Multinomial Processing Tree Inferred from Age-Related Memory-Error Probabilities: Possibility of Inferring More if Response Times were Available

--Manuscript Draft--

Highlights

Older and younger adults studied paired-associates.

Some pairs, such as "sheep-doctor" were easily integrated.

A Multinomial Processing Tree (MPT) accounted for response probabilities.

Factors can selectively influence vertices in an MPT.

To test with response probability and time, only two MPTs need be considered.

Under certain conditions selectively influenced vertex order is undetermined.

Abstract

Use of Multinomial Processing Tree (MPT) models is illustrated by fitting one to data of Dhir (2017). Her experiment examined age and association type in a paired-associate recall task. Age and Pair-Type had interactive effects on probability of a correct response. A natural interpretation of the interaction would be that both factors impact the same mental process. However, fitting an MPT leads to the conclusion that Age and Pair-Type selectively influence two separate processes, one following the other. A possible interpretation of these is as attempts at specific (verbatim) retrieval and knowledge supported (gist) processing, selectively influenced by Age and Pair-Type, respectively. The order of these processes is not determined by the response probabilities. In a further section of the paper, we show that if response times or other measures had also been available, they could have resolved the process order, but might have left it undetermined. We give necessary and sufficient conditions for two factors to selectively influence two ordered vertices in an MPT, with either order of the vertices accounting for both response probability and response time. They do so if and only if the MPT is equivalent to a special processing tree, not necessarily an MPT itself.

A Multinomial Processing Tree Inferred from Age-Related Memory-Error Probabilities: Possibility of Inferring More if Response Times were Available

Richard Schweickert, Purdue University

Pritha Dhir, City, University of London

Xiaofang Zheng, Purdue University

Marie Poirier, City, University of London

Author Note

Richard Schweickert, Department of Psychological Sciences, Purdue University; Pritha Dhir, Department of Psychology, City, University of London; Xiaofang Zheng, Department of Psychological Sciences, Purdue University; Marie Poirier, Department of Psychology, City, University of London.

Acknowledgements: This work is supported in part by a scholarship from China Scholarship Council (CSC) to Zheng.

Correspondence concerning this article should be sent to Richard Schweickert, Department of Psychological Sciences, Purdue University, 703 Third St., West Lafayette, IN 47907, schweick@purdue.edu.

18 July 2020

A Multinomial Processing Tree Inferred from Age-Related Memory-Error Probabilities:

Possibility of Inferring More if Response Times were Available

One of the many lasting contributions of William H. Batchelder was the establishment of Multinomial Processing Trees (MPTs) as general-purpose models in psychology. Through his work and friendly encouragement of work of many others, he ensured use of MPTs would spread and continue. At this time, they are models in many domains, including perception (e.g., Bishara & Payne, 2009), memory (e.g., Chechile , 1977), social cognition (e.g., Klauer & Wegener, 1998), and psychological assessment (e.g., Batchelder, 1998). For reviews, see Batchelder and Riefer (1999), Erdfelder, Auer, Hilbig, Aβfalg, Moshagen and Nadarevic (2009), and Hütter and Klauer (2016).

This paper begins with a brief description of Multinomial Processing Trees. One application is in studies of aging (e.g., Greene & Naveh-Benjamin, in press) and we fit an MPT to data from a paired associate learning experiment by Dhir (2017) with factors of age and paired associate type. Analysis indicates that age and pair-type have effects on two different processes, represented by two vertices in an MPT. The data lead to an MPT in which the processes represented by the two vertices are executed one after the other, but their order is not determined. The data were probabilities of correct responses. Some investigators incorporate response time and other measures in MPTs (e.g., Heck & Erdfelder, 2016; Hu, 2001, Klauer & Kellen, 2018; Link, 1982; Schweickert & Zheng, 2018; Wollschläger & Diederich, 2012). Accordingly, the last part of the paper considers whether response times or other measures, if available, could reveal more than response probabilities do about the form of an MPT, in particular whether they could determine the order of vertices unresolved by response probabilities.

Multinomial Processing Trees

In a Multinomial Processing Tree, a mental process such as memory retrieval is represented with a point, called a vertex. (An example will be given later.) A possible outcome of a process, such as successful retrieval, is represented by a line, called an arc, descending from the corresponding vertex (see Figures 1-3). The first process to be executed is represented by the *source vertex* of the tree, a vertex with no arcs entering it. On each trial, every outcome of the source vertex has some probability of occurring, and one outcome occurs. The sum of the probabilities of the arcs descending from a vertex is 1. When an outcome occurs, the arc representing this outcome is traversed and the vertex at the lower end of the arc is reached. That vertex represents a further process, which is executed. One of its outcomes occurs, with a certain probability. These steps are repeated until a vertex that has no arcs descending from it is reached. Such a vertex is a *terminal vertex*, and at it a response is made. Responses are in classes such as correct or incorrect, or high, medium and low confidence. The probability of a path from one vertex to another, going along each arc in the descending direction, is the product of the probabilities associated with the arcs on the path. The probability of a response in a certain class is the sum of the probabilities of all the paths from the source vertex to those terminal vertices associated with that class.

In some MPT models there is more than one source vertex, for example, one for a recall trial and another for a recognition trial. Usually to analyze a given situation, say a recognition trial, only one source vertex is relevant and other source vertices and arcs that follow them can be ignored.

An experimental factor that changes parameters associated with arcs that descend from one and only one vertex in an MPT is said to *selectively influence* that vertex. For example, in an immediate memory experiment by Poirier, Schweickert and Oliver (2005), the serial position of items selectively influenced a vertex representing the degradation of representations in memory, while list length selectively influenced a vertex representing the redintegration (reconstruction) of degraded representations. For details and a review of selective influence in MPT models see Schweickert, Fisher and Sung (2012).

Suppose an experiment has two factors. For notation in this paper we assume throughout that one factor, Φ, has levels $i = 1, \ldots, I$ and the other factor, Ψ, has levels $j = 1, \ldots, J$. With response probabilities, if there are two response classes (e.g., correct and incorrect) one can test whether each factor selectively influences a different vertex in an MPT. One need only test MPTs of two forms (Schweickert & Chen, 2008). The first is in Figure 1. Suppose Factor Φ selectively influences the vertex with probabilities indexed by i in the MPT in Figure 1. Suppose Factor Ψ selectively influences the vertex with probabilities indexed by j. A correct response can be made by following either of two paths. The probability of a correct response when Factor Φ is at level i and Factor Ψ is at level j is

$$p(i,j) = p_D(i) + p_B(i)p_F(j). \tag{1}$$

In the equation above, the second term on the right side produces an interaction between the two factors.

Similarly, the MPT in Figure 2 predicts the probability of a correct response to be

$$p(i,j) = p_A p_D(i) + p_B p_F(j). \tag{2}$$

The MPT in Figure 2 predicts additive effects of the factors because the first term on the right-hand side depends only on the level i of Factor Φ and the second term depends only on the level j of Factor Ψ.

The difference between the two forms of MPTs is straightforward. In the MPT in Figure 1 there is a path (arcs *B* and *F*) descending from the vertex selectively influenced by a first factor to the vertex selectively influenced by a second factor, and on this path there is an arc whose parameter values change when the level of the second factor changes. We say the vertices are *ordered by the factors,* or for short, *ordered*. In the MPT in Figure 2 there is no such path and we say the vertices are *unordered*. Either there is such a path or there is not. If there are more than two response classes, the MPTs in Figures 1 and 2 can be extended in a straightforward way. Figure 3, for example, extends the MPT in Figure 1 to more than two response classes. It may be that a participant in an experiment executes processes in a rather complicated MPT. Importantly, if each of two factors selectively influences a different vertex in an MPT, then it turns out that the MPT the participant used makes exactly the same predictions for response probabilities as either the MPT in Figure 1 or that in Figure 2, or to the extension of one of these to more than two response classes (Schweickert & Chen, 2008; Schweickert & Xi, 2011). If these are rejected, then no MPT is possible in which each factor selectively influences a different vertex. An MPT that accounts for response probabilities may be possible, but not one in which the factors selectively influence different vertices.

Age and Integrative Relations in Paired Associate Learning

Here we give an example of selective influence of factors with data from a study on aging and associative memory, taken from the first experiment in Dhir (2017; Dhir & Poirier, 2015). An old controversy in the study of aging was about whether aging impairs all cognitive processes and in the same way (e.g., Fisk & Fisher, 1994; Cerella & Hale, 1994). There is now considerable evidence that some processes have little or no impairment with age, and efforts have turned to finding which these are, such as automatic processing (e.g., Hasher & Zacks,

1979) and processing based on word knowledge (e.g., Horn, 1982; Salthouse, 1991). For review see Salthouse (2010).

Recently Greene and Naveh-Benjamin (in press) considered whether older adults are impaired in memory for specifics but not in memory of gist. The distinction comes from Fuzzy Trace Theory (Brainerd, Reyna, & Mojardin, 1999), in which an episode leaves two traces, one specific (called “verbatim”) and another fuzzier trace often based on knowledge (called “gist”). At test, separate retrieval attempts can be made from the “verbatim” trace or from the "gist" trace. With an MPT model, Greene and Naveh-Benjamin (in press) compared younger and older adults on recognition memory for face-place pairs. Their MPT model modifies those of Brainerd, Reuna and Mojardin (1999) and Stahl and Klauer (2008). The MPT analysis found that younger and older adults differed in probability of specific "verbatim" retrieval; however, there was no age effect when the probability of recognition depended on a knowledge-based "gist" trace.

We note that the vertex labels "verbatim" and "gist" of the MPT model of Greene and Naveh-Benjamin (in press) come from earlier papers and may not apply aptly to entities in a particular experiment. For example, stimuli in their experiment were pictures of a face and a place, which would not form a literally verbatim trace. In what follows, when speaking of an MPT, we will say, as they do, "verbatim" and "gist" when referring to certain vertices. When speaking of what a participant does we will say “specific retrieval” and “knowledge-supported processing” when referring to the corresponding mental processes.

Experiment 1 of Dhir (2017) gives an opportunity to see, with different stimuli and knowledge support varied in a different way, whether specific retrieval will be impaired in older adults while knowledge-supported processing is age invariant. Moreover, while Greene and

Naveh-Benjamin (in press) used a recognition paradigm, Dhir (2017) used a paired-associate recall task; this allows us to test MPTs in the context of a recall task. The original purpose of Dhir's experiment was to learn whether the difficulty older adults have in learning new word associations can be alleviated when they are easily integrated (e.g., *sheep-doctor*). Three types of word pair were used, described below. Because the pair-types differ in how meaningful and familiar they are and younger and older adults were studied, the experiment is well suited to examine the effect of age on specific memory and knowledge-supported processing.

Background and Experiment

Episodic memory decline is one of the hallmarks of normal cognitive aging. It is well established that this decline is related to growing difficulties with retrieving associative relationships, a view known as the associative deficit hypothesis (ADH; Naveh-Benjamin, 2000; Naveh-Benjamin & Mayr, 2018). According to the ADH, the age-related episodic memory problems are in part caused by a decline in the capacity to encode and retrieve new associations between the features of an episode (Naveh-Benjamin, 2000; Bayen, Phelps, & Spaniol, 2000; Chalfonte & Johnson, 1996). The hypothesis applies to different types of new associations, including links between two items, between an item and its source, and between an item and its context. These effects are robust and have been reported in a variety of studies, including meta-analyses (e.g., Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995). An important fact in the present context is that the associative memory deficit is known to be alleviated if the new learning can be supported by prior knowledge (Naveh-Benjamin, 2000; Badham & Maylor, 2015).

Moreover, using a paired-associate learning task, Badham, Estes and Maylor (2012) suggested that older adults (OA) can form new associations as readily as younger adults (YA) if they are presented with items that make sense when considered together (e.g., *sheep-doctor*). Such items pairs are said to be *integrative* (Estes, Golonka, & Jones, 2011) in that the first word of the pair specifies or defines the second.

Badham et al (2012) compared memory performance of younger and older adults with pairs that were unrelated (*pillow-candle*), integrative (*lemon-cake*) or semantically related (*article-book*); importantly, the integrative pairs were chosen to have a very low-level of prior association—i.e. they represented new associations. The authors reported that the integrative pairs alleviated the age-related deficit just as well as the semantic pairs did, despite the integrative pairs being “unassociated and semantically dissimilar”. However, Dhir (2017; Dhir & Poirier, 2015) argued that closer inspection of the stimuli revealed that many of the integrative pairs had pre-established associations (e.g. *herb-garden, winter-sports*), suggesting this may have supported performance for older adults. Dhir hence set out to replicate and extend the study by Badham et al. She compared the performance of young and old adults for integrative pairs that had pre-established associative links, integrative pairs that did not have prior associative links, and unrelated pairs. For example, *border-land* is an integrative pair with pre-established associative link, *dinosour-land* is an integrative pair without prior associative link, and *stripe-land* is an unrelated pair. Her aim was to offer a stringent test of the proposal that integration was a sufficient condition to produce an improvement in the associative memory deficit of OA. Accordingly, she used a design with two age levels crossed with three pair-types (see Dhir, 2017, for details).

