

DERIVING ULTRAMETRIC TREE STRUCTURES FROM PROXIMITY DATA CONFOUNDED BY DIFFERENTIAL STIMULUS FAMILIARITY

WAYNE S. DESARBO, RABIKAR CHATTERJEE, AND JUYOUNG KIM

SCHOOL OF BUSINESS ADMINISTRATION
THE UNIVERSITY OF MICHIGAN

This paper presents a new procedure called TREEFAM for estimating ultrametric tree structures from proximity data confounded by differential stimulus familiarity. The objective of the proposed TREEFAM procedure is to quantitatively “filter out” the effects of stimulus unfamiliarity in the estimation of an ultrametric tree. A conditional, alternating maximum likelihood procedure is formulated to simultaneously estimate an ultrametric tree, under the unobserved condition of complete stimulus familiarity, and subject-specific parameters capturing the adjustments due to differential unfamiliarity. We demonstrate the performance of the TREEFAM procedure under a variety of alternative conditions via a modest Monte Carlo experimental study. An empirical application provides evidence that the TREEFAM outperforms traditional models that ignore the effects of unfamiliarity in terms of superior tree recovery and overall goodness-of-fit.

Key words: hierarchical clustering, maximum likelihood estimation, familiarity, consumer psychology.

1. Introduction

An ultrametric or hierarchical tree can be defined as a rooted tree in which a positive value is assigned to each node such that (a) the terminal nodes have zero value, (b) the root possesses the largest value, and (c) the values attached to the nodes on the path from any terminal node to the root constitute a strictly increasing sequence (De Soete, 1984). Such ultrametric trees are quite useful for representing the discrete structure in proximity data since they define a hierarchical clustering on the stimulus set. Hartigan (1967), Jardine, Jardine, and Sibson (1967), and Johnson (1967) have all independently shown that a necessary and sufficient condition for the existence of an exact ultrametric tree representation is the ultrametric inequality. Letting $\Delta = ((\delta_{jk}))$ be a square symmetric matrix containing the empirical pairwise nonnegative dissimilarities between N stimuli, the ultrametric inequality holds if

$$\delta_{jk} \leq \max(\delta_{jl}, \delta_{kl})$$

for all j, k, l triples, and an exact ultrametric tree representation can thus be constructed. Given that such empirical proximities rarely satisfy this property, Hartigan (1967), Chandon, Lemaire, and Pouget (1980), Carroll and Pruzanski (1975), De Soete (1984), and DeSarbo, Manrai, and Burke (1990), among others, have proposed a variety of procedures to “optimally” derive such ultrametric tree representations. These more recent developments augment the huge battery of available hierarchical clustering heuristics commonly available in most major statistical software packages (Hartigan, 1975; Jain & Dubes, 1988).

A potential problem in the application of any of these hierarchical clustering or optimal tree fitting procedures is the effect of subjects’ differential familiarity with the

Requests for reprints should be directed to Wayne S. DeSarbo, School of Business Administration, The University of Michigan, Ann Arbor, MI 48109-1234.

stimuli under investigation on their proximity judgments. That is, how do different subjects respond when asked to judge the proximity of stimuli j and k , when they are not completely familiar with j and/or k ? And, if the subjects' responses are influenced by the degree of stimulus familiarity, how might these responses distort the resulting ultrametric tree structure derived from such proximity data? Traditionally (for example, in a multidimensional scaling study), the subject's familiarity with the stimuli is verified prior to obtaining the proximity data, and subjects not familiar with all the stimuli are dropped from the analysis. This is reasonable in situations where some small subset of the subjects may be completely unfamiliar with the stimuli. However, consider, for example, an application where the stimuli are competing brands in some frequently purchased product category (such as beer). Subjects are likely to be very familiar with the brand(s) they consume most frequently, somewhat less familiar with other brands that they may have tried occasionally, and much less familiar with untried brands for which subjects have only "second hand" knowledge via word-of-mouth and advertisements. There are two points to note in this type of situation, characterized by relatively complex multiattribute stimuli. First, *familiarity may not be dichotomous* (familiar/unfamiliar), *but may vary instead along some continuum* (i.e., there are degrees of familiarity). Second, it is most likely that most (if not all) of the subjects *will not be completely familiar with all the stimuli in the set*. In such circumstances, rather than simply dropping data points involving unfamiliar and partially familiar stimuli, it may be fruitful to consider instead an approach that explicitly models the impact of stimulus unfamiliarity (including partial familiarity) on proximity judgments, and thereby controls for the potential distortions on account of stimulus unfamiliarity while deriving the ultrametric tree structure. This is the motivation for our research focus on a new procedure which aims to "filter out" the effect of stimulus unfamiliarity while estimating hierarchical tree structures.

We first review the empirical and theoretical literature that motivates our quantitative model of the effect of differential stimulus familiarity on subjects' interobject proximity judgments. While we could find no research in the psychometric literature that *directly* addresses this effect on proximity judgments, as evoked on some response scale (e.g., an itemized rating scale with the end points labeled as "extremely similar" and "extremely dissimilar"), there is a substantial body of related empirical work concerning the impact of incomplete information on attitudes. This stream of literature suggests that subjects tend to infer typical or neutral (i.e., average) values for missing attributes (Jaccard & Wood, 1988; Meyer, 1981; Slovic & MacPhillamy, 1974; Yagamichi & Hill, 1983; Yates, Jagacinski, & Faber, 1978). Sanbonmatsu, Kardes, and Herr (1992) observe that overall judgments move toward a relatively neutral position "because a moderate stance is easily justified and is readily altered as additional information becomes available" (p. 77). While the literature referenced above focuses on the effect on attitudes, the underlying *process* leading to attitude formation suggests that subjects' proximity judgments may similarly be moderated in the presence of unfamiliarity with stimuli. Stimulus unfamiliarity implies less information and greater uncertainty about the stimulus. Under complete unfamiliarity, subjects might tend to *anchor* their response, perhaps at some subject-specific "moderate" value. With increasing familiarity, there might be an *adjustment* in response from the anchor toward the "true" proximity value. This adjustment is analogous to the finding that the accuracy and speed of recognition (of, for example, faces) improves with familiarity (Valentine & Bruce, 1986) or, equivalently, with better developed schema (Goldstein & Chance, 1980).

The notion of the anchoring of a subject's proximity response under unfamiliarity at some moderate scale value also follows from a conceptualization of the interobject

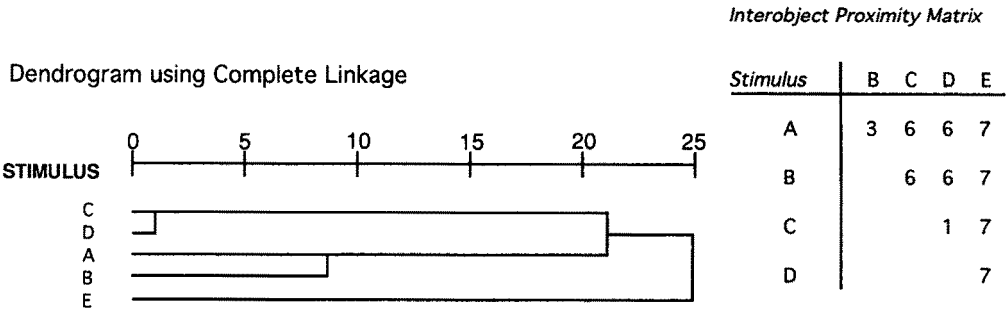
proximity judgment task as an examination of the extent to which the stimuli share (or do not share) common features (Tversky, 1977). This is the basis for the branching of the ultrametric tree: each node typically represents some attribute (or feature) and each branch from the node represents a set of stimuli sharing that feature (or attribute level), while the branches differ from each other on that attribute. In this context, we may view unfamiliarity as implying that the subject is uncertain about which features of the unfamiliar stimulus are common and which are distinct vis-a-vis the other (referent) stimulus. Under these circumstances, stimulus unfamiliarity could have a moderating effect on stated interobject proximities. A related phenomenon with a moderating effect is the tendency of subjects to avoid the ends of the scale reflecting extreme judgments (e.g., "extremely similar" or "extremely dissimilar") under low confidence in judgments about stimuli induced by unfamiliarity (Sanbonmatsu et al., 1992).

While the above discussion suggests that stimulus unfamiliarity would tend to have a *moderating* effect on subjects' proximity judgments, it is also possible that, for some subjects, responses under stimulus unfamiliarity may be anchored *away* from, say, an average interobject distance or the neutral point of the scale (which are both plausible candidates for a "moderate scale value" anchor). For example, if increasing familiarity makes feature differences between stimuli disproportionately more salient (relative to their feature commonalities), then subjects would tend to bias their anchor toward the "similarity" end of the scale. There is some evidence for this type of bias, such as the other-race effect in the face recognition literature: individuals report that members of other tribes or races look alike (Goldstein & Chance, 1980). In a study by R. N. Shepard (reported in Green, Tull, and Albaum, 1988), a nonmetric MDS map of the United States based on subjective proximity judgments on pairs of all 48 continental states collected from a sample of long-term residents of Boston exaggerated the Northeastern region and shrank distant states: respondents tended to inflate the size of areas near Boston and attenuate those farther away. On the other hand, there is also evidence that subjects may have "a response bias to describe unfamiliar category items as atypical, using the heuristic 'if I don't know about it, it must be weird'" (Boster, 1988, p. 259). In this case, one might expect a bias toward the "dissimilarity" end of the scale. Finally, the uncertainty in the judgment task involving unfamiliar stimuli may tend to make a subject's responses less patterned (Pereboom, 1971). Thus, a subject's response under unfamiliarity might be expected to show greater variation (around the anchor) than in the case of judgments involving familiar stimuli.

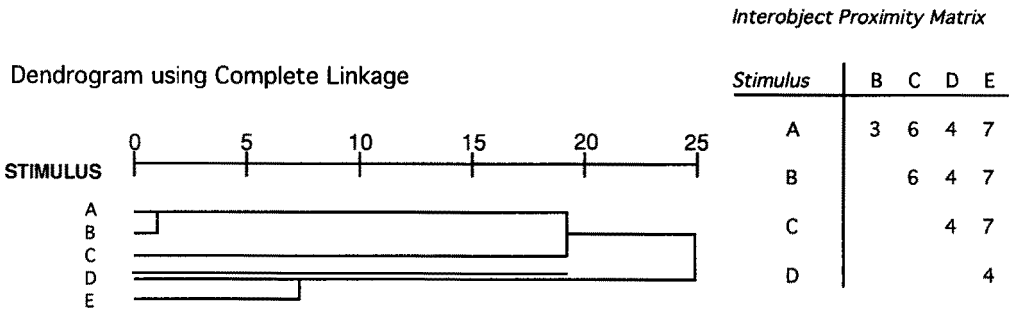
In summary, when asked to provide interobject similarity judgments for a pair of stimuli one (or both) of which is unfamiliar, subjects may anchor their responses, possibly at some central scale value or, in some cases, *away* from a central scale value. Further, given the uncertainty, their response may display some random variation around this anchor. With increasing familiarity, a subject's proximity response might be expected to adjust toward the "true" value (i.e., the response if the subject were completely familiar with both stimuli in the pair under evaluation), and also show less random variation. These premises guide our quantitative model of the impact of stimulus unfamiliarity on interobject proximity responses.

An illustration of the extent of distortion in the resulting tree structure that can potentially be caused by stimulus unfamiliarity is provided in Figure 1. The hierarchical tree in Panel (a), derived using complete linkage, is based on the synthetic proximity data matrix (perfectly satisfying the ultrametric inequality) on the right side of the panel representing the case of complete familiarity. Next, we consider the case of unfamiliarity with Stimulus D, with the implication that all proximity judgments involving Stimulus D are affected. We assume for this illustration that, under complete unfamiliarity, the response corresponds to the mid-point of the range of proximity scores, that

(a) Complete familiarity with all stimuli



(b) Complete unfamiliarity with Stimulus D



(c) Complete familiarity with Stimuli A and D

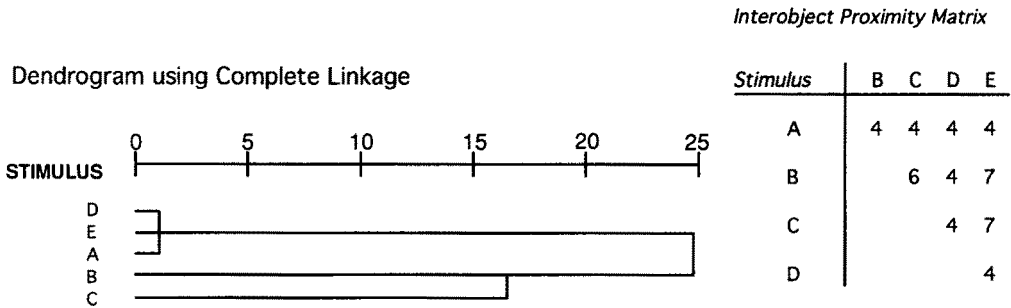


FIGURE 1.
An example of distortion in tree structure as a result of stimulus unfamiliarity.

is, a score of 4, given the 1–7 range. The modified proximity matrix and the resulting tree are shown in Panel (b) of Figure 1. Finally, we consider unfamiliarity with Stimulus A as well as with Stimulus D in Panel (c), in which case the tree is distorted beyond recognition.

We propose a new procedure, called TREEFAM, for estimating hierarchical (ultrametric) tree structures from proximity data confounded by differential stimulus familiarity. Our objective is to quantitatively “filter out” the effect of stimulus unfamiliarity in deriving an ultrametric (hierarchical) tree. The TREEFAM procedure simultaneously estimates an ultrametric tree under the (unobserved) condition of com-

plete familiarity, as well as individual-level model parameters capturing the "adjustment" due to unfamiliarity. The input data are proximity ratings for stimulus pairs and familiarity ratings for each stimulus in evaluation set collected from a common sample of respondents. The TREEFAM model specification, discussed in the next section, is sufficiently flexible to accommodate any type of anchoring discussed above.

Chatterjee and DeSarbo (1992) have recently proposed a spatial vector or scalar products MDS model and estimation procedure for accommodating the effects of stimulus unfamiliarity in the multidimensional scaling of preference data. Their empirical application demonstrates that accounting for unfamiliarity can have a considerable impact on the orientation of subjects' preference vectors in the resulting joint space, suggesting that stimulus unfamiliarity can significantly distort resulting joint space analyses. The TREEFAM procedure, in contrast, analyzes proximity rather than preference data, and is based on an ultrametric tree structure rather than a spatial model.

This paper is organized as follows. In the next section, we present the TREEFAM procedure, starting with a formulation of the model and followed by a description of the estimation procedure. Next, in section 3, we discuss a Monte Carlo study, in which synthetic data are created to evaluate the performance of the proposed estimation procedure. In section 4, we describe an empirical application of the TREEFAM procedure based on a small scale study in which graduate students provided dissimilarity judgments on various pairs of brands of beer. We conclude, in section 5, with a summary of the implications of our methodology and suggestions for future research.

2. The TREEFAM Procedure

The Model

Let:

- i = subject index, $i = 1, \dots, I$;
- j, k = object (stimulus) index, $j, k = 1, \dots, N$;
- δ_{ijk} = subject i 's *stated* dissimilarity between stimuli j and k ;
- f_{ij} = subject i 's *stated* familiarity with stimulus j , $0 \leq f_{ij} \leq 1$ (0: totally unfamiliar, 1: totally familiar).

We model a subject's stated dissimilarity rating as a function of the ultrametric tree distance (the "true" interobject distance) and stimulus familiarity as follows:

$$\delta_{ijk} = (f_{ijk})^{\alpha_i} d_{jk} + [1 - (f_{ijk})^{\alpha_i}] [\beta_i + u_{ijk}] + e_{ijk}, \quad (1)$$

where

- $f_{ijk} = f_{ij} \cdot f_{ik}$, $j \neq k$;
- d_{jk} = the ultrametric tree distance between brands j and k assumed common across subjects under complete familiarity, satisfying the one-mode ultrametric inequality: $d_{jk} \leq \max(d_{jl}, d_{lk})$, $\forall j, k, l$;
- α_i = subject i 's "exponent" parameter ($\alpha_i \geq 0$);
- β_i = subject i 's "anchor" parameter;
- u_{ijk} = error due to unfamiliarity; for given i , u_{ijk} iid $N(0, \sigma_i^2)$;
- e_{ijk} = measurement error; for given i , e_{ijk} iid $N(0, \phi_i^2)$; e_{ijk} and u_{ijk} independent.

