

Purdue University Purdue e-Pubs

Department of Psychological Sciences Faculty
Publications

Department of Psychological Sciences

2010

Additive factors and stages of mental processes in task networks.

Richard Schweickert

Purdue University, schweick@purdue.edu

Donald L. Fisher

William M. Goldstein

Follow this and additional works at: <http://docs.lib.purdue.edu/psychpubs>



Part of the [Psychology Commons](#)

Recommended Citation

Schweickert, Richard; Fisher, Donald L.; and Goldstein, William M., "Additive factors and stages of mental processes in task networks." (2010). *Department of Psychological Sciences Faculty Publications*. Paper 33.
<http://dx.doi.org/10.1016/j.jmp.2010.06.004>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Additive Factors and Stages of Mental Processes in Task Networks

Richard Schweickert

Purdue University

Donald L. Fisher

University of Massachusetts

William M. Goldstein

University of Chicago

Running Head: Stages of Mental Processes

Submitted to *Journal of Mathematical Psychology* July 2010

Abstract. To perform a task a subject executes mental processes. An experimental manipulation, such as a change in stimulus intensity, is said to selectively influence a process if it changes the duration of that process leaving other process durations unchanged. For random process durations a definition of a factor selectively influencing a process by increments is given in terms of stochastic dominance (also called “the usual stochastic order”). A technique for analyzing reaction times, Sternberg's Additive Factor Method, assumes all the processes are in series. When all processes are in series, each process is called a stage. With the Additive Factor Method, if two experimental factors selectively influence two different stages by increments, the factors will have additive effects on reaction time. An assumption of the Additive Factor Method is that if two experimental factors interact, then they influence the same stage. We consider sets of processes in which some pairs of processes are sequential and some are concurrent (i. e., the processes are partially ordered). We propose a natural definition of a stage for such sets of processes. For partially ordered processes, with our definition of a stage, if two experimental factors selectively influence two different processes by increments, each within a different stage, then the factors have additive effects. If each process selectively influenced by increments is in the same stage, then an interaction is possible, although not inevitable.

In many situations it is natural to assume that one mental process follows another, for example, that a response movement follows response selection. Such serial processing is assumed in the classic paper of Donders (1868), and in modern work by Sternberg (1969), Ratcliff (1978) and others. The popularity of the assumption is largely due to the elegance of the test for it, Sternberg's (1969) Additive Factor Method. See Thomas (2006) for recent applications.

To use the method to analyze a task, the experimenter manipulates two factors, such as stimulus intensity and response movement difficulty. Suppose the task is carried out with a series of processes, say, perception, followed by response selection, followed by movement. When all processes are in series, each process is called a *stage*. Suppose the response time is the sum of the durations of the stages. When the stimulus is made less intense, suppose the duration of perception increases but durations of other stages are unchanged. When response movement difficulty is increased, suppose movement duration increases, but durations of other stages are unchanged. Then the combined effect on response time of making the stimulus less intense and of increasing movement difficulty will be the sum of their individual effects. Factors with additive effects on response time are called *additive factors*; otherwise they are said to interact.

A factor that changes the duration of a single stage, leaving durations of other stages unchanged, is said to *selectively influence* the stage. A deduction from the assumptions is that if each of two factors selectively influences a different stage, the factors will be additive. Hence, additive factors found in an experiment provide empirical support for the assumption that the factors selectively influence different stages. With the Additive Factor Method data are interpreted as follows (Sternberg, 1969). If two factors are additive, then they influence different stages in a series of stages, and if two factors interact then they influence the same stage (and there may or may not be other stages). The conclusions are plausible, but not deductions from the assumptions. That is, additivity of two factors can be explained by saying they selectively influence two different stages, and an interaction can be

explained by saying they selectively influence the same stage, but the explanations are not logical implications.

Here we consider additivity of factors when both sequential and concurrent processes are executed. In dual tasks it is often assumed that some processes, such as peripheral processes, are concurrent while others, such as central processes, are sequential (Davis, 1957; Ehrenstein, Schweickert, Choi, Proctor, 1997; Meyer & Kieras, 1997a, 1997b; Pashler & Johnston, 1989; Welford, 1959, 1967). For example, a model in which one process is concurrent with two sequential processes is in Figure 1; each arc represents a process. This model was used by Schweickert, Fortin and Sung (2007) for a dual time production and visual search task. In the model, two experimental factors selectively influencing the two sequential processes may be additive or not, depending on whether the process concurrent with them is short or long (Schweickert & Townsend, 1989).

The model in Figure 1 is a directed acyclic network. Models in which all processes are in series or all are in parallel (e. g., Townsend, 1972) are special cases of directed acyclic networks. Directed acyclic networks are especially useful for models with both sequential and concurrent processing. These include the dual task models mentioned above, as well as models for detection (Townsend & Nozawa, 1995), for search (Sung, 2008; Van Zandt & Townsend, 1993), for stimulus-response compatibility (Kornblum, Hasbroucq & Osman, 1990), and for aging (Fisher & Glaser, 1996). An example in human factors is the model for a telephone operator task by Gray, John & Atwood, 1993; see also Schweickert, Fisher & Proctor, 2003. Directed acyclic network models of tasks are themselves a special case of Order-of-Processing Diagrams, which are useful for deriving expressions for probability distribution functions and their moments (Fisher & Goldstein, 1983; Fisher, Saisi & Goldstein, 1985; Goldstein & Fisher, 1991, 1992).

A *directed network* consists of vertices and, for some pairs of vertices, an arc directed from one vertex to the other. A *starting vertex of a network* has no arc directed into it. In a cognitive task, processing begins with presentation of a stimulus at a starting vertex of the network. A mental process is represented by an arc directed from one vertex to another. The *starting vertex of the arc*, at the tail of the arc, represents the starting point of the process. The *ending vertex of the arc*, at the head of the arc, represents the finishing point of the process. An *ending vertex of the network* is a vertex with no arc directed from it. Responses are made at the ending vertices of the network. We will assume the network has a single starting vertex and a single ending vertex.

A *path* from a vertex u to a vertex z consists of the vertex u , followed by an arc directed from u to a vertex v , followed by an arc directed from v to a vertex w , and so on, with the last arc having ending vertex z . We consider a single vertex to be a path. A *nontrivial path* has at least one arc. To indicate that one process immediately precedes another, the head of the arc representing the first process is incident with the tail of the arc representing the second. If one process precedes another (not necessarily immediately), there will be a path from the head of the arc representing the first process to the tail of the arc representing the second; the path goes along arcs in the direction indicated by the arrows. A vertex preceding a process, and so on, are defined similarly, with the requirement that for one vertex to precede another there must be a nontrivial path from one to the other.

A *cycle* is a nontrivial path that goes from a vertex u to the same vertex u . An *acyclic network* has no cycles, so a vertex or process does not precede itself. Thus, precedence is *irreflexive*. We also assume precedence is *transitive*, that is, if process x precedes process y , and process y precedes process z , then x precedes z . Two processes are *sequential* if one precedes the other; otherwise they are *concurrent*. (Note that concurrent processes are not necessarily executed simultaneously in time.) Arcs in a directed acyclic network are partially ordered by precedence; that is, precedence is irreflexive and transitive.

A process whose starting vertex is at an *AND gate* begins execution as soon as *all* processes immediately preceding it are finished. A process whose starting vertex is an *OR gate* begins execution as soon as *any* process immediately preceding it is finished. Some processes have their ending vertex at a response. If the response is made as soon as *all* these processes are finished, the response is at an AND gate, and if the response is made as soon as *any* of these processes are finished, the response is at an OR gate. Here we assume that except for the starting vertex, every vertex is an AND gate or every vertex is an OR gate. In the former case the network is called an *AND network* and the latter case an *OR network*. In scheduling theory, AND networks are known as *PERT* (Program Evaluation and Review Technique) networks or *critical path* networks (Kelley & Walker, 1959; Malcom, Roseboom, Clark & Fazar, 1959).

Directed Acyclic Task Networks: The Durations

Associated with every arc z in the network is a nonnegative random variable, $D(z)$, the duration of the process the arc represents. A value the random variable $D(z)$ can take on is denoted $d(z)$. The duration of a path is the sum of the durations of all the arcs on it. The duration of a vertex is 0. The network with the durations omitted is called a *graph*.

