AGENDAS AND CONSUMER CHOICE

by

John R. Hauser

Massachusetts Institute of Technology

M.I.T. Working Paper No. 1641-85

January 1985

ACKNOWLEDGEMENTS

This paper has benefited tremendously from the insights and suggestions of Amos Tversky who, for the past five years, has collaborated with me on the interesting topic of agenda setting. In particular, the ideas of top-down versus bottom-up agendas, the generalization of the hierarchical elimination model, and the behavioral hypotheses are fruits of these discussions. Of course, any errors in the mathematical proofs or the conceptualizations are my responsibility.

This paper has benefited from seminars and presentations at M.I.T., INSEAD, the University of Chicago, the Duke University Conference on Choice Theory, and the 1981 Association of Consumer Research Conference.

AGENDAS AND CONSUMER CHOICE

ABSTRACT

Suppose a consumer is asked first to choose among "foreign" vs. "domestic" automobiles and then, depending on his choice, to choose within the class of "foreign" (or "domestic") automobiles. Will such a restriction influence choice probabilities?

This paper investigates the relationships between such hierarchical constraints, cognitive processing rules, the attribute (or aspect) structure of choice alternatives, and choice probabilities. We base our analyses on three well-known and empirically documented probablistic choice models: the constant ratio model (CRM), elimination-by-aspects (EBA), and a generalization of the hierarchical elimination model (GEM). An agenda (either implicit or imposed) is then a constraining hierarchy. It can be "top down" (as above), "bottom up" (as a tournament), or more general We provide behavioral hypotheses suggesting when various types of self-imposed agendas will be used by consumers.

Our first set of analytic results identifies which choice rules are affected by agendas and on which aspect structures. Our second set of results illustrates how agendas can be used to enhance target products. And our third set of results shows that two commonly accepted axioms of choice behavior, dominance and regularity, need not apply on agendas and that such violations make intuitive sense and are useful managerially.

Throughout this paper we provide examples and discuss the implications of our results for marketing management and for behavioral research.

1. MOTIVATION

Agendas influence choice. For example, academic hiring decisions are sometimes made as sequential decisions. If there are many good candidates for a junior faculty position and we do not wish to interview them all, we may first decide on a subfield, say information processing, and then search within that subfield. We make our decision on subfields based on our prior experience which tells us what to expect from the type of candidates that choose to enter that subfield. Contrast this "top down" sequential decision (subfield, then candidate within subfield), with a "bottom up" sequential decision such as might be used in a field where it is feasible to interview all of the potentially good candidates. For example, if there are relatively few candidates, we may interview them all, decide who is the best "model builder", who is the best "consumer behavioralist", and who is the best "managerial analyst". We would then contrast the best with the best taking into consideration all of the candidates' characteristics as well as their fields of interest.

Unlike voting agendas, where the agenda setter seeks to influence the outcome (e.g., see Farquharson 1969), hiring committees may or may not wish to influence the outcome by their choice of agenda. Members of hiring committeees may not even vary in terms of preferences. Often the sequential decision rule, i.e., the agenda, is chosen to structure or simplify the task. The committee may be unaware of whether the agenda influences the probabilities of outcomes.

Sequential decision rules may also apply to individual choice. Consider the following print advertisements that appeared recently:

PITNEY BOWES - If you have $5,000 or $6,000 earmarked for a Savin desk copier or a Xerox 3100, you can spend it on a full-sized copier (Pitney Bowes) and still have money left over.

SAVIN - If Savin couldn't make a copying system better than the Xerox 4500, we wouldn't have bothered.

XEROX - $2,995 for a what? (With visuals of the machine and the Xerox trademark).

One can interpret these advertisements as an attempt by the manufacturers to influence the sequential decision process used by buyers of copying machines. Presumably, Pitney Bowes believes it benefits by influencing consumers to make a decision of {Pitney Bowes, Savin, Xerox} vs. {all others}, while Savin prefers the decision of {Savin, Xerox} vs. {all others}. Xerox tries to prevent such sequential decisions.

The effect of an advertising campaign is not identical to the effect of the sequential decision process used by hiring committees. Advertising can, at best, influence the decision process, not set it. Furthermore, it is an empirical question whether or not the consumer actually follows a sequential decision process. Finally, an agenda effect may be only one of many outcomes of an advertising campaign. Nonetheless, there are similarities and it would be useful to know whether sequential decision processes influence consumer choice. Such knowledge would help us analyze and design communication strategies, such as advertising, which appear to influence sequential decision processes.