Nested within (1) is the following model that ignores the effect of the degree of stimulus familiarity:

$$\delta_{ijk} = d_{jk} + e_{ijk}, \quad (2)$$

which serves as a baseline for comparison with the “full” model. Note that this model in (2) is equivalent to the formulation in Carroll and Pruzanski (1975, 1980) and DeSoete (1984) with their penalty function based least-squares estimation procedures, where no parametric specification is given for e_{ijk} . Also, observe that if $f_{ijk} = 1$, that is, the subject is completely familiar with both stimuli involved in the interobject dissimilarity judgment ($f_{ij} = f_{ik} = 1$), then the “full” model (1) reduces to (2). In other words, under complete familiarity, the stated interobject dissimilarity is assumed to be the “true” interobject dissimilarity (represented by the ultrametric tree distance between the stimuli) plus measurement error. On the other hand, if the subject is completely unfamiliar with at least one of the stimuli, so that $f_{ijk} = 0$, the model reduces to

$$\delta_{ijk} \{f_{ijk} = 0\} = \beta_i + u_{ijk} + e_{ijk}. \quad (3)$$

In this extreme case of complete unfamiliarity, the subject’s stated dissimilarity consists of a reference response with a subject-specific expected value of β_i and a random component u_{ijk} , plus measurement error e_{ijk} . For intermediate levels of familiarity ($0 < f_{ijk} < 1$), it may be observed from (1) that the stated dissimilarity δ_{ijk} is a weighted average of the extreme cases in (2) and (3), where the relative weights are determined by the subject’s familiarity with the two stimuli in the pair. More specifically, the weighting factor is formulated as $(f_{ij} \cdot f_{ik})^{\alpha_i}$ to capture the interactive nature in which familiarity levels for the two stimuli are assumed to affect a subject’s dissimilarity judgment.

Given the lack of unambiguous evidence about the location of the anchor point under unfamiliarity based on our review of the literature, our model specification is flexible on this issue. More specifically, we model a subject’s response under complete unfamiliarity (*excluding* measurement error) stochastically, with an expected value, or anchor, that is a subject-specific free parameter to be estimated from the data. The stochasticity captures the notion that an unfamiliar stimulus may produce some random variation in response (i.e., a less patterned reaction; Pereboom, 1971), and also recognizes other factors not specified in our model that could potentially influence the response. The model specification thus implies that a subject makes his/her proximity judgment under complete unfamiliarity (with at least one stimulus in the pair) *as if* he/she were drawing this value from some distribution, assumed normal, with subject-specific mean β_i (the anchor) and variance σ_i^2 (capturing uncertainty in judgment due to stimulus unfamiliarity). Our model places no specific constraints on β_i , although our previous discussion would suggest that the condition: $\min_i \delta_{ijk} \leq \beta_i \leq \max_i \delta_{ijk}$ should hold. The α_i parameter determines how a subject’s proximity judgment evolves as a function of stimulus familiarity. Note that the weight $(f_{ij} \cdot f_{ik})^{\alpha_i}$ decreases in α_i , so that, given f_{ij} and f_{ik} , a subject’s proximity judgment is closer to the “true” response under complete familiarity when α_i is smaller. Thus, a smaller α_i value implies a more rapid adjustment toward the “true” response with increasing stimulus familiarity. Finally, the observed proximity judgment is specified as containing a random (measurement) error component that is assumed independent of stimulus familiarity.

Estimation Procedure

We develop a maximum likelihood based estimation procedure to estimate this model. Noting that u_{ijk} and e_{ijk} are assumed to be independent and each normally distributed with zero mean and variance σ_i^2 and ϕ_i^2 respectively, it may be observed

from (1) that subject i 's dissimilarity rating for stimulus pair (j, k) , δ_{ijk} , will be normally distributed with mean μ_{ijk} and variance θ_{ijk}^2 , where:

$$\mu_{ijk} = (f_{ijk})^{\alpha_i} d_{jk} + [1 - (f_{ijk})^{\alpha_i}] \beta_i; \tag{4}$$

$$\theta_{ijk}^2 = [1 - (f_{ijk})^{\alpha_i}]^2 \sigma_i^2 + \phi_i^2. \tag{5}$$

Thus, given data $\Delta = ((\delta_{ijk}))$ and $F = ((f_{ij}))$, and assuming independence over i, j , and k (i.e., over subjects and stimuli as in Ramsay, 1977), we can form the likelihood function:

$$L = \prod_i \prod_{j < k} (2\pi \theta_{ijk}^2)^{-1/2} \exp \left[\frac{-(\delta_{ijk} - \mu_{ijk})^2}{2\theta_{ijk}^2} \right], \tag{6}$$

and the log likelihood function is:

$$\ln L = \sum_i \sum_{j < k} \left[\frac{-\ln(2\pi)}{2} - \ln(\theta_{ijk}) - \frac{(\delta_{ijk} - \mu_{ijk})^2}{2\theta_{ijk}^2} \right]. \tag{7}$$

The objective of the maximum likelihood procedure is therefore to:

$$\text{Maximize } \ln L \tag{8}$$

$\alpha_i, \beta_i, \sigma_i, \phi_i, d_{jk}$

subject to the one-mode ultrametric inequality:

$$d_{jk} \leq \max(d_{jl}, d_{lk}), \quad \forall j, k, l, \tag{9}$$

and $\alpha_i, \sigma_i, \phi_i > 0$, given $\Delta = ((\delta_{ijk}))$ and $F = ((f_{ij}))$.

The constraint implied by the ultrametric inequality is accommodated by forming the augmented log likelihood function:

$$Z = \ln L - \lambda P(\mathbf{D}), \tag{10}$$

where

$$P(\mathbf{D}) = \frac{A}{B};$$

$$A = \sum_{l < k, j \neq l, k} \sum_j \sum_k w_{lk}^j (d_{lj} - d_{jk})^2;$$

$$w_{lk}^j = \begin{cases} 1 & \text{if } d_{lk} \leq \min(d_{lj}, d_{jk}), \\ 0 & \text{else;} \end{cases}$$

$$B = \sum_{j < k} (d_{jk} - \bar{d})^2;$$

$$\bar{d} = \left[\frac{2}{N(N-1)} \right] \sum_{j < k} d_{jk};$$

λ = the penalty parameter.

A conditional, alternating maximum likelihood estimation procedure is developed and implemented via the TREEFAM algorithm, which is summarized in Figure 2. The user

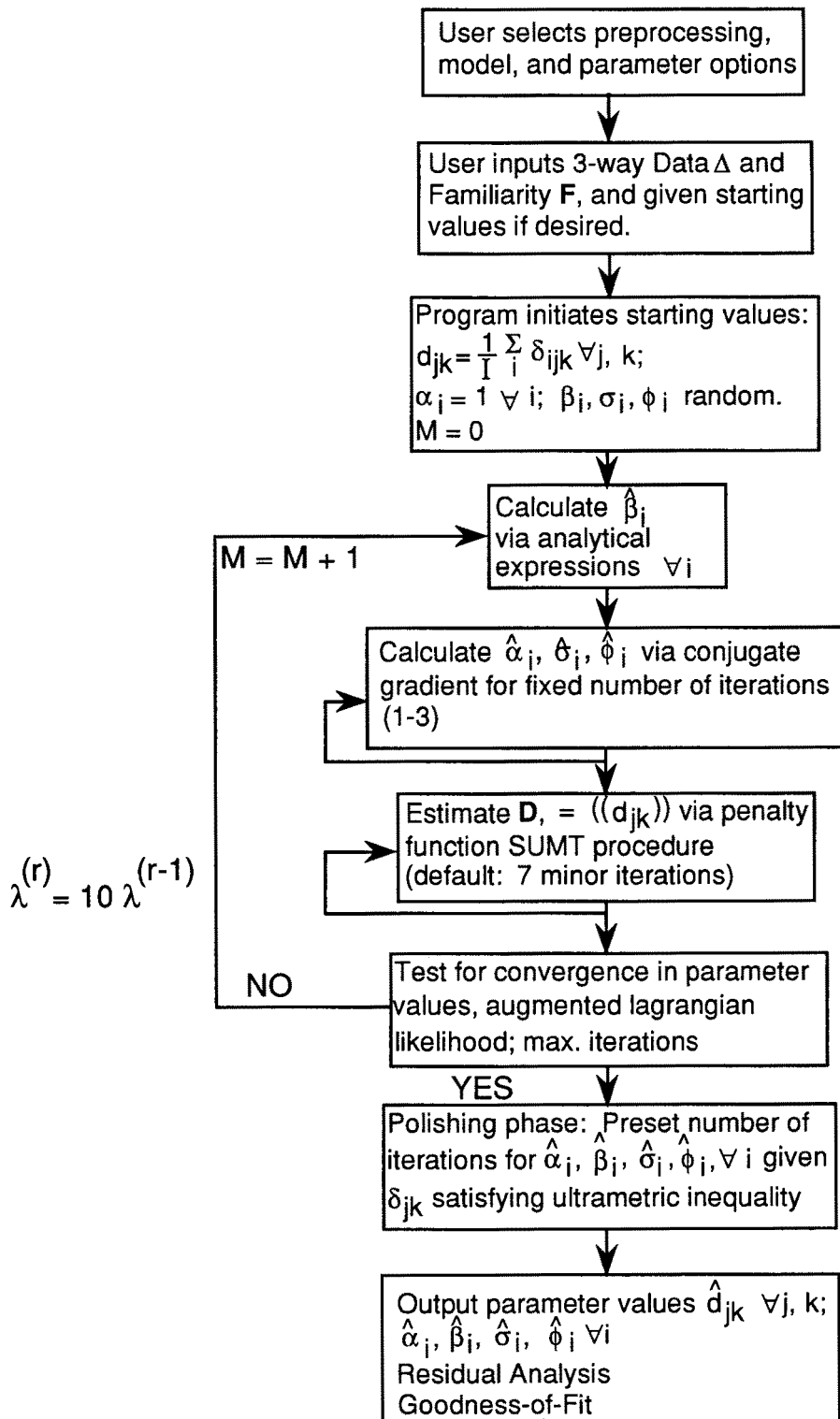


FIGURE 2.
TREEFAM algorithm flowchart.

first specifies a vector of control parameters that includes options for preprocessing of the 3-way dissimilarities input data matrix (no preprocessing, slice standardization/normalization, array standardization/normalization), analysis (internal, where the tree is to be estimated, vs. external, where the tree is prespecified), model parameters (to be estimated vs. given and fixed) and algorithm parameters specifying the maximum number of iterations, step size, convergence tolerance, etc. Along with the data matrices Δ and F , the user may also optionally specify starting values. (In the absence of specified values, the TREEFAM program initiates starting values as follows: $d_{jk} = (\sum_i \delta_{ijk})/I$, $\alpha_i = 1, \forall i$; $\beta_i, \sigma_i, \phi_i$ random.)

In each cycle of the iterative estimation procedure, each parameter set is estimated in turn, while other parameter sets are held fixed. The stationary equations for the various parameter sets are as follows:

For α_i :

$$\frac{\partial Z}{\partial \alpha_i} = 0 = \sum_{j < k} \sum \left\{ \left(\frac{\sigma_i^2}{\theta_{ijk}^2} \right) [1 - (f_{ijk})^{\alpha_i}] (f_{ijk})^{\alpha_i} \ln f_{ijk} - \left[\frac{\sigma_i^2 (\delta_{ijk} - \mu_{ijk})^2}{\theta_{ijk}^4} \right] \right. \\ \left. \times [1 - (f_{ijk})^{\alpha_i}] (f_{ijk})^{\alpha_i} \ln f_{ijk} + \left[\frac{\delta_{ijk} - \mu_{ijk}}{\theta_{ijk}^2} \right] (f_{ijk})^{\alpha_i} \ln f_{ijk} (d_{jk} - \beta_i) \right\}; \quad (11)$$

For β_i :

$$\frac{\partial Z}{\partial \beta_i} = 0 = \sum_{j < k} \sum \frac{[\delta_{ijk} - [1 - (f_{ijk})^{\alpha_i}] \beta_i - (f_{ijk})^{\alpha_i} d_{jk}] [1 - (f_{ijk})^{\alpha_i}]}{[1 - (f_{ijk})^{\alpha_i}]^2 \sigma_i^2 + \phi_i^2}; \quad (12)$$

For σ_i :

$$\frac{\partial Z}{\partial \sigma_i} = 0 = \sum_{j < k} \sum - \frac{[1 - (f_{ijk})^{\alpha_i}]^2 \sigma_i}{[[1 - (f_{ijk})^{\alpha_i}]^2 \sigma_i^2 + \phi_i^2]} + \frac{[1 - (f_{ijk})^{\alpha_i}]^2 [\delta_{ijk} - \mu_{ijk}]^2 \sigma_i}{[[1 - (f_{ijk})^{\alpha_i}]^2 \sigma_i^2 + \phi_i^2]^2}; \quad (13)$$

For ϕ_i :

$$\frac{\partial Z}{\partial \phi_i} = 0 = \sum_{j < k} \sum - \frac{\phi_i}{[[1 - (f_{ijk})^{\alpha_i}]^2 \sigma_i^2 + \phi_i^2]} + \frac{[\delta_{ijk} - \mu_{ijk}]^2 \phi_i}{[[1 - (f_{ijk})^{\alpha_i}]^2 \sigma_i^2 + \phi_i^2]^2}. \quad (14)$$

A closed form analytical expression exists only for the estimator for β_i , as follows:

$$\hat{\beta}_i = \frac{\sum_{j < k} \sum \frac{[\delta_{ijk} - (f_{ijk})^{\alpha_i} d_{jk}] [1 - (f_{ijk})^{\alpha_i}]}{[1 - (f_{ijk})^{\alpha_i}]^2 \sigma_i^2 + \phi_i^2}}{\sum_{j < k} \sum \frac{[1 - (f_{ijk})^{\alpha_i}]^2}{[1 - (f_{ijk})^{\alpha_i}]^2 \sigma_i^2 + \phi_i^2}}. \quad (15)$$

While β_i is estimated by (15), estimates of α_i, σ_i and $\phi_i, \forall i$, are obtained by a conjugate gradient method, using a fixed number of minor iterations. Next, $\mathbf{D} = ((d_{jk}))$ is estimated via a penalty function SUMT procedure (see De Soete, 1984) in which the value of the penalty function λ is increased by a factor of 10 after each major iteration. The stationary equation for d_{jk} is:

$$\frac{\partial Z}{\partial d_{jk}} = 0 = \sum_i \left\{ (\delta_{ijk} - \mu_{ijk}) \frac{(f_{ijk})^{\alpha_i}}{\theta_{ijk}^2} \right\} - \left(\frac{2\lambda}{B^2} \right) \\ \times \left\{ B \sum_l [(w_{lk}^j + w_{lj}^k)d_{jk} - w_{lk}^j d_{lj} - w_{lj}^k d_{lk}] - A(d_{jk} - \bar{d}) \left[1 - \frac{2}{N(N-1)} \right] \right\}. \quad (16)$$

This completes one cycle or major iteration of the estimation procedure. The cycle is repeated (the iterations are indexed by M in Figure 2) until convergence criteria (for parameter estimates and the augmented maximum likelihood) are met or a preset maximum number of iterations is reached. Convergence is followed by an optional "polishing" phase, in which estimates for α_i , β_i , σ_i and ϕ_i , $\forall i$, are improved through a further (prespecified) number of iterations, given the estimates of $((d_{jk}))$ satisfying ultrametric inequality.

The TREEFAM output consists of estimates of the ultrametric distance matrix $\mathbf{D} = ((d_{jk}))$ and the parameters α_i , β_i , σ_i and ϕ_i , $\forall i$, along with an analysis of the residuals and various goodness-of-fit statistics. These statistics include the values at convergence of the log likelihood function, the penalty function $\lambda P(\mathbf{D})$ (which is a measure of the extent to which the ultrametric inequality constraint is violated), and the complete objective function (10), as well as the percentage of the variance in the data Δ accounted for by the model.

3. Monte Carlo Analysis

Objective and Experimental Design

The objective of this modest Monte Carlo study was to evaluate the performance of the proposed estimation procedure in terms of its ability to recover known tree structures, as well as subject-level model parameters, under a variety of conditions. Tree recovery here was also evaluated relative to the "naive" (restricted) model (2), which corresponds conceptually to the model underlying traditional ultrametric tree fitting methods that ignore stimulus unfamiliarity effects. Synthetic data sets were generated under a variety of conditions created by systematically manipulating five "factors"—the parameters α_i , β_i , σ_i and ϕ_i , and the stimulus familiarity ratings f_{ij} —at two levels each, as shown in Table 1. Given the parameters, familiarity values, and \mathbf{D} , Δ was generated according to equation (1) as described below.

