf{U}_{m+1} = \omega_{m+1}[\mathbf{P}_1, ..., \mathbf{P}_m; \Phi_{m+1}^0] \\ \cdots \\ \mathbf{U}_n = \omega_n[\mathbf{P}_1, ..., \mathbf{P}_m; \Phi_n^0] \end{bmatrix}.$$

By Lemma 6, this representation is homeomorphic, hence it satisfies Definition 1*, and the statement of the theorem holds. It remains to consider the case when all regular copular subbases of $\{\mathbf{X}_1, ..., \mathbf{X}_n\}(\Phi)$ contain $\mathbf{U}_1$. Then none of the random variables $\{\mathbf{U}_{m+1}, ..., \mathbf{U}_n\}$ is a function of $\mathbf{U}_1$, for any Φ. Indeed, if $\mathbf{U}_i$, $i = m+1, ..., n$, were a function of $\mathbf{U}_1$, then it would have been a continuously differentiable function of $\{\mathbf{U}_1, ..., \mathbf{U}_m\}$, and the subvector $\{\mathbf{U}_2, ..., \mathbf{U}_m, \mathbf{U}_i\}$ would have been a regular copular subbase (for it would have possessed a continuous density), contrary to the assumption that all regular copular subbases contain $\mathbf{U}_1$. We have therefore

$$\mathbf{U}_i = \omega_i(\mathbf{U}_2, ..., \mathbf{U}_m; \Phi), \qquad i = m+1, ..., n,$$

where all the functions are continuously differentiable. The proof of Theorem 5 (under Definition 1) can now be applied to $\{\mathbf{U}_2, ..., \mathbf{U}_m\}(\Phi')$ to transform

$$\mathbf{U}_i = \xi_i(\mathbf{P}_1, \mathbf{P}_2, ..., \mathbf{P}_n; \Gamma_i), \qquad i = 2, ..., m,$$

into a homeomorphism

$$\mathbf{U}_i = \zeta_i(\mathbf{P}_2, ..., \mathbf{P}_m; \Gamma_i), \qquad i = 2, ..., m.$$

This completes the proof.

REFERENCES

Colonius, H. (1990). Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology*, **34**, 253–275.

Cortese, J. M., & Dzhafarov, E. N. (1996). Empirical recovery of response time decomposition rules. II. Discriminability of serial and parallel architectures. *Journal of Mathematical Psychology*, **40**, 203–218.

Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology*, **36**, 235–268.

Dzhafarov, E. N. (1997). Process representations and decompositions of response times. In A. A. J. Marley (Ed.), *Choice, decision and measurement: Essays in honor of R. Duncan Luce* (pp. 255–278). Mahwah, NJ: Erlbaum.

Dzhafarov, E. N. (1999). Conditionally selective dependence of random variables on external factors. *Journal of Mathematical Psychology*, **43**, 123–157.

Dzhafarov, E. N., & Cortese, J. M. (1996). Empirical recovery of response time decomposition rules. I. Sample-level decomposition tests. *Journal of Mathematical Psychology*, **40**, 185–202.

Dzhafarov, E. N., & Schweickert, R. (1995). Decompositions of response times: An almost general theory. *Journal of Mathematical Psychology*, **39**, 285–314.

Nelsen, R. B. (1999). *An introduction to copulas* (Lecture Notes in Statistics, Vol. 139). New York: Springer.

Rüschendorf, L., Schweizer, B., & Taylor, M. D. (1996). *Distributions with fixed marginals and related topics*. (Lecture Notes Monograph Series, Vol. 28). Hayward, CA: Institute of Mathematical Statistics.

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology*, **28**, 363–400.

Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, **33**, 309–327.

Townsend, J. T., & Thomas, R. D. (1994). Stochastic dependencies in parallel and serial models: Effects on systems factorial interactions. *Journal of Mathematical Psychology*, **38**, 1–34.

Received: May 17, 1999; published online April 4, 2001