On a particular trial, we assume that values for the durations of the processes are chosen according to the joint probability distribution of the process durations. In an AND network, all the processes must be completed for the task to be completed. The time required to complete the task is the time required to complete every process on the longest path from the starting vertex of the network to the ending vertex. This path is called the *critical path*, and the response time is the duration of the critical path. In an AND network, the duration of the longest path from a vertex u to a vertex v on a particular trial is denoted $d(u,v)$. When the notation $d(u,v)$ is used here, vertex u always precedes vertex v . The path which is the longest path from u to v may vary from trial to trial. In the population

of trials, the duration of the longest path from u to v is a random variable denoted $D(u,v)$. Note when the symbol d or D is used to denote a function of a single argument, the argument is an arc, and when it is used to denote a function of two arguments, the arguments are vertices, with the first vertex listed preceding the second.

Townsend and Nozawa (1995) have shown that both AND networks and OR networks are useful as models of mental processes. In an OR network, the time required to complete the task, the response time, is the time required to complete every process on the shortest path from the starting vertex to the ending vertex of the network. We use the same notation for OR networks and AND networks, with different terminology. In an OR network, the duration of the shortest path from a vertex u to a vertex v on a particular trial is denoted $d(u,v)$. In the population of trials, the duration of the shortest path from u to v is a random variable, denoted $D(u,v)$.

Effects of Factors Selectively Influencing Processes

If all the processes are sequential, we say the processes are in *series*. In that case if a certain process is prolonged by some amount the response time will increase by that amount. The situation is more complicated when concurrent processes are present, but their presence makes effects of factors potentially very informative about the form of the network. In this section we summarize some results relevant here from earlier papers on Task Network Inference, the construction, from observed effects of factors, of a directed acyclic network in which the factors selectively influence processes (Schweickert, 1978, 1992; Schweickert & Wang, 1993).

In the beginning of the paper, processes with fixed durations are considered and a simple notion of what it means for a factor to selectively influence a process will suffice. Later in the paper, random process durations are considered. A definition of selective influence by Dzhafarov and his colleagues

is discussed. The notion of selective influence from the beginning of the paper is generalized to the notion of selective influence by increments. Complexities easily arise (Townsend & Thomas, 1994), and for further discussion see Townsend and Schweickert (1989), Dzhafarov (2003), Dzhafarov and Gluhovsky (2006), Dzhafarov and Kujala (2010), Dzhafarov, Schweickert & Sung (2004), and Kujala and Dzhafarov (2008).

Selective influence of a single process. Suppose the duration of every process z in the network is assigned a real number value $d(z) \geq 0$. Now suppose the duration of a particular process x is prolonged by an amount $a \geq 0$. Consider an AND network; results are analogous for an OR network. If x is on the critical path, the response time increases by a . If x is not on the critical path, then there is some amount of time by which x can be prolonged without increasing the response time. The largest amount of time by which x can be prolonged without increasing the response time is called the *total slack* for x . Let r denote the ending vertex of the network (typically where the response is made). The total slack for x is denoted $s(x,r)$. Suppose process x is prolonged by amount a . If $a \leq s(x,r)$ there is no increase in response time. If $a > s(x,r)$ the increase in response time is $a - s(x,r)$. Let $[w]^+ = 0$ if $w \leq 0$, and let $[w]^+ = w$ if $w > 0$. Then the increase in response time is $[a - s(x,r)]^+$.

Selective influence of two processes. Suppose process x precedes process y . Again, consider an AND network. The combined effect of prolonging both x and y is complicated by the fact that prolonging x may make y start late, and the further effect on response time of prolonging y depends on whether it starts late or not. Let y' denote the starting vertex of y . The largest amount of time by which x can be prolonged without delaying the time at which y begins is the slack from x to y' , denoted $s(x,y')$. More detail is in Schweickert (1978) for AND networks and in Schweickert and Wang (1993) for OR networks.

Here, we usually need only consider process durations as fixed nonnegative numbers, and consider a factor selectively influencing a process as increasing its duration by a fixed nonnegative

number. Suppose a factor selectively influences process x and another factor selectively influences process y . Suppose when both factors are at level 1, each process z has a duration $d(z)$. Slacks for these values of the process durations can be calculated as follows (Schweickert, 1978). Let o denote the starting vertex of the network (typically stimulus onset). Let z' and z'' denote, respectively, the starting and ending vertices of a process z . Then

$$s(x,r) = d(o,r) - d(o,x') - d(x) - d(x'',r), \text{ and}$$

$$s(x,y') = d(o,y') - d(o,x') - d(x) - d(x'',y').$$

The first equation says the total slack for x is the duration of the longest path from the starting vertex, o , of the network to the ending vertex, r , minus the duration of the longest such path going through x .

The second equation is obtained as follows. Remove all vertices and edges that follow y' . Then y' is the ending vertex of the network. When the first equation is applied to the remaining network, the result is the second equation.

Another useful quantity is obtained by reversing the directions of all the arcs in the network, keeping the arc durations the same. The resulting network is called the *converse network*. If x precedes y in the original network, then y precedes x in the converse network. In the converse network, let the slack from y to x be $s^*(y,x'')$. Then by applying the second equation above, an expression for $s^*(y,x'')$ can be found in terms of path durations in the original network,

$$s^*(y,x'') = d(x'',r) - d(x'',y') - d(y) - d(y'',r).$$

Define the *coupled slack* between x and y in an AND network as

$$k(x,y) = s(x,r) - s(x,y') = s(y,r) - s^*(y,x'') = d(o,r) - d(o,y') - d(x'',r) + d(x'',y').$$

In an OR network, we use the same notation for an analogous term. In an OR network, we define the *coupled slack* between x and y , with x preceding y , as

$$k(x,y) = d(o,r) - d(o,y') - d(x'',r) + d(x'',y').$$

(The symbols are the same as for an AND network, but recall that in an OR network, $d(u,v)$ denotes the

duration of the shortest path between vertices u and v .)

The following notation is the same for both AND and OR networks. Let t_{11} denote the response time when the factor selectively influencing process x and the factor selectively influencing process y are both at level 1; then $t_{11} = d(o,r)$. When the factor selectively influencing process x is at level 2, and the factor selectively influencing y is at level 1, suppose the duration of process x is prolonged by amount a , to become $d(x) + a$, but all other process durations are unchanged. Denote the new response time as t_{21} .

When the factor selectively influencing process y is at level 2, and the factor selectively influencing x is at level 1, suppose the duration of process x is returned to $d(x)$, as before, and suppose the duration of process y is prolonged by amount b , to become $d(y) + b$, with all other process durations unchanged. Denote the new response time as t_{12} . Finally, when both factors are at level 2, suppose x is prolonged by a and y is prolonged by b , with the duration of every other process z unchanged from its original value $d(z)$. Denote the new response time as t_{22} .

Suppose in addition to processes x and y , there are n other processes, z_1, \dots, z_n . The interaction contrast $t_{22} - t_{21} - t_{12} + t_{11}$ is a function of the process durations $d(x), d(y), d(z_1), \dots, d(z_n)$ and prolongations a and b . (Note that $t_{22} - t_{21} = t_{12} - t_{11}$ if there is no interaction, i.e., $t_{22} - t_{21} - t_{12} + t_{11} = 0$). Call the interaction contrast function h , that is,

$$t_{22} - t_{21} - t_{12} + t_{11} = h(a,b,d(x),d(y),d(z_1), \dots, d(z_n)).$$

In Schweickert (1978) it is shown that for an AND network,

$$h(a,b,d(x),d(y),d(z_1), \dots, d(z_n)) = [b - s(y,r) - [a - s(x,r)]^+ + [a - s(x,y')]^+]^+ - [b - s(y,r)]^+. \quad (1)$$

The above expression takes a simple form when $a > s(x,r)$, $s(x,y')$ and $b > s(y,r)$, $s^*(y,x'')$. The combined effect of prolonging x by a and y by b is then equal to the coupled slack; that is, the interaction contrast is

$$t_{22} - t_{21} - t_{12} + t_{11} = s(x,r) - s(x,y') = k(x,y),$$

(Schweickert, 1978). The same expression holds for OR networks (Schweickert & Wang, 1993).

A factor selectively influencing x and a factor selectively influencing y with the parameters just described will be additive if the coupled slack $k(x,y) = 0$. With AND and OR networks, it is possible that $k(x,y)$ is negative, zero or positive. With an AND network, a negative value can only occur if the network contains a subnetwork in the form of a Wheatstone bridge (Schweickert, 1978). Likewise, with an OR network, a positive value of $k(x,y)$ can only occur if the network contains a subnetwork in the form of a Wheatstone bridge (Schweickert & Wang, 1993). (For illustration, in Figure 2 vertex c and the part of the network to the left of vertex c form a Wheatstone bridge.)

Additive Factors and Stages

After these preliminary remarks, we take up additivity of factors selectively influencing processes in AND or OR networks. We give a natural definition of stages in such networks; these stages are sets of processes. Each stage can be considered a superprocess, and the superprocesses are in series. The superprocesses in series have properties analogous to processes in series.