Results and Discussion

Figure 4 with data from Dhir (2017) summarizes mean performance for each group and each type of list; it suggests that integrative-associative pairs were recalled the best and unrelated pairs the worst. In addition, the age difference appeared to be largest with the unrelated pairs, and smallest with the integrative associative pairs. Data were analyzed with a 2 (age: young, old) x 3 (pair-type: integrative associative, integrative non associative, unrelated) mixed ANOVA which confirmed these observations. There was a main effect of age ($F(1, 38) = 33.7$, MSE = 0.7, $p<.001$) such that YA recalled more items than OA and that pair-type affected performance. Importantly, there was a main effect of pair-type ($F(2, 76) = 230.1$, MSE = 1, $p<.001$) and a significant interaction between age and pair-type ($F(2, 76) = 16.3$, MSE = 0.1, $p<.001$). To clarify the source of the interaction, further ANOVAs were run. The first was 2 (age: young, old) x 2 (pair-type: integrative associative, integrative non-associative). It showed an effect of pair-type ($F(1, 38) = 153.4$, MSE = .4, $p<.001$), age ($F(1, 38) = 18.$, MSE = .2, $p<.001$), as well as a significant interaction ($F(1, 38) = 7.9$, MSE<0 , $p = .008$). Further analyses showed that as Figure 4 suggests, OA benefited more than YA when going from non-associative integrative to associative integrative pairs, although YA performed significantly better in both cases. Post-hoc *t*-tests indicated a significant age difference (YA>OA) for both the integrative non-associative ($t(38) = 7$, p<.001) and the integrative associative ($t(38) = 4.2$, $p<.001$) conditions. A further 2 (age: young vs old) x 2 (pair-type : integrative non-associative vs unrelated) ANOVA produced a main effect of pair-type ($F(1, 38) = 123.3$, MSE = .6, $p<.001$), age ($F(1, 38) = 36.9$, MSE = .8, $p<.001$), and a further significant interaction ($F(1, 38) = 10.9$, MSE = .1, $p = .002$). Figure 4 suggests that OA benefitted more than YA when going from unrelated to the integrative non-associative pairs. Here also, post-hoc independent *t*-tests showed an age difference (YA>OA) for the integrative non associative ($t(38) = 7$, $p<.001$), as well as for the unrelated ($t(38) = 9$, p<.001) word pairs.

The results reproduce the age-related deficit in associative memory, most obvious in the relative difficulties of OA participants in the unrelated pairs condition. The original purpose of the experiment was to test the hypothesis that this age-related deficit can be significantly alleviated when the to-be-remembered association is imbedded within an integrative structure – even if the association is new or unfamiliar. This hypothesis was supported by the findings: When memory for unrelated items was compared to memory for unfamiliar but integrative pairs, there was a disproportionate advantage for the OA. Finally, if item pairs can be integrated and are familiar (integrative associative pairs) OA again benefit more than YA, relative to the performance observed with integrative non-associative pairs. This indicates that both integration and prior knowledge or familiarity with item-pairs contribute to reducing the age-related deficit in associative memory. We now turn to our current purpose, examining the data for what they reveal about the effect of age on specific retrieval and knowledge-supported processing.

A Multinomial Processing Tree for Dhir (2017) Experiment 1

How do age and pair-type combine to produce the results of Dhir (2017)? Because the factors interact, it is tempting to suggest they both affect some memory process, with the effect of one factor on the process depending on the level of the other. But it is instructive to examine the MPT model of Greene and Naveh-Benjamin (in press) for effect of age in an associative recognition memory task.

There were two major processes in the MPT model, an attempt to retrieve a specific memory representation (represented by a "verbatim" vertex in the MPT), and if that failed, an attempt at knowledge-supported processing (represented by a "gist" vertex in the MPT), followed by minor processes of guessing. Based on estimates of parameters in the model, the authors concluded that young adults had higher probability of successful specific memory retrieval, but did not differ in probability of successful knowledge-supported processing. The results encourage testing for selective influence of age in an MPT for the experiment of Dhir (2017).

Let's consider an MPT. If age and pair-type selectively influence different vertices, then, as we said earlier, the MPT must be equivalent to either that in Figure 1 or that in Figure 2. If neither of these MPTs fits the data it is not possible that the two factors selectively influence different vertices in any MPT (Schweickert & Chen, 2008). The MPT in Figure 2 predicts additive effects of age and pair-type, contrary to the interaction in the results (see the Appendix for discussion), so we consider the MPT in Figure 1. Consider the hypothesis that when the cue is presented, the first vertex is influenced by age but not by pair-type and the second vertex is influenced by pair-type but not by age.

The MPT is in Figure 1. Denote the levels of age as $i = 1$ (YA) and $i = 2$ (OA), and the levels of pair-type as $j = 1$ (Unrelated), $j = 2$ (Integrative NonAssociative) and $j = 3$ (Integrative Associative). In the MPT, at the first vertex an outcome produces a correct response with probability $p_D(i)$, an outcome leads to further processing with probability $p_B(i)$, and an outcome produces an incorrect response with probability $1 - p_D(i) - p_B(i)$. Probabilities $p_D(i)$ and $p_B(i)$ depend on the age level i of the participant, but not on the pair-type. If the outcome of the first vertex leads to further processing, the further processing produces a correct response with probability $p_F(j)$. Probability $p_F(j)$ depends on the pair-type j, but not on the age of the participant.

With the MPT, the probability of correct recall of the target when the cue is presented is, as in Equation (1),

$$p(i,j) = p_D(i) + p_B(i)p_F(j).$$

Correct recall data are in Table 1. For this model, arc probability parameters were estimated to minimize the likelihood ratio statistic, G^2, a procedure that produces the maximum likelihood estimates (Bishop, Feinberg, & Holland, 1975). Results are in Table 1. The goodness of fit, $G^2 = .12$, has approximately a chi square distribution with 1 df, and is not significant, indicating a good fit. One can see in Table 1 that predicted and observed values are quite close. For details of parameter estimation, see Appendix.

The formula for the probability of a correct response has a product term, $p_B(i)p_F(j)$, which leads to an interaction of the effect of age and pair-type. One multiplier depends on age, the other on pair type, so although the factors interact, they do not both influence the same process. Instead, they selectively influence different processes.

In another version of the model, the first vertex is influenced by the pair-type, and if an outcome of the first vertex requires further processing, the success of further processing is influenced by the age of the participant. Then the probability of correct recall for age i pair-type j is

$$p(i,j) = p^*_D(j) + p^*_B(j)p^*_F(i). \qquad (3)$$

The indices i and j are interchanged in Equations (1) and (3). It turns out that this version of the model has just enough free parameters to exactly account for the data from the experiment, see Appendix. Because this version of the model fits perfectly, the versions cannot be meaningfully compared by goodness of fit alone. In the Appendix we compare models by taking the number of free parameters into account with the Akaike Information Criterion (Akaike, 1973). We arrive at a slight preference for the model in Equation (1). In any case, each version of the model demonstrates that a Multinomial Processing Tree in which each factor selectively influences a different process can account for the data.

In a Multinomial Processing Tree the vertices represent processes. But it is not necessary to know what the processes are doing for the MPT to account for response probabilities. In other words, an MPT, i. e., a tree with vertices, arcs and parameters, is not committed to a particular interpretation of its vertices. Data from Dhir (2017) Experiment 1 can be accounted for by an MPT in which one vertex is selectively influenced by participants' age and another vertex is selectively influenced by pair-type. Essentially the meaning of a parameter is operationally defined by the factor that changes the values of the parameter. Knowing that a factor selectively influences a vertex constrains possibilities for what the process the vertex represents could be doing, but does not specify the function of the process uniquely. Knowing, for example, that a change in age changes the processing represented by a vertex does not allow us to know what the process is as such.

Keeping in mind that an MPT can accurately account for data, despite an erroneous interpretation of its vertices, let's consider a possible interpretation in terms of the similar model of Greene and Naveh-Benjamin (in press), in which the source vertex represents an attempt at specific retrieval and the second represents an attempt at retrieval based on knowledge-supported processing. In the experiment of Dhir (2017) overall performance of older adults is worse than that of younger adults, and the extent of the deficit is greater for unrelated pairs than for the two types of integrative pairs. In terms of the model, relative to YA, associations available for OA are more difficult to retrieve specifically. But the effect of pair-type -- i.e. of support afforded by knowledge – is equivalent for both YA and OA, because, knowledge-supported processing is affected by pair-type but it is unaffected by age. As proposed by the ADH, age has a detrimental effect on the processing of associative information. Although the MPT of Green and Naveh-Benjamin differs from ours in several ways, for example, the former for recognition, ours for recall, comparing them suggests a possibility for which cognitive process is impaired by age and which is not.

What Response Times and Response Probabilities Together Reveal

The preceding MPT analysis was based on response probabilities. Earlier we said some MPT models incorporate an additional measure such as response time (Hu, 2001; Link, 1982) or distance (Rosenbaum, 1980), and recently additional measures have become of interest again (Heck & Erdfelder, 2016, 2017; Klauer& Kellen, 2018; Schweickert & Zheng, 2018, 2019a). Suppose an additional measure is observed. What would we learn beyond what we learn from response probability alone?

What we learn depends on the technique used. We recommend the excellent papers by Heck and Erdfelder (2016, 2017) and by Klauer and Kellen (2018) for information gained with their techniques, and continue here to discuss the technique of manipulating factors that selectively influence processes. This technique has been fruitfully used for response time analysis from the pioneering work of Sternberg (1969) until now (e.g., Reimer, Strobach, & Schubert, 2017; Sung & Gordon, 2018).

With an additional measure we would obtain further tests of an MPT model and estimates of the other measure's parameter values. Such tests and estimates are discussed in earlier papers (Schweickert & Zheng, 2018, 2019a, 2019b, in press). Here we consider whether we could learn anything about the form of an MPT not already revealed by the response probabilities.

Suppose each of two factors selectively influences a different vertex in an MPT. Two questions about form arise. First, are the selectively influenced vertices ordered or not? If response probabilities are the only observations, either the MPT in Figure 1 (for ordered vertices) or that in Figure 2 (for unordered vertices) will suffice for two response classes (Schweickert & Chen, 2008). With more than two response classes, the MPTs in Figures 1 and 2 suffice if extended in a straightforward way (Schweickert & Xi, 2011); see Figure 3 for the extension of the MPT in Figure 1. If there is an additional measure, must some additional form be considered? According to a previous paper (Schweickert & Zheng, 2019a), the answer is no, except for an unsettled case, which we consider in a moment.

The second question about form arises if the vertices are ordered: What is their order? Sometimes response probabilities alone do not determine the order of the vertices on which the factors have their influences. For example, in the experiment discussed in the first part of this paper (Dhir, 2017), the Akaike Information Criterion favored the order expressed in Equation (1) over the order expressed in Equation (3), but not strongly. Suppose response probabilities alone do not establish the order of vertices selectively influenced by factors. Is it possible for an additional measure to resolve the question of order?

Notation and Assumptions for an Additional Measure

For an additional measure we need additional assumptions and notation. We usually say the additional measure is response time, but it could be some other quantity.

If the starting vertex of an arc L in an MPT is reached, we assume, as before, that arc L has a probability p_L of being selected and we now assume that selection of arc L takes time t_L. Here we make the simple assumptions that the probability p_L associated with arc L is a fixed number and the time t_L is the expected value of a random variable T_L. These assumptions can be weakened so response time distributions can be considered (Schweickert & Zheng, 2018, in press), but they are beyond the scope of this paper. Suppose a factor, say Factor Φ, selectively influences the starting vertex of arc L. Then when the factor is at level i, the probability and time associated with arc L are denoted $p_L(i)$ and $t_L(i)$, respectively. (The level i could be removed from the symbol $p_L(i)$ or $t_L(i)$ if the probability or time associated with arc L does not change when the level i changes. In this paper, it is not necessary to do so.) As before, the probability of a path from one vertex to another going along arcs in a descending direction is the product of the probabilities associated with the arcs on the path. The time for such a path is the sum of the times associated with the arcs on the path.

Each terminal vertex has an associated response class k, $k = 1, , , , K$. A response of class k is made by following a path from the source vertex to a terminal vertex of class k. As before, the probability of a response in class k is the sum of the probabilities of all the possible paths from the source vertex to a terminal vertex of class k. The response time for a response of class k is obtained from all the paths leading to a response of class k. We explain this with an example.

Consider the MPT in Figure 1. Parameters on the arcs are labeled with the levels i and j of the factors selectively influencing vertices in the MPT; those labels will be relevant later. There are two paths from the source vertex to a response in the class of correct responses. One path consists simply of the arc D. The probability a correct response is made via this path is $p_D(i)$. The other path consists of the arcs B and F. The probability of a correct response made via this path is $p_B(i)p_F(f)$. Hence, the probability of a correct response is, as in Equation (1),

$$p(i,j) = p_D(i) + p_B(i)p_F(j).$$

To calculate the time for a correct response, we need conditional probabilities. Given that a correct response is made, the probability it is made via the path consisting of arc D is $p_D(i)/p(i,j)$. The time to make a response via the path consisting of arc D is $t_D(i)$. Given that a correct response is made, the probability it is made via the path consisting of arcs B and F is $p_B(i)p_F(j)/p(i,j)$. The time to make a response via the path consisting of arcs B and F is $t_B(i) + t_F(j)$. Now let $t(i,j)$ denote the time to make a correct response when the factors are at their levels i and j. Then

$$t(i,j) = \frac{p_D(i)}{p(i,j)} t_D(i) + \frac{p_B(i)p_F(j)}{p(i,j)} [t_B(i) + t_F(j)].$$

The above equation can be put in a more useful form,

$$p(i,j)t(i,j) = p_D(i)t_D(i) + p_B(i)p_F(j)[t_B(i) + t_F(j)].$$

The expression on the left hand side above is useful because it combines both accuracy and time in a natural way.