The subject level parameters α_i and β_i were varied to examine the effect of heterogeneity across subjects with respect to these parameters. Level 1 for α_i was a constant value (equal to 1) across subjects, whereas at Level 2, α_i was drawn independently for each subject from a Uniform (0, 2) distribution. For β_i , the mean for both levels was assumed to be the average interobject distance, while the variance of the distribution from which the β_i 's were drawn was higher for Level 2. The variances of the unfamiliarity and measurement error terms, ϕ_i^2 and σ_i^2 respectively, were assumed equal across subjects, and each was set low in one treatment level and high in the other. While in both cases the "high" condition implies greater random error, it is clear from the model in (1) that the two error terms have somewhat different implications. Mathematically, the effect of the unfamiliarity error term u_{ijk} is moderated by the unfamiliarity weight $[1 - (f_{ijk})^{\alpha_i}]$. Conceptually, the two levels of σ_i correspond to different degrees of within-subject variability in the "anchor" response under unfamiliarity, while the two levels of ϕ_i correspond to different levels of measurement error. Finally, the two levels of f_{ij} reflect different degrees of heterogeneity in stimulus familiarity

TABLE 1

Factors and Their Levels in Monte Carlo Study

Factor	Level 1	Level 2
α_i	constant across subjects; $\alpha_i = 1 \forall i$	varying across subjects; α_i drawn from $U(0, 2)$
β_i	low variance across subjects; β_i drawn from $N(\bar{\beta}, 0.01\bar{\beta}^2)$; $\bar{\beta} = 2(\sum_{j < k} \delta_{ij}) / [N(N-1)]$	high variance across subjects; β_i drawn from $N(\bar{\beta}, 0.25\bar{\beta}^2)$
σ_i	low ($\sigma_i = 0.5$)	high ($\sigma_i = 2.5$)
ϕ_i	low ($\phi_i = 0.5$)	high ($\phi_i = 2.5$)
f_{ij}	homogeneous across subjects; $f_{ij} = f'_j \forall i$; f'_j drawn from $U(0.01, 0.99)$	heterogeneous across subjects; $f_{ij} = \begin{cases} f'_j & i = 1, \dots, I/2 \\ 1 - f'_j & i = I/2+1, \dots, I \end{cases}$

across subjects. Note that the within subject variation in familiarity across stimuli was the same for both treatment levels. For each stimulus, a value f'_j was drawn from a Uniform (0.01, 0.99) distribution. For Level 1, $f_{ij} = f'_j$ for all subjects (homogeneity across subjects), while for Level 2, $f_{ij} = f'_j$ for half the subjects and $f_{ij} = 1 - f'_j$ for the other half (heterogeneity across subjects). Here, the case of low variation in familiarity across stimuli is less interesting since the distortion in the data (and consequently in the derived ultrametric tree) is a result of differential stimulus familiarity. Thus, we manipulated the variation (heterogeneity) across subjects only, while maintaining high variation across stimuli at both levels.

An exact ultrametric tree with distance matrix $D = ((d_{jk}))$ was initially created for 6 hypothetical stimuli. For each of the $2^5 = 32$ cells in the full factorial design, a three-way dissimilarities matrix $\Delta = (((\delta_{ijk})))$ was generated by the model (1), given D and values for α_i , β_i , σ_i , ϕ_i , and f_{ij} selected (or drawn from a distribution) for each cell as indicated in Table 1 for $I = N = 6$. The data thus generated were used to estimate the full model (1) as well as the restricted model (2). There were two replications per cell, yielding 64 trials in all. Several model performance measures relating to

recovery of model parameters (i.e., the ultrametric tree as well as the subject-level and error variance parameters) and overall model goodness-of-fit were computed for each trial. Specifically, tree recovery was measured by the percentage of the variance in the prespecified ultrametric tree accounted for by the estimated tree (abbreviated as VAF(D)), while recovery of the parameters α_i , β_i , σ_i , and ϕ_i was assessed by the root mean square errors of the estimates versus actual values. For overall goodness-of-fit, we examined the log likelihood value at convergence, since the order of Δ was constant across trials.

Results

Several analyses of variance were performed for the 2^5 full factorial model, using different performance measures as the dependent variable in turn, with the objective of assessing which factors significantly affect model estimation. The within cell replications permitted estimation of the saturated model (with all interaction effects). The resulting ANOVA tables are presented in Tables 2 through 8, and the results are discussed below for significant effects (with p -value $\leq .05$).

Tree recovery. The analysis of variance with VAF(D) (after Fisher's z transformation) as the dependent variable is presented in Table 2. The only significant main effect is that of the f_{ij} factor ($p < .01$); VAF(D) is significantly higher when subjects are heterogenous in f_{ij} . The $f_{ij} \times \alpha_i$ interaction is also significant ($p < .05$). Given the significant interaction, we examined the *conditional* (main) effects of f_{ij} and α_i , that is, the main effect of each factor *given* a particular level of the other (interacting) factor. The increase in VAF(D) (or improvement in tree recovery) when subjects are heterogenous in f_{ij} is significant when α_i is at Level 1 (constant α_i across subjects) but not at Level 2 (heterogeneity in α_i). Conversely, the conditional main effect of α_i (when f_{ij} is at Level 2, i.e., heterogeneity in f_{ij}) becomes significant ($p < .01$). When subjects are heterogenous in f_{ij} , tree recovery is significantly better when α_i is constant across subjects. In essence, the improvement in tree recovery with increasing heterogeneity in the pattern of stimulus familiarity is moderated by the effect of heterogeneity in α_i across subjects: the improvement is greatest when subjects tend to be homogenous in α_i . Greater variation in f_{ij} across subjects implies more information in the data and therefore better estimates. Observe from (1), however, that the information for estimation really depends on the variation across subjects in $(f_{ij} \cdot f_{ik})^{\alpha_i}$. Variation in the exponent parameter α_i has a moderating influence on the across-subject variation of $(f_{ij} \cdot f_{ik})^{\alpha_i}$, and thus on the effect of heterogeneity in f_{ij} .

Parameter recovery. The ANOVA tables with the RMSEs for α_i , β_i , σ_i , and ϕ_i as the dependent variable in turn are presented in Tables 3 through 6. We first describe the significant effects for each of the four parameters, and then summarize and interpret the results.

The significant main effects in the ANOVA on RMSE (α_i) in Table 3 suggest that recovery of α_i is significantly better when subjects are more heterogenous in β_i and f_{ij} , and more homogenous in α_i . However, the $\beta_i \times f_{ij}$ interaction is also significant at the .05 level. Examining the conditional main effects of β_i and f_{ij} (given the level of the other factor), we find that β_i is significant for both levels of f_{ij} (i.e., the β_i main effect is unconditionally significant), whereas the f_{ij} effect is significant only at Level 1 for β_i

¹ Instead of the raw VAF(D) measure, it is more appropriate to use the transformed measure $z = \ln [(1 + \sqrt{\text{VAF(D)}})/(1 - \sqrt{\text{VAF(D)}})]/2$, which tends to have a normal distribution, as the dependent variable in an ANOVA (Morrison, 1983).

TABLE 2
ANOVA Table for VAF(D)*

Source of Variation	Sum of Squares	d. f.	Mean Square	F	Significance of F (p-value)
Main Effects	5.331	5	1.066	2.866	.030
α_i	1.210	1	1.210	3.252	.081
β_i	0.932	1	0.932	2.507	.123
σ_i	0.001	1	0.001	0.003	.954
ϕ_i	0.000	1	0.000	0.000	.985
f_{ij}	3.188	1	3.188	8.569	.006
2-Way Interactions	4.738	10	0.474	1.274	.286
$\alpha_i \times \beta_i$	0.078	1	0.078	0.210	.650
$\alpha_i \times \sigma_i$	0.063	1	0.063	0.169	.684
$\alpha_i \times \phi_i$	0.170	1	0.170	0.457	.504
$\alpha_i \times f_{ij}$	1.697	1	1.697	4.562	.040
$\beta_i \times \sigma_i$	0.884	1	0.884	2.377	.133
$\beta_i \times \phi_i$	0.022	1	0.022	0.060	.809
$\beta_i \times f_{ij}$	0.155	1	0.155	0.417	.523
$\sigma_i \times \phi_i$	0.457	1	0.457	1.228	.276
$\sigma_i \times f_{ij}$	0.428	1	0.428	1.151	.291
$\phi_i \times f_{ij}$	0.783	1	0.783	2.104	.157
3-Way Interactions	1.253	10	0.125	0.337	.964
$\alpha_i \times \beta_i \times \sigma_i$	0.111	1	0.111	0.298	.589
$\alpha_i \times \beta_i \times \phi_i$	0.322	1	0.322	0.866	.359
$\alpha_i \times \beta_i \times f_{ij}$	0.050	1	0.050	0.135	.716
$\alpha_i \times \sigma_i \times \phi_i$	0.009	1	0.009	0.023	.880
$\alpha_i \times \sigma_i \times f_{ij}$	0.280	1	0.280	0.753	.392
$\alpha_i \times \phi_i \times f_{ij}$	0.075	1	0.075	0.200	.657
$\beta_i \times \sigma_i \times \phi_i$	0.214	1	0.214	0.574	.454
$\beta_i \times \sigma_i \times f_{ij}$	0.084	1	0.084	0.225	.639
$\beta_i \times \phi_i \times f_{ij}$	0.041	1	0.041	0.111	.741
$\sigma_i \times \phi_i \times f_{ij}$	0.068	1	0.068	0.182	.673
4-Way Interactions	0.477	5	0.095	0.257	.933
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i$	0.260	1	0.260	0.698	.410
$\alpha_i \times \beta_i \times \sigma_i \times f_{ij}$	0.186	1	0.186	0.500	.485
$\alpha_i \times \beta_i \times \phi_i \times f_{ij}$	0.001	1	0.001	0.004	.952
$\alpha_i \times \sigma_i \times \phi_i \times f_{ij}$	0.029	1	0.029	0.078	.782
$\beta_i \times \sigma_i \times \phi_i \times f_{ij}$	0.001	1	0.001	0.004	.953
5-Way Interactions	0.017	1	0.017	0.045	.833
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i \times f_{ij}$	0.017	1	0.017	0.045	.833
Explained	11.816	31	0.381	1.025	.472
Residual	11.904	32	0.372		
Total	23.719	63	0.376		

*Dependent variable used: $z = \ln [(1 + \sqrt{\text{VAF(D)}}) / (1 - \sqrt{\text{VAF(D)}})] / 2$

(low heterogeneity). In other words, homogeneity in f_{ij} has a negative impact on the recovery of α_i only when subjects are relatively homogenous in β_i as well.

Examining the ANOVA on RMSE (β_i) in Table 4, the only significant effect (at the .05 level) is the f_{ij} main effect, suggesting recovery of β_i is better when subjects are

TABLE 3
ANOVA Table for RMSE (α_i)

Source of Variation	Sum of Squares	d. f.	Mean Square	F	Significance of F (p-value)
Main Effects	2.069	5	0.414	10.775	.000
α_i	1.078	1	1.078	28.074	.000
β_i	0.814	1	0.814	21.190	.000
σ_i	0.000	1	0.000	0.002	.963
ϕ_i	0.009	1	0.009	0.231	.634
f_{ij}	0.168	1	0.168	4.378	.044
2-Way Interactions	0.453	10	0.045	1.180	.340
$\alpha_i \times \beta_i$	0.003	1	0.003	0.067	.797
$\alpha_i \times \sigma_i$	0.002	1	0.002	0.044	.836
$\alpha_i \times \phi_i$	0.014	1	0.014	0.353	.556
$\alpha_i \times f_{ij}$	0.090	1	0.090	2.340	.136
$\beta_i \times \sigma_i$	0.031	1	0.031	0.818	.372
$\beta_i \times \phi_i$	0.085	1	0.085	2.209	.147
$\beta_i \times f_{ij}$	0.166	1	0.166	4.335	.045
$\sigma_i \times \phi_i$	0.048	1	0.048	1.255	.271
$\sigma_i \times f_{ij}$	0.014	1	0.014	0.377	.544
$\phi_i \times f_{ij}$	0.000	1	0.000	0.004	.951
3-Way Interactions	0.512	10	0.051	1.334	.255
$\alpha_i \times \beta_i \times \sigma_i$	0.126	1	0.126	3.291	.079
$\alpha_i \times \beta_i \times \phi_i$	0.080	1	0.080	2.078	.159
$\alpha_i \times \beta_i \times f_{ij}$	0.038	1	0.038	0.978	.330
$\alpha_i \times \sigma_i \times \phi_i$	0.022	1	0.022	0.572	.455
$\alpha_i \times \sigma_i \times f_{ij}$	0.017	1	0.017	0.454	.505
$\alpha_i \times \phi_i \times f_{ij}$	0.007	1	0.007	0.182	.673
$\beta_i \times \sigma_i \times \phi_i$	0.013	1	0.013	0.350	.558
$\beta_i \times \sigma_i \times f_{ij}$	0.120	1	0.120	3.131	.086
$\beta_i \times \phi_i \times f_{ij}$	0.002	1	0.002	0.063	.803
$\sigma_i \times \phi_i \times f_{ij}$	0.086	1	0.086	2.243	.144
4-Way Interactions	0.086	5	0.017	0.449	.811
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i$	0.011	1	0.011	0.289	.595
$\alpha_i \times \beta_i \times \sigma_i \times f_{ij}$	0.015	1	0.015	0.384	.540
$\alpha_i \times \beta_i \times \phi_i \times f_{ij}$	0.005	1	0.005	0.122	.729
$\alpha_i \times \sigma_i \times \phi_i \times f_{ij}$	0.012	1	0.012	0.322	.574
$\beta_i \times \sigma_i \times \phi_i \times f_{ij}$	0.043	1	0.043	1.127	.296
5-Way Interactions	0.019	1	0.019	0.483	.492
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i \times f_{ij}$	0.019	1	0.019	0.483	.492
Explained	3.139	31	0.101	2.637	.004
Residual	1.229	32	0.038		
Total	4.368	63	0.069		

heterogenous in f_{ij} . Since the $\beta_i \times f_{ij}$ interaction approaches significance ($p = .098$), we analyzed the conditional main effects to find that the f_{ij} effect is significant only when β_i is in Level 1. Further, β_i has a significant effect ($p < .05$) when f_{ij} is at Level 1. Hence, homogeneity in $f_{ij}(\beta_i)$ has an adverse impact on the recovery of β_i only when subjects are relatively homogenous in $\beta_i(f_{ij})$.