Sometimes factors selectively influencing processes x and y will have, or not have, additive effects depending on the duration of some other process. See, for example, processes x and y in the AND network of Figure 1. With a short duration of process z , prolonging x and prolonging y will produce additive effects. But with a long duration of z , additivity is not produced by prolonging x and y . Little can be concluded from the fact that additivity occurs for one certain set of process durations and prolongations of x and y . Let us say that *additivity follows from network structure* if for all assignments of numbers as process durations and as prolongations produced by the factors, the two factors selectively influencing the processes x and y have additive effects.

In more detail, suppose there are n process in addition to processes x and y . Consider an assignment of nonnegative real numbers $d(x)$, $d(y)$, $d(z_1)$, \dots , $d(z_n)$ as durations of processes x , y , z_1 , \dots

\dots, z_n , respectively. As before, let t_{11} denote the duration of the task with this assignment. Consider nonnegative real numbers a and b as prolongations of x and y , respectively. In the expression t_{ij} , subscript $i = 1$ when x has duration $d(x)$ and $i = 2$ when x has duration $d(x) + a$. Subscript $j = 1$ when y has duration $d(y)$ and $j = 2$ when y has duration $d(y) + b$. Additivity follows from network structure if for all nonnegative $a, b, d(x), d(y), d(z_1), \dots, d(z_n)$,

$$t_{22} - t_{21} - t_{12} + t_{11} = h(a, b, d(x), d(y), d(z_1), \dots, d(z_n)) = 0.$$

By considering the possibilities in Equation (1), it can be seen that the interaction contrast $h(a, b, d(x), d(y), d(z_1), \dots, d(z_n)) = 0$ for all values of its arguments iff the coupled slack $k(x, y) = 0$ for all process durations.

Now consider the network in Figure 2. Vertex c is called a *cut vertex* (sometimes called a *vertex of articulation*); a definition will be given in a moment. It is obvious that if one factor prolongs a single process to the left of vertex c and another factor prolongs a single process to the right of c , the factors will have additive effects on mean response time. We will show that additivity of factors selectively influencing x and y is a consequence of network structure if and only if there is a cut vertex between x and y .

We define the part of an acyclic task network from the starting vertex of the network to the first cut vertex as the first stage, the part from the first cut vertex to the second cut vertex as the second stage, and so on. With this natural definition of stages, we will see that two factors selectively influencing different processes have additive effects following from network structure if and only if the processes are in different stages. With this definition, if two factors selectively influence processes x and y , and the factors sometimes interact, then processes x and y are in the same stage.

Further Definitions and Assumptions

The directed acyclic network for a task is assumed to be *weakly connected*, that is, it is possible

to go along arcs from any given vertex to any other vertex, ignoring the directions of the arcs. A *cut vertex* of a weakly connected directed graph is a vertex v such that if v and all arcs incident with v (i.e., arcs with v as the starting or ending vertex) are removed, the resulting directed graph is not weakly connected. It is easy to see that in a directed acyclic graph with one starting vertex o and one finishing vertex r , a vertex v different from o and r is a cut vertex if and only if every directed path from o to r contains v .

Let $f[d(a,b), \dots, d(m,n)]$ and $g[d(u,v), \dots, d(w,z)]$ be functions of the indicated path durations $d(a,b)$, and so on. If for all possible values of the path durations in the arguments of f and g , $f[d(a,b), \dots, d(m,n)] = g[d(u,v), \dots, d(w,z)]$ then we write

$$f[d(a,b), \dots, d(m,n)] \equiv g[d(u,v), \dots, d(w,z)].$$

Clearly \equiv is an equivalence relation; that is, it is reflexive, symmetric and transitive. (The relation is sometimes called identity.)

Additivity Following from Network Structure

If processes x and y are concurrent, and selectively influenced by two factors, there are probability distributions of the durations which yield a negative interaction (Schweickert, 1978; Schweickert & Townsend, 1989, Theorem 1). Hence, for additivity to follow from network structure, the selectively influenced processes must be sequential. We have seen that factors selectively influencing x and y have additive effects following from network structure if and only if the coupled slack $k(x,y) \equiv 0$; that is, the value of the coupled slack on a particular trial is 0 no matter what the durations of the processes are for that trial. According to the following theorem, this occurs if and only if there is a cut vertex between x and y .

Theorem. *Let x and y be two sequential arcs in a directed acyclic AND network or OR network. The coupled slack $k(x,y) \equiv 0$ if and only if there is a cut vertex between x and y .*

Before proving the theorem, we prove the following lemma.

Lemma. *Suppose the following vertices are listed in their order of precedence in an AND or OR network: $u, v_1, v_2, \dots, v_n, w$. Vertices v_1, v_2, \dots, v_n are on every path from u to w if and only if $d(u,w) \equiv d(u,v_1) + d(v_1,v_2) + \dots + d(v_n,w)$.*

Proof of Lemma: Clearly, if v_1, v_2, \dots, v_n are on every path from u to w , the equation holds.

Suppose $d(u,w) \equiv d(u,v_1) + d(v_1,v_2) + \dots + d(v_n,w)$. Suppose some path P from u to w does not contain v_h , for some $h, 1 \leq h \leq n$. Let h be the smallest integer for which v_h is not on P . Let v_0 denote the vertex u and let v_{n+1} denote the vertex w . Let k be the first integer to follow h for which v_k is on the path P . Consider an AND network; the argument for an OR network is similar. One of the arcs, say a , on the section of the path P from v_{h-1} to v_k can be prolonged to increase the duration of the path P without increasing any of the durations $d(v_{h-1}, v_h), d(v_h, v_{h+1}), \dots, d(v_{k-1}, v_k)$. The arc a follows v_{h-1} , so prolonging a does not increase any of the durations $d(v_{i-1}, v_i)$ for $i \leq h - 1$. Likewise, the arc a precedes v_k , so prolonging a does not increase any of the durations $d(v_i, v_{i+1})$ for $i \geq k$. Then the new duration $d(u,w) > d(u,v_1) + \dots + d(v_{h-2}, v_{h-1}) + d(v_{h-1}, v_h) + \dots + d(v_{k-1}, v_k) + d(v_k, v_{k+1}) + \dots + d(v_n, w)$. This contradicts the original assumption. QED

Before starting the proof, recall that the starting and ending vertices of an arc x are denoted x' and x'' , respectively. Let $d(u,v)$ denote the duration of the path from u to v of longest duration, if all vertices are AND gates, or of shortest duration, if all vertices are OR gates. Recall that in either case, $k(x,y) = s(x,r) - s(x,y') = s(y,r) - s^*(y,x'') = d(o,r) - d(o,y') - d(x'',r) + d(x'',y')$ (Schweickert, 1978; Schweickert & Wang, 1993).

Proof of Theorem: Without loss of generality, relabel the arcs so x precedes y . Suppose c is a cut vertex between x and y . It is easy to see that $s(x,r) \equiv s(x,y') \equiv s(x,c)$. Then $k(x,y) \equiv s(x,r) - s(x,y') \equiv 0$.

For the proof in the other direction, suppose $k(x,y) \equiv 0$. Then $d(o,r) - d(o,y') - d(x'',r) + d(x'',y')$

$\equiv 0$. There are three cases.

Case 1. Suppose y' is on every path from x'' to r . Then $d(x'',r) \equiv d(x'',y') + d(y',r)$, so

$$0 \equiv k(x,y) \equiv d(o,r) - d(o,y') - d(x'',y') - d(y',r) + d(x'',y'),$$

$$\text{so } d(o,r) \equiv d(o,y') + d(y',r).$$

Then by the lemma above, y' is on every path from o to r , so y' is a cut vertex between x and y .

Case 2. Suppose x'' is on every path from o to y' . Then $d(o,y') \equiv d(o,x'') + d(x'',y')$, so

$$0 \equiv k(x,y) \equiv d(o,r) - d(o,x'') - d(x'',y') - d(x'',r) + d(x'',y'),$$

$$\text{so } d(o,r) \equiv d(o,x'') + d(x'',r).$$

Then by the lemma above, x'' is on every path from o to r , so x'' is a cut vertex between x and y .

Case 3. If neither case 1 nor case 2 applies, then there is some path from x'' to r not containing y' and there is some path from o to y' not containing x'' . Then x'' and y' are not the same vertex. Since x precedes y , there is at least one nontrivial path C from x'' to y' .

Let $\{p_1, p_2, \dots, p_m\}$ be the vertices on every path from x'' to y' and also on every path from x'' to r . This set is nonempty because x'' is such a vertex. The vertices in the set are on all paths in the same order, suppose in the order as listed. Denote p_m by p .