New product strategies also can influence and/or depend upon consumers' sequential decision processes. For example, consider market entry strategy. Selecting the "right" market to enter greatly enhances the likelihood and magnitude of a new product success. Selecting the wrong market can doom a new product to failure. One component of market selection is the identification of market structure such as the simplified hypothetical market structure for televisions shown in figure 1. If, as shown, the consumer first chooses or can be influenced to choose among 'consoles' and 'portables' then among 'color' and 'black & white', a manufacturer of color consoles might favor color portables over black & white consoles so he can compete in both "markets" and not cannibalize his existing product line. Once such market structure is understood, the innovating firm can better select a market on the basis of sales potential, penetration, scale, input, reward, risk, and match to the organization's capabilities (e.g., see Urban and Hauser 1980).

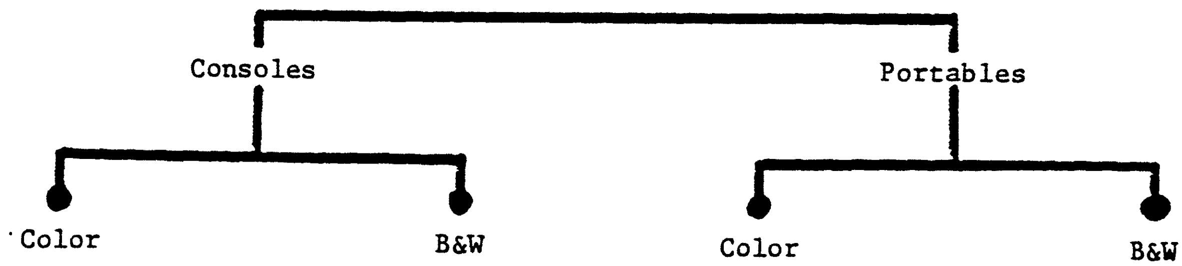


Figure 1: Simplified Hypothetical Hierarchy for Television Markets

There are a myriad of other interesting, relevant situations where individual consumers are faced with sequential decision processes. In shopping behavior a consumer limits his set of product options by choosing one particular shopping center. The industrial salesman often tries to influence customer's consideration sets; supermarkets use displays to set one brand away from others; package goods manufacturers often use similar packaging (compare Colgate to Close-Up toothpaste) or distinct packaging; and some advertisements mention competitive products, others do not. Some of these strategies are planned, some dictated by environmental constraints, and some serendipitous. But in each case the sequence of decisions among products is influenced and in each case the modification of the decision sequence may or may not affect choice outcomes.

We can better understand and model these marketing strategies if we first understand the effects of sequential processing on choice. By drawing on and extending recent developments in probabilistic choice theory, this paper explores a mathematical analysis of sequential decision processes, i.e., agendas, and their ability to influence consumers' probabilities of choice. We begin with a short discussion of research that addresses issues related to the issues we address in this paper.

In marketing, agendas have been studied by researchers seeking to influence market entry strategy with normative models that identify market structure. See review by Day, Shocker, and Srivastava (1980). Perhaps the best known example is the Hendry system, described in Kalwani and Morrison (1977), which defines market structure relative to theoretical probabilities of switching among products. The identified agenda is based on a comparison of theoretical and observed switching probabilities. In a related method, Rao and Sabavala (1981) use hierarchical cluster analysis on switching probabilities with the assumption that switching is greatest at the lowest levels in the hierarchy. Srivastava, Shocker, and Leone (1981) apply a similar method to data which measures substitution-in-use. Urban, Johnson, and Hauser (1984) make the same switching assumption but use forced switching data. They remove a consumer's most preferred product from a simulated store and observe subsequent behavior. In each of these methodologies, the concern is a description of the market, not necessarily the individual consumer.

In social choice economics, researchers have dealt primarily with voting agendas. Plott and Levine (1978) develop a descriptive model of "naive" voter behavior and illustrate that it fits empirical data quite well. McKelvey (1981) adapts this model to identify a "best" agenda via dynamic programming. In alternative models, Farquharson (1969) and McKelvey and Niemi (1978) assume "sophisticated" voters who have perfect information about every other voters' preferences. These authors show that unique alternatives will be selected for binary agendas.