TABLE 4
ANOVA Table for RMSE (β_i)

Source of Variation	Sum of Squares	d. f.	Mean Square	F	Significance of F (p-value)
Main Effects	56.698	5	11.340	2.206	.078
α_i	0.178	1	0.178	0.035	.853
β_i	7.509	1	7.509	1.461	.236
σ_i	16.076	1	16.076	3.128	.087
ϕ_i	3.175	1	3.175	0.618	.438
f_{ij}	29.760	1	29.760	5.790	.022
2-Way Interactions	45.344	10	4.534	0.882	.559
$\alpha_i \times \beta_i$	0.090	1	0.090	0.017	.896
$\alpha_i \times \sigma_i$	1.561	1	1.561	0.304	.585
$\alpha_i \times \phi_i$	0.304	1	0.304	0.059	.810
$\alpha_i \times f_{ij}$	1.134	1	1.134	0.221	.642
$\beta_i \times \sigma_i$	6.152	1	6.152	1.197	.282
$\beta_i \times \phi_i$	3.272	1	3.272	0.637	.431
$\beta_i \times f_{ij}$	14.946	1	14.946	2.908	.098
$\sigma_i \times \phi_i$	7.881	1	7.881	1.533	.225
$\sigma_i \times f_{ij}$	7.909	1	7.909	1.539	.224
$\phi_i \times f_{ij}$	2.097	1	2.097	0.408	.528
3-Way Interactions	34.381	10	3.438	0.669	.744
$\alpha_i \times \beta_i \times \sigma_i$	2.177	1	2.177	0.424	.520
$\alpha_i \times \beta_i \times \phi_i$	0.906	1	0.906	0.176	.677
$\alpha_i \times \beta_i \times f_{ij}$	0.035	1	0.035	0.007	.935
$\alpha_i \times \sigma_i \times \phi_i$	0.775	1	0.775	0.151	.700
$\alpha_i \times \sigma_i \times f_{ij}$	2.100	1	2.100	0.408	.527
$\alpha_i \times \phi_i \times f_{ij}$	0.431	1	0.431	0.084	.774
$\beta_i \times \sigma_i \times \phi_i$	8.949	1	8.949	1.741	.196
$\beta_i \times \sigma_i \times f_{ij}$	10.589	1	10.589	2.060	.161
$\beta_i \times \phi_i \times f_{ij}$	4.846	1	4.846	0.943	.339
$\sigma_i \times \phi_i \times f_{ij}$	3.574	1	3.574	0.695	.411
4-Way Interactions	14.661	5	2.932	0.570	.722
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i$	0.022	1	0.022	0.004	.948
$\alpha_i \times \beta_i \times \sigma_i \times f_{ij}$	0.603	1	0.603	0.117	.734
$\alpha_i \times \beta_i \times \phi_i \times f_{ij}$	0.438	1	0.438	0.085	.772
$\alpha_i \times \sigma_i \times \phi_i \times f_{ij}$	0.116	1	0.116	0.023	.881
$\beta_i \times \sigma_i \times \phi_i \times f_{ij}$	13.482	1	13.482	2.623	.115
5-Way Interactions	0.777	1	0.777	0.151	.700
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i \times f_{ij}$	0.777	1	0.777	0.151	.700
Explained	151.861	31	4.899	0.953	.553
Residual	164.487	32	5.140		
Total	316.348	63	5.021		

We next consider recovery of the unfamiliarity error parameter σ_i . From Table 5, the significant main effects are σ_i and ϕ_i . The direction of the effects suggest that recovery of σ_i is better at lower levels of σ_i and ϕ_i (i.e., when the unfamiliarity and measurement error variances are small). There are, however, three significant 2-way

TABLE 5
ANOVA Table for RMSE (σ_i)

Source of Variation	Sum of Squares	d. f.	Mean Square	F	Significance of F (p-value)
Main Effects	4.238	5	0.848	8.080	.000
α_i	1.119	1	1.119	1.138	.294
β_i	0.292	1	0.292	2.780	.105
σ_i	3.061	1	3.061	29.177	.000
ϕ_i	0.766	1	0.766	7.302	.011
ϵ_{ij}	0.000	1	0.000	0.002	.963
2-Way Interactions	2.684	10	0.268	2.559	.021
$\alpha_i \times \beta_i$	0.006	1	0.006	0.059	.810
$\alpha_i \times \sigma_i$	0.837	1	0.837	7.981	.008
$\alpha_i \times \phi_i$	0.040	1	0.040	0.378	.543
$\alpha_i \times \epsilon_{ij}$	0.019	1	0.019	0.178	.676
$\beta_i \times \sigma_i$	0.507	1	0.507	4.832	.035
$\beta_i \times \phi_i$	0.001	1	0.001	0.011	.916
$\beta_i \times \epsilon_{ij}$	0.192	1	0.192	1.827	.186
$\sigma_i \times \phi_i$	0.852	1	0.852	8.125	.008
$\sigma_i \times \epsilon_{ij}$	0.228	1	0.228	2.176	.150
$\phi_i \times \epsilon_{ij}$	0.002	1	0.002	0.022	.884
3-Way Interactions	0.966	10	0.097	0.921	.527
$\alpha_i \times \beta_i \times \sigma_i$	0.000	1	0.000	0.004	.952
$\alpha_i \times \beta_i \times \phi_i$	0.265	1	0.265	2.531	.121
$\alpha_i \times \beta_i \times \epsilon_{ij}$	0.040	1	0.040	0.378	.543
$\alpha_i \times \sigma_i \times \phi_i$	0.007	1	0.007	0.070	.793
$\alpha_i \times \sigma_i \times \epsilon_{ij}$	0.336	1	0.336	3.204	.083
$\alpha_i \times \phi_i \times \epsilon_{ij}$	0.013	1	0.013	0.127	.724
$\beta_i \times \sigma_i \times \phi_i$	0.227	1	0.227	2.167	.151
$\beta_i \times \sigma_i \times \epsilon_{ij}$	0.010	1	0.010	0.091	.765
$\beta_i \times \phi_i \times \epsilon_{ij}$	0.017	1	0.017	0.161	.691
$\sigma_i \times \phi_i \times \epsilon_{ij}$	0.050	1	0.050	0.474	.496
4-Way Interactions	0.523	5	0.105	0.997	.435
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i$	0.084	1	0.084	0.800	.378
$\alpha_i \times \beta_i \times \sigma_i \times \epsilon_{ij}$	0.022	1	0.022	0.211	.649
$\alpha_i \times \beta_i \times \phi_i \times \epsilon_{ij}$	0.003	1	0.003	0.031	.861
$\alpha_i \times \sigma_i \times \phi_i \times \epsilon_{ij}$	0.253	1	0.253	2.412	.130
$\beta_i \times \sigma_i \times \phi_i \times \epsilon_{ij}$	0.161	1	0.161	1.533	.225
5-Way Interactions	0.010	1	0.010	0.095	.760
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i \times \epsilon_{ij}$	0.010	1	0.010	0.095	.760
Explained	8.421	31	0.272	2.590	.005
Residual	3.357	32	0.105		
Total	11.778	63	0.187		

interaction effects involving σ_i : $\alpha_i \times \sigma_i$, $\beta_i \times \sigma_i$, and $\sigma_i \times \phi_i$. As before, we examined the conditional main effects for each interaction. While the main effect of σ_i is robust, the effect of ϕ_i is not significant when σ_i is at Level 2; that is, when the unfamiliarity error variance is large, an increase in the measurement error variance

TABLE 6
ANOVA Table for RMSE (ϕ_i)

Source of Variation	Sum of Squares	d. f.	Mean Square	F	Significance of F (p-value)
Main Effects	5.662	5	1.132	30.359	.000
α_i	0.350	1	0.350	9.380	.004
β_i	0.167	1	0.167	4.490	.042
σ_i	0.595	1	0.595	15.947	.000
ϕ_i	4.529	1	4.529	121.431	.000
f_{ij}	0.020	1	0.020	0.546	.465
2-Way Interactions	2.177	10	0.218	5.836	.000
$\alpha_i \times \beta_i$	0.000	1	0.000	0.011	.918
$\alpha_i \times \sigma_i$	0.049	1	0.049	1.321	.259
$\alpha_i \times \phi_i$	0.039	1	0.039	1.056	.312
$\alpha_i \times f_{ij}$	0.147	1	0.147	3.933	.056
$\beta_i \times \sigma_i$	0.038	1	0.038	1.022	.320
$\beta_i \times \phi_i$	0.255	1	0.255	6.830	.014
$\beta_i \times f_{ij}$	0.055	1	0.055	1.471	.234
$\sigma_i \times \phi_i$	1.531	1	1.531	41.039	.000
$\sigma_i \times f_{ij}$	0.062	1	0.062	1.672	.205
$\phi_i \times f_{ij}$	0.000	1	0.000	0.000	.989
3-Way Interactions	0.464	10	0.046	1.243	.303
$\alpha_i \times \beta_i \times \sigma_i$	0.059	1	0.059	1.583	.217
$\alpha_i \times \beta_i \times \phi_i$	0.000	1	0.000	0.002	.963
$\alpha_i \times \beta_i \times f_{ij}$	0.006	1	0.006	0.161	.691
$\alpha_i \times \sigma_i \times \phi_i$	0.093	1	0.093	2.494	.124
$\alpha_i \times \sigma_i \times f_{ij}$	0.087	1	0.087	2.325	.137
$\alpha_i \times \phi_i \times f_{ij}$	0.090	1	0.090	2.413	.130
$\beta_i \times \sigma_i \times \phi_i$	0.032	1	0.032	0.866	.359
$\beta_i \times \sigma_i \times f_{ij}$	0.019	1	0.019	0.505	.482
$\beta_i \times \phi_i \times f_{ij}$	0.028	1	0.028	0.759	.390
$\sigma_i \times \phi_i \times f_{ij}$	0.049	1	0.049	1.324	.258
4-Way Interactions	0.272	5	0.054	1.459	.230
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i$	0.074	1	0.074	1.983	.169
$\alpha_i \times \beta_i \times \sigma_i \times f_{ij}$	0.045	1	0.045	1.211	.279
$\alpha_i \times \beta_i \times \phi_i \times f_{ij}$	0.010	1	0.010	0.276	.603
$\alpha_i \times \sigma_i \times \phi_i \times f_{ij}$	0.125	1	0.125	3.350	.077
$\beta_i \times \sigma_i \times \phi_i \times f_{ij}$	0.018	1	0.018	0.476	.495
5-Way Interactions	0.056	1	0.056	1.512	.228
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i \times f_{ij}$	0.056	1	0.056	1.512	.228
Explained	8.631	31	0.278	7.464	.000
Residual	1.194	32	0.037		
Total	9.825	63	0.156		

does not contribute to further deterioration in the recovery of σ_i . On the other hand, conditional effects of α_i (σ_i at Level 2) and β_i (σ_i at Level 1) are significant at $p < .05$. Heterogeneity in β_i appears to improve the recovery of σ_i when the unfamiliarity error variance is small, while homogeneity in α_i seems to have a positive effect on recovery when unfamiliarity error variance is large.

Finally, from Table 6, the significant main effects on the recovery of ϕ_i appear to be α_i , β_i , σ_i and ϕ_i . For each factor, the recovery deteriorates in moving from Level 1 to 2. Again, however, there are significant 2-way interactions that must be analyzed: $\alpha_i \times f_{ij}$, $\beta_i \times \phi_i$, and $\sigma_i \times \phi_i$. The main effects of ϕ_i and σ_i are robust, in that they hold unconditionally. That is, the recovery of ϕ_i deteriorates as the unfamiliarity and measurement error variances increase. However, α_i has a significant effect only when f_{ij} is at Level 1, that is, under low heterogeneity in familiarity; and β_i has a significant effect only when ϕ_i is at Level 2, that is, when the measurement error variance is high. In both cases, the recovery deteriorates as heterogeneity in the subject-level parameter (α_i or β_i) increases.

We may summarize the key effects on parameter recovery (i.e., the quality of the estimates) as follows. Low across subject heterogeneity in stimulus familiarity f_{ij} adversely affects parameter recovery for α_i and β_i , when heterogeneity in β_i is low. As in the case of tree recovery, across-subject variation in f_{ij} provides more information for the estimation of α_i or β_i . Variation in β_i has a moderating impact on the f_{ij} effect. (Note from the structure of the model in (1) that f_{ij} and β_i appear multiplicatively.) Recovery for α_i is facilitated when there is heterogeneity in β_i (variation in β_i influences the amount of information available for estimating α_i), but is adversely affected by heterogeneity in α_i itself, which is not surprising. Recovery of the error parameters σ_i and ϕ_i generally deteriorates as the magnitude of either of these parameters increases; that is, when there is greater error (due to measurement and/or uncertainty). This is to be expected, given the model structure: in the absence of perfect partitioning between the two components of error, an increase in error (from either source) will adversely impact the estimation of the variance parameter for both components. In addition, under certain conditions (detailed above), greater across subject heterogeneity in α_i and lower heterogeneity in β_i adversely affect recovery of σ_i , while greater heterogeneity in either α_i or β_i adversely affects recovery of ϕ_i .

Goodness-of-fit. The ANOVA table with the log likelihood value at convergence (LLCONV) as the dependent measure is presented in Table 7. The σ_i and ϕ_i main effects are dominant, with LLCONV decreasing (becoming more negative) in σ_i and ϕ_i , but the $\sigma_i \times \phi_i$ and $\alpha_i \times \phi_i$ interactions are also highly significant. The $\sigma_i \times \phi_i$ interaction is negative in the sense that σ_i and ϕ_i have a moderating effect on each other; however, the individual effects of σ_i and ϕ_i are highly significant under all conditions. The effect of α_i is significant ($p < .05$) when ϕ_i is at Level 1, indicating that when the measurement error variance is small, greater heterogeneity in α_i has an adverse impact on the log likelihood value at convergence. The magnitude of this impact, however, is small relative to the dominant main effects of error variances. The impact of the unfamiliarity and measurement error variances on the log likelihood value at convergence is, of course, entirely expected. For a more meaningful evaluation of the impact of these factors on the goodness-of-fit, we adjusted LLCONV by subtracting the value of the log likelihood function based on the *actual* parameter values used to create Δ (LLACTUAL). Using (LLCONV – LLACTUAL) as the dependent measure, none of the factors has a significant impact, as evidenced by the ANOVA table in Table 8. In effect, this adjustment controls for the obvious distorting effects of the variances of the error terms. The lack of significant effects suggests that the deviation between the log likelihood value at convergence and the log likelihood based on the actual parameter values is not affected by the factors as manipulated in our study, which implies that our procedure is relatively robust, at least under the range of conditions covered in our experimental design.

TABLE 7
ANOVA Table for LLCONV (Log Likelihood Value at Convergence)

Source of Variation	Sum of Squares	d. f.	Mean Square	F	Significance of F (p-value)
Main Effects	39237.110	5	7847.422	147.582	.000
α_i	14.484	1	14.484	0.272	.605
β_i	72.004	1	72.004	1.354	.253
σ_i	10339.280	1	10339.280	194.445	.000
ϕ_i	28810.734	1	28810.734	541.829	.000
f_{ij}	0.609	1	0.609	0.011	.915
2-Way Interactions	2128.147	10	212.815	4.002	.001
$\alpha_i \times \beta_i$	24.248	1	24.248	0.456	.504
$\alpha_i \times \sigma_i$	199.982	1	199.982	3.761	.061
$\alpha_i \times \phi_i$	334.963	1	334.963	6.299	.017
$\alpha_i \times f_{ij}$	113.774	1	113.774	2.140	.153
$\beta_i \times \sigma_i$	15.239	1	15.239	0.287	.596
$\beta_i \times \phi_i$	5.628	1	5.628	0.106	.747
$\beta_i \times f_{ij}$	66.933	1	66.933	1.259	.270
$\sigma_i \times \phi_i$	1307.799	1	1307.799	24.595	.000
$\sigma_i \times f_{ij}$	6.975	1	6.975	0.131	.720
$\phi_i \times f_{ij}$	52.606	1	52.606	0.989	.327
3-Way Interactions	740.340	10	74.034	1.392	.228
$\alpha_i \times \beta_i \times \sigma_i$	67.289	1	67.289	1.265	.269
$\alpha_i \times \beta_i \times \phi_i$	49.646	1	49.646	0.934	.341
$\alpha_i \times \beta_i \times f_{ij}$	7.223	1	7.223	0.136	.715
$\alpha_i \times \sigma_i \times \phi_i$	15.727	1	15.727	0.296	.590
$\alpha_i \times \sigma_i \times f_{ij}$	43.920	1	43.920	0.826	.370
$\alpha_i \times \phi_i \times f_{ij}$	150.952	1	150.952	2.839	.102
$\beta_i \times \sigma_i \times \phi_i$	100.501	1	100.501	1.890	.179
$\beta_i \times \sigma_i \times f_{ij}$	166.500	1	166.500	3.131	.086
$\beta_i \times \phi_i \times f_{ij}$	138.345	1	138.345	2.602	.117
$\sigma_i \times \phi_i \times f_{ij}$	0.237	1	0.237	0.004	.947
4-Way Interactions	327.906	5	65.581	1.233	.317
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i$	94.639	1	94.639	1.780	.192
$\alpha_i \times \beta_i \times \sigma_i \times f_{ij}$	80.564	1	80.564	1.515	.227
$\alpha_i \times \beta_i \times \phi_i \times f_{ij}$	32.402	1	32.402	0.609	.441
$\alpha_i \times \sigma_i \times \phi_i \times f_{ij}$	56.776	1	56.776	1.068	.309
$\beta_i \times \sigma_i \times \phi_i \times f_{ij}$	63.525	1	63.525	1.195	.283
5-Way Interactions	15.245	1	15.245	0.287	.596
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i \times f_{ij}$	15.245	1	15.245	0.287	.596
Explained	42448.748	31	1369.314	25.752	.000
Residual	1701.541	32	53.173		
Total	44150.289	63	700.798		

Full versus restricted model. Although the synthetic data were generated by the full model (1) itself (with error), it is nevertheless instructive to compare the ability of the full model to recover the prespecified tree structure with that of the more parsimonious restricted model (2). $VAF(\mathbf{D})$ for the full model was higher than $VAF(\mathbf{D})$ for the

TABLE 8
ANOVA Table for (LLCONV – LLACTUAL)