Let $\{q_1, q_2, \dots, q_n\}$ be the set of all vertices on every path from o to y' and also on every path from x'' to y' . This set is nonempty because y' is such a vertex. These vertices are on all paths in the same order, say, in the order listed. Denote q_1 by q .

Vertices p and q are on a path together since there is at least one path from x'' to y' , and both p and q are on it. If a vertex u precedes a vertex v different from u , we write $u < v$. There are three cases, $p < q$, $p = q$, and $q < p$.

Case 3a. Suppose $p < q$. Then there is a path from p to q , so x and y are in a Wheatstone bridge on opposite sides of the bridge. Then if the network is a longest path network, there is an assignment of the process durations such that $k(x,y) < 0$. If the network is a shortest path network, there is an

assignment of the process durations such that $k(x,y) > 0$. Each contradicts the hypothesis of the theorem.

Case 3b. Suppose $p = q$. From the lemma above and the definition of p and q ,

$$d(x'',y') \equiv d(x'',p) + d(p,y')$$

$$d(x'',r) \equiv d(x'',p) + d(p,r)$$

$$d(o,y') \equiv d(o,p) + d(p,y').$$

Then, by substitution,

$$\begin{aligned} 0 \equiv k(x,y) &\equiv d(o,r) - d(o,y') - d(x'',r) + d(x'',y') \\ &\equiv d(o,r) - d(o,p) - d(p,y') - d(x'',p) - d(p,r) + d(x'',p) + d(p,y') \\ &\equiv d(o,r) - d(o,p) - d(p,r). \end{aligned}$$

Then $d(o,r) \equiv d(o,p) + d(p,r)$ and by the lemma p is on every path from o to r . Then p is a cut vertex between x and y .

Case 3c. Suppose $q < p$. Consider any path from x'' to p . Since $p < y'$, this path can be extended to a path from x'' to y' . By the definition of q , q must be on this extended path, and since $q < p$, q must precede p on this path. Then q is on every path from x'' to p . Then $d(x'',p) \equiv d(x'',q) + d(q,p)$. Since p is on every path from x'' to r ,

$$d(x'',r) \equiv d(x'',p) + d(p,r) \equiv d(x'',q) + d(q,p) + d(p,r).$$

Likewise, p is on every path from q to y' , so $d(q,y') \equiv d(q,p) + d(p,y')$.

Then since q is on every path from o to y'

$$d(o,y') \equiv d(o,q) + d(q,y') = d(o,q) + d(q,p) + d(p,y').$$

Finally, by definition, p is on every path from x'' to y' , so

$$d(x'',y') \equiv d(x'',p) + d(p,y') = d(x'',q) + d(q,p) + d(p,y').$$

Then, by substitution,

$$0 \equiv k(x,y) \equiv d(o,r) - d(o,y') - d(x'',r) + d(x'',y')$$

$$\begin{aligned}
&\equiv d(o,r) - d(o,q) - d(q,p) - d(p,y') - d(x'',q) - d(q,p) \\
&\quad - d(p,r) + d(x'',q) + d(q,p) + d(p,y') \\
&\equiv d(o,r) - d(o,q) - d(q,p) - d(p,r).
\end{aligned}$$

Then $d(o,r) \equiv d(o,q) + d(q,p) + d(p,r)$ and by the lemma p is on every path from o to r , as is q , so p and q are both cut vertices between x and y . QED.

It immediately follows that prolonging x and prolonging y will have additive effects following from network structure if and only if there is a cut vertex between x and y .

Random Process Durations

We now consider process durations which are nonnegative random variables. On the one hand, if there is a cut vertex c between process x and process y , then no matter what the distribution of process durations is, the response time is the sum of two random variables, $D(o,c) + D(c,r)$, where o is the starting vertex of the network and r is the ending vertex. For all but exotic definitions of what it means for a factor to selectively influence a process, a factor selectively influencing process x changes the first random variable and a factor selectively influencing process y changes the second random variable. It is straightforward that the factors are additive.

On the other hand, if there is no cut vertex between processes x and y , then according to the above theorem there exists at least one set of values of the process durations such that the coupled slack $k(x,y)$ is not 0. If there were only one such set of values for which $k(x,y) \neq 0$, one would be concerned that in an experiment these particular values would occur with probability 0, so an interaction would not be observed. (Recall that for long prolongations, $k(x,y)$ equals the interaction.) In the following corollary we show that if $k(x,y) \neq 0$ for one set of process durations then $k(x,y)$ has the same sign throughout a region of positive volume in the space of process durations. The second result of the

corollary shows it is theoretically possible for an experiment to produce a joint probability distribution of process durations and prolongations with positive probability over a region that leads to an interaction. Because the result is intuitively clear, the proof is not written out for all cases; its main purpose is to make explicit the details that need to be checked.

Selective influence. We now discuss factors selectively influencing process durations, when process durations are random variables. One technical problem is that typically on a given trial in an experiment a factor is either at one level or another, but not both. For a particular process, it is natural to define the duration of a process as one random variable when a factor influencing it is at one level, and as another random variable when the factor is at another level. This is sufficient for some purposes, but a problem arises when one wants to combine the random variables by, say, addition, because they are not necessarily defined on the same probability space. (One could add the reaction time for the first trial at level 1 of a factor to the reaction time for the first trial at level 2. But this pairing of the trials at the two levels is an arbitrary choice from many possibilities.) We discuss this problem after explaining notation.

A capital letter, e. g., W , denotes a random variable, and the corresponding small letter, w , denotes a real number value the random variable takes on. A random vector with n components is an ordered list of n random variables, $\mathbf{W} = \langle W_1, \dots, W_n \rangle$ that have a joint cumulative probability distribution. A vector of values taken on by the components of \mathbf{W} is denoted \mathbf{w} .

The *joint cumulative probability distribution function* of a random vector \mathbf{W} with n components is the function $F_{\mathbf{W}}(\mathbf{w})$ defined for every vector \mathbf{w} with n real components,

$$F_{\mathbf{W}}(\mathbf{w}) = P(\mathbf{W} \leq \mathbf{w}) = P(W_1 \leq w_1, \dots, W_n \leq w_n).$$

The *marginal cumulative probability distribution* for a single random variable W_k is $F_k(w_k) = P(W_k \leq w_k)$.

If two random vectors \mathbf{W} and $\hat{\mathbf{W}}$ have the same joint cumulative probability distribution

function, we write $\mathbf{W} \approx \hat{\mathbf{W}}$. Note that this relation requires that \mathbf{W} and $\hat{\mathbf{W}}$ have the same number of components, but it does not require that they be defined on the same probability space. For example, suppose we set \mathbf{W} to 1 for a head in a coin toss and to 2 for a tail, and we set $\hat{\mathbf{W}}$ to 1 for an odd number in a roll of a die and to 2 for an even number. Then \mathbf{W} and $\hat{\mathbf{W}}$ have the same probability distribution, but they are not defined on the same probability space.

Let i denote the level of Factor \mathcal{A} , with $i = 1, \dots, I$. Likewise, let j denote the level of Factor \mathcal{B} , with $j = 1, \dots, J$. Let the processes be denoted x, y, z_1, \dots, z_n . When Factor \mathcal{A} is at level i and Factor \mathcal{B} is at level j , let the duration of an arbitrary process w be denoted $W(ij)$; it is a random variable. In our earlier notation, the random variable duration of process w was denoted $D(w)$; to reduce multiple parentheses we modify the notation for this section. Note that w is the label of a process and also denotes a real number value that the duration of this process takes on; the meaning is determined by context.

A useful and general definition of factors selectively influencing random variables has been developed by Dzhafarov and his colleagues (Dzhafarov, 2003; Dzhafarov & Gluhovsky, 2006; Dzhafarov & Kujala, 2010; Kujala & Dzhafarov, 2008). Among other things, it resolves the problem of having different probability spaces for different levels of factors. After explaining this definition, we explain a stronger assumption needed for our purposes, selective influence by increments. If selective influence by increments occurs, selective influence in the sense of Dzhafarov and his colleagues also occurs.

Suppose when Factor \mathcal{A} is at level i and Factor \mathcal{B} is at level j there is a probability space with the random variables for process durations defined on it; that is, the multivariate random variable $\langle X(ij), Y(ij), Z_1(ij), \dots, Z_n(ij) \rangle$ is defined on this probability space. Suppose the random vector $\langle X, Y, Z_1, \dots, Z_n \rangle$ is a member of the family of random vectors $\{ \langle X(ij), Y(ij), Z_1(ij), \dots, Z_n(ij) \rangle \mid i = 1, \dots, I; j = 1, \dots, J \}$.