Our present approach differs from the approaches in these literatures. Unlike the social choice theorists, we deal with descriptive, probabilistic models of individual choice where the individual either is faced with a constrained agenda or uses an implicit agenda.

Unlike the market structure literature, we do not begin by defining agendas relative to the likelihood of switching among products. Rather, we investigate how such definitions and other properties are related to more primitive assumptions of cognitive processing and its relationship to sequential and random decision structures.

Our purpose is to construct and investigate a mathematical representation, or theory, of how sequential cognitive decision rules affect individual choice probabilities. We take as our basis, existing psychological models of choice behavior, models that have been used successfully (e.g., Tversky 1972 and Tversky and Sattath 1979) to describe and explain how individuals make choices among sets of choice objects. Because we are using these models in new ways, we need to make some generalizations, but we attempt to limit our generalizations so that our model clearly reduces to existing models when the choice situation is identical to that which has been studied previously. Furthermore, we do investigate a series of models which together bracket the types of hypotheses likely to be made about individual choice behavior.

We investigate and extend three decision rules: the constant ratio model (CRM), elimination by aspects (EBA), and the hierarchical elimination model (HEM). We investigate their relationship to decision sequences and provide testable hypotheses as to when a decision rule would be used by an individual. Our analyses indicate when alternative decision rules provide different choice probabilities and when they do not, when and how agendas affect choice probabilities, and how all of these results depend on the structure of interrelationships among choice objects.

Because we rely heavily on the recent literature in mathematical psychology, section 2 reviews the models we use in this paper. Section 3 defines an agenda and investigates its relationship to the choice probabilities. Section 4 explores the relationships among agendas and product attributes (aspects). Section 5 provides some illustrative results on when agendas are effective strategically. Section 6 discusses dominance and regularity and section 7 concludes with a discussion of limitations, extensions, and suggested experiments.

Throughout the text, we motivate the results with examples. All mathematical proofs are in the appendix.

2. PROBABILISTIC MODELS OF INDIVIDUAL CHOICE BEHAVIOR

Our review is by necessity concise. We seek to provide a basic review (1) because some readers may not be familiar with the details of these models, (2) our generalizations must be put in perspective, and (3) the remainder of this paper depends upon these models. Those readers wishing greater detail are referred to Luce (1959, 1977), Tversky (1972a, b) and particularly, Tversky and Sattath (1979). Those readers already familiar with these models may wish to skip this section pausing only to review notation.

We use lower case English letters r, s, t, v, w, x, y, and z to denote choice objects, e.g., automobiles, restaurants, and televisions. We use upper case English letters, A, B, C, ..., to denote sets of choice objects, e.g., $A = \{x, y, z\}$. The total finite set of all choice objects being studied is denoted by T, and the null set is denoted by $\emptyset$.

Let $P(x|A)$ denote the probability that object x is chosen when the choice set is A. Naturally we assume $P(x|A) \geq 0$ and the sum of all $P(x|A)$ for all x in A is equal to 1.0.

We describe choice objects as a collection of aspects. For example, an automobile may be described by aspects such as 'sporty', 'high mpg', 'sedan', 'front wheel drive', 'Chevrolet', etc. An aspect is a binary descriptor of a choice object, e.g., 'sedan', in the sense that a choice object either has an aspect or it does not. We use lower case Greek letters, α, β, γ, ..., to denote aspects. For continuous attributes such as mpg, we define ranges. For example, an automobile with the attribute of 30 or more mpg, would have the aspect 'high mpg'. An aspect in this analysis can be a collection of more elementary aspects, e.g., 'high mpg, front wheel drive, sporty, and red', or it can be that which is unique to a choice object, e.g., 'Honda Accordness'.

Let $x' = \{\alpha, \beta, ...\}$ be the set of aspects associated with choice alternative x. For any set of choice alternatives, A, let A' be the set of aspects that belong to at least one alternative in A; i.e., $A' = \{\alpha | \alpha \varepsilon x'$ for some $x \varepsilon A\}$. For any aspect, α, and set of choice alternatives, A, let A_α denote the set of all choice alternatives in A that have aspect α, i.e., $A_\alpha = \{x | x \varepsilon A, \alpha \varepsilon x'\}$. Note that A' is a set of aspects and A_α is a set of choice objects. For example, if A = {Honda Civic, Honda Accord, Chevrolet Chevette, Chevrolet Citation}, then A' = {'compact', 'front wheel drive', 'Honda', 'Chevrolet', ...}. If α = 'Honda', then A_α = {Honda Civic, Honda Accord}.