Source of Variation	Sum of Squares	d. f.	Mean Square	F	Significance of F (p-value)
Main Effects	96.519	5	19.304	0.444	.814
α_i	47.626	1	47.626	1.096	.303
β_i	1.209	1	1.209	0.028	.869
σ_i	8.246	1	8.246	0.190	.666
ϕ_i	35.630	1	35.630	0.820	.372
f_{ij}	3.808	1	3.808	0.088	.769
2-Way Interactions	273.909	10	27.391	0.630	.777
$\alpha_i \times \beta_i$	54.179	1	54.179	1.247	.272
$\alpha_i \times \sigma_i$	4.249	1	4.249	0.098	.757
$\alpha_i \times \phi_i$	48.318	1	48.318	1.112	.300
$\alpha_i \times f_{ij}$	6.956	1	6.956	0.160	.692
$\beta_i \times \sigma_i$	49.525	1	49.525	1.140	.294
$\beta_i \times \phi_i$	33.824	1	33.824	0.778	.384
$\beta_i \times f_{ij}$	0.999	1	0.999	0.023	.880
$\sigma_i \times \phi_i$	47.353	1	47.353	1.090	.304
$\sigma_i \times f_{ij}$	11.247	1	11.247	0.259	.614
$\phi_i \times f_{ij}$	17.259	1	17.259	0.397	.533
3-Way Interactions	283.441	10	28.344	0.652	.758
$\alpha_i \times \beta_i \times \sigma_i$	5.709	1	5.709	0.131	.719
$\alpha_i \times \beta_i \times \phi_i$	16.582	1	16.582	0.382	.541
$\alpha_i \times \beta_i \times f_{ij}$	1.707	1	1.707	0.039	.844
$\alpha_i \times \sigma_i \times \phi_i$	35.350	1	35.350	0.814	.374
$\alpha_i \times \sigma_i \times f_{ij}$	22.538	1	22.538	0.519	.477
$\alpha_i \times \phi_i \times f_{ij}$	64.603	1	64.603	1.487	.232
$\beta_i \times \sigma_i \times \phi_i$	3.273	1	3.273	0.075	.785
$\beta_i \times \sigma_i \times f_{ij}$	113.638	1	113.638	2.615	.116
$\beta_i \times \phi_i \times f_{ij}$	1.069	1	1.069	0.025	.876
$\sigma_i \times \phi_i \times f_{ij}$	18.971	1	18.971	0.437	.513
4-Way Interactions	229.311	5	45.862	1.056	.403
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i$	6.964	1	6.964	0.160	.692
$\alpha_i \times \beta_i \times \sigma_i \times f_{ij}$	22.280	1	22.280	0.513	.479
$\alpha_i \times \beta_i \times \phi_i \times f_{ij}$	28.624	1	28.624	0.659	.423
$\alpha_i \times \sigma_i \times \phi_i \times f_{ij}$	135.894	1	135.894	3.128	.087
$\beta_i \times \sigma_i \times \phi_i \times f_{ij}$	35.550	1	35.550	0.818	.372
5-Way Interactions	0.379	1	0.379	0.009	.926
$\alpha_i \times \beta_i \times \sigma_i \times \phi_i \times f_{ij}$	0.379	1	0.379	0.009	.926
Explained	883.560	31	28.502	0.656	.878
Residual	1390.373	32	43.449		
Total	2273.933	63	36.094		

restricted model on *each* of the 64 trials; the mean VAF(D) across all trials was .893 for the full model versus .738 for the restricted model. We also extended the ANOVA with VAF(D) as the dependent variable by including the model as the sixth factor at two levels (Full and Restricted) in a 2^6 factorial design. The resulting ANOVA table is

TABLE 9
ANOVA Table for VAF(D) with Model as a Factor*

Source of Variation	Sum of Squares	d. f.	Mean Square	F	Significance of F (p-value)
Main Effects	21.181	6	3.530	17.637	.000
α_i	1.268	1	1.268	6.335	.014
β_i	0.622	1	0.622	3.106	.083
σ_i	0.020	1	0.020	0.098	.756
ϕ_i	0.000	1	0.000	0.002	.968
f_{ij}	2.149	1	2.149	10.738	.002
Model	17.122	1	17.122	85.541	.000
2-Way Interactions	4.801	15	0.320	1.599	.099
$\alpha_i \times \beta_i$	0.003	1	0.003	0.015	.904
$\alpha_i \times \sigma_i$	0.136	1	0.136	0.679	.413
$\alpha_i \times \phi_i$	0.167	1	0.167	0.835	.364
$\alpha_i \times f_{ij}$	1.110	1	1.110	5.543	.022
$\alpha_i \times \text{Model}$	0.184	1	0.184	0.921	.341
$\beta_i \times \sigma_i$	0.394	1	0.394	1.970	.165
$\beta_i \times \phi_i$	0.013	1	0.013	0.066	.798
$\beta_i \times f_{ij}$	0.075	1	0.075	0.376	.542
$\beta_i \times \text{Model}$	0.333	1	0.333	1.664	.202
$\sigma_i \times \phi_i$	0.243	1	0.243	1.214	.275
$\sigma_i \times f_{ij}$	0.162	1	0.162	0.807	.372
$\sigma_i \times \text{Model}$	0.036	1	0.036	0.180	.673
$\phi_i \times f_{ij}$	0.822	1	0.822	4.106	.047
$\phi_i \times \text{Model}$	0.001	1	0.001	0.006	.940
$f_{ij} \times \text{Model}$	1.121	1	1.211	5.601	.021
3-Way Interactions	2.834	20	0.142	0.708	.804
4-Way Interactions	0.920	15	0.061	0.306	.993
5-Way Interactions	0.268	6	0.045	0.223	.968
Explained	30.004	62	0.484	2.418	.000
Residual	13.011	65	0.200		
Total	43.015	127	0.339		

*Dependent variable used: $z = \ln [(1 + \sqrt{\text{VAF(D)}}) / (1 - \sqrt{\text{VAF(D)}})] / 2$

presented in Table 9. (For the sake of brevity, since none of the individual 3- or higher-way interaction effects were significant, we have shown in detail only the main effects and 2-way interaction effects individually.) It may be observed that the model is the dominant main effect ($p < .001$), accounting for 81% of the total main effects sum of squares and 57% of the total sum of squares explained by the saturated model. The model $\times f_{ij}$ interaction is significant, which is expected, since the f_{ij} main effect is significant in the case of the full model whereas there is no such effect in the restricted model by definition (f_{ij} is ignored in the restricted model specification). In sum, *the full model does significantly better in recovering the tree than the restricted model* which, as mentioned earlier, is equivalent to the Carroll and Pruzansky (1975, 1980) and De Soete (1984) tree fitting procedures that ignore unfamiliarity effects.

Discussion

The Monte Carlo study yields some interesting and important insights into the performance of the TREEFAM procedure, evaluated in terms of its ability to recover the prespecified ultrametric tree and model parameters used to generate synthetic data. An important factor affecting the estimation of the tree as well as the subject-level parameters, α_i and β_i , is the pattern of stimulus familiarity across subjects. Recovery of the tree, as well as the parameters α_i and β_i , tends to be better when subjects are heterogeneous in terms of their stimulus familiarity profiles. Heterogeneity implies that, for brand j , f_{ij} tends to vary over subjects, whereas homogeneity results in f_{ij} being at some consistent level across subjects for each brand. As discussed, the former is better in terms of information available in the data for estimation. It is also worth noting that the error variances, σ_i^2 and ϕ_i^2 , appear to have little impact on the quality of estimates of the ultrametric distance matrix (i.e., the tree structure) or the parameters α_i and β_i (possibly excepting the adverse effect of σ_i^2 on the estimate of β_i , which approaches significance). The size of the error variances do, of course, affect estimates of the variance parameters themselves. However, the variance parameters are incidental: the estimates of the tree and the subject-level parameters of interest appear to be quite robust to variation in the magnitude of error. The analysis of variance with the log likelihood value at convergence adjusted by the log likelihood based on actual parameter values as the dependent variable also points to the robustness of the estimation procedure. While the full model was used to create the synthetic data and therefore the full model might naturally be expected to perform better than the restricted model, the simulations demonstrate that ignoring unfamiliarity can potentially result in significantly poorer fits and serious distortions in the estimated ultrametric tree.

The various ANOVA analyses have examined the *relative* performance of the proposed methodology as a number of independent factors have been experimentally varied. To examine *absolute* performance, we present Table 10 which displays the means and standard deviations for the parameter recovery measures over the 32 cells of the Monte Carlo experiment. Concerning recovery of β_i , only two of the cells of the design (12111 and 12112) show inordinately large RMSE values. Consistent fitting and parameter recovery is witnessed with respect to the entire design in the case of the RMSE measures for α_i , σ_i , and ϕ_i . Finally, with the possible exception of one cell (11112), the ultrametric tree appears to be consistently recovered in the design. In summary, then, the true parameters used to generate the data for these 32 experimental trials appear to be consistently recovered in this Monte Carlo analysis.

The Monte Carlo study provides a basis for evaluating the estimation procedure under a variety of conditions since it allows comparison of estimates with actual (synthetic) values and systematic manipulation of the conditions. An obvious limitation to such an experimental approach is that the magnitude (and consequently the statistical significance) of effects depends on the treatment levels selected, as well as the total number of trials. Clearly, more thorough Monte Carlo testing is appropriate utilizing different number of stimuli, more replications, different tree structures, etc. Finally, it must be noted that the data are synthetically generated by the model (1), given the parameter values for the trial. Thus, the Monte Carlo evaluation of the TREEFAM procedure is *conditioned* on the assumption that model (1) is theoretically correct.

² The p -value corresponding to the main effect σ_i on $\text{RMSE}(\beta_i)$ is .087 (see Table 4). When σ_i is large, $\text{RMSE}(\beta_i)$ increases.

TABLE 10
 Absolute Performance Measures for Model Parameter Recovery

Factor Levels*					Cell Means (Standard Deviations) of Parameter Recovery Measures				
β_i	σ_i	ϕ_i	f_{ij}	α_i	RMSE(α_i)	RMSE(β_i)	RMSE(σ_i)	RMSE(ϕ_i)	VAF(D)
1	1	1	1	1	.549 (.126)	.918 (.271)	1.043 (.388)	.283 (.037)	.833 (.035)
1	1	1	1	2	.833 (.048)	1.091 (.484)	.466 (.070)	.442 (.049)	.702 (.018)
1	1	1	2	1	.393 (.179)	.612 (.196)	1.009 (.527)	.261 (.010)	.934 (.064)
1	1	1	2	2	.925 (.028)	.726 (.161)	.792 (.172)	.485 (.066)	.883 (.004)
1	1	2	1	1	.468 (.170)	1.284 (.520)	1.653 (.268)	1.002 (.174)	.800 (.085)
1	1	2	1	2	.879 (.261)	2.786 (1.869)	1.364 (.364)	1.116 (.085)	.865 (.034)
1	1	2	2	1	.280 (.028)	.636 (.060)	1.094 (.036)	1.149 (.068)	.917 (.056)
1	1	2	2	2	.539 (.033)	.850 (.196)	1.509 (.327)	1.244 (.193)	.912 (.039)
1	2	1	1	1	.577 (.225)	7.572 (6.246)	1.298 (.051)	.755 (.161)	.866 (.037)
1	2	1	1	2	.908 (.016)	6.396 (5.801)	1.454 (.176)	1.116 (.037)	.829 (.033)
1	2	1	2	1	.176 (.069)	.951 (.270)	1.403 (.125)	.779 (.256)	.986 (.001)
1	2	1	2	2	.517 (.194)	.718 (.085)	1.593 (.134)	.814 (.171)	.923 (.012)
1	2	2	1	1	.664 (.032)	3.383 (.512)	.797 (.238)	.769 (.041)	.922 (.040)
1	2	2	1	2	.621 (.020)	1.284 (.264)	1.439 (.002)	1.167 (.156)	.914 (.056)
1	2	2	2	1	.484 (.216)	.759 (.230)	1.403 (.125)	.989 (.116)	.966 (.003)
1	2	2	2	2	.547 (.081)	.821 (.002)	1.618 (.185)	.746 (.167)	.919 (.001)
2	1	1	1	1	.314 (.081)	1.237 (.917)	.752 (.323)	.280 (.079)	.826 (.163)
2	1	1	1	2	.373 (.245)	1.663 (.982)	.822 (.190)	.323 (.008)	.895 (.092)
2	1	1	2	1	.176 (.041)	.793 (.263)	.478 (.041)	.460 (.033)	.990 (.004)
2	1	1	2	2	.350 (.073)	.578 (.293)	.517 (.059)	.296 (.031)	.965 (.026)
2	1	2	1	1	.372 (.071)	1.598 (.001)	1.265 (.136)	1.107 (.075)	.945 (.010)
2	1	2	1	2	.530 (.242)	.709 (.254)	.775 (.383)	1.249 (.173)	.833 (.080)
2	1	2	2	1	.155 (.061)	.738 (.111)	.954 (.053)	1.284 (.176)	.967 (.026)
2	1	2	2	2	.438 (.164)	1.068 (.271)	.863 (.296)	1.409 (.190)	.823 (.169)
2	2	1	1	1	.215 (.018)	1.081 (.359)	1.252 (.018)	.592 (.010)	.827 (.033)
2	2	1	1	2	.313 (.085)	1.380 (.470)	1.497 (.215)	.902 (.006)	.873 (.104)
2	2	1	2	1	.129 (.003)	1.142 (.620)	1.171 (.009)	.639 (.381)	.973 (.003)
2	2	1	2	2	.620 (.176)	2.013 (.949)	1.557 (.250)	1.251 (.232)	.875 (.036)
2	2	2	1	1	.291 (.043)	2.521 (.673)	1.142 (.468)	1.061 (.075)	.861 (.112)
2	2	2	1	2	.470 (.103)	1.313 (.765)	1.804 (.013)	1.484 (.028)	.978 (.001)
2	2	2	2	1	.236 (.077)	.925 (.200)	1.449 (.113)	1.341 (.043)	.951 (.037)
2	2	2	2	2	.769 (.305)	1.068 (.356)	1.476 (.069)	1.073 (.057)	.841 (.127)

*See Table 1 for description of Levels 1 and 2 for each factor.

4. Illustrative Application

Setting

The data for our empirical illustration of the TREEFAM procedure consist of proximity and familiarity judgments on a set of 14 beer brands (Beck's, Budweiser, Corona, Dos Equis, Heineken, Kaliber, Kingsbury, Labatt's Blue, Lowenbrau, Michelob, Miller Genuine Draft, Signature, St. Pauli Girl, and Stroh's)³ collected from 50 graduate students. Specifically, similarity ratings for each brand pair and familiarity (confidence in judging the brand) ratings for each brand were collected from the subjects on 7-point bipolar scales ("very dissimilar—very similar" and "not at all confident"—"very confident" respectively).

³ These names are registered trademarks of their respective companies.

The degree of familiarity with the brand corresponds to the extent of knowledge the consumer has about the brand. The extent of knowledge, in turn, affects the confidence in the consumer's judgments (in this case, proximity judgments) involving the brand. Alba and Hutchinson (1987) define product familiarity as "the number of product related experiences that have been accumulated by the consumer," including "... advertising exposures, information search, interactions with salespersons, choice and decision making, purchasing, and product usage in various situations". This definition applies in our context, except that we are concerned with familiarity at the *brand*, rather than the product [category], level.

Analysis and Results

Test of the premise of anchoring under unfamiliarity. Before applying our methodology to the data, we wanted to test the hypothesis that subjects tend to moderate their proximity judgments under stimulus familiarity—that is, whether there is a tendency for subjects to anchor their response in case of unfamiliarity near the mean value. We proceeded as follows. For each subject i , the average dissimilarity for stimulus j from the other $N - 1$ stimuli is

$$\delta_{ij}^* = \left[\frac{1}{N-1} \right] \sum_{k \neq j} \delta_{ijk}, \quad (17)$$

while the mean interobject dissimilarity is

$$\bar{\delta}_i = \left[\frac{2}{N(N-1)} \right] \sum_{j < k} \delta_{ijk}. \quad (18)$$

We can then define the deviation from the mean for stimulus j by:

$$\text{DEV}_{ij} = |\delta_{ij}^* - \bar{\delta}_i|. \quad (19)$$

Note that DEV_{ij} consists of a stimulus-specific component and (if there is an anchoring effect) a component that is a function of stimulus unfamiliarity. Thus, in order to test the anchoring premise, we need to investigate the association between f_{ij} and DEV_{ij} , *controlling* for possible stimulus effects. We therefore estimated the following regression model:

$$\text{DEV}_{ij} = b_0 + \sum_k b_k x_k + b_N f_{ij} + v_{ij}, \quad (20)$$

where $x_k = 1$ if $k = j$, 0 otherwise, for $k = 1, \dots, N - 1$ (i.e., $j = N$ is the baseline), and v_{ij} is the error term, assumed iid $N(0, \xi^2)$. The stimulus-specific dummy variables control for stimulus effects. The anchoring hypothesis corresponds to a test of whether $b_N > 0$ (i.e., whether the partial association between f_{ij} and DEV_{ij} is positive). The results of the analysis yielded an R^2 of 0.32 ($p < .0001$). The estimate of b_N was 0.0336 ($p < .0005$), supporting the hypothesis of an anchoring effect in the presence of stimulus unfamiliarity.