With analogous reasoning for the MPT in Figure 2, the probability of a correct response when the factors are at levels i and j, is in Equation (2),

$$p(i,j) = p_A p_D(i) + p_B p_F(j).$$

And the time for a correct response satisfies the equation

$$p(i,j)t(i,j) = p_A p_D(i)[t_A + t_D(i)] + p_B p_F(j)[t_F(j)].$$

A key feature of the MPT in Figure 2 is that it predicts the factors to have additive effects on both $p(i,j)$ and $p(i,j)t(i,j)$. Consequently, the MPTs in Figures 1 and 2 can readily be distinguished in data by the presence of an interaction (Figure 1) or its absence (Figure 2).

The equations above were derived for fixed quantities and a reader may wonder if they apply to response times, which are variable. Under plausible assumptions, the equations above may be considered as applying to expected values of random variables. Details are not needed here, but for further discussion and tests based on response time cumulative distributions, see Schweickert and Zheng (2018, in press).

Suppose Factor Φ has levels $i = 1, \ldots, I$ and Factor Ψ has levels $j = 1, \ldots, J$ and suppose these factors change parameters in two MPTs. The two MPTs are *equivalent* for these factors and these levels if they make the same predictions for $p(i,j)$ and $t(i,j)$ for all i and j.

We assume each factor is *effective*, that is, it is not the case that for every level i $p(i,j)$ and $t(i,j)$ never change when level j changes, and the analogous statement holds for every level j.

Further, we assume that in the MPT in Figure 1 there is at least one change of the level i of Factor Φ that changes the probability or time associated with arc B. That is, it is not the case that for every i, $p_B(i) = p_B$ and $t_B(i) = t_B$. Without this assumption, the MPT in Figure 1 may be equivalent to the MPT in Figure 2, complicating discussion. We return to the question of what an additional measure could reveal.

Does Response Time Reveal Additional MPT Form?

Consider an arbitrary MPT. Suppose Factor Φ and Factor Ψ each selectively influence a different vertex. Suppose there is a path directed from the vertex selectively influenced by Factor Φ to the vertex selectively influenced by Factor Ψ and on this path there is an arc whose parameters change value when the level of Factor Φ changes. An earlier result (Schweickert & Zheng, 2019a, Theorem 10) considered such an MPT with the restriction that response probabilities $p(i,j)$ are strictly between 0 and 1. With that restriction, such an MPT is equivalent to the Standard Tree for Ordered Processes (Figure 1) or its extension, the K-Class Standard Tree for Ordered Processes (Figure 3). Here we give an alternate derivation that allows response probabilities equal to 0 or 1.

For consistency, notation and reasoning here follow that in Schweickert and Zheng (2019a) as far as possible. When Factor Φ is at level i and Factor Ψ is at level j, the probability of a response in class k is denoted $p(i,j)$, $0 \leq p(i,j) \leq 1$, and $p(i,j)$, is the entry in row i and colunn j of a matrix, $\mathbf{P_k}$. Likewise, the measure produced by a response in class k is denoted $t(i,j)$ and is the entry in row i and column j of a matrix, $\mathbf{T_k}$. Although we usually speak of the measure as time, it could be voltage or payoff, so no assumption is made about the sign of $t(i,j)$.

Theorem 1 (Schweickert & Zheng, 2019a). *Suppose there are K response classes. Suppose for every class k, probability matrix* $\mathbf{P_k}$ *and measure matrix* $\mathbf{T_k}$ *are produced by Factors* Φ *and* Ψ

each selectively influencing a different vertex in a Multinomial Processing Tree with K response classes. Suppose there is a path λ from the source to the vertex selectively influenced by Factor Φ, through an arc whose parameter values change when the level of Factor Φ changes, to the vertex selectively influenced by Factor Ψ. Then for every class k, $\mathbf{P_k}$ *and* $\mathbf{T_k}$ *are produced by Factors Φ and Ψ each selectively influencing vertices ordered by the factors in an equivalent K-class Standard Tree for Ordered Processes, with the vertex selectively influenced by Factor Φ preceding the vertex selectively influenced by Factor Ψ.*

Proof. Suppose in an arbitrary MPT with K response classes Factor Φ selectively influences one vertex and Factor Ψ selectively influences a different vertex. Suppose there is a path λ from the source to the vertex selectively influenced by Factor Φ, through an arc whose parameter values change when the level of Factor Φ changes, to the vertex selectively influenced by Factor Ψ.

A response of class k is made by following a path from the source vertex to a terminal vertex of class k. Such a path can be formed in one of three ways (see Figure 5). (1) Both the vertex selectively influenced by Factor Φ and the vertex selectively influenced by Factor Ψ are on the path. (2) The vertex selectively influenced by Factor Φ is on the path, but the vertex selectively influenced by Factor Ψ is not. (3) Neither of the selectively influenced vertices is on the path. We consider each way in turn. Denote the vertex selectively influenced by Factor Φ as v_1 and the vertex selectively influenced by Factor Ψ as v_2.

(1) Consider a response of class k produced by following a path from the source vertex to a terminal vertex of class k with both v_1 and v_2 on the path. We can divide such a path into parts. There is a single path α from the source vertex to vertex v_1. Denote the probability of this path as p_α and the measure of this path as t_α. Descending from vertex v_1 is an arc e_1 that is on the path to vertex v_2 and whose probability and time parameter depend on the level i of Factor Φ.

Denote the probability of arc e_1 as $p_{e1}(i)$ and the time for arc e_1 as $t_{e1}(i)$. Following the end vertex of arc e_1 there is a single path β to vertex v_2. Denote the probability of this path as p_β and the measure of this path as t_β.

Descending from vertex v_2 are arcs that precede a terminal vertex of class k. Denote these as $f1, \ldots, fq, \ldots, fQ$. (If an arc descending from vertex v_2 does not precede a terminal vertex of class k, it is not included in this list. Note that to avoid a subscript to a subscript, an arc is denoted with two symbols, e.g., fq.) One of these arcs, say fq, is selected to be on the path to a response of class k. Because vertex v_2 is selectively influenced by Factor Ψ, the probability and measure for arc fq may change when the level j of Factor Ψ changes. Denote the probability of arc fq as $p_{fq}(j)$ and the measure for arc fq as $t_{fq}(j)$. (If the probability or measure of an arc fq does not change when the level j changes, the j in the notation is superfluous, but the reasoning is unaffected.) Following the end vertex of arc fq there are paths to a terminal vertex of class k. No arcs on these paths depend on the level i of Factor Φ or the level j of Factor Ψ. Denote the probability over all these paths of reaching a terminal vertex of class k from the end vertex of arc fq as π_{fq} and the measure to reach a terminal vertex of class k as τ_{fq}.

Denote the probability a response of class k is made in way (1), for levels i and j of the factors, as $p_k^{(1)}(i,j)$. Assembling the parts of the paths, we find the probability a response of class k is made in way (1) is, for every i and j,

$$p_k{}^{(1)}(i,j) = p_\alpha p_{e1}(i) p_\beta \sum_{q=1}^{Q} p_{fq}(j)\pi_{fq}.$$

Denote the measure for making a response of class k by way (1) as $t_k^{(1)}(i,j)$. We find the measure for making a response of class k made in way (1) satisfies the following equation, for every i and j,

$$p_k{}^{(1)}(i,j)t_k{}^{(1)}(i,j) = p_\alpha p_{e1}(i) p_\beta \sum_{q=1}^{Q} p_{fq}(j)\pi_{fq}\left[t_\alpha + t_{e1}(i) + t_\beta + t_{fq}(j) + \tau_{fq}\right].$$

(2) Consider a response of class k produced by following a path from the source vertex to a terminal vertex of class k with v_1 on the path but not v_2. We can divide such a path into parts. As before, the single path α from the source vertex to vertex v_1 has probability p_α and measure t_α. Descending from vertex v_1 are arcs that precede a terminal vertex of class k. Denote these as $e1, \ldots, em, \ldots, eM$. One of these arcs, say em, is selected to be on the path to a response of class k. Because vertex v_1 is selectively influenced by Factor Φ, the probability and measure for arc em may change when the level i of Factor Φ changes. Denote the probability of arc em as $p_{em}(i)$ and the measure for arc em as $t_{em}(i)$. (If the probability or measure of an arc em does not change when the level i changes, the i in the notation is superfluous, but the reasoning is unaffected.) Following the end vertex of arc em there may be paths to a terminal vertex of class k that do not go through vertex v_2. No arcs on these paths depend on the level i of Factor Φ or the level j of Factor Ψ. Denote the probability over all these paths of reaching a terminal vertex of class k from the end vertex of arc em as π_{em} and the measure to reach a terminal vertex of class k as τ_{em}.

Denote the probability a response of class k is made in way (2), for levels i and j of the factors, as $p_k^{(2)}(i,j)$. Assembling the parts of the paths relevant to way (2), we find

$$p_k{}^{(2)}(i,j) = p_\alpha \sum_{m=1}^{M} p_{em}(i)\pi_{em}.$$

Denote the measure for making a response of class k in way (2), for levels i and j of the factors, as $t_k^{(2)}(i,j)$. We find the measure for making a response of class k made in way (2) satisfies the following equation, for every i and j,

$$p_k{}^{(2)}(i,j)t_k{}^{(2)}(i,j) = p_\alpha \sum_{m=1}^{M} p_{em}(i)\pi_{em}\,[t_\alpha + t_{em}(i) + \tau_{em}].$$

Note in the above expressions that neither $p_k{}^{(2)}(i,j)$ nor $p_k{}^{(2)}(i,j)p_k{}^{(2)}(i,j)$ depend on the level j.

(3) Consider a response of class k produced by following a path from the source vertex to a terminal vertex of class k with neither v_1 and v_2 on the path. No arc on any such path depends on the level i of Factor Φ or the level j of Factor Ψ. Denote the probability over all such paths of reaching a terminal vertex of class k as p_3 and the measure for reaching a terminal vertex of class k as t_3.

Denote the probability a response of class k is made in way (3), for levels i and j of the factors, as $p_k^{(3)}(i,j)$. Then

$$p_k^{(3)}(i,j) = p_3.$$

Denote the measure to make a response of class k in way (3), for levels i and j of the factors, as $t_k^{(3)}(i,j)$. Then

$$p_k^{(3)}(i,j)t_k^{(3)}(i,j) = p_3t_3.$$

Note in the above expressions that neither $p_k^{(3)}(i,j)$ nor $p_k^{(3)}(i,j)t_k^{(3)}(i,j)$ depend on the level i or j.

The three ways of making a response of class k are mutually exclusive and jointly exhaustive. Hence, the probability of making a response of class k is found by adding the probabilities of making such a response in each way. For every i and j the probability of making a response of class k is

$$\begin{aligned} p(i,j) &= p_k^{(1)}(i,j) + p_k^{(2)}(i,j) + p_k^{(3)}(i,j) \\ &= p_\alpha p_{e1}(i) p_\beta \textstyle\sum_{q=1}^{Q} p_{fq}(j)\pi_{fq} + p_\alpha \textstyle\sum_{m=1}^{M} p_{em}(i)\pi_{em} + p_3. \qquad (4) \end{aligned}$$

Further, for every i and j the measure $t(i,j)$ to make a response of class k satisfies the equation

$$\begin{aligned} p(i,j)t(i,j) &= p_k^{(1)}(i,j)t_k^{(1)}(i,j) + p_k^{(2)}(i,j)t_k^{(2)}(i,j) + p_k^{(3)}(i,j)t_k^{(3)}(i,j) \\ &= p_\alpha p_{e1}(i) p_\beta \textstyle\sum_{q=1}^{Q} p_{fq}(j)\pi_{fq}\left[t_\alpha + t_{e1}(i) + t_\beta + t_{fq}(j) + \tau_{fq}\right] \\ &\quad + p_\alpha \textstyle\sum_{m=1}^{M} p_{em}(i)\pi_{em}\left[t_\alpha + t_{em}(i) + \tau_{em}\right] + p_3\, t_3. \qquad (5) \end{aligned}$$

The above equations are for an arbitrary MPT in which Factors *Φ* and *Ψ* selectively influence two vertices, with the vertex selectively influenced by Factor *Φ* preceding the vertex selectively influenced by Factor *Ψ*. We now find parameter values for an equivalent *K*-Class Standard Tree for Ordered Processes will account for the response probabilities.

Let

$$p_{D.k}(i) = p_\alpha \sum_{m=1}^{M} p_{em}(i)\pi_{em} + p_3. \qquad (6)$$

Let

$$p_B(i) = p_\alpha p_{e1}(i)p_\beta.$$

Note that no parameter on the right side of the equation above is associated with an arc that precedes only terminal vertices of class *k*. Hence, the value of $p_B(i)$ does not depend on the response class *k*.

Let

$$p_{F.k}(j) = \sum_{q=1}^{Q} p_{fq}(j)\pi_{fq}.$$

From Equation (4).

$$p(i,j) = p_{D.k}(i) + p_B(i)p_{F.k}(j).$$

Further, because for every *i* and *j*, $0 \leq p(i,j) \leq 1$, it must be that for every *i* and *j* $0 \leq p_{D.k}(i),\ p_B(i),\ p_{F.k}(j) \leq 1$. Hence the parameter values are suitable as probabilities.

We now find measure parameter values for the arcs of a *K*-Class Standard Tree for Ordered Processes that accounts for response measures and is equivalent to the arbitrary MPT we started with.

If $p_{D.k}(i) = 0$, let $t_{D.k}(i) = 0$.