The best way to think of an aspect is as a unit of analysis. If attribute-like aspects, such as 'compact', 'front wheel drive', etc. are sufficient to describe the choice process for strategic managerial understanding, then we prefer to work with them since they have intuitive appeal to the manager. However, it is not necessary that we be able to name the aspects. We can also think of aspects as descriptors of unique components of sets of choice objects. Aspects can be 'that which is unique to a Chevy Chevette', or 'that which is shared by a Chevy Chevette and a Chevy Citation and no other car', etc. In such a specification, we could have an aspect for every possible subset of T, that is, 2^n-2 aspects for n objects. Both interpretations of aspects are consistent with our mathematical analysis. In fact, they would also be semantically equivalent if the manager could only articulate more elementary descriptors of 'that which is shared by a Chevy Chevette and a Chevy Citation and no other car'. Throughout the paper we use attribute-like aspects for our examples recognizing that our results apply equally well to the interpretation based on subsets of choice objects.

Constant Ratio Model (CRM). Perhaps the most commonly used probabilitic choice model in marketing is the constant ratio model. See Bell, Keeney, Little (1975); Hauser and Urban (1977); Jeuland, Bass, Wright (1980); Johnson (1975); Luce (1959, 1977); McFadden (1980); Pessemier (1977); Punj and Staelin (1978); Silk and Urban

(1978); and others. The basic assumption underlying the CRM is that the ratio of (non-zero) choice probabilities for two objects is independent of the choice set. Mathematically, $P(x|A)/P(y|A) = P(x|B)/P(y|B)$ for all A and B such that the probabilities are non-zero. It is easily shown e.g., Luce (1977), that CRM implies there exist some scale values, u(x), for the objects, x, such that:

$$P(x|A) = \frac{u(x)}{\sum_{y \varepsilon A} u(y)} \quad (1)$$

where $y \varepsilon A$ denotes all objects, y, contained in choice set A. The simplicity of equation 1 has led to its wide acceptance. Furthermore, in many situations, it approximates behavior quite well. For example, Silk and Urban (1978) have forecast market share to within one or two percentage points on over 400 new products using a model that incorporates equation 1. See evidence in Urban and Katz (1983).

However, several authors (Debreu 1960; Luce and Suppes 1965; Restle 1961; Rumelhart and Greeno 1971; Tversky 1972b) have presented conceptual and empirical evidence that CRM fails to account for the similarity among choice objects. For example, consider four automobiles, x_1, x_2, y_1, y_2. Assume x_1 and x_2 are identical except for an unimportant aspect, say the type of electric clock, analog or digital, respectively. Automobile y_1 is quite different from x_1 and x_2, but y_2 is the same as y_1 except for a very popular feature such as cruise control. All four automobiles are priced identically. Suppose that an individual is indifferent between x_1 and y_1. Then, since x_1 and x_2 are virtually identical, we expect $P(x_1|\{x_1, x_2\}) = 1/2$, $P(x_1|\{x_1, y_1\}) = 1/2$, and $P(x_2|\{x_2, y_1\}) = 1/2$, but $P(y_2|\{y_1, y_2\}) = 1$. It follows from CRM (equation 1) tha $P(y_1|\{x_1, x_2, y_1\}) = 1/3$. However, common sense suggests the choice is more likely among x-cars and y-cars. Consequently, we expect $P(y_1|\{x_1, x_2, y_1\})$ to be close to 1/2 and the other two trinary probabilities close to 1/4 rather than all equal to 1/3 as predicted by CRM.

Furthermore, CRM implies if two automobiles are substitutable in one context they are substitutable in all contexts. Thus, if $P(y_2|\{y_1,y_2\}) = 1$, then CRM implies $P(y_2|\{x_1,y_2\}) = 1$. Again, common sense suggests such a result is implausible. The addition of cruise control is unlikely to eliminate all conflict among automobile x, say the Toyota Celica, and automobile y, say the Buick Skyhawk.