$\dots, J\}$. Then $\langle X, Y, Z_1, \dots, Z_n \rangle$ is said to *depend on* Factor \mathcal{A} and Factor \mathcal{B} .

As needed here, the definition of factors selectively influencing random variables by Dzhafarov and his colleagues can be put in the following form. The original papers give more general formulations and details (Dzhafarov, 2003; Dzhafarov & Gluhovsky, 2006; Dzhafarov & Kujala, 2010; Kujala & Dzhafarov, 2008). To start, consider two random variables. Suppose the random vector $\langle X, Y \rangle$ depends on Factor \mathcal{A} and Factor \mathcal{B} . Random variables X and Y are *selectively influenced* by Factor \mathcal{A} and Factor \mathcal{B} , respectively, if there is a random vector \mathbf{C} defined on a probability space such that for every level i of Factor \mathcal{A} there is a function $g_{1,i}$ and for every level j of Factor \mathcal{B} there is a function $g_{2,j}$ with

$$\langle X, Y \rangle \approx \langle g_{1,i}(\mathbf{C}), g_{2,j}(\mathbf{C}) \rangle.$$

Now suppose there are several random variables. Suppose the random vector $\langle X, Y, Z_1, \dots, Z_n \rangle$ depends on Factor \mathcal{A} and Factor \mathcal{B} . Random variables X and Y are selectively influenced by Factor \mathcal{A} and Factor \mathcal{B} , respectively, and random variables Z_1, \dots, Z_n are not influenced by Factor \mathcal{A} or Factor \mathcal{B} if there is a random vector \mathbf{C} defined on a probability space such that for every level i of Factor \mathcal{A} there is a function $g_{1,i}$ and for every level j of Factor \mathcal{B} there is a function $g_{2,j}$ and further there are functions g_3, \dots, g_{n+2} with

$$\langle X, Y, Z_1, \dots, Z_n \rangle \approx \langle g_{1,i}(\mathbf{C}), g_{2,j}(\mathbf{C}), g_3(\mathbf{C}), \dots, g_{n+2}(\mathbf{C}) \rangle.$$

Generalization is straightforward to a finite number of more factors and random variables.

Note that functions g_3, \dots, g_{n+2} do not depend on i and j . Hence for every pair of levels i and j

$$Z_1(ij) \approx g_3(\mathbf{C}),$$

so the distribution of $Z_1(ij)$ does not depend on level i or level j ; the same is true for $Z_2(ij), \dots, Z_n(ij)$

Also, because function $g_{1,i}$ does not depend on j , for every pair of levels i and j ,

$$X(ij) \approx g_{1,i}(\mathbf{C}),$$

so the distribution of $X(ij)$ depends on the level i of Factor \mathcal{A} but not the level j of Factor \mathcal{B} . We can write $X(ij)$ simply as $X(i)$. Likewise, we can write $Y(ij)$ simply as $Y(j)$. If the above definition of factors selectively influencing processes holds, the random vector \mathbf{C} is defined on the same probability space for all levels i of Factor \mathcal{A} and j of Factor \mathcal{B} . The result is that although random variables such as $X(11)$ and $Y(22)$ may not be defined on the same probability space, each is equivalent to a random variable and the random variables they are equivalent to are defined on the same probability space (because they are functions of the same random vector, \mathbf{C}).

For purposes here, when a factor selectively influences a process it must satisfy the definition above. But another assumption is needed, that as the level of a factor selectively influencing a process increases, the duration of the process becomes longer.

Selective influence by increments. The basic notion is simple. Consider one factor and one process. Roughly speaking, when Factor \mathcal{A} is at level 1, the duration of process x is a random variable $X(1)$. When Factor \mathcal{A} is at level 2, the duration of process x is this random variable plus a random variable increment U_1 . When Factor \mathcal{A} is at level 3, the duration of process x is the random variable $X(2)$ at level 2 plus a further random variable increment U_2 , and so on. But to be precise, random variables cannot be added unless they are defined on the same probability space, so the notion and notation become more complicated. We will begin with a set of random variables all defined on the same probability space and having a joint distribution. One of these, \widehat{X}_{base} , has the same distribution as $X(1)$, that is, $X(1) \approx \widehat{X}_{base}$. Another of these, \widehat{U}_2 , is the amount by which the duration of process x is incremented when the level of Factor \mathcal{A} is increased from level 1 to level 2, that is, $X(2) \approx \widehat{X}_{base} + \widehat{U}_2$. Continuing in this way, $X(i) \approx \widehat{X}_{base} + \widehat{U}_2 + \dots + \widehat{U}_i$.

Now suppose random vector $\langle X, Y, Z_1, \dots, Z_n \rangle$ depends on Factor \mathcal{A} and Factor \mathcal{B} . Suppose there is a random vector

$$\mathbf{C} = \langle \widehat{U}_2, \dots, \widehat{U}_I, \widehat{V}_2, \dots, \widehat{V}_J, \widehat{X}_{base}, \widehat{Y}_{base}, \widehat{Z}_1, \dots, \widehat{Z}_n \rangle \geq \mathbf{0}.$$

Note that all random variables, including the increments, are nonnegative. Suppose for every level i of Factor \mathcal{A} and every level j of Factor \mathcal{B}

$$\langle X(ij), Y(ij), Z_1, \dots, Z_n \rangle \approx \langle \widehat{X}_{base} + \widehat{U}_2 + \dots + \widehat{U}_i, \widehat{Y}_{base} + \widehat{V}_2 + \dots + \widehat{V}_j, \widehat{Z}_1, \dots, \widehat{Z}_n \rangle.$$

Then Factor \mathcal{A} and Factor \mathcal{B} *selectively influence by increments* random variables X and Y , respectively. Generalization to a finite number of more factors and random variables is straightforward.

Clearly, if Factor \mathcal{A} and Factor \mathcal{B} selectively influence by increments random variables X and Y , then they selectively influence X and Y according to the definition of Dzhafarov and his colleagues.

The definition of selective influence by increments leads to the following. As before, let the processes be denoted x, y, z_1, \dots, z_n . When Factor \mathcal{A} is at level i and Factor \mathcal{B} is at level j , recall that the duration of an arbitrary process w is the random variable $W(ij)$. Suppose Factor \mathcal{A} selectively influences process x and Factor \mathcal{B} selectively influences process y . Then there exist nonnegative random variables $\widehat{U}_2, \widehat{V}_2, \widehat{X}_{base}, \widehat{Y}_{base}, \widehat{Z}_1, \dots, \widehat{Z}_n$, all defined on the same probability space, such that

(1) when both factors are at level 1

$$\langle X(11), Y(11), Z_1(11), \dots, Z_n(11) \rangle \approx \langle \widehat{X}_{base}, \widehat{Y}_{base}, \widehat{Z}_1, \dots, \widehat{Z}_n \rangle;$$

(2) when Factor \mathcal{A} is at level 2 and Factor \mathcal{B} is at level 1

$$\langle X(21), Y(21), Z_1(21), \dots, Z_n(21) \rangle \approx \langle \widehat{X}_{base} + \widehat{U}_2, \widehat{Y}_{base}, \widehat{Z}_1, \dots, \widehat{Z}_n \rangle;$$

(3) when Factor \mathcal{A} is at level 1 and Factor \mathcal{B} is at level 2

$$\langle X(12), Y(12), Z_1(12), \dots, Z_n(12) \rangle \approx \langle \widehat{X}_{base}, \widehat{Y}_{base} + \widehat{V}_2, \widehat{Z}_1, \dots, \widehat{Z}_n \rangle; \text{ and}$$

(4) when both factors are at level 2,

$$\langle X(22), Y(22), Z_1(22), \dots, Z_n(22) \rangle \approx \langle \hat{X}_{base} + \hat{U}_2, \hat{Y}_{base} + \hat{V}_2, \hat{Z}_1, \dots, \hat{Z}_n \rangle.$$

In short, when the difficulty of process x is increased, the duration of process x is increased by the addition of a nonnegative random variable and when the difficulty of process y is increased, the duration of process y is increased by the addition of a nonnegative random variable. Note that the added random variables \hat{U}_2 and \hat{V}_2 may be correlated with each other, and with the durations of other processes. In particular, the random variable \hat{U}_2 may be positively or negatively correlated with the random variable \hat{X}_{base} . This allows a variety of possible correlations between process durations. For example, \hat{X}_{base} and \hat{Y}_{base} (durations of processes x and y when they are easy) can be positively correlated although $\hat{X}_{base} + \hat{U}_2$ and $\hat{Y}_{base} + \hat{V}_2$ (durations of processes x and y when they are difficult) are negatively correlated.