If $p_{D.k}(i) \neq 0$ let

$$t_{D.k}(i) = \{p_\alpha \sum_{m=1}^{M} p_{em}(i)\pi_{em}[t_\alpha + t_{em}(i) + \tau_{em}] + p_3\, t_3\}/p_{D.k}(i). \qquad (7)$$

Note that if it is possible to make a response of class k by way (3), then the probability p_3 of making such a response by way (3) is greater than 0, so $p_{D.k}(i) \neq 0$ and can be a divisor.

If $p_B(i) = 0$, let $t_B(i) = 0$.

If $p_B(i) \neq 0$ let

$$t_B(i) = \{p_\alpha p_{e1}(i) p_\beta [t_\alpha + t_{e1}(i) + t_\beta]\}/p_B(i).$$

Note that no parameter on the right side of the equation above is associated with an arc that precedes only terminal vertices of class k. Hence, the value of $t_B(i)$ does not depend on the response class k.

If $p_{F.k}(j) = 0$, let $t_{F.k}(j) = 0$.

If $p_{F.k}(j) \neq 0$ let

$$t_{F.k}(j) = \{\sum_{q=1}^{Q} p_{fq}(j) \pi_{fq} [t_{fq}(j) + \tau_{fq}]\}/p_{F.k}(j).$$

With these assignments, for every i and j, from Equation (5)

$$p_{D.k}(i) t_{D.k}(i) + p_B(i) p_{F.k}(j) [t_B(i) + t_{F.k}(j)] = p(i,j) t(i,j).$$

Then for every class k, $\mathbf{P_k}$ and $\mathbf{T_k}$ are produced by Factors Φ and Ψ each selectively influencing vertices ordered by the factors in an equivalent K-class Standard Tree for Ordered Processes, with the vertex selectively influenced by Factor Φ preceding the vertex selectively influenced by Factor Ψ.

QED

The conclusion is straightforward. Suppose two factors each selectively influence a different vertex in an arbitrary MPT, and there is a path from one vertex to the other with an arc on the path influenced by one of the factors. Then there is an equivalent K-Class Standard Tree for Ordered Processes that accounts for both response probabilities and response measures.

Response measures do not reveal anything about the form of the equivalent MPT that is not already revealed by the response probabilities.

Remarks. Suppose measures associated with the arcs are nonnegative, as times are. Then all quantities $t_\alpha, t_{em}(i), \tau_{em}, t_3,\ t_\beta, t_{fq}(j),$ and τ_{fq} are nonnegative. It follows immediately from assignments that for every *i*, $t_{D.k}(i), t_B(i) \geq 0$ and for every *j*, $t_{F.k}(j) \geq 0$; that is, these measures are nonnegative also.

In an arbitrary MPT, a response in class *k* may be made in way 3, via a path from the source to a terminal vertex of class *k* with no vertex selectively influenced by either Factor *Φ* or Factor *Ψ* on the path. No such path is needed in an equivalent *K*-Class Standard Tree for Ordered Processes. The probability associated with such paths can simply be added to the probability of arc *D.k* as in Equation (6). The measures associated with such paths can be included in the time of arc *D.k* as in Equation (7).

Does Response Time Reveal the Order of Selectively Influenced Vertices?

Suppose each of two factors selectively influences a different vertex in an MPT. If the vertices are ordered, a further question arises: Which vertex is first? For the order to be reversible, response probabilities must satisfy certain conditions and response times must satisfy additional conditions. Response times may provide information in addition to that provided by response probabilities. We consider the conditions in turn for two response classes.

Response probabilities. Suppose Factors *Φ* and *Ψ* selectively influence two vertices in the Standard Tree for Ordered Processes. Suppose Factor *Φ* selectively influences the source vertex. Then for every level *i* of Factor *Φ* and every level *j* of Factor *Ψ* there are probabilities $p_D(i)$, $p_B(i)$ and $p_F(j)$ such that the probability of a correct response is

$$p(i,j) = p_D(i) + p_B(i)p_F(j). \quad (8)$$

Now suppose there is another Standard Tree for Ordered Processes in which the factors influence vertices in the reverse order. That is, in this second MPT Factor Ψ selectively influences the source vertex. Then for every level i of Factor Φ and every level j of Factor Ψ there are probabilities $p^*_D(j)$, $p^*_B(j)$ and $p^*_F(i)$ such that the probability of a correct response is

$$p(i,j) = p^*_D(j) + p^*_B(j)p^*_F(i). \qquad (9)$$

The equations above lead to conditions for both orders to be possible (Schweickert & Chen, 2008). A qualitative condition can be quickly checked. Put probabilities $p(i,j)$ in a matrix with rows indexed by i and columns indexed by j. Then it must be possible to permute the rows and columns so that $p(i,j)$ monotonically increases across columns and monotonically increases down rows. To see this, order the levels j of Factor Ψ so that if $j < j'$ then $p_F(j) \leq p_F(j')$. Then by Equation (8) for every level i of Factor Φ, $p(i,j) \leq p(i,j')$. Likewise, order the levels i of Factor Φ so that if $i < i'$ then $p^*_F(i) \leq p^*_F(i')$. Then by Equation (9) for every level j of Factor Ψ, $p(i,j) \leq p(i',j)$. As an example, in Table 1 probabilities of a correct response in the experiment of Dhir (2017) increase across rows and down columns.

A quantitative condition is necessary and sufficient for two orders of the selectively influenced vertices to be possible (Schweickert & Chen, 2008). Order the levels i and j as described above. Parameter values can be transformed to convenient values for which for $i = 1$, $p^*_F(1) = 0$ and for $j = 1$, $p_F(1) = 0$. Using Equations (8) and (9) we find

$$\begin{aligned} &p(i,j) - p(1,j) - p(i,1) + p(1,1) \\ &\qquad = p_B(i)p_F(j) - p_B(1)p_F(j) = p^*_B(j)p^*_F(i) - p^*_B(1)p^*_F(i). \end{aligned}$$

Then for every i and j,

$$[p_B(i) - p_B(1)]p_F(j) = [p^*_B(j) - p^*_B(1)]p^*_F(i), \qquad (10)$$

so

$$\frac{p_F(j)}{p^*{}_B(j) - p^*{}_B(1)} = \frac{p^*{}_F(i)}{p_B(i) - p_B(1)},$$

provided denominators are not 0.

Expressions on the left side of the above equation depend only on j and do not change when i changes. Hence, the right side must take the same value for every value of i. That is, there is a constant c such that for every i and j

$$c = p_F(j)/[p^*{}_B(j) - p^*{}_B(1)] = p^*{}_F(i)/[p_B(i) - p_B(1)], \qquad (11)$$

provided denominators are not 0.

Consider the case of a 0 denominator, say for i', $p_B(i') - p_B(1) = 0$. Then by Equation (10), $0 = [p^*{}_B(j) - p^*{}_B(1)]p^*{}_F(i')$. By our general assumptions, it is not true that $p^*{}_B(j) - p^*{}_B(1) = 0$ for every j. So $p^*{}_F(i') = 0$. Likewise, if $p^*{}_B(j) - p^*{}_B(1) = 0$ then $p_F(j) = 0$. Then we can generalize Equation (11) to every i and every j,

$$p_F(j) = c[p^*{}_B(j) - p^*{}_B(1)] \text{ and } p^*{}_F(i) = c[p_B(i) - p_B(1)]. \qquad (12)$$

For response probabilities the existence of a constant c such that the above equation is true turns out to be the necessary and sufficient for the order in which the vertices selectively influenced by Factors Φ and Ψ are reversible in the Standard Tree for Ordered Processes (Schweickert & Chen, 2008, Theorem 13).

The constant c must be between upper and lower bounds, see Schweickert and Chen (2008) for details. For our purposes, note that $p_F(j) \geq 0$ and by Equation (12) with the order we have chosen for j, $p^*{}_B(j) \geq p^*{}_B(1)$. Then $c \geq 0$. But $c = 0$ is not possible, because then for every j, $p_F(j) = 0$ and then Factor Ψ is ineffective for response probabilities. Hence, $c > 0$.

If there is a number c such that Equation (12) is true, then response probabilities do not determine the order in which the factors influence vertices. Can response times settle the question?

Response times. Assume as before that Factors Φ and Ψ selectively influence two vertices in each of two Standard Trees for Ordered Processes. In one MPT Factor Φ selectively influences the source vertex and in the other MPT Factor Ψ selectively influences the source vertex. Suppose both MPTs account for response probabilities and response times.

As before, for response probabilities for every level i of Factor Φ and every level j of Factor Ψ there are arc probabilities $p_D(i)$, $p_B(i)$ and $p_F(j)$ such that Equation (8) is true, and there are arc probabilities $p^*_D(j)$, $p^*_B(j)$ and $p^*_F(i)$ such that Equation (9) is true. Further, for response times, for every level i of Factor Φ and every level j of Factor Ψ there are arc times $t_D(i)$, $t_B(i)$ and $t_F(j)$ such that the time $t(i,j)$ for a correct response satisfies the equation

$$p(i,j)t(i,j) = p_D(i)t_D(i) + p_B(i)p_F(j)[t_B(i) + t_F(j)]. \qquad (13)$$

And, for every level i of Factor Φ and every level j of Factor Ψ there are arc times $t^*_D(j)$, $t^*_B(j)$ and $t^*_F(i)$ such that the time for a correct response satisfies the equation

$$p(i,j)t(i,j) = p^*_D(j)t^*_D(j) + p^*_B(j)p^*_F(i)[t^*_B(j) + t^*_F(i)]. \qquad (14)$$

Table 2 gives an example of probabilities and times for correct responses that can be accounted for by two MPTs in which each of two factors selectively influences a different vertex. In one MPT the vertex selectively influenced by Factor Φ comes first, in the other MPT it is second.

The interested reader can check that Eq. (8) for probabilities and Eq. (13) for times are satisfied when the vertex selectively influenced by Factor Φ comes first in an MPT with parameter values $p_D(1) = .20$, $p_D(2) = .22$, $p_D(3) = .24$, $p_B(1) = .3$, $p_B(2) = .4$, $p_B(3) = .5$, $p_F(1) = 0$, $p_F(2) = .2$, $p_F(3) = .4$, $t_D(1) = 60$, $t_D(2) = 80$, $t_D(3) = 100$, $t_B(1) = 90$, $t_B(2) = 80$, $t_B(3) = 75$, $t_F(1) = 100$, $t_F(2) = 80$, $t_F(3) = 50$. Also, Eq. (9) and Eq. (14) are satisfied when the vertex selectively influenced by Factor Φ comes second in an MPT with parameter values $p^*_D(1) = .20$, $p^*_D(2) =$

.26, $p^*_D(3) = .32$, $p^*_B(1) = .2$, $p^*_B(2) = .4$, $p^*_B(3) = .6$, $p^*_F(1) = 0$, $p^*_F(2) = .1$, $p^*_F(3) = .2$, $t^*_D(1) =60$, $t^*_D(2) = 81.92$, $t^*_D(3) = 82.50$, $t^*_B(1) = 110$, $t^*_B(2) = 90$, $t^*_B(3) = 60$, $t^*_F(1)= 12$, $t^*_F(2) = 15.1$, $t^*_F(3) = 18.3$.

Two such MPTs are possible only under certain conditions. Table 3 gives an example of probabilities and times for correct responses that can be accounted for by an MPT in which two factors selectively influence ordered vertices, but the order of the vertices cannot be reversed. To see this, we derive qualitative conditions that are violated by the reverse order.

Consider an MPT in which a Factor Φ, with levels indexed by i, selectively influences a vertex followed by a vertex selectively influenced by a Factor Ψ, with levels indexed by j. Then Eq. (8) for probabilities and Eq. (13) for times both hold. Renumber the levels j, if necessary, so $p_F(j)$ monotonically increases with j. Conveniently, one can always find parameters such that $p_F(1) = 0$ (Schweickert & Chen, 2008). Then from Eq. (8), for any i and j,

$$p(i,j) - p(i,1) = p_D(i) + p_B(i)p_F(i) - p_D(i) = p_B(i)p_F(j).$$

Likewise from Eq. (13), for any i and j,

$$p(i,j)t(i,j) - p(i,1)t(i,1) = p_D(i)t_D(i) + p_B(i)p_F(j)[t_B(i) + t_F(j)] - p_D(i)t_D(i) = p_B(i)p_F(j)[t_B(i) + t_F(j)].$$

From the above two equations,

$$\frac{p(i,j)t(i,j) - p(i,1)t(i,1)}{p(i,j) - p(i,1)} = t_B(i) + t_F(j),$$

provided the denominator is not 0.

One consequence is that if $t_B(i)$ and $t_F(j)$ are times, the expression on the left hand side of the above equation must be nonnegative. Another consequence is that over the values of i and j it must be possible to order expressions on the left hand side so they increase monotonically with i and increase monotonically with j, whether or not $t_B(i)$ and $t_F(j)$ are nonnegative. For the example in Table 3, expressions on the left hand side are in the top panel of Table 4; they satisfy the conditions. However, if the order of the selectively influenced vertices is reversed, the roles of i and j in the above equation change and it becomes

$$\frac{p(i,j)t(i,j) - p(1,j)t(1,j)}{p(i,j) - p(1,j)} = t^*_B(j) + t^*_F(i).$$

For the example in Table 3, expressions on the left hand side are in the bottom panel of Table 4; they do not satisfy the conditions. Some are negative. Further, expressions in the first column increase with i but expressions in the other two columns decrease with i. No MPT is possible in which measures $t(i,j)$ can be accounted for by an MPT in which a vertex selectively influenced by Factor Ψ precedes a vertex selectively influenced by Factor Φ, although such an MPT can account for probabilities.