Profile of the brands. Table 11 provides a description of the 14 brands constituting the stimulus set. The summary measures in the table consist of the mean familiarity for each brand (calculated over subjects), the across subject mean of δ_{ij}^* (which is the subject level average dissimilarity for stimulus j from all other stimuli as defined in (17)),

TABLE 11

A Profile of the Brands*

j	Brand Name	Mean (f_{ij})	Mean (δ_{ij}^*)	Variance (DEV_{ij})
1	St. Pauli Girl	4.66	3.935	0.0414
2	Beck's	4.84	3.923	0.0448
3	Miller G' Draft	5.74	3.771	0.0678
4	Kaliber	3.38	2.708	0.6312
5	Lowenbrau	4.96	3.872	0.0810
6	Dos Equis	4.58	3.543	0.0800
7	Signature	4.38	3.825	0.0527
8	Budweiser	6.12	3.646	0.0779
9	Stroh's	5.58	3.846	0.0596
10	Labatt's Blue	5.18	3.831	0.0600
11	Kingsbury	3.42	2.543	0.6811
12	Heineken	5.76	4.052	0.0816
13	Corona	5.68	3.535	0.1055
14	Michelob	5.78	3.972	0.0569

*The brands are defined in terms of summary measures (means and variance) calculated across subjects.

and the variance (across subjects) of the deviation measure DEV_{ij} defined in (19). The two least familiar brands, Kaliber and Kingsbury, are both nonalcoholic beers that had recently been introduced to the market at the time of the study. Note that Variance (DEV_{ij}) is exceptionally large for these two cases, suggestive of less patterned responses associated with these relatively unfamiliar brands (although the variance is across, not within, subjects). The mean (δ_{ij}^*) scores for these two brands are also lower than for the other brands, suggesting that the anchor response under unfamiliarity (at least in this instance) may be somewhat lower than the overall average dissimilarity

⁴ Calculation of an appropriate within-subject variance would require replication within subjects which was not available in this case.

rating. That is, the anchor appears to be biased toward the “similar” end of the proximity scale, relative to the overall average rating or the midpoint of the scale (which is 4 on the 1–7 scale used in this study). Interestingly, if the presence (or absence) of alcohol is viewed as a basic dimension in subjects’ perceptual structure of this set of beers (and our analysis below suggests that this is the case), then the “true” value of δ_{ij}^* (under familiarity) should be *larger* for Kaliber and Kingsbury relative to other brands. The distorting effect of stimulus unfamiliarity seems to be in evidence.

Model estimation and evaluation. Ideally, one would like to know the *true* structure of the tree against which the estimated trees can be compared. In the absence of any strictly objective basis for determining the actual (i.e., true) underlying tree structure, we constructed a tree for the 50 subjects using only those proximity judgments that involved brands rated as completely familiar to serve as the “truth” (at least in the perception of this group of subjects). This tree, shown in Figure 3, was constructed by the complete linkage method in the SPSS-X (release 3.0) hierarchical cluster analysis program.

Examining Figure 3, the structure appears to have substantial face validity. The brands are first classified on the basis of alcohol content: alcoholic versus nonalcoholic (Kaliber and Kingsbury are the nonalcoholic brands). Next, the alcoholic beers are classified as imported or domestic. The imports are next classified into Mexican (Dos Equis and Corona) and non-Mexican brands. The non-Mexican brands then split into Lowenbrau (which is actually manufactured in the U.S.) and others. The classification of Lowenbrau along with foreign (rather than domestic) brands appears to be a function of the brand’s market positioning, and its impact on consumer perceptions. The others (which are genuinely foreign brands) are further classified by country of origin into European (Beck’s and St. Pauli Girl, which are German, and Heineken, which is Dutch) and Canadian (Labatt’s). The domestic brands break into Signature (a super-premium brand with low market share) and others (the more mainstream brands). The latter, in turn, are classified as draft (Miller Genuine Draft) and nondraft (Budweiser, Michelob, and Stroh’s).

For this illustrative application, we focussed on four subjects for detailed analysis. This focus enables us to examine the implications of the TREEFAM procedure in greater depth, particularly relative to approaches that ignore stimulus unfamiliarity. The complete linkage trees (ignoring the effects of unfamiliarity) for these four subjects (# 1, 2, 7, and 41) are shown in Figure 4 (Panels A, B, C, and D). The structures do not appear to be logically interpretable, except in the case of Subject # 2, whose structure reveals the initial alcoholic versus nonalcoholic split, but is unclear thereafter. When the four subjects are considered as a sample, and a common tree derived (i.e., based on proximity data averaged over the subjects), the structure is equally uninterpretable (see Figure 5).

We applied the TREEFAM procedure to the data from these four subjects, varying program options to generate six different models as follows:

- A: Conditional (external) restricted model: estimate ϕ_i , $\forall i$, given the 4-subject complete linkage tree in Figure 5;
- B: Conditional (external) full model: estimate α_i , β_i , σ_i , and ϕ_i , $\forall i$, given the 4-subject complete linkage tree in Figure 5;

⁵ The criterion for inclusion was a maximum familiarity rating of 7 for both brands in the pair. However, there were some brand pairs that did not meet this criterion for any of the 50 subjects. To avoid missing data, in such cases brands with familiarity ratings of 6 and (in a few cases) 5 were included.

Dendrogram using Complete Linkage

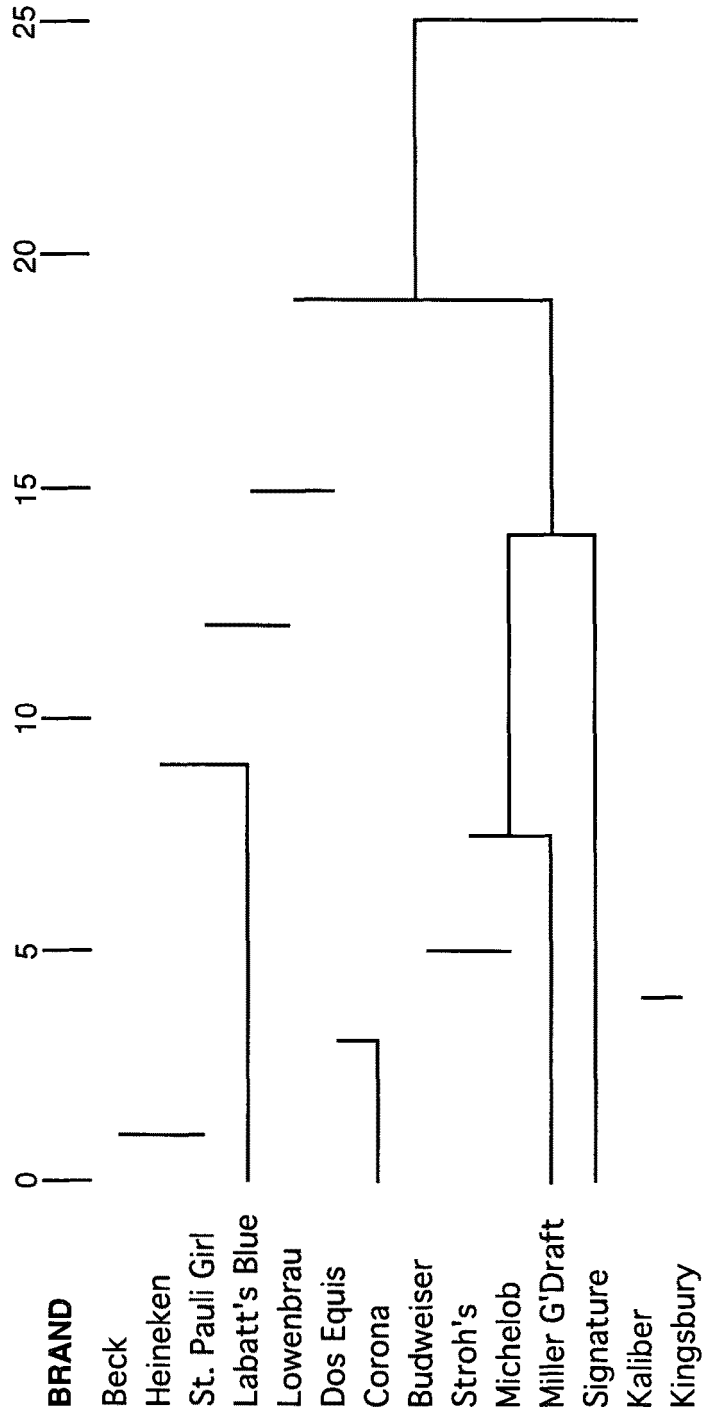


FIGURE 3.
Ultrametric tree based on proximity judgments for completely familiar brands only: The "Truth."

Dendrogram using Complete Linkage

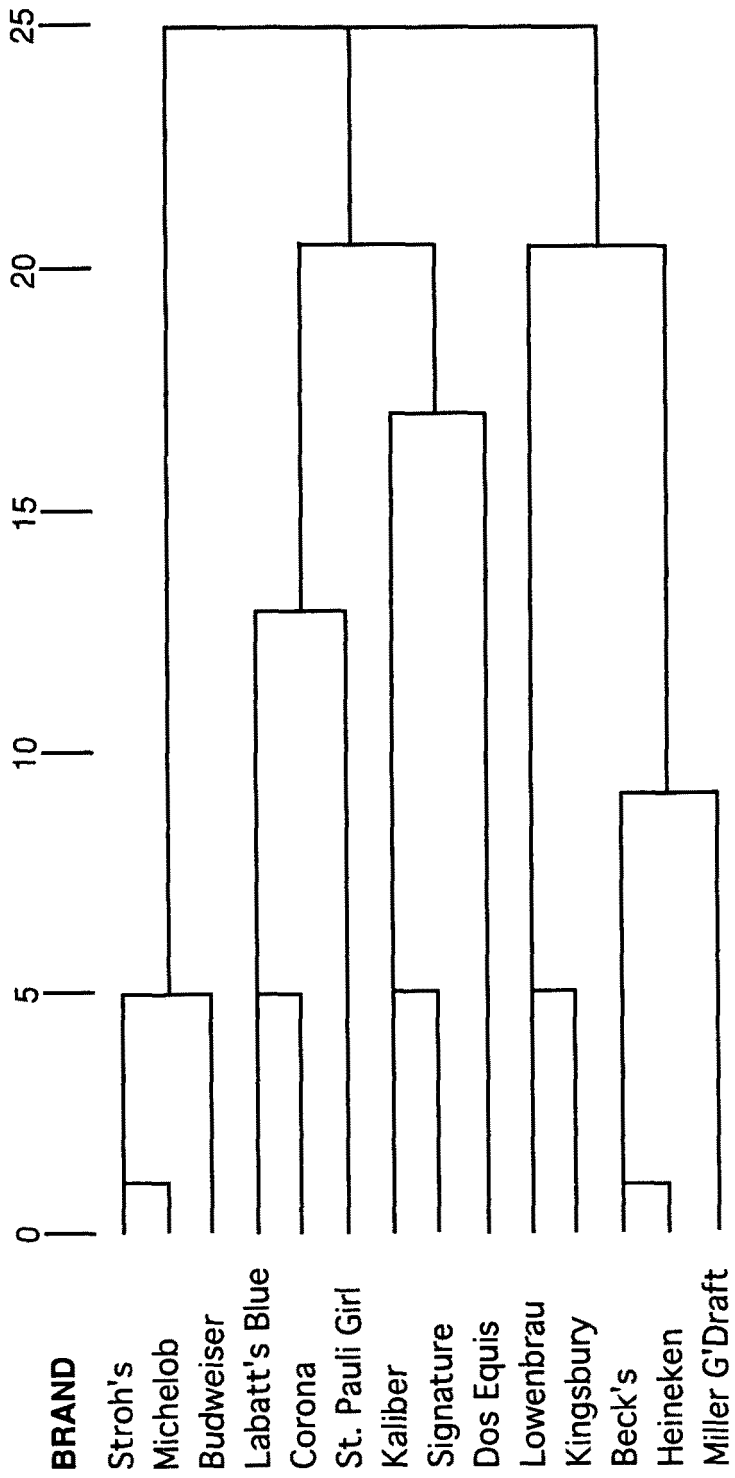


FIGURE 4A.
Ultrametric tree for Subject #1.

Dendrogram using Complete Linkage

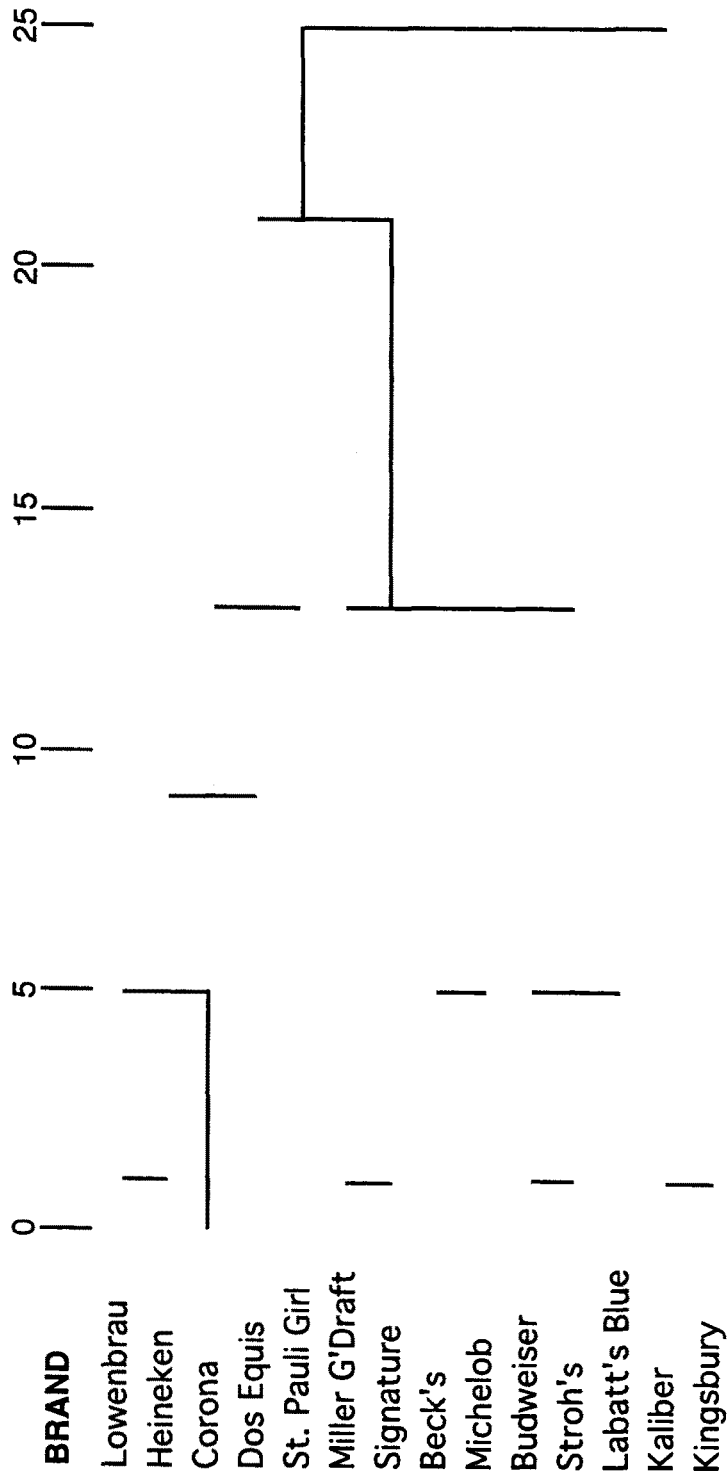


FIGURE 4B.
Ultrametric tree for Subject #2.