If Factor \mathcal{A} and Factor \mathcal{B} selectively influence by increments random variables X and Y they order the process durations in a way we now explain.

Factors ordering random vectors. An intuitive way to say a task is more difficult when Factor \mathcal{A} is at level 2 than when it is at level 1 is to say that any monotonic function of the durations of the processes in the task has greater expected value when Factor \mathcal{A} is at level 2 than when it is at level 1; and, if the influence of Factor \mathcal{A} is selective for process x , then the joint distribution of the durations of all processes except x is the same whether Factor \mathcal{A} is at level 1 or level 2.

Consider the following situation for Factor A and Factor B each with two levels.

Condition (1): For any function f of $n + 2$ arguments, f monotonically increasing in each argument separately,

$$\begin{aligned} E[f(X(22), Y(22), Z_1(22), \dots, Z_n(22))] \\ \geq E[f(X(12), Y(12), Z_1(12), \dots, Z_n(12))], \quad E[f(X(21), Y(21), Z_1(21), \dots, Z_n(21))] \end{aligned}$$

$$\geq E[f(X(11), Y(11), Z_1(11), \dots, Z_n(11))].$$

Further,

Condition (2): For every j the joint probability distribution of $Y(ij), Z_1(ij), \dots, Z_n(ij)$ is the same for every i , and for every i , the joint probability distribution of $X(ij), Z_1(ij), \dots, Z_n(ij)$ is the same for every j .

It follows immediately from Condition (2) that for every i and every j , the joint probability distribution of $Z_1(ij), \dots, Z_n(ij)$ is the same.

The assumption in Condition (1) is that certain random variables have an ordering called *stochastic dominance* in Economics and called *the usual stochastic order* in Operations Research. Müller and Stoyan (2002) and Shaked and Shanthikumar (2007) provide useful reviews and Townsend (1990) provides a useful introduction. Let Re denote the real numbers. Random vector \mathbf{S} is *smaller than* random vector \mathbf{W} in the *usual stochastic order* if they have the same number of components, n , and if $Ef(\mathbf{S}) \leq Ef(\mathbf{W})$ for all monotonically increasing functions f from Re^n into Re , for which both expectations exist. In that case, we write $\mathbf{S} \leq_{st} \mathbf{W}$. For random vectors the relation \leq_{st} is stronger than the assumption that the corresponding joint cumulative distribution functions are ordered by \leq . That is, for random vectors \mathbf{S} and \mathbf{W} , if $\mathbf{S} \leq_{st} \mathbf{W}$, then $F_{\mathbf{S}}(\mathbf{t}) \leq F_{\mathbf{W}}(\mathbf{t})$, for all \mathbf{t} , but the converse is not true. However, the situation is simpler for univariate random variables S and W (considered as random vectors with one component); $S \leq_{st} W$ if and only if $F_S(t) \leq F_W(t)$, for all t (e. g., Müller & Stoyan, 2002, Theorem 1.2.8; Townsend & Schweickert, 1989).

The usual stochastic ordering is equivalent to another useful condition, which we restate here (see, e.g., Müller & Stoyan, Theorem 3.3.5):

The following two statements are equivalent for random vectors \mathbf{S} and \mathbf{W} .

Statement (1): $\mathbf{S} \leq_{st} \mathbf{W}$.

Statement (2): There exist random vectors $\hat{\mathbf{S}}$ and $\hat{\mathbf{W}}$ defined on the same probability space, with

$\mathbf{S} \approx \hat{\mathbf{S}}$ and $\mathbf{W} \approx \hat{\mathbf{W}}$, such that $P(\hat{\mathbf{S}} \leq \hat{\mathbf{W}}) = 1$.

The usefulness of Statement (2) is subtle when first encountered, so here is an example.

Suppose distributions of the heights of men and women have the stochastic dominance relation in Statement (1). Directly from this statement one cannot define a random variable that is the difference between men's and women's heights because there is no (nonarbitrary) way to line up the men and women, i.e., the elementary experiments (selecting a man vs. selecting a woman) refer to different probability spaces. However, by Statement (2) one can define random variables M and F on a single probability space, whose distributions match the heights of males and females respectively. In this common probability space, each elementary experiment (i.e., trial) gives both a value of M and a value of F , so one can compute the difference.

Statement (2) above is equivalent to another statement useful for our purposes. Suppose Statement (2) above is true. Random vectors $\hat{\mathbf{S}}$ and $\hat{\mathbf{W}}$ are defined on the same probability space, so we can take their difference $\hat{\mathbf{U}} = \hat{\mathbf{W}} - \hat{\mathbf{S}}$. Denote a vector of 0's with the same number of components as $\hat{\mathbf{U}}$ by $\mathbf{0}$. Because $P(\hat{\mathbf{S}} \leq \hat{\mathbf{W}}) = 1$, $P(\mathbf{0} \leq \hat{\mathbf{W}} - \hat{\mathbf{S}}) = 1$, that is, $P(\mathbf{0} \leq \hat{\mathbf{U}}) = 1$. On the other hand, suppose $\hat{\mathbf{U}}$ is a random vector defined on the same probability space as $\hat{\mathbf{S}}$ and $\hat{\mathbf{W}}$, every component of $\hat{\mathbf{U}}$ is nonnegative, and $\hat{\mathbf{W}} = \hat{\mathbf{S}} + \hat{\mathbf{U}}$. Then clearly $P(\hat{\mathbf{S}} \leq \hat{\mathbf{W}}) = 1$. Hence, the following statement is equivalent to Statement (2) above, and thereby equivalent to Statement (1) above.

Statement (3): There exist random vectors $\hat{\mathbf{S}}$ and $\hat{\mathbf{U}}$ defined on the same probability space, with $\hat{\mathbf{U}} \geq \mathbf{0}$, $\mathbf{S} \approx \hat{\mathbf{S}}$ and $\mathbf{W} \approx \hat{\mathbf{S}} + \hat{\mathbf{U}}$.

For related work on univariate random variables see Townsend and Schweickert (1989).

As before, for every level i of Factor \mathcal{A} and every level j of Factor \mathcal{B} suppose there is a probability space on which the random vector of process durations $\mathbf{D}_{ij} = \langle X(ij), Y(ij), Z_1(ij), \dots, Z_n(ij) \rangle$ is defined. Suppose whenever $i' \leq i$ and $j' \leq j$, $\mathbf{D}_{i'j'} \leq_{\text{st}} \mathbf{D}_{ij}$. Then we say Factor \mathcal{A} and Factor \mathcal{B} order the random vectors $\{\mathbf{D}_{ij}\}$ with the usual stochastic order.

By Statement (3) for every pair that is ordered, $\mathbf{D}_{i'j'} \leq_{\text{st}} \mathbf{D}_{ij}$, there exist random vectors $\widehat{\mathbf{D}}_{i'j'}$ and $\widehat{\mathbf{U}}$ defined on the same probability space, with $\widehat{\mathbf{U}} \geq \mathbf{0}$, $\mathbf{D}_{i'j'} \approx \widehat{\mathbf{D}}_{i'j'}$ and $\mathbf{D}_{ij} \approx \widehat{\mathbf{D}}_{i'j'} + \widehat{\mathbf{U}}$. This assures a common probability space for every pair of ordered random vectors, but it turns out it does not assure that the same probability space will serve for every pair, i. e., different pairs may require different probability spaces, see Fill and Machida (2001). (We are grateful to E.N, Dzhafarov for pointing this out, personal communication, January 10, 2010.) Consequently, if Factor \mathcal{A} and Factor \mathcal{B} order the process durations with the usual stochastic order, the definition of selective influence by Dzhafarov and his colleagues may fail to hold. Factors \mathcal{A} and \mathcal{B} may not satisfy the definition of selective influence by increments.

The following condition is stronger than the usual stochastic order alone (Fill & Machida, 2001). Random vectors $\{ \mathbf{D}_{ij} \}$ are *realizably monotone* if there exist random variables $\{ \widehat{\mathbf{D}}_{ij} \}$ all defined on the same probability space such that for every i and j , $\mathbf{D}_{ij} \approx \widehat{\mathbf{D}}_{ij}$, and whenever $i' \leq i$ and $j' \leq j$, $\mathbf{D}_{i'j'} \leq_{\text{st}} \mathbf{D}_{ij}$.

In summary, suppose for every level i of Factor \mathcal{A} and every level j of Factor \mathcal{B} random vector $\mathbf{D}_{ij} = \langle X(ij), Y(ij), Z_1(ij), \dots, Z_n(ij) \rangle$ is defined on a probability space, and Factor \mathcal{A} and Factor \mathcal{B} selectively influence by increments random variables X and Y , respectively. Then Factor \mathcal{A} and Factor \mathcal{B} selectively influence random variables X and Y , respectively, according to the definition of Dzhafarov and his colleagues. Further, Factors \mathcal{A} and \mathcal{B} order the random vectors $\{ \mathbf{D}_{ij} \}$ with the usual stochastic order, and they are realizably monotone.