The interested reader can check that arc times producing the numbers in Table 3 are $t_D(1) = 200$, $t_D(2) = 50$, $t_D(3) = 100$, $t_B(1) = 100$, $t_B(2) = 1000$, $t_B(3) = 150$, $t_F(1) = 0$, $t_F(2) = 300$, $t_F(3) = 500$. Arc probabilities are the same as those used for Table 2.

We have discussed qualitative conditions. The following theorem gives quantitative conditions. Parameters for two such MPTs are possible only if the arc probabilities of one MPT are related to those of the other MPT via a constant c, and time parameters of one MPT are related to those of the other MPT via a constant e.

Details about the parameters are somewhat complicated, but the key idea is relatively simple. One can easily see that two factors selectively influence two ordered vertices in an MPT, but the order of the vertices is not determined, if the MPT is equivalent to a tree with the special form in Figure 6. In this tree for every i and j there are parameters α, β_i, and γ_j such that

$$p(i,j) = \alpha + \beta_i\gamma_j$$

and there are parameters δ, ε_i, and ζ_j such that

$$p(i,j)t(i,j) = \alpha\delta + \beta_i\gamma_j(\varepsilon_i + \zeta_j).$$

In the above equations parameters with subscripts i and j commute, so the equations do not specify an order for the two factors indexed by i and j. It turns out, according to the following theorem, that the only way two factors selectively influence two ordered vertices in an MPT, but the vertex order is not determined, is when the MPT is equivalent to the commutative tree in Figure 6.

We note that the tree in Figure 6 is not necessarily an MPT because parameters α, β_i, and γ_j might not be probabilities. Further, parameters δ, ε_i, and ζ_j might not be times. For example, α might negative, β_i might be greater than 1 for some j, or δ might be negative.

Although we have usually considered response time as a dependent variable to account for, processing trees are also used as models for payoffs, lengths, and other dependent variables that need not be nonnegative. In what follows we do not restrict $t(i,j)$ to be nonnegative, as it would be if it were response time, and we refer to it as a measure.

To prove the following theorem, we need to know when probabilities $p(i,j)$ and measures $t(i,j)$ can be produced by the Standard Tree for Ordered Processes with the vertex selectively influenced by Factor Φ preceding the vertex selectively influenced by Factor Ψ. Three conditions, from Schweickert and Zheng (2019, Theorem 5) are sufficient (and necessary).

1. The levels of Factor Ψ can be numbered so $j > j'$ implies that for every i,

$$p(i,j) \geq p(i,j'). \tag{15}$$

2. There are levels i^* and j^* and for every level i of Factor Φ there is a number $r_i \geq 0$ such that for every level j of Factor Ψ

$$p(i,j) - p(i,j^*) = r_i[p(i^*,j) - p(i^*,j^*)]. \tag{16}$$

3. There is a level n of Factor Ψ and for every level i of Factor Φ there is a number s_i such that the following is true. Let $\max\{r_i\} = r_h$. For every i and j,

$$r_h r_i s_i[p(h,j) - p(h,n)]$$

$$= r_h[p(i,j)t(i,j) - p(i,n)t(i,n)] - r_i[p(h,j)t(h,j) - p(h,n)t(h,n)]. \quad (17)$$

If the two factors selectively influence two ordered vertices, but the order is the reverse of that considered above, then Equations (15), (16), and (17) hold, but with the roles of i and j switched.

The following theorem gives necessary and sufficient conditions for Factors Φ and Ψ to selectively influence two ordered vertices in two MPTs, with the vertex selectively influenced by Factor Φ being first in one MPT and the vertex selectively influenced by Factor Ψ being first in the other MPT.

Theorem 2. *For every level i of Factor Φ and for every level j of Factor Ψ denote the probability and measure for a correct response as* $p(i,j)$ *and* $t(i,j)$, *respectively. The following three statements are equivalent.*

(1) *For every i and j*, $p(i,j)$ *and* $t(i,j)$ *are produced by Factors Φ and Ψ selectively influencing vertices ordered by the factors in two MPTs. In one MPT the vertex selectively influenced by Factor Φ precedes the vertex selectively influenced by Factor Ψ and in the other MPT the order is reversed.*

(2) *For every level i of Factor Φ there are probabilities* $p_B(i)$, $p_D(i)$, *and values* $t_B(i)$, *and* $t_D(i)$, *and for every level j of Factor Ψ there are probabilities* $p_F(j)$ *and values* $t_F(j)$ *such that Equations* (8) *and* (13) *are true. For every level i of Factor Φ there are further probabilities* $p^*_F(i)$, *and values* $t^*_F(i)$, *and for every level j of Factor Ψ there are further probabilities* $p^*_B(j)$, $p^*_D(j)$, *and values* $t^*_B(j)$, *and* $t^*_F(j)$ *such that Equations* (9) *and* (14) *are true.*

The parameters satisfy the following conditions.

There is a number c such that for every i and j

$$p_F(j) = c[p^*_B(j) - p^*_B(1)] \text{ and } p^*_F(i) = c[p_B(i) - p_B(1)].$$

There is a number e such that for every i

$$e = \frac{p_B(i)t_B(i) - p_B(1)t_B(1)}{p_B(i) - p_B(1)} - t^*_F(i) \tag{18}$$

and for every j

$$e = \frac{p^*_B(j)t^*_B(j) - p^*_B(1)t^*_B(1)}{p^*_B(j) - p^*_B(1)} - t_F(j), \tag{19}$$

where each equation holds provided the denominator is not 0.

If $p_B(i) = p_B(1) \neq 0$, then $t_B(i) = t_B(1)$. *If* $p^*_B(j) = p^*_B(1) \neq 0$, then $t^*_B(j) = t^*_B(1)$.

Further, $c \neq 0$ *and*

if $c < 0$, *then* $\frac{1}{(\min\{p_B(i) - p_B(1)\})} \leq c$ *and* $\frac{1}{\min\{p^*_B(j) - p^*_B(1)\}} \leq c$

if $c > 0$, *then* $c \leq 1/\max\{p_B(i) - p_B(1)\}$ *and* $c \leq 1/\max\{p^*_B(j) - p^*_B(1)\}$.

(3) *For every i and j, p(i,j) and t(i,j) are produced by Factors Φ and Ψ selectively influencing vertices in a commutative tree in which* β_i *has the same sign for all i and* γ_j *has the same sign for all j.*

Proof: I. Suppose Statement (1) is true. We show that Statement (2) is true. Consider the MPT in which the vertex selectively influenced by Factor *Φ* precedes the vertex selectively influenced by Factor *Ψ*. This MPT is equivalent to a Standard Tree for Ordered Processes in which Factor *Φ* selectively influences the source vertex (as in Figure 1). With this Standard Tree, for every level *i* of Factor *Φ* there are parameter values $p_B(i)$, $p_D(i)$, $t_B(i)$, and $t_D(i)$, and for every level *j* of Factor *Ψ* there are parameter values $p_F(j)$ and $t_F(j)$ such that Equations (8) and (13) are true.

Order the levels j of Factor Ψ so that if $j < j'$ then $p_F(j) \leq p_F(j')$. Transform the arc probability values (if necessary) so for $j = 1$, $p_F(1) = 0$. (This is always possible, Schweickert & Chen, 2008.)

Likewise, the MPT in which the vertex selectively influenced by Factor Ψ precedes the vertex selectively influenced by Factor Φ is equivalent to a Standard Tree for Ordered Processes in which Factor Ψ selectively influences the source vertex. With this Standard Tree, for every level i of Factor Φ there are parameter values $p^*_F(i)$, and $t^*_F(i)$, and for every level j of Factor Ψ there are parameter values $p^*_B(j)$, $p^*_D(j)$, $t^*_B(j)$, and $t^*_D(j)$ such that Equations (9) and (14) are true.

Order the levels i of Factor Φ so that if $i < i'$ then $p^*_F(i) \leq p^*_F(i')$. Transform the arc probability values (if necessary) so for $i = 1$, $p^*_F(1) = 0$.

From Equation (12) there is a number c such that for every i and j

$$p_F(j) = c[p^*_B(j) - p^*_B(1)] \text{ and } p^*_F(i) = c[p_B(i) - p_B(1)]$$

(Schweickert & Chen, 2008). If $c = 0$, then $p_F(j) = 0$ for every j. But then by Eq. (8) $p(i,j)$ does not change when j changes, contrary to the assumption that each factor is effective. Hence, $c \neq 0$.

The lower and upper bounds on c follow from a little algebra and the requirement that $p_F(j)$ and $p^*_F(i)$ are probabilities.

Using Equations (13) and (14), noting that their left hand sides are the same, we find

$$\begin{aligned} &p(i,j)t(i,j) - p(1,j)t(1,j) - p(i,1)t(i,1) + p(1,1)t(1,1) \\ &\quad = p_B(i)p_F(j)[t_B(i) + t_F(j)] - p_B(1)p_F(j)[t_B(1) + t_F(j)] \\ &\quad = p^*_B(j)p^*_F(i)[t^*_B(j) + t^*_F(i)] - p^*_B(1)p^*_F(i)[t^*_B(1) + t^*_F(i)]. \end{aligned}$$

Then

$[p_B(i)t_B(i) - p_B(1)t_B(1)]p_F(j) + [p_B(i) - p_B(1)]p_F(j)t_F(j)$

$= [p^*_B(j)t^*_B(j) - p^*_B(1)t^*_B(1)]p^*_F(i) + [p^*_B(j) - p^*_B(1)]p^*_F(i)t^*_F(i).$ (20)

Consider i such that $p_B(i) = p_B(1)$. Then $p^*_F(i) = c[p_B(i) - p_B(1)] = 0$. Then the equation above becomes $[p_B(1)t_B(i) - p_B(1)t_B(1)]p_F(j) = 0$. This must be true when $p_F(j) \neq 0$, so $p_B(1)[t_B(i) - t_B(1)] = 0$. Then either $p_B(1) = 0$ or $t_B(i) = t_B(1)$.

Likewise, for j such that $p^*_B(j) = p^*_B(1)$, either $p^*_B(1) = 0$ or $t^*_B(j) = t^*_B(1)$.

Now choose i such that $p^*_F(i) \neq 0$ and choose j such that $p_F(j) \neq 0$. Dividing each side of Eq. (20) by $p_F(j)p^*_F(i)$, we obtain

$$\frac{p_B(i)t_B(i) - p_B(1)t_B(1)}{p^*_F(i)} + \frac{p_B(i) - p_B(1)}{p^*_F(i)}t_F(j)$$

$$= \frac{p^*_B(j)t^*_B(j) - p^*_B(1)t^*_B(1)}{p_F(j)} + \frac{p^*_B(j) - p^*_B(1)}{p_F(j)}t^*_F(i).$$

From Equation (12), $p^*_F(i) = c[p_B(i) - p_B(1)]$ and $p_F(j) = c[p^*_B(j) - p^*_B(1)]$. Substituting these values and cancelling c from the two sides of the equation we obtain

$$\frac{p_B(i)t_B(i) - p_B(1)t_B(1)}{p_B(i) - p_B(1)} + t_F(j)$$

$$= \frac{p^*_B(j)t^*_B(j) - p^*_B(1)t^*_B(1)}{p^*_B(j) - p^*_B(1)} + t^*_F(i).$$

Then

$$\frac{p_B(i)t_B(i) - p_B(1)t_B(1)}{p_B(i) - p_B(1)} - t^*_F(i)$$

$$= \frac{p^*_B(j)t^*_B(j) - p^*_B(1)t^*_B(1)}{p^*_B(j) - p^*_B(1)} - t_F(j).$$

The left hand side of the above equation does not depend on j and the right hand side does not depend on i. Hence there is a constant e such that for every i

$$e = \frac{p_B(i)t_B(i) - p_B(1)t_B(1)}{p_B(i) - p_B(1)} - t^*_F(i)$$

and for every j

$$e = \frac{p^*_B(j)t^*_B(j) - p^*_B(1)t^*_B(1)}{p^*_B(j) - p^*_B(1)} - t_F(j),$$

provided the denominators are not 0.

At this point we have shown that Statement (1) implies Statement (2).

II. Suppose Statement (2) is true. We show that Statement (3) is true.

With probabilities described in Statement (2)

$$p(i,j) = p_D(i) + p_B(i)p_F(j) = p^*_D(j) + p^*_B(j)p^*_F(i). \qquad (21)$$

Order the levels i of Factor Φ so that if $i < i'$ then $p^*_F(i) \leq p^*_F(i')$. Transform the arc probability values (if necessary) so for $i = 1$, $p^*_F(1) = 0$. Likewise, order the levels j of Factor Ψ so that if $j < j'$ then $p_F(j) \leq p_F(j')$. Transform the arc probability values (if necessary) so for $j = 1$, $p_F(1) = 0$. (These orderings and transformations are always possible, Schweickert & Chen, 2008.)

In Eq. (21), set $i = 1$. Then because $p^*_F(1) = 0$, $p^*_D(j) = p_D(1) + p_B(1)p_F(j)$.

In Eq. (21), substitute this value for $p^*_D(j)$ and substitute the value $p^*_F(i) = c[p_B(i) - p_B(1)]$ from Statement (2). We obtain

$$p(i,j) = p_D(1) + p_B(1)p_F(j) + p^*_B(j)c[p_B(i) - p_B(1)]$$
$$= p_D(1) + p_B(1)p_F(j) + p^*_B(j)cp_B(i) - p^*_B(j)cp_B(1).$$

From Statement (2), $p^*_B(j) = p_F(j)/c + p^*_B(1)$. Substitute this value for the second occurrence of $p^*_B(j)$ in the equation above for $p(i,j)$. We obtain

$$p(i,j) = p_D(1) - cp_B(1)p^*_B(1) + cp_B(i)p^*_B(j).$$

In the above equation, $p(i,j)$ is produced by a commutative tree. Further, $cp_B(i)$ has the same sign for all i and $p^*_B(j)$ has the same sign for all j.