Dendrogram using Complete Linkage

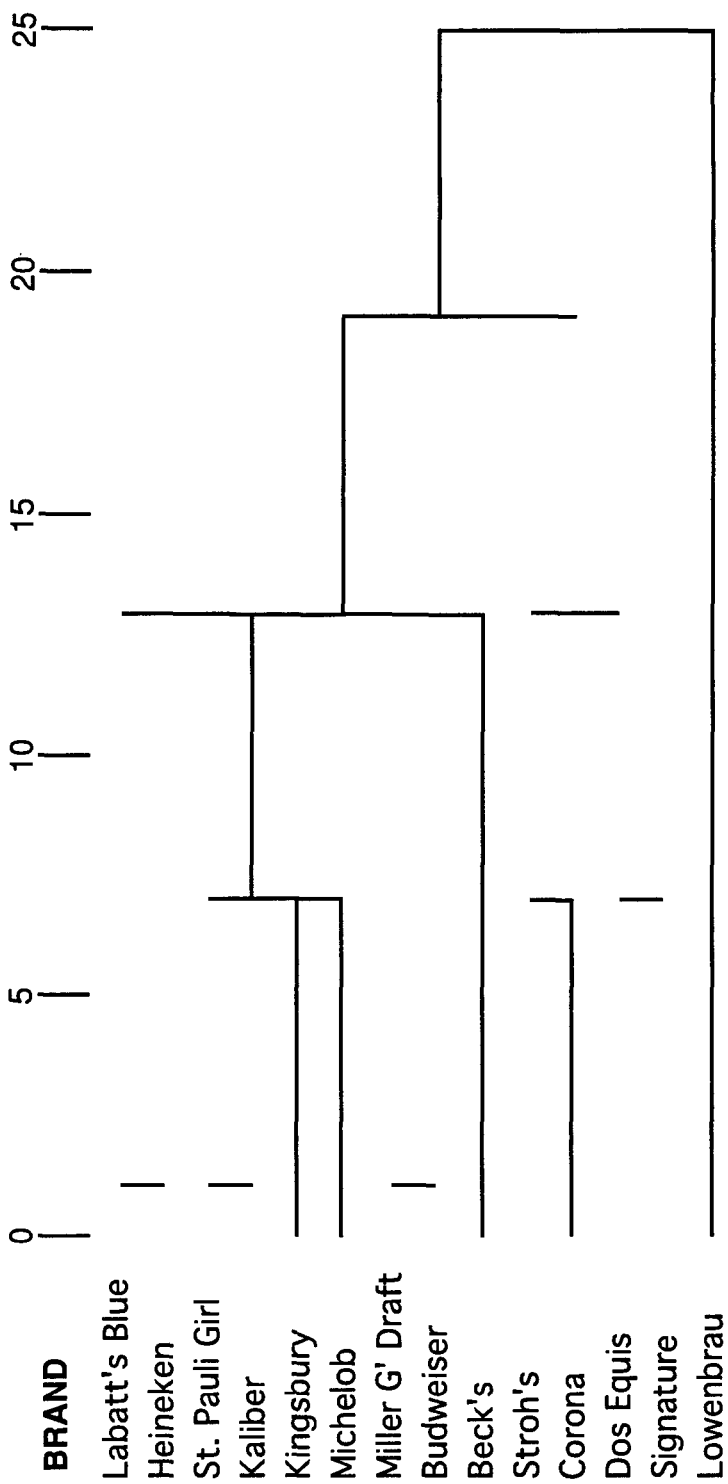


FIGURE 4C.
Ultrametric tree for Subject #7.

Dendrogram using Complete Linkage

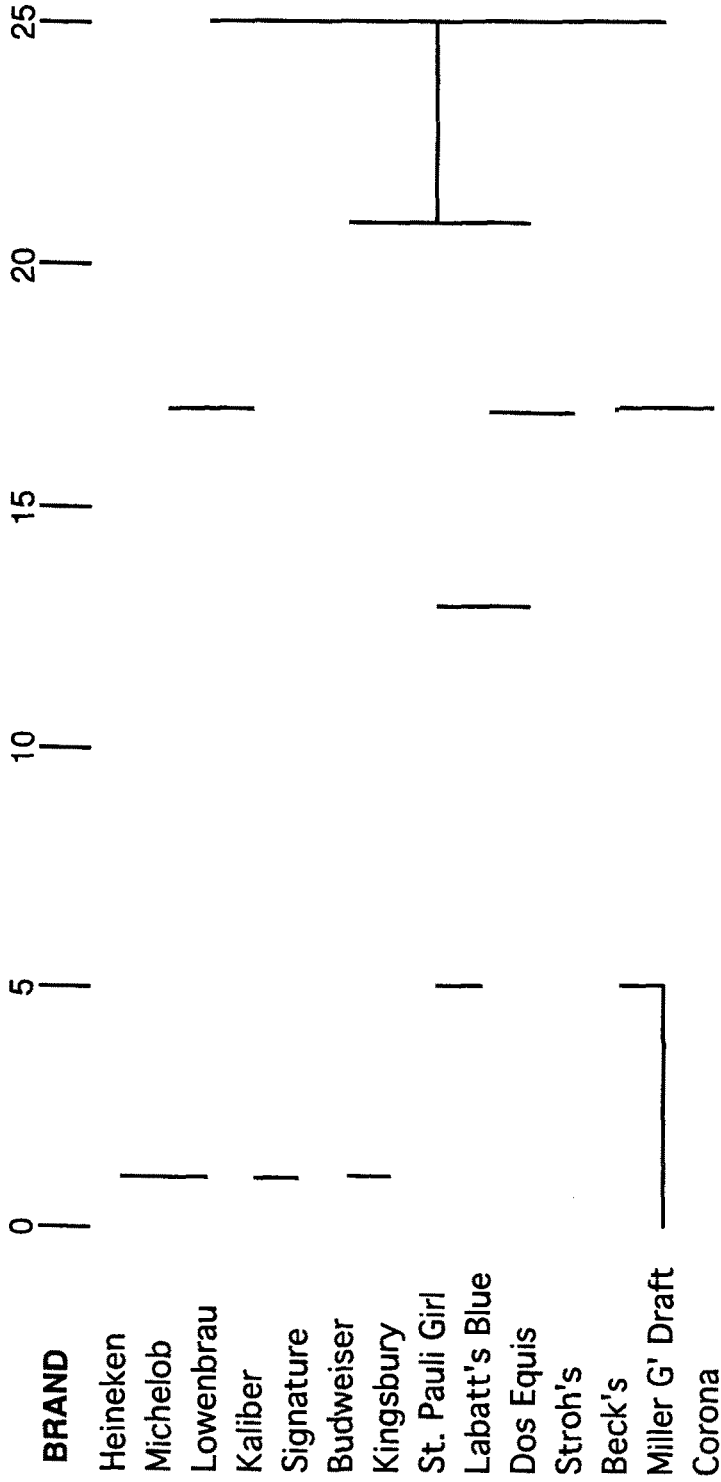


FIGURE 4D.
Ultrametric tree for Subject #41.

Dendrogram using Complete Linkage

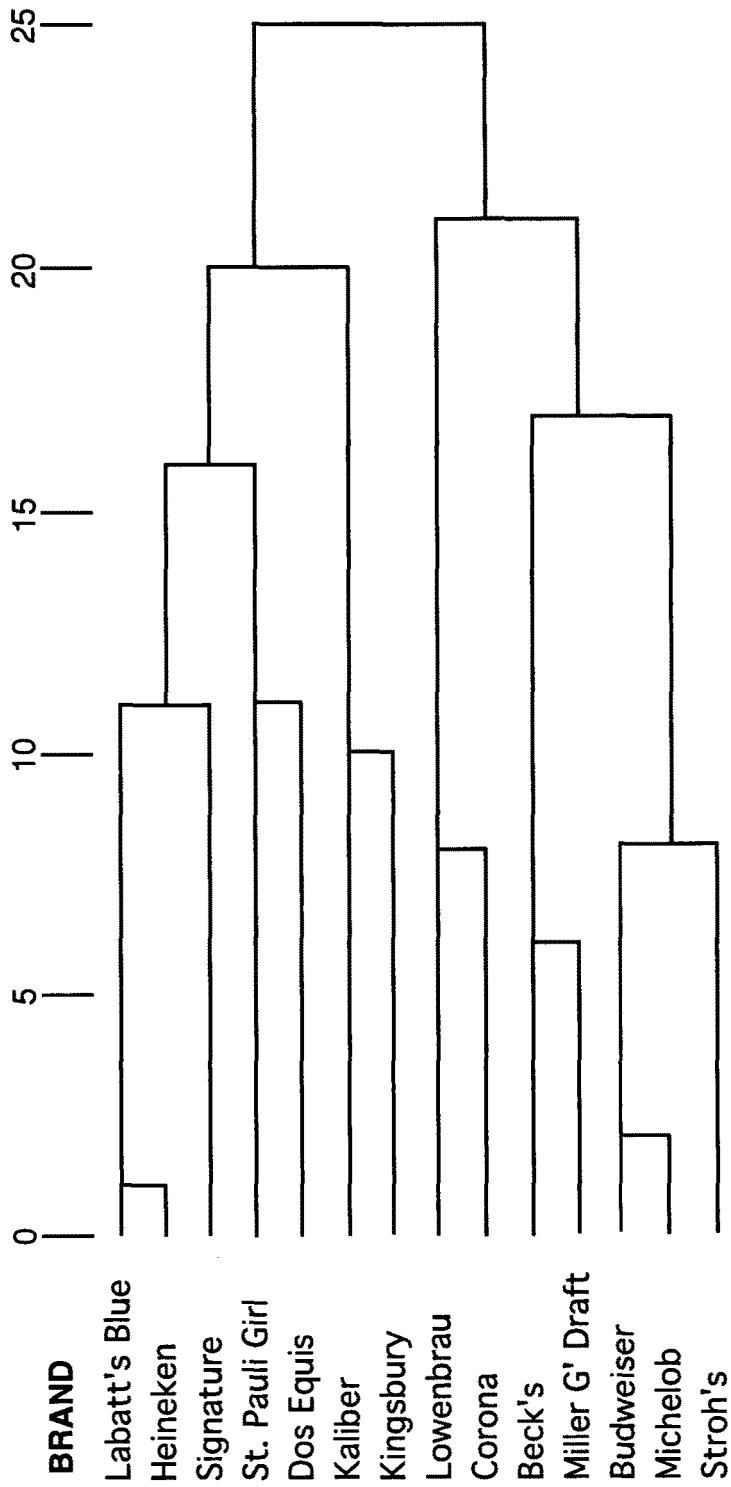


FIGURE 5.
 Ultrametric tree based on proximity judgments averaged over the four subjects.

- C: Conditional (external) restricted model: estimate $\phi_i, \forall i$, given the “true” tree in Figure 3;
- D: Conditional (external) full model: estimate $\alpha_i, \beta_i, \sigma_i$, and $\phi_i, \forall i$, given the “true” tree in Figure 3;
- E: Unconditional (internal) restricted model: estimate tree and $\phi_i, \forall i$;
- F: Unconditional (internal) full model: estimate tree, and $\alpha_i, \beta_i, \sigma_i$, and $\phi_i, \forall i$.

Model F is the full model specified in (1) with the tree and all subject-level parameters estimated by the TREEFAM procedure. The other models (A through E) serve as bases for comparative evaluation of the unconditional full model. In selecting these models, we essentially manipulated two “factors”: treatment of stimulus unfamiliarity and the type of analysis. Stimulus familiarity was either ignored, in line with the formulation in (2) (Models A, C and E) or explicitly considered as per specification in (1) (Models B, D and F). The type of analysis was either *external*, where the tree was prespecified (Models A through D), or *internal*, where the tree was estimated by the TREEFAM procedure (Models E and F). In the case of external analysis, the given tree was either the 4-subject solution in Figure 5 (Models A and B) or the “truth” (Models C and D). If our methodology is truly appropriate, we would expect a significant improvement in goodness-of-fit moving from ignoring unfamiliarity to considering it explicitly, and also from moving from conditional (external) to unconditional (internal) analysis. The tree resulting from the internal analysis should, however, be close to the “true” tree in Figure 5.

In Table 12, the models are compared in terms of their goodness-of-fit to data. The improvement in moving from the restricted to the full specification, and from conditional on a given tree to an unconditional version in which the tree is freely estimated is apparent. Akaike’s (1973, 1974) information criterion (AIC) adjusts the log likelihood measure to reflect the number of parameters. The Bayesian information criterion, or BIC (Schwarz, 1978) corrects for the tendency of AIC to select overspecified models (Koehler and Murphree, 1988), while the consistent AIC, or CAIC (Bozdogan, 1987), penalizes overparameterization more strongly than does AIC or BIC. Based on all three criteria (AIC, BIC, and CAIC), the unconditional full Model F dominates the other models. Model F is also superior when compared to the other (nested) models, using a likelihood ratio test ($p < .001$ in each comparison).

The trees corresponding to the unconditional restricted Model E and the unconditional full Model F are shown in Figures 6 and 7. The correlation between the ultrametric distances derived from the restricted Model E and the distances based on the “true” tree (Figure 3) is .508; in comparison, the correlation between the ultrametric distances derived from the full Model F and the “true” tree is .874. Visually, the tree in Figure 6 does not appear to have an interpretable structure. On the other hand, the tree from the full model in Figure 7 does a much better job of recovering the structure of the “true” tree in Figure 3, although they are not identical. The principal anomaly in the structure is the relatively early branching out of Dos Equis, a Mexican beer brand that had very limited availability in the area where study was conducted.

The subject level parameter estimates for the unconditional full Model F are reported in Table 13. It is interesting to note that the values of the anchor parameter β_i for the four subjects are close to the mean (across stimulus pairs) of their stated dissimilarities (2.74, 2.75, 2.64, and 2.72 respectively). At least for these subjects, the anchor response (under complete unfamiliarity) appears to correspond to an average

⁶ AIC = $-2 \ln L + 2k$, BIC = $-2 \ln L + k (\ln n)$, while CAIC = $-2 \ln L + k (\ln n + 1)$, where k = number of free parameters in the model and n = sample size. In our application, $n = I \times N (N - 1)/2 = 364$, since $I = 4$ and $N = 14$.

TABLE 12

Goodness-of-Fit Statistics for Alternative Models

Model	Log Likelihood	Free Parameters	AIC	BIC	CAIC
A	- 870.65	4	1749.30	1764.89	1768.89
B	- 464.89	16	961.78	1024.13	1040.13
C	- 819.76	4	1647.52	1663.11	1667.11
D	- 439.85	16	911.70	974.05	990.05
E	- 449.37	17	932.74	998.99	1015.99
F	- 249.89	29	557.78*	670.80*	699.80*

*denotes minimum AIC, BIC, and CAIC.

interobject dissimilarity rating. Relative to β_i , there is wide variation across subjects with respect to the exponent parameter α_i , which is particularly small in the case of Subject # 2. What this suggests in substantive terms is that, for a given level of brand familiarity, this subject's proximity judgments are nearer to corresponding judgments under complete brand familiarity, compared to another subject with a larger α_i (since $\alpha_1 < \alpha_2$ implies that the familiarity weight $(f_{ijk})^{\alpha_1} > (f_{ijk})^{\alpha_2}$). Thus, Subject # 2's low exponent parameter value is consistent with the observation that his/her tree structure is closer to the "truth" than that of the other subjects (compare Panel B with the other panels in Figure 4).

In order to assess the impact of the problem of local optimum solutions, which is a potential source of concern with our methodology, we performed five additional analyses for the full model on this beer data set using different random starts. Table 14 presents the detailed results of these five analyses in comparing computational requirements, goodness of fit, and comparative parameter recovery. As shown, the five values of the objective function are within $\pm 3\%$ of each other. The parameter values are quite close across the five analyses, especially in the cases of β_i and ϕ_i . In particular, as noted by the correlations of the five ultrametric tree distances, the recovered tree structures are nearly identical. Nevertheless, given the possibility of local optima, it is prudent to repeat the analysis using different random starts.