Regions leading to observable interactions. We now consider real number values that can be taken on by the random variables $\widehat{U}, \widehat{V}, \widehat{X}_{base}, \widehat{Y}_{base}, \widehat{Z}_1, \dots, \widehat{Z}_n$. We consider two levels of each factor; generalization to more levels is straightforward. Let Re_+ denote the nonnegative real numbers. There

are $n + 4$ random variables, all defined on the same probability space, so their values are real numbers in Re_+^{n+4} . In the previous section, values of durations of processes were denoted by small letters. We now return to our previous notation, and denote a value of the duration of a process z_k by $d(z_k)$.

Consider process durations as a vector $\mathbf{d} = \langle d(x), d(y), d(z_1), \dots, d(z_n) \rangle \in Re_+^{n+2}$. Consider prolongations (increments) of processes x and y as a vector $\mathbf{u} = \langle u, v \rangle \in Re_+^2$, and process durations together with prolongations u and v as a vector $\mathbf{e} = \langle u, v, d(x), d(y), d(z_1), \dots, d(z_n) \rangle \in Re_+^{n+4}$. We will replace the list of arguments of the function h in Equation (1) with the vector \mathbf{e} whose components are those arguments.

Corollary. *Suppose there is no cut vertex between sequential arcs x and y in an AND or OR network. Then*

(a) *there is a region $R \subseteq Re_+^{n+2}$ of positive volume such that $k(x,y)$ is nonzero, with the same sign, for all $\mathbf{d} \in R$, and*

(b) *there is a region $S \subseteq Re_+^{n+4}$ of positive volume such that $h(\mathbf{e})$ is nonzero, with the same sign, for all $\mathbf{e} \in S$.*

Proof. Suppose there is no cut vertex between sequential arcs x and y . Consider an OR network; reasoning is similar for an AND network. Let \mathbf{d}_0 be a vector of process durations for which $k(x,y)$ is not 0. For these durations $k(x,y) = d(o,r) - d(o,y') - d(x'',r) + d(x'',y')$, with $d(u,v)$ denoting the duration of the shortest path between vertices u and v .

We consider arcs one by one. An increase in the duration of arc z_1 may change one or more path duration terms in the equation for $k(x,y)$. We will show that for each such path duration term m we can define a coefficient c_m that determines the contribution of a small increase in the duration of z_1 to the path duration term, and an upper limit L_m such that the coefficient applies when the increase is less than L_m .

We start with the process z_1 , and for this process we start with the term $d(o,r)$. If z_1 is not on a path of shortest duration between o and r , any increase in the duration of z_1 will not change the duration of the shortest path between o and r . Let $c_1 = 0$ and L_1 be ∞ .

If there is more than one path of shortest duration between o and r , and z_1 is on one of them, an increase in the duration of z_1 will not change the duration of some shortest path between o and r . Let $c_1 = 0$ and L_1 be ∞ .

Suppose there is exactly one path of shortest duration between o and r , and z_1 is on it. A small increase e in the duration of z_1 will increase the duration of the path of shortest duration between o and r by e . Let $c_1 = 1$. A large increase in the duration of z_1 may result in some path not containing z_1 becoming the shortest path between o and r , so that further increases in the duration of z_1 have no effect on the duration of the shortest path from o to r . Let L_1 denote the largest amount by which the duration of z_1 can be increased and still remain on the shortest path between o and r .

In all cases, an increase e , $0 \leq e < L_1$, in the duration of z_1 will produce an increase in $d(o,r)$ of $c_1 e$.

The next term in the equation for $k(x,y)$ is the duration of the shortest path between o and y' . In the analogous way, a coefficient c_2 and limit L_2 are defined such that if the duration of z_1 is increased by e , $0 \leq e < L_2$, the contribution to $k(x,y)$ of the change in the shortest path between o and y' is $c_2 e$.

The value of c_2 is either 0 or -1 because the term $d(o,y')$ is negative in the equation for $k(x,y)$.

For the last two terms in the equation for $k(x,y)$, in the analogous way, coefficients c_3 and c_4 and limits L_3 and L_4 are chosen. Then, suppose the duration of z_1 is increased by e , $0 \leq e < L_1, L_2, L_3, L_4$. The value of $k(x,y)$ is changed by $e \sum c_m$ to become $k(x,y) + e \sum c_m$.

Suppose for the process durations in \mathbf{d}_0 , $k(x,y) > 0$. The reasoning if $k(x,y) < 0$ is similar.

If $\sum c_m \geq 0$, the new value of $k(x,y)$ is positive.

If, instead, $\sum c_m < 0$, the new value of $k(x,y)$ is positive if $e < -k(x,y) / \sum c_m$. Note that the

expression on the right hand side is positive.

Consequently, if $d_0(z_1)$ denotes the duration of process z_1 in \mathbf{d}_0 , $k(x,y)$ is positive for any duration of z_1 in an interval of positive length having lower and upper limits

$$d_0(z_1) < d_0(z_1) + \min\{L_1, L_2, L_3, L_4, -k(x,y)/\Sigma c_m\}.$$

Suppose a value for the duration of z_1 is chosen from the above interval. Then $k(x,y) > 0$.

Reasoning as before, for process z_2 an interval of positive length can be found such that if the duration of z_2 is in its interval, then $k(x,y) > 0$. At this point, for any values of the durations of z_1 and z_2 within their respective intervals so constructed, $k(x,y) > 0$. These steps can be repeated for each process. An interval of positive length is found for each process duration, such that if we let R denote the Cartesian product of these intervals, and each process duration is within its interval, then $k(x,y) > 0$.

Reasoning for other cases is similar, completing the proof of (a).

For (b), let R be a region of positive volume such that for all $\mathbf{d} \in R$, $k(x,y) > 0$.

For given process durations $\mathbf{d} \in R$, values of $s(x'',r)$, $s(x'',y')$, $s(y'',r)$ and $s^*(y,x'')$ can be calculated. For prolongations u and v , if $u > s(x'',r)$, $s(x'',y')$ and $v > s(y'',r)$, $s^*(y,x'')$, then $h(\mathbf{e}) > 0$.

Let $L_a = \max\{s(x'',r), s(x'',y')\}$ and $L_b = \max\{s(y'',r), s^*(y,x'')\}$.

Let S be the region of Re_+^{n+4} defined as $S = \{\mathbf{e} \mid \mathbf{d} \in R, a > L_a, b > L_b\}$.

Region S has positive volume, and for all $\mathbf{e} \in S$, $h(\mathbf{e}) > 0$.

Reasoning in the case where $k(x,y) < 0$ is similar, completing the proof of (b).

QED

To use the theorem above to infer the existence of a cut vertex between x and y , one must know whether $k(x,y) \equiv 0$. If $k(x,y) \equiv 0$, then $s(x,r) \equiv s(x,y')$ and it follows that the interaction contrasts are 0 for all levels of the factors. But in practice, one would only know that a nonzero interaction was never found in the available data. In practice, it is reasonable to conclude that a cut vertex between x and y exists if nonsignificant interactions are found in studies with high power, using several levels of the

factors selectively influencing processes x and y , and under a variety of circumstances likely to affect the durations of processes other than x and y .

Superarcs

Every arc in a stage is sequential with all the arcs outside that stage, so the arcs in a stage behave as a unit with respect to arcs outside the stage. A superarc (or superprocess) is a useful generalization of this notion. Suppose M is a subnetwork of the directed acyclic network N . A vertex is *incident* with an arc if the vertex is either the starting or ending vertex of the arc. A *vertex of attachment* of M is a vertex of M which is incident with an arc of N not in M . A *starting vertex* of M is a vertex s of M such that no arc of M has s as its ending vertex. An *ending vertex* of M is a vertex t of M such that no arc of M has t as its starting vertex. A *superarc* in a directed acyclic network N is a subnetwork with exactly one starting vertex, exactly one ending vertex, which is different from the starting vertex, and no vertex of attachment other than the starting or the ending vertex. In a task network, the arcs represent processes and we say a superarc represents a superprocesses.