We turn to the times. With arc probabilities and arc times described in Statement (2),

$$p(i,j)t(i,j) = p_D(i)t_D(i) + p_B(i)p_F(j)[t_B(i) + t_F(j)]$$

$$= p^*_D(j)t^*_D(j) \ + p^*_B(j)p^*_F(i)[t^*_B(j) + t^*_F(i)]. \qquad (22)$$

Levels of the factors can be ordered and probability and time values can be transformed so if $i < i'$ then $p^*_F(i) \le p^*_F(i')$, $p^*_F(1) = 0$, if $j < j'$ then $p_F(j) \le p_F(j')$, and $p_F(1) = 0$. (These orderings and transformations are always possible, Schweickert & Chen, 2008; Schweickert & Zheng, 2019b).

Set $j = 1$ in Eq. (22). Because $p_F(1) = 0$,

$$p_D(i)t_D(i) = p^*_D(j)t^*_D(j) \ + p^*_B(j)p^*_F(i)[t^*_B(j) + t^*_F(i)],$$

so

$$p^*_D(j)t^*_D(j) = p_D(i)t_D(i) - p^*_B(j)p^*_F(i)[t^*_B(j) + t^*_F(i)]. \qquad (23)$$

Similarly, by setting $i = 1$ in Eq. (22), because $p^*_F(1) = 0$, we obtain

$$p_D(i)t_D(i) = p^*_D(j)t^*_D(j) - p_B(i)p_F(j)[t_B(i) + t_F(j)]. \qquad (24)$$

By setting $i = 1$ and j = 1 in Eq. (22), we obtain

$$p_D(1)t_D(1) = p^*_D(1)t^*_D(1).$$

Then by setting $j = 1$ in Eq. (23) and $i = 1$ in Eq. (24),

$$p_D(i)t_D(i) - p^*_B(1)p^*_F(i)[t^*_B(1) + t^*_F(i)] = p^*_D(j)t^*_D(j) - p_B(1)p_F(j)[t_B(1) + t_F(j)].$$

From the above equation we obtain two expressions we will use in substitutions below.

$$p_D(i)t_D(i) = p^*_B(1)p^*_F(i)[t^*_B(1) + t^*_F(i)] + p^*_D(j)t^*_D(j) - p_B(1)p_F(j)[t_B(1) + t_F(j)] \quad (25)$$

and

$$p^*_D(j)t^*_D(j) = p_B(1)p_F(j)[t_B(1) + t_F(j)] + p_D(i)t_D(i) - p^*_B(1)p^*_F(i)[t^*_B(1) + t^*_F(i)]. \quad (26)$$

We now derive an equation for $p(i,j)t(i,j)$, starting from the first equation of Eq. (22). Substitute the value of $p_D(i)t_D(i)$ from Eq. (25).

$$p(i,j)t(i,j) \ = p^*_B(1)p^*_F(i)[t^*_B(1) + t^*_F(i)] + p^*_D(j)t^*_D(j) - p_B(1)p_F(j)[t_B(1) + t_F(j)]$$
$$+ p_B(i)p_F(j)[t_B(i) + t_F(j)].$$

In the above equation, substitute the values of $p^*_F(i)$ and $t^*_F(i)$ given in Statement (2), and the value of $p^*_D(j)t^*_D(j)$ from Eq. (26). After a little algebra, we obtain

$$\begin{aligned} p(i,j)t(i,j) &= p_D(1)t_D(1) - cp_B(1)p^*_B(1)[t_B(1) + t^*_B(1) - \mathrm{e}] \\ &\quad + cp_B(i)p^*_B(1)[t_B(i) + t^*_B(1) - \mathrm{e}] \\ &\quad + p_B(i)p_F(j)[t_B(i) + t_F(j)]. \end{aligned} \tag{27}$$

Finally, in the above equation substitute the values of $p_F(j)$ and $t_F(j)$ given in Statement (2). After a little algebra, we obtain

$$\begin{aligned} p(i,j)t(i,j) &= p_D(1)t_D(1) - cp_B(1)p^*_B(1)[t_B(1) + t^*_B(1) - \mathrm{e}] \\ &\quad + cp_B(i)p^*_B(j)[t_B(i) + t^*_B(j) - \mathrm{e}]. \end{aligned} \tag{28}$$

In the above equation, $p(i,j)t(i,j)$ is produced by a commutative tree.

We now consider the situation where a denominator is 0 in Eq. (18) or Eq. (19). There are several cases to consider.

(1) Suppose $p^*_B(j') = p^*_B(1) \neq 0$ and $p_B(i') \neq p_B(1)$.

Equation (27) holds. From Statement (2), $p^*_B(j') = p^*_B(1) \neq 0$ implies $p_F(j') = 0$ and $t^*_B(j') = t^*_B(1)$.

Then Eq. (27) becomes Eq. (28), the equation of the commutative tree.

(2) Suppose $p^*_B(j') = p^*_B(1) \neq 0$ and $p_B(i') = p_B(1) \neq 0$.

From Statement (2), $p_F(j') = 0$ and $p^*_F(i') = 0$. Recall that $p_F(1) = 0$ and $p^*_F(1) = 0$.

We have

$$p(i',1) = p_D(i') + p_B(i')p_F(1) = p_D(i'),$$

because $p_F(1) = 0$.

Also,

$$p(i',1) = p^*_D(1) + p^*_B(1)p^*_F(i') = p^*_D(1),$$

because $p^*_F(i') = 0$.

Then

$$p_D(i') = p^*_D(1). \tag{29}$$

Analogous reasoning shows

$$p^*_D(j') = p_D(1). \tag{30}$$

For every j,

$$\begin{aligned} p(i',j)t(i',j) &= p^*_D(j)t^*_D(j) + p^*_B(j)p^*_F(i')[\, t^*_B(j) + t^*_F(i')] \\ &= p^*_D(j)t^*_D(j), \qquad (31) \end{aligned}$$

because $p^*_F(i') = 0$.

When $j = 1$ the above equation becomes $p(i',1)t(i',1) = p^*_D(1)t^*_D(1)$.

We also have

$$p(i',1)t(i',1) = p_D(i')t_D(i') + p_B(i')p_F(1)[t_B(i') + t_F(1)] = p_D(i')t_D(i'),$$

because $p_F(1) = 0$.

From the lines immediately above $p_D(i')t_D(i') = p^*_D(1)t^*_D(1)$. Then from Eq. (29), $t_D(i') = t^*_D(1)$.

Analogous reasoning shows, $t^*_D(j') = t_D(1)$.

From Eq. (31) we have from the line above and Eq. (30),

$$p(i',j')t(i',j') = p^*_D(j')t^*_D(j')$$

$$= p_D(1)t_D(1).$$

On the other hand, from Statement (2), $p^*_B(j') = p^*_B(1) \neq 0$ implies $t^*_B(j') = t^*_B(1)$ and $p_B(i') = p_B(1) \neq 0$ implies $t_B(i') = t_B(1)$. Making these substitutions in Eq. (28) we obtain the result above,

$$p(i',j')t(i',j') = p_D(1)t_D(1).$$

Hence, $p(i',j')t(i',j')$ is obtained from the equation for the commutative tree.

Reasoning is similar in the remaining cases.

Hence, Statement (2) implies Statement (3).

III. Suppose Statement (3) is true. We show Statement (1) is true by showing Eq. (15), (16) and (17) hold.

Suppose for every i and j there are parameters α, β_i, and γ_j such that

$$p(i,j) = \alpha + \beta_i\gamma_j.$$

Number the levels j so that $\gamma_1 \leq \gamma_2 \leq \ldots$. Then for every i if $j > j'$ then $p(i,j) \geq p(i,j')$, so Eq. (15) holds.

If $\beta_i = 0$ for every i, Factor Φ is ineffective. Choose h so max $\{|\beta_i|\} = |\beta_h| \neq 0$. For every i and j,

$$p(i,j) - p(i,1) = \beta_i\gamma_j - \beta_i\gamma_1$$

$$= \frac{\beta_i}{\beta_h}\beta_h(\gamma_j - \gamma_1)$$

$$= \frac{\beta i}{\beta h}[p(h,j) - p(h,1)]$$

Let $r_i = \beta_i/\beta_h$. Because β_i has the same sign for all i, $r_i \geq 0$. Then Eq. (16) holds,, with $i^* = h$ and $j^* = 1$.

We turn to Eq. (17). By Statement (3), there are parameters δ, ε_i, and ζ_j such that

$$p(i,j)t(i,j) = \alpha\delta + \beta_i\gamma_j(\varepsilon_i + \zeta_j).$$

Let n be any level j. Note that max $\{r_i\} = r_h = 1$. If $r_i = 0$, let $s_i = 0$, otherwise let $s_i = \varepsilon_i - \varepsilon_1$.

In the case that $r_i = 0$, the left hand side of Eq. (17) is 0. Also, if $r_i = 0$, $\beta_i = 0$. Then $p(i,j)t(i,j) - p(n,j)t(i,n) = \beta_i\gamma_j(\varepsilon_i + \zeta_j) - \beta_i\gamma_n(\varepsilon_i + \zeta_n) = 0$. So the right hand side of Eq. (17) is also 0.

Consider the case that $r_i \neq 0$. The left hand side of Eq. (17) is

$$\frac{\beta_h}{\beta_h}\frac{\beta_i}{\beta_h}[\epsilon_i - \varepsilon_h][\beta_h\gamma_j - \beta_h\gamma_n] = \beta_i[\gamma_j - \gamma_n][\epsilon_i - \varepsilon_h].$$

The right hand side of Eq. (17) is

$$\frac{\beta_h}{\beta_h}[\beta_i\gamma_j(\epsilon_i - \zeta_j) \quad - \beta_i\gamma_n(\epsilon_i - \zeta_n)]$$

$$- \frac{\beta_i}{\beta_h}[\beta_h\gamma_j(\epsilon_h - \zeta_j) \quad - \beta_h\gamma_n(\epsilon_h - \zeta_n)]$$

$$= \beta_i[\gamma_j - \gamma_n][\epsilon_i - \varepsilon_h].$$

Hence, Eq. (17) holds.

Because Eq. (15), (16) and (17) hold $p(i,j)$ and $t(i,j)$ are produced by the Standard Tree for Ordered Processes, with the vertex selectively influenced by the factor whose levels are denoted i preceding the vertex selectively influenced by the factor whose levels are denoted j. Similar reasoning shows the order of the vertices can be reversed. Hence, Statement (3) implies Statement (1).

QED

Remark. Given parameter values for parameters of one of the Standard Trees described in the theorem, values for the numbers c and e, and a value for $p^*_B(1)$ it is straightforward to find parameter values for the other Standard Tree.

Suppose Statement (2) is true. Suppose for every level i of Factor Φ we have values $p_B(i)$, $p_D(i)$, $t_B(i)$, and $t_D(i)$, and for every level j of Factor Ψ we have values $p_F(j)$ and $t_F(j)$. Transform these, if necessary, so $p_F(1) = 0$. Equation (12) directly gives the value $p^*_F(i) = c[p_B(i) - p_B(1)]$. Because $p^*_F(i) \geq 0$, $c \geq 0$. If $c = 0$, the factor whose levels are indexed by i is ineffective. We can divide by c in Equation (12) and we find $p^*_B(j) = p_F(j)/c + p^*_B(1)$. Note that $p^*_F(1) = 0$. Then by Equation (9), $p^*_D(j) = p^*_D(j) + p^*_B(j)p^*_F(1) = p(1,j)$.

We now find values for arc time parameters $t^*_F(j)$, $t^*_B(j)$, and $t^*_D(j)$.

Equation (18) holds for every i for which the denominator is not 0. For every such i Equation (18) can be solved for $t^*_F(i)$. Let

$$t^*_F(i) = \frac{p_B(i)t_B(i) - p_B(1)t_B(1)}{p_B(i) - p_B(1)} - e.$$

Suppose for a level i the denominator of Equation (18) is 0. Then $p_B(i) - p_B(1) = 0$. Then by Equation (12), $p^*_F(i) = 0$. By Equation (14), the value of $p(i,j)t(i,j)$ does not depend on $t^*_F(i)$ and we can let $t^*_F(i) = 0$.

Equation (19) holds for every j for which the denominator is not 0. For every such j it can be solved for $t^*_B(j)$.

$$\begin{aligned} p^*_B(j)t^*_B(j) - p^*_B(1)t^*_B(1) &= [t_F(j) + e][p^*_B(j) - p^*_B(1)] \\ &= [t_F(j) + e]p_F(j)/c. \end{aligned}$$

So

$$t^*_B(j) = \frac{p^*_B(1)t^*_B(1) + p_F(j)[t_F(j)+e]/c}{p^*_B(j)} = \frac{cp^*_B(1)t^*_B(1) + p_F(j)[t_F(j)+e]}{p_F(j) + cp^*_B(1)}.$$

Note that division by $p*_B(j)$ is justified for such a j, because by the order of the levels j, $p*_B(j) \geq p*_B(1)$. If $p*_B(j) = 0$, then $p*_B(1) = 0$. Then the denominator of Equation (19) would be 0 and j would not be in the case we are considering.

Now consider the case of j for which the denominator of Equation (19) is 0. Then $p^*_B(j) - p^*_B(1) = 0$, and by Equation (12), $p_F(j) = 0$. Then solving Equation (18), with $p^*_B(j) = p^*_B(1)$, we find $t^*_B(j) = t^*_B(1)$.