Dendrogram using Complete Linkage

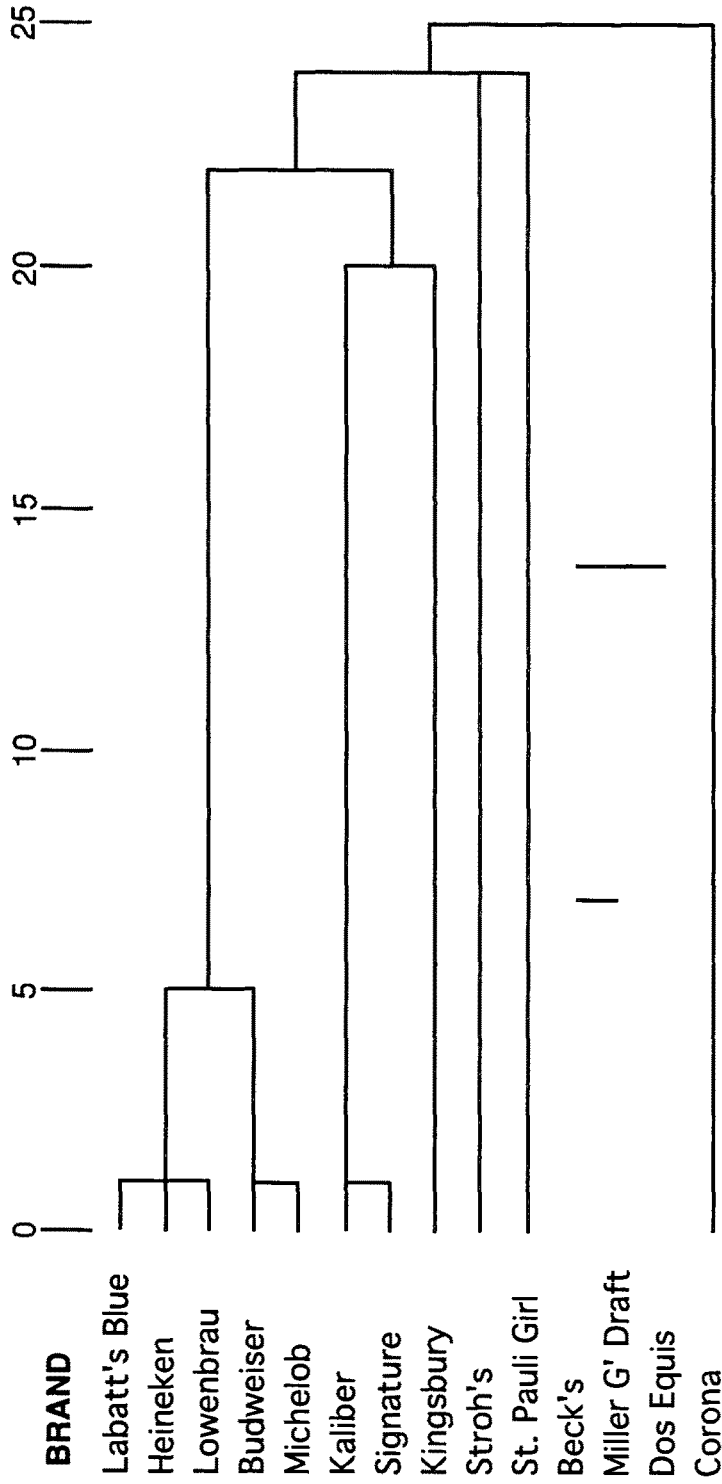


FIGURE 6.
Ultrametric tree based on restricted Model E.

Dendrogram using Complete Linkage

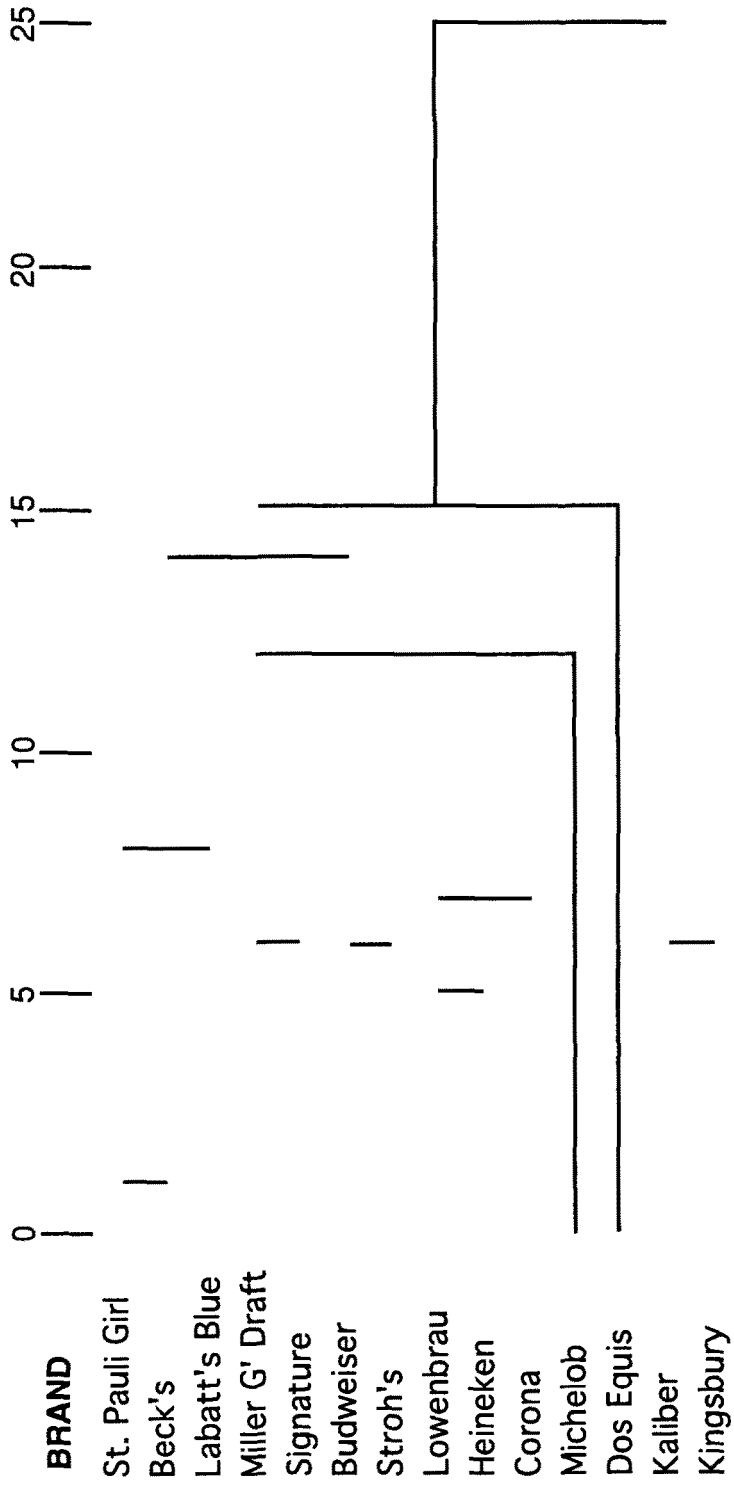


FIGURE 7.
Ultrametric tree based on full Model F.

TABLE 13

Subject Level Parameter Estimates for Model F

i^*	α_i	β_i	σ_i	ϕ_i
1	0.60	2.82	0.80	0.75
2	0.22	3.15	1.33	0.05
3	1.22	2.62	0.16	0.01
4	0.95	2.85	0.44	0.33

* $i = 1, 2, 3,$ and 4 correspond to Subject # $1, 2, 7,$ and 41 in the total sample of 50 subjects

Discussion

The illustrative empirical study provides a demonstration of the TREEFAM procedure as applied to empirical data. The full model (1) with all subject-level parameters and the tree estimated (Model F in our study) fits the data better, based on likelihood ratio tests, as well as AIC and CAIC criteria. Further, assuming that the tree constructed by using only "maximum familiarity" data across the 50 subjects represents the "truth" (this tree appears to have face validity), the unconditional full model solution recovers the "truth" better than the other models, except for those treating the "true" tree as the given solution (Models C and D). The recovery is not exactly complete; nevertheless, given the noisy data (from the 4 subjects), the degree of resolution relative to the true structure must be considered satisfactory. In summary, while this small-scale application is by no means a rigorous validation of the model, it indicates the potential benefits of the TREEFAM procedure in situations of differential brand familiarity.

5. Conclusions

We have developed and applied (to both synthetic and empirical data) a new procedure for capturing and separating the effects of unfamiliarity on proximity judgments and consequently on derived tree structures. The application demonstrates the potential usefulness of the TREEFAM procedure when differential stimulus familiarity is encountered. In terms of inputs from subjects, our procedure requires stimulus familiarity ratings in addition to pairwise interobject proximity judgments. The respondent task appears to be relatively simple and intuitive.

We recognize that there are limitations to our methodology that suggest directions

TABLE 14
Detailed Results of Five Random Starts

Trial	No. of Function Evaluations	Number of Iterations	Ln L	Penalty	Objective Function	Correlation (model vs. data)
1	541	150	-247.685	.00332	-247.688	.492
2	517	159	-244.955	.00317	-244.958	.494
3	540	165	-248.944	.00380	-248.947	.486
4	571	150	-241.231	.00281	-241.234	.522
5	506	148	-244.695	.00245	-244.697	.486

Parameter	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
α_1	0.66	0.79	0.77	0.52	0.72
α_2	0.85	0.32	0.54	0.05	0.42
α_3	1.40	1.22	1.34	1.06	1.27
α_4	0.45	0.88	1.34	1.07	1.09
β_1	2.72	2.74	2.75	2.74	2.72
β_2	2.42	2.05	2.29	1.75	2.28
β_3	2.60	2.59	2.61	2.59	2.59
β_4	2.43	2.71	2.78	2.80	2.76
σ_1	0.30	0.13	0.20	0.35	0.10
σ_2	0.35	1.34	1.27	1.59	0.25
σ_3	0.14	0.15	0.14	0.17	0.14
σ_4	1.22	0.30	0.03	0.12	0.08
ϕ_1	0.21	0.30	0.25	0.25	0.30
ϕ_2	0.11	0.18	0.13	0.12	0.18
ϕ_3	0.01	0.01	0.01	0.01	0.01
ϕ_4	0.27	0.18	0.31	0.29	0.30

Correlations of the Recovered Ultrametric Tree Distances					
Trial	1	2	3	4	5
1	1.000				
2	0.993	1.000			
3	0.992	0.991	1.000		
4	0.993	0.993	0.992	1.000	
5	0.991	0.990	0.992	0.990	1.000

for future research. The model, as specified in (1), does not explicitly include a separate error term for the familiarity ratings, f_{ij} . The reason for this is mathematical tractability, given the complex nonlinear specification. First, note that unfamiliarity in the context of interobject dissimilarity is captured by f_{ijk} which is the product of f_{ij} and f_{ik} , subject i 's familiarity with stimuli j and k respectively. Since f_{ij} and f_{ik} are evoked by the same subject, they need not be independent. Further, observe, from (1), that if f_{ij} were modeled as stochastic, we would need to ascertain the distribution of $(f_{ij} \cdot f_{ik})^{\alpha_i}$ and then of $(f_{ij} \cdot f_{ik})^{\alpha_i} \cdot u_{ijk}$. Except under very special distributional assumptions for f_{ij} , this would be analytically intractable. Finally, the distribution of the expression $(f_{ij} \cdot f_{ik})^{\alpha_i} \cdot u_{ijk}$ would need to be convoluted with the normal distribution of the error term e_{ijk} . Thus, explicitly incorporating measurement error for familiarity ratings would require a different modeling approach, and we leave this issue for future re-

search. In the present formulation, the additive error term e_{ijk} implicitly captures, at least in part, the measurement error on the familiarity ratings (in addition to capturing measurement error on the dissimilarity ratings). As another issue for future research, a more extensive Monte Carlo study than that conducted here will allow a more rigorous test of the performance of the procedure.

In further applications of the model, it may be interesting to investigate whether the subject level parameters α_i and β_i are related to subject background variables (for example, product category usage in the application provided). Future research should also be directed toward more rigorous tests of the behavioral premises underlying the model, possibly in an experimental setting in which the level of familiarity can be manipulated by the researcher. A possible limitation of the model is that it may oversimplify the effect of brand unfamiliarity on proximity judgments. For example, when comparing a familiar stimulus with an unfamiliar one, a subject's proximity judgment is likely to be influenced by the features of the familiar stimulus, an aspect that is not explicitly incorporated in our model. However, given our objectives in developing the proposed methodology, the consideration of behavioral richness must be traded off against the need for parsimony in model specification.

We also assume that, under complete familiarity, all subjects have a common tree structure. A potential extension of the model would be to allow for segments with different tree structures (under complete stimulus familiarity). A latent class approach would permit the procedure to identify such segments endogenously. A further extension would allow for hybrid tree/spatial representation of stimuli, where the sample is first hierarchically defined on the basis of discrete and well known attributes, and then, at each terminal node of the tree, a spatial map is used to describe the relative positions of stimuli associated with the node, where the map captures abstract and finer aspects of the stimuli.

References

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In N. Petrov & F. Csadki (Eds.), *Second International Symposium on Information Theory* (pp. 267–281). Budapest: Akademiai Kiado.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *AC-19*, 716–723.
- Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research*, *13*, 411–454.
- Boster, J. S. (1988). Natural sources of internal category structure: Typicality, familiarity, and similarity of birds. *Memory and Cognition*, *16*, 258–270.
- Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52*, 345–370.
- Carroll, J. D., & Pruzanski, S. (1975, August). *Fitting of hierarchical tree structure (HTS) models, mixtures of HTS models, and hybrid models via mathematical programming and alternating least squares*. Paper presented at U.S.-Japan Seminar on Multidimensional Scaling, University of California at San Diego, La Jolla, California.
- Carroll, J. D., & Pruzanski, S. (1980). Discrete and hybrid scaling models. In E. D. Lantermann & H. Feger (Eds.), *Similarity and choice* (pp. 48–69). Bern: Hans Huber.
- Chandon, J. L., Lemaire, L., & Pouget, J. (1980). Construction de l'ultramétrie la plus proche d'une dissimilarité au sens des moindres carrés [Construction of least-squares ultrametric trees from dissimilarity data]. *R. A. I. R. O., Recherche Operationelle*, *14*, 157–170.
- Chatterjee, R., & DeSarbo, W. S. (1992). Accommodating the effects of brand unfamiliarity in the multidimensional scaling of preference data. *Marketing Letters*, *3*, 85–99.

⁷ It is not clear from the literature what the impact on the proximity judgment might be. In our model, the unfamiliarity error term (representing stochasticity around the anchor) recognizes the presence of this and other effects that are not explicitly specified.

- DeSarbo, W. S., Manrai, A. K., & Burke, R. R. (1990). A nonspatial methodology for the analysis of two-way proximity data incorporating the distance-density hypothesis. *Psychometrika*, *55*, 229–253.
- De Soete, G. (1984). A least squares algorithm for fitting an ultrametric tree to a dissimilarity matrix. *Pattern Recognition Letters*, *2*, 33–137.
- Goldstein, A. G., & Chance, J. E. (1980). Memory for faces and schema theory. *Journal of Psychology*, *105*, 47–59.
- Green, P. E., Tull, D. S., & Albaum, G. (1988). *Research for marketing decisions*. Englewood Cliffs, NJ: Prentice-Hall.
- Hartigan, J. A. (1967). Representation of similarity matrices by trees. *Journal of the American Statistical Association*, *62*, 1140–1158.
- Hartigan, J. A. (1975). *Clustering algorithms*. New York: Wiley.
- Jaccard, J., & Wood, G. (1988). The effects of incomplete information on the formation of attitudes toward behavioral alternatives. *Journal of Personality and Social Psychology*, *54*, 580–591.
- Jain, A. K., & Dubes, R. C. (1988). *Algorithms for clustering data*. Englewood Cliffs, NJ: Prentice-Hall.
- Jardine, C. J., Jardine, N., & Sibson, R. (1967). The structure and construction of taxonomic hierarchies. *Mathematical Bioscience*, *1*, 173–179.
- Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, *32*, 241–254.
- Koehler, A. B., & Murphree, E. S. (1988). A comparison of the Akaike and Schwarz criteria for selecting model order. *Applied Statistics*, *37*, 187–195.
- Meyer, R. J. (1981). A model of multiattribute judgments under attribute uncertainty and informational constraint. *Journal of Marketing Research*, *18*, 428–441.
- Morrison, D. F. (1983). *Applied linear statistical methods*. Englewood Cliffs, NJ: Prentice-Hall.
- Pereboom, A. C. (1971). Some fundamental problems in experimental psychology: An overview. *Psychological Reports*, *28*, 439–455.
- Ramsay, J. O. (1977). Maximum likelihood estimation in multidimensional scaling. *Psychometrika*, *42*, 241–266.
- Sanbonmatsu, D. M., Kardes, F. R., & Herr, P. M. (1992). The role of prior knowledge and missing information in multiattribute evaluation. *Organizational Behavior and Human Decision Processes*, *51*, 76–91.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.
- Slovic, P., & MacPhillamy, D. (1974). Dimensional commensurability and cue utilization in comparative judgment. *Organizational Behavior and Human Performance*, *11*, 172–194.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.
- Valentine, T., & Bruce, V. (1986). Recognizing familiar faces: The role of distinctiveness and familiarity. *Canadian Journal of Psychology*, *40*, 300–305.
- Yagamichi, T., & Hill, C. T. (1983). Initial impressions versus missing information as explanations of set-size effect. *Journal of Personality and Social Psychology*, *44*, 942–951.
- Yates, J. F., Jagacinski, C. M., & Faber, M. D. (1978). Evaluation of partially described multi-attribute options. *Organizational Behavior and Human Performance*, *21*, 240–251.

Manuscript received 10/27/92

Final version received 10/21/93