A superarc M has the following three properties (Schweickert, 1983), which make the arcs in it behave as a unit with respect to the arcs outside of it. (a) It is weakly connected; that is, given any two vertices u and v of M one can follow arcs of M from u to v , ignoring the directions of the arcs. (b) The arcs of M are *convex*; that is, if x and z are arcs of M , and there is an arc y such that x precedes y and y precedes z , then y is an arc of M . (c) The arcs of M are *partitive*; that is, an arc outside M is sequential with an arc inside M if and only if it is sequential with all arcs of M , and an arc outside of M is concurrent with an arc of M if and only if it is concurrent with all the arcs of M . (Arcs in a partitive set of arcs may not all be in the same superarc. A partitive set of elements is sometimes called a *module*.) For some purposes, because of these properties, a superarc M can be replaced by a single arc in the network with the only loss of information being about relations of arcs within M to each other. A stage

as we define it is a superarc, but not all superarcs are stages.

Conclusion

A large body of studies report that two factors had additive effects on reaction time in an experiment (see Sanders, 1980). The additivity is easily explained by saying the factors selectively influence different processing stages in series. But other studies, and the anatomy of the brain, suggest that concurrent processing must be prevalent. A natural way to reconcile the findings is to say that processing can be cut at various places; processes separated by a cut behave as if in series, and processes not separated by a cut do not. This notion can be made precise by assuming the processes are partially ordered; that is, representable in a directed acyclic network. A cut vertex in such a network is a vertex that is on every path from a starting vertex to the ending vertex of the network. If we suppose the network has a single starting vertex and a single ending vertex, then it is natural to define a stage as the set of processes between two cut vertices, or between a cut vertex and either the starting or ending vertex of the network. An indication that the definition is useful is that stages defined in this way turn out to be analogous to stages as assumed in the Additive Factor Method. With our definition two factors selectively influencing processes will have additive effects if the processes are in different stages, and the factors can be made to interact if the processes are in the same stage.

References

- Davis, R. (1957). The human operator as a single channel information system. *Quarterly Journal of Experimental Psychology*, 9, 119-129.
- Donders, F. C. (1868). Die Schnelligkeit Psychischer Prozesse. *Archiv fur Anatomie und Physiologie*, 657-681. [On the speed of mental processes. In W. G. Koster (Ed. and Trans.), (1969), *Attention and performance II* (pp. 412-431). Amsterdam: North Holland.
- Dzhafarov, E. N. (2003). Selective influence through conditional independence. *Psychometrika*, 68, 7-26.
- Dzhafarov, E. N., & Gluhovsky, I. (2006). Notes on selective influence, probabilistic causality, and probabilistic dimensionality. *Journal of Mathematical Psychology*, 50, 390-401.
- Dzhafarov, E. N., & Kujala, J. V. (2010). The joint distribution criterion and the distance tests for selective probabilistic causality. Manuscript submitted for publication.
- Dzhafarov, E. N., Schweickert, R., & Sung, K. (2004). Mental architectures with selectively influenced but stochastically interdependent components. *Journal of Mathematical Psychology*, 48, 51-64.
- Ehrenstein, A., Schweickert, R., Choi, S., Proctor, R. W. (1997). Scheduling processes in working memory: Instructions control the order of memory search and mental arithmetic. *The Quarterly Journal of Experimental Psychology*, 50A, 766-802.
- Fill, J. A., & Machida, M. (2001). Stochastic monotonicity and realizable monotonicity. *The Annals of Probability*, 29, 938-978.
- Fisher, D. L., & Glaser, R. A. (1996). Molar and latent models of cognitive slowing: Implications for aging, dementia, depression, development, and intelligence. *Psychonomic Bulletin & Review*, 4, 458-480.

- Fisher, D. L., & Goldstein, W. M. (1983). Stochastic PERT networks as models of cognition: Derivation of the mean, variance, and distribution of reaction time using order-of-processing (OP) diagrams. *Journal of Mathematical Psychology*, 27, 121-151.
- Fisher, D. L., Saisi, D., & Goldstein, W. M. (1985). Stochastic PERT networks: OP diagrams, critical paths, and the expected completion time. *Computers and Operations Research*, 12, 471-482.
- Goldstein, W. M., & Fisher, D. L. (1991). Stochastic networks as models of cognition: Derivation of response time distributions using the order-of-processing method. *Journal of Mathematical Psychology*, 35, 214-241.
- Goldstein, W. M., & Fisher, D. L. (1992). Stochastic networks as models of cognition: Deriving predictions for resource-constrained mental processing. *Journal of Mathematical Psychology*, 36, 129-145.
- Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world task performance. *Human-Computer Interaction*, 8, 237-309.
- Kelley, J. E., & Walker, M. R. (1959). Critical path planning and scheduling. *Proceedings of the Eastern Joint Computer Conference* (pp. 160-173). Boston, MA.
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility--A model and taxonomy. *Psychological Review*, 97, 253-270.
- Kujala, J., & Dzhamfarov, E. N. (2008). Testing for selectivity in the dependence of random variables on external factors. *Journal of Mathematical Psychology*, 52, 128-144.
- Malcolm, D. G., Roseboom, J. H., Clark, C. E., & Fazar, W. (1959). Applications of a technique for research and development program evaluation. *Operations Research*, 7, 646-669.

- Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance. Part 1. Basic mechanisms. *Psychological Review*, 104, 3-65.
- Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance. Part 2. Accounts of Psychological Refractory-Period Phenomena. *Psychological Review*, 104, 749–791.
- Müller, A., & Stoyan, D. (2002). Comparison methods for stochastic models and risks. NY: Wiley.
- Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology*, 41A, 19-45.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Sanders, A. F. (1980). Stage analysis of reaction processes. In Stelmach G E, Requin J (eds.) *Tutorials in motor behavior*. Amsterdam: North Holland.
- Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology*, 18, 105-139.
- Schweickert, R. (1983). Synthesizing partial orders given comparability information: Partitive sets and slack in critical path networks. *Journal of Mathematical Psychology*, 27, 261-276.
- Schweickert, R. (1992). Constructing a stochastic critical path network given the slacks: Representation. *Mathematical Social Sciences*, 23, 343-366.
- Schweickert, R., Fisher, D. L., & Proctor, R. W. (2003). Steps toward building mathematical and computer models from cognitive task analysis. *Human Factors*, 45, 77-103.
- Schweickert, R., Fortin, C., & Sung, K. (2007.). Parallel visual search and time reproduction. *Journal of Mathematical Psychology*, 51, 99-121.
- Schweickert, R., & Townsend, J. T. (1989). A trichotomy: Interactions of factors prolonging

sequential and concurrent mental processes in stochastic discrete mental (PERT) networks.

Journal of Mathematical Psychology, 33, 328-347.

Schweickert, R., and Wang, Z. (1993). Effects on response time of factors selectively influencing processes in acyclic task networks with OR gates. *British Journal of Mathematical and Statistical Psychology*, 46, 1-40.

Shaked, M., & Shanthikumar, J. G. (2007). *Stochastic orders*. NY: Springer.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II*, *Acta Psychologica*, 30, 276-315.

Sung, K. (2008). Serial and parallel attentive visual searches: Evidence from cumulative distribution functions of response times. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1372-1388.

Thomas, R. D. (2006). Processing time predictions of current models of perception in the classic additive factors paradigm. *Journal of Mathematical Psychology*, 50, 441-455.

Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, 25, 168-197.

Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inferences. *Psychological Bulletin*, 108, 551-567.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, 39, 321-359.

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.

Van Zandt, T., & Townsend, J. T. (1993). Self-terminating versus exhaustive processes in rapid visual

and memory search: An evaluative review. *Perception & Psychophysics*, 53, 563-580.

Welford, A. T. (1959). Evidence of a single-channel decision mechanism limiting performance in a

serial reaction task. *Quarterly Journal of Experimental Psychology*, 11, 193-210.

Welford, A. T. (1967). Single channel operation in the brain. *Acta Psychologica*, 27, 5-22.

Author Note

Richard Schweickert, Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, swike@psych.purdue.edu. Donald L. Fisher, Department of Mechanical and Industrial Engineering, University of Massachusetts, Amherst, MA 01003, fisher@zonker.umass.edu. William M. Goldstein, Department of Psychology, University of Chicago, gold@uchicago.edu.

Portions of this study were funded by NIMH Grant 1 RO1 MH41452, NSF Grant DBS-9123865, AFOSR Grant No. FA9550-06-0383 to Schweickert and AFOSR Grant No. FA9550-09-1-0252 to Schweickert and Dzhafarov. Claudette Fortin provided a stimulating environment for writing at Université Laval, Québec. We thank Ehtibar Dzhafarov for helpful comments.

Correspondence about this article should be addressed to Richard Schweickert, Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, swike@psych.purdue.edu.

Figure Captions

Figure 1. Processes x and y are sequential, each is concurrent with process z . Starting vertex of this directed acyclic task network is o , ending vertex is r .

Figure 2. Cut vertex c divides the directed acyclic task network into two stages.