Finally, we obtain the value of $t^*_D(j)$. Above, we found $p^*_D(j) = p(1,j)$. By Equation (13),

$$p(1,j)t(1,j) = p^*_D(j)t^*_D(j) + p^*_B(j)p^*_F(1)[t^*_B(j) + t^*_F(1)] = p^*_D(j)t^*_D(j).$$

Because $p(1,j) = p^*_D(j)$, we let $t^*_D(j) = t(1,j)$.

In this section, we have shown that response times may resolve a question about the form of an MPT that response probabilities do not. Suppose Factors Φ and Ψ selectively influence two vertices in the Standard Tree for Ordered Processes with Factor Φ selectively influencing the source vertex. If there is a number c such that Equation (12) is true, then response probabilities can be accounted for by another Standard Tree for Ordered Processes in which Factors Φ and Ψ selectively influence two vertices, but the order of the selectively influenced vertices is reversed. If further, there is a number e such that Equations (18) and (19) are true, then response times can be accounted for by another Standard Tree for Ordered Processes in which Factors Φ and Ψ selectively influence two vertices, but the order of the selectively influenced vertices is reversed. Because the existence of the number c does not imply the existence of the number e, response times may reveal an aspect of form that response probabilities alone do not.

Conclusion

This paper began with discussion of an experiment by Dhir (2017) on the effect of participant age on learning of word pairs. After study, participants were cued with the first member of a pair and asked to recall the second. Three types of pairs varied in how much support word knowledge affords recall. A Multinomial Processing Tree gives a good account of the response probabilities. In the MPT age and pair-type selectively influence two different ordered vertices. Drawing on an MPT model by Greene and Naveh-Benjamin (in press), we – propose as a possibility that the vertex selectively influenced by age as an attempt at retrieval of a specific verbatim trace and the vertex selectively influenced by pair-type as an attempt at recall via a fuzzier gist trace. Processing at the second vertex is unaffected by age, consistent with earlier findings that word knowledge is not impaired by age. The MPT gives resolution about which processes are impaired by age and which are not. A question unresolved by the response probability data is the order of the vertices. Two orders are possible and both lead to good fits. In the second part of the paper we considered whether observation of an additional measure, such as response time, could resolve the question of order. We showed that indeed an additional measure can imply an order on selectively influenced processes even when response probabilities alone do not.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F.Csaki (Eds.), *Second international symposium on information theory* (pp. 267-281). Budapest: Akademiai Kiado.

Badham, S.P., Estes, Z., & Maylor, E.A. (2012). Integrative and semantic relations equally alleviate age-related associative memory deficits. *Psychology and Aging, 27* (1), 141-152

Badham, S.P. & Maylor, E.A. (2015). What you know can influence what you are going to know (especially for older adults). *Psychonomic Bulletin & Review*, *22* (1), pp. 141-146.

Batchelder, W. H. (1998). Multinomial processing tree models and psychological assessment. *Psychological Assessment, 10,* 331-344.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6*, 57-86.

Bayen, U. J., Phelps, M. P., & Spaniol, J. (2000). Age-related differences in the use of contextual information in recognition memory: A global matching approach. *The Journals of Gerontology: Series B, Psychological Sciences and Social Sciences, 55*, 131–141.

Bishara, A. J., & Payne, B. K. (2009). Multinomial process tree models of control and automaticity in weapon misidentification. *Journal of Experimental Social Psychology, 45,* 524-534.

Bishop, Y. M. M., Feinberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis.* Cambridge: MIT Press.

Brainerd, C. J., Reyna, V. F., & Mojardin, A. H. (1999). Conjoint recognition. *Psychological Review, 106*, 160–179.

Cerella, J., & Hale, S. (1994). The rise and fall in information-processing rates over the life span. *Acta Psychologica,* 109-197.

Chalfonte, B.I., & Johnson, M.K. (1996). Feature memory and binding in young and older adults. *Memory and Cognition, 24* (4), 403-416

Chechile, R. (1977). Storage-retrieval analysis of acoustic similarity. *Memory & Cognition, 5*, 535–540.

Dhir, P. (2017). *Uncovering the mechanisms which underlie the integrative advantage in healthy aging* (Doctoral dissertation, City, University of London). https://openaccess.city.ac.uk/id/eprint/17688/

Dhir, P. & Poirier, M. (2015). The Integrative Advantage in Healthy Aging: How Can It Be Explained? Annual Meeting of the Psychonomic Society, Chicago, USA

Erdfelder, E., Auer, T. S., Hilbig, B. E., Aβfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models. *Zeitschrift für Psychologie/Journal of Psychology, 217*, 108-124.

Estes, Z., Golonka, S., & Jones, L. L. (2011). Thematic thinking: The apprehension and consequences of thematic relations. In B. Ross (Ed.), *Psychology of learning and motivation, vol 54* (pp. 249–294). Burlington. MA: Academic Press.

Fisk, A. D., & Fisher, D. L. (1994). Brindly plots and theories of aging: The explicit, muddled, and implicit debates. *Journal of Gerontology: PSYCHOLOGICAL SCIENCES, 49,* 81-89.

Greene, N. R., & Naveh-Benjamin, M. (in press). A Specificity Principle of Memory: Evidence from Aging and Associative Memory. *Psychological Science.*

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356-388.

Heck, D. W., & Erdfelder, E. (2016). Extending multinomial processing tree models to measure the relative speed of cognitive processes. *Psychonomic Bulletin & Review*, *23*, 1440-1465. DOI 10.3758/s13423-016-1025-6

Heck, D. W., & Erdfelder, E. (2017). Linking process and measurement models of recognition-based decisions. *Psychological Review, 124,* 442-471. http://dx.doi.org/10.1037/rev0000063

Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In E I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 237-278). New York: Plenum.

Hu, X. (2001). Extending General Processing Tree models to analyze reaction time experiments. *Journal of Mathematical Psychology, 45*, 603-634.

Hütter, M., & Klauer, K. C. (2016). Applying processing trees in social psychology. *European Review of Social Psychology, 27,* 116–159. doi:10.1080/10463283.2016.1212966

Klauer, K. C., & Kellen, D. (2018). RT-MPTs: Process models for response-time distributions based on multinomial processing trees with applications to recognition memory. *Journal of Mathematical Psychology, 82,* 111-130.

Klauer, K. C., & Wegner, I. (1998). Unraveling social categorization in the "Who said what" paradigm. *Journal of Personality and Social Psychology, 75,* 1155-1178.

Link, S. W. (1982). Correcting response measures for guessing and partial information. *Psychological Bulletin, 92*, 469-486.

Naveh-Benjamin, M. (2000). Adult-age differences in memory performance: Tests of an associative deficit hypothesis. Journal of Experimental Psychology: Learning, *Memory and Cognition, 26*, 1170-1187.

Naveh-Benjamin, M., & Mayr, U. (2018). Age-related differences in associative memory: Empirical evidence and theoretical perspectives. *Psychology and Aging*, *33*(1), 1-6.

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23*, 104 –118.

Poirier, M., Schweickert, R., & Oliver, J. (2005). Silent reading rate and memory span. *Memory, 13*, 380–387.

Reimer, C. B., Strobach, T., & Schubert, T. (2017). Concurrent deployment of visual attention and response selection bottleneck in a dual-task: Electrophysiological and behavioral evidence. *The Quarterly Journal of Experimental Psychology,* https://doi.org/10.1080/17470218.2016.1245348

Rosenbaum, D. A. (1980). Human movement initiation--specification of arm, direction, and extent. *Journal of Experimental Psychology: General, 109,* 444-474.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society, 16,* 754-760.

Schweickert, R., & Chen, S. (2008). Tree inference with factors selectively influencing processes in a processing tree. *Journal of Mathematical Psychology, 52*, 158–183.

Schweickert, R., Fisher, D. L., & Sung, K. (2012). *Discovering cognitive architecture by selectively influencing mental processes.* Singapore: World Scientific.

Schweickert, R., & Xi, Z. (2011). Multiplicatively interacting factors selectively influencing parameters in multiple response class processing and rate trees. *Journal of Mathematical Psychology, 55*, 348-364.

Schweickert, R., & Zheng, X. (2018). Tree Inference: Selective influence in multinomial processing trees with supplementary measures such as response time. *Journal of Mathematical Psychology, 86,* 10-29.

Schweickert, R., & Zheng, X. (2019a). Tree Inference: Uniqueness of Multinomial Processing Trees representing response time when two factors selectively influence processes. *Journal of Mathematical Psychology, 88*, 58-77. https://doi.org/10.1016/j.jmp.2018.12.001.

Schweickert, R., & Zheng, X. (2019b). Tree Inference: Response time in Multinomial Processing Trees, representation and uniqueness of parameters. *Journal of Mathematical Psychology*, *91*, 51-69.

Schweickert, R., & Zheng, X. (in press). Multinomial processing trees with response times: Changing speed and accuracy by selectively influencing a vertex. *Journal of Mathematical Psychology.* https://doi.org/10.1016/j.jmp.2019.02.005.

Singmann, H., & Kellen, D. (2013). MPTinR: Analysis of multinomial processing tree models in R. *Behavior Research Methods, 45*(2), 560-575. http://doi.org/10.3758/s13428-012-0259-0

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging, 10*, 527–539.

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., & Gordon, B. (2018). Transcranial direct current stimulation (tDCS) facilitates overall visual search response times but does not interact with visual search task factors. *PLoS ONE, 13*, e0194640. https://doi.org/10.1371/journal.pone.0194640

Wollschläger, L. M., & Diederich, A. (2012). The 2*N*-ary choice tree model for *N* alternative preferential choice. *Frontiers in Psychology, 3,* Article 189. doi: 10.3389/fpsyg.2012.00189

Appendix

Model Fitting and Comparisons

When estimating probability values in Equations (1), (2) and (3), two probability values in each equation can be set to arbitrary values before estimating the remaining ones. The reason, in brief, is that some of the probability values are on an interval scale. Such values can be transformed with two free parameters, a slope and intercept. Alternatively, two values can be set arbitrarily. For details see Schweickert and Chen (2008).

There are two levels of age and three levels of pair-type. For the model in Equation (1), there are seven probability values, $p_D(1)$, $p_D(2)$, $p_B(1)$, $p_B(2)$, $p_F(1)$, $p_F(2)$, and $p_F(3)$. To fit the model, two values were set arbitrarily ahead of time, $p_F(1) = 0$ and $p_F(3) = 1$. The other values were then chosen to minimize G^2, by using Solver in Excel. Results are in Table 1, with G^2 = .12.

For the model in Equation (3), there are eight probability values, $p_D(1)$, $p_D(2)$, $p_D(3)$, $p_B(1)$, $p_B(2)$, $p_B(3)$, $p_F(1)$, and $p_F(2)$. To fit the model, two values would be set arbitrarily ahead of time. The six remaining probability values would then be used to fit the six observed values of correct response probability. The fit would be perfect, so the model cannot be rejected on the basis of goodness of fit.

We can compare the models in Equations (1) and (3) by taking the number of parameters into consideration with the Akaike Information Criterion (Akaike, 1973). For a particular model,

$$AIC = G^2 + 2S,$$

where S is the number of estimated parameters of the model (see, e.g., Singmann & Kellen, 2013). For the model in Equation (1),

$$AIC = .12 + 2(5) = 10.12.$$

For the model in Equation (3), the perfect fit indicates $G^2 = 0$. So for this model,

$$AIC = 0 + 2(6) = 12.$$

The slightly lower AIC for the model in Equation (1) leads to preferring it, although not strongly.

For the model in Equation (2), there are five probability values, p_A, $p_D(1)$, $p_D(2)$, $p_F(1)$, $p_F(2)$, and $p_F(3)$. To fit the model, two values were set arbitrarily ahead of time, $p_F(1) = 0$ and $p_F(3) = 1$. The other values were then chosen to minimize G^2, by using Solver in Excel. For this model, $G^2 = 3.63$. The distribution of G^2 is approximately chi square, with 2 degrees of freedom in this case. The obtained value is not significant, indicating a good fit. An argument against this model is that it predicts additive effects of Age and Pair-Type, contrary to the interaction found with an ANOVA. For comparing this model with the other two,

$$AIC = 3.63 + 2(4) = 11.63.$$

Although the *AIC* values are not very different from each other, the smallest *AIC* is that for the model in Equation (1), so it is the preferred model, although not strongly preferred.

Table 1

Correct and Incorrect Recall Frequencies of Young and Old Participants For Three Types of Paired Associates in Experiment 1 of Dhir (2017)

Age i			Pair-Type j		
			Unrelated $j = 1$	Integrative Non-Assoc $j = 2$	Integrative Assoc $j = 3$
			Correct Recall Frequency		
Old	$i = 1$	Observed	90	150	198
		Predicted	90.61	148.45	198.74
Young	$i = 2$	Observed	162	201	240
		Predicted	161.01	202.87	239.26
			Incorrect Recall Frequency		
Old	$i = 1$	Observed	210	150	102
		Predicted	209.39	151.55	101.26
Young	$i = 2$	Observed	138	99	60
		Predicted	138.99	97.13	60.74

Note. Predicted values are from Equation (1) with the following parameter values: $p_D(1) = .30$, $p_B(1) = .36$; $p_D(2) = .54$, $p_B(2) = .26$; $p_F(1) = 0$, $p_F(2) = .53$, $p_F(3) = 1$.

Table 2

Numerical Example of Response Probabilities and Times Predicted by Factors Φ and Ψ Selectively Influencing Ordered Vertices in an MPT But Order is Not Determined

Level of Factor Φ	Level of Factor Ψ		
	$j = 1$	$j = 2$	$j = 3$
		Probability $p(i, j)$	
$i = 1$	0.20	0.26	0.32
$i = 2$	0.22	0.30	0.38
$i = 3$	0.24	0.34	0.44
		$p(i,j)t(i, j)$	
$i = 1$	12.0	21.3	26.4
$i = 2$	15.1	26.7	32.7
$i = 3$	18.3	32.3	39.3

Table 3

Numerical Example of Response Probabilities and Times Predicted by Factors Φ and Ψ Selectively Influencing Ordered Vertices in an MPT And Order is Determined

Level of Factor Φ	Level of Factor Ψ		
	$j = 1$	$j = 2$	$j = 3$
		Probability $p(i, j)$	
$i = 1$	0.20	0.26	0.32
$i = 2$	0.22	0.30	0.38
$i = 3$	0.24	0.34	0.44
		$p(i,j)t(i, j)$	
$i = 1$	40	64	112
$i = 2$	11	115	251
$i = 3$	24	69	154

Table 4

Values of $t_B(i) + t_F(j)$ and $t^*_B(j) + t^*_F(i)$ Derived from Table 3

In Top Panel Factor Φ Selectively Influences the Source Vertex: Possible

In Bottom Panel Factor Ψ Selectively Influences the Source Vertex: Impossible

Level of Factor Φ	Level of Factor Ψ $j = 1$	$j = 2$	$j = 3$
		$t_B(i) + t_F(j)$	
$i = 1$		400	600
$i = 2$		1300	1500
$i = 3$		450	650
		$t^*_B(j) + t^*_F(i)$	
$i = 1$			
$i = 2$	-1450.0	1275.0	2316.7
$i = 3$	-400.0	62.5	350.0

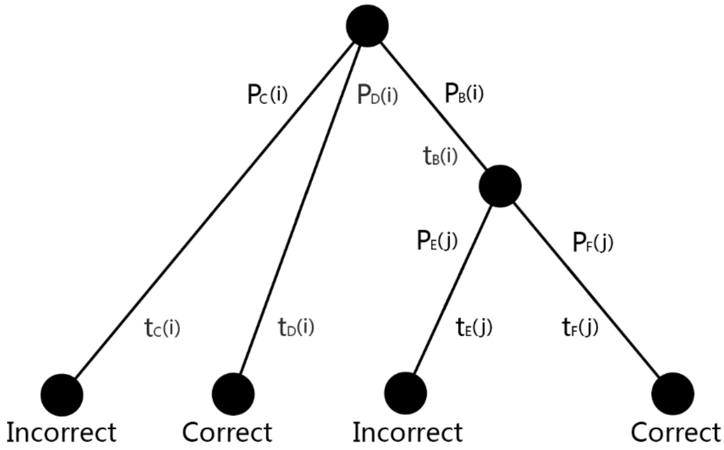

Figure 1. A Multinomial Processing Tree for an experiment with two factors. Probability values $p_D(i)$ and $p_B(i)$ depend on the level i of one of the factors and probability value $p_F(j)$ depends on the level j of the other factors. The vertex with probabilities indexed by i precedes the vertex with probabilities indexed by j on a path. A parameter such as $t_B(i)$ on an arc is the time required for the outcome represented by the arc to occur. A terminal vertex results in a Correct response an Incorrect response.

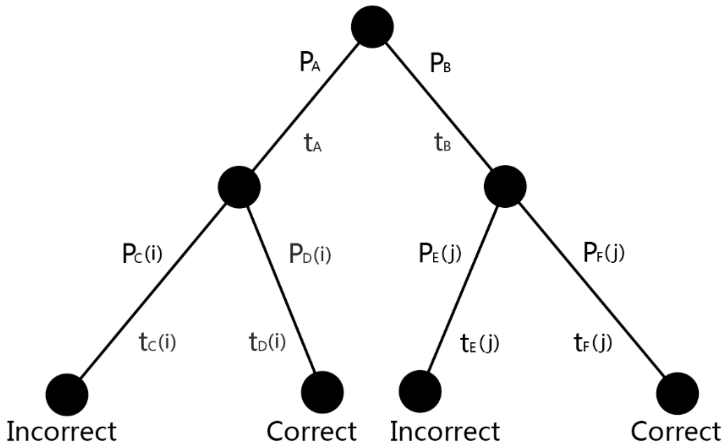

Figure 2. A Multinomial Processing Tree for an experiment with two factors. Probability value $p_D(i)$ depends on the level i of one of the factors and probability value $p_F(j)$ depends on the level j of the other factor. The vertex with probabilities indexed by i is not on a path with the vertex with probabilities indexed by j.

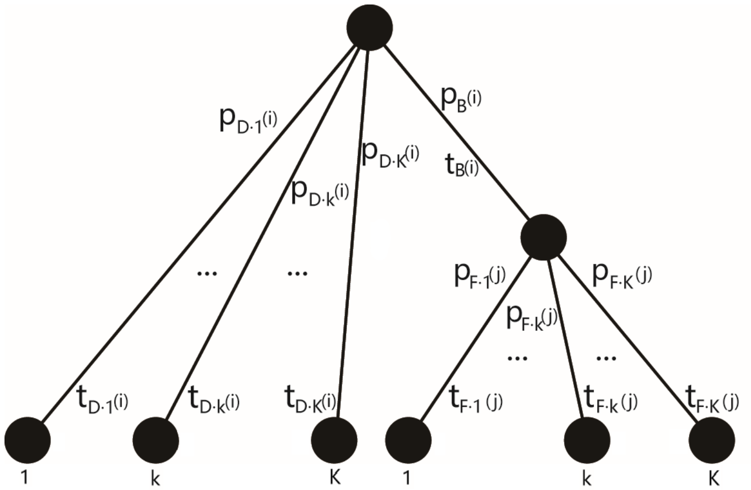

Figure 3. The *K*-Class Standard Tree for Ordered Processes

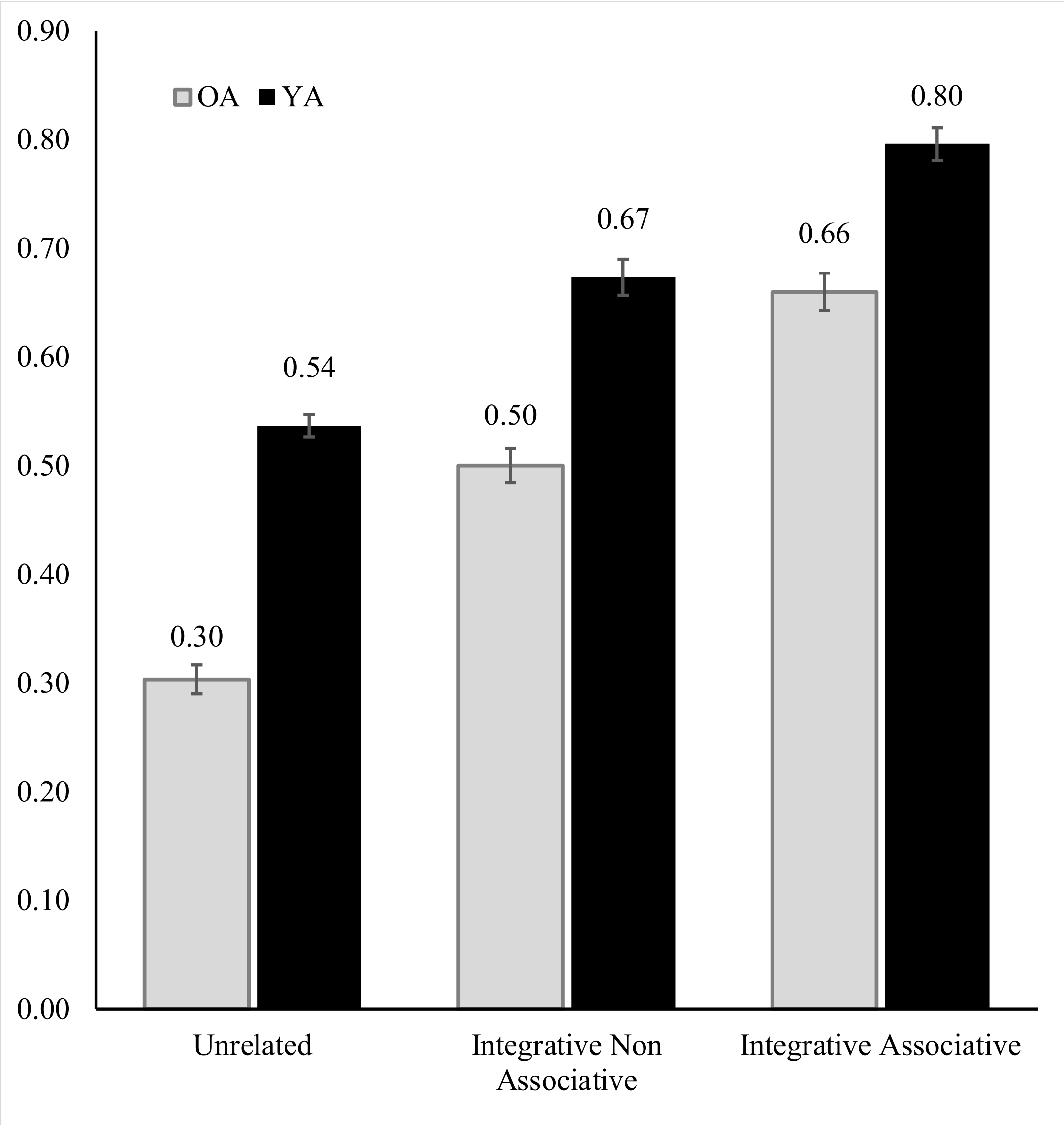

Figure 4. Proportion of integrative associative, integrative non-associative and unrelated targets correctly recalled across young and old adults. Error bars represent means to ±1 standard error.

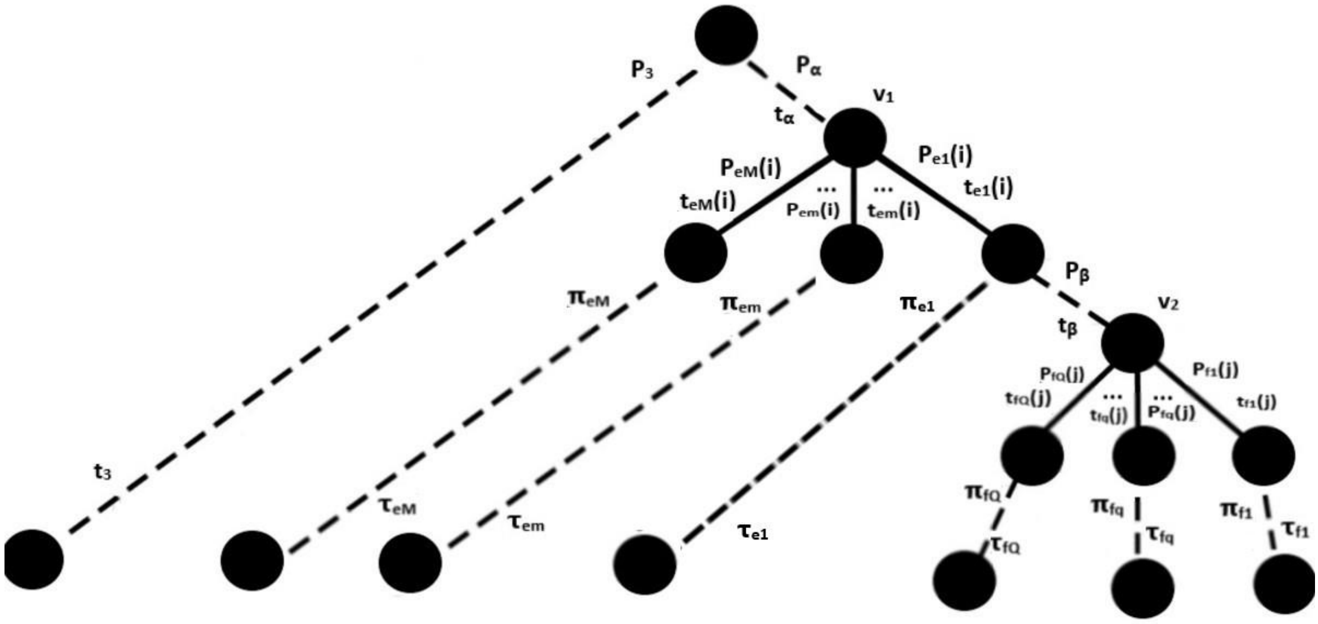

Figure 5. Paths from the source of an arbitrary MPT to a terminal vertex. A factor with levels indexed by *i* selectively influences one vertex and a factor with levels indexed by *j* selectively influences a following vertex. All possible paths to a vertex of one particular response class are illustrated.

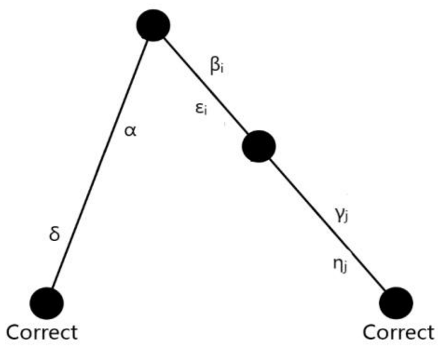

Figure 6. The commutative tree. Only terminal vertices for correct responses are illustrated. One arc has parameters indexed by *i*; another arc has parameters indexed by *j*. The order of these arcs makes no difference for predictions.