Elimination by Aspects (EBA). The above criticisms are formulated with appeals to common sense based on similarities among the characteristics, or aspects, of the choice objects. Elimination by aspects addresses these criticisms by postulating that choices among sets of objects depend upon the aspects of the objects not the objects per se. This assumption is analogous to the economic models of Lancaster (1971) and his colleagues and to the multiattributed models in marketing as reviewed by Wilkie and Pessemier (1973).

EBA postulates that an individual chooses among all aspects in the offered set, A', with probability proportional to the scale value of the aspect. He then eliminates all choice objects not having the chosen aspect and continues choosing aspects and eliminating objects until one choice object is left.

If $\boldsymbol{\alpha}$ represents the scale value of aspect α, then EBA is given mathematically by the following recursive equation:

$$P(x|A) = \sum_{\alpha \epsilon x'} \frac{\boldsymbol{\alpha}}{\sum_{\beta \epsilon A'} \boldsymbol{\beta}} \cdot P(x|A_\alpha) \qquad (2)$$

It is easy to show that any aspect common to all alternatives in a choice set, e.g., the aspect 'automobiles', does not affect choice probabilities and will, therefore, be discarded. See Tversky (1972a, 1972b).

To illustrate EBA, we consider four automobiles: x = Honda Civic, y = Honda Accord, z = Chevrolet Chevette, and w = Chevrolet Citation. Let α = 'Automobile', β = 'Japanese', γ = 'American', δ = aspects unique to the Honda Civic, μ = aspects unique to the Honda Accord, λ = aspects unique to the Chevrolet Chevette, and η = aspects unique to the Chevrolet Citation. According to our definitions:

$$x' = \{\alpha, \beta, \quad \delta\}$$
$$y' = \{\alpha, \beta, \quad \mu\}$$
$$z' = \{\alpha, \quad \gamma, \quad \lambda\}$$
$$w' = \{\alpha, \quad \gamma, \quad \eta\}$$

The aspect, α, is common to all objects and can be ignored. Alternative x will then be chosen if δ is the first aspect chosen or if β is the first aspect chosen and then δ is chosen from the set $\{\delta, \mu\}$. Specifically:

$$P(x|\{x, y, z, w\}) = \delta + \beta \cdot P(x|\{x, y\}) = \delta + \beta \cdot \frac{\delta}{\delta + \mu} \quad (3)$$

where, without loss of generality, we have set $\beta + \gamma + \delta + \mu + \lambda + \eta = 1$.

To illustrate how EBA resolves the conceptual criticisms of CRM, set α = unique aspects of Toyota Celica, β = unique aspects of Buick Skyhawk, γ_1 = digital clock, γ_2 = analog clock, and δ = cruise control. Then,

$$x_1' = \{\alpha, \quad \gamma_1\}$$
$$x_2' = \{\alpha, \quad \gamma_2\}$$
$$y_1' = \{\beta\}$$
$$y_2' = \{\beta, \quad \delta\}$$

Suppose $\alpha = \beta$, $\gamma_1 = \gamma_2$, and γ_1, γ_2, and δ are small compared to α and β. Then equation 2 yields: $P(y_1|\{x_1, x_2, y_1\}) = \beta/(\beta + \alpha + \gamma_1 + \gamma_2) \cong \beta/(\alpha + \beta) = 1/2$. Similarly, $P(y_2|\{x_1, y_2\}) = (\beta + \delta)/(\beta + \delta + \alpha + \gamma_1) \cong \beta/(\alpha + \beta) = 1/2$. We leave it to the reader to show that all other binary and trinary probabilities are as expected by common sense.

EBA is a generalization of CRM because equation 2 reduces to equation 1, with $u(x) = \sum_{\alpha \in x'} \alpha$, when choice objects are disjoint, i.e., $x' \cap y' = \emptyset$ for all $x, y \in A$.

Finally, note that EBA does not imply a fixed sequential decision process. Rather, EBA implies that the individual probabilistically selects aspects for consideration and, hence, EBA is a randomized sequential decision process.

Preference Trees

By 1985 there were over 160 different automobiles available. The number of aspects necessary to fully describe and distinguish these automobiles could be very large. For example, if there were an aspect for every possible subset, a fully specified EBA model would require approximately 1.5×10^{48} aspects. At the other extreme is CRM, which applies when objects have no shared aspects. CRM requires only as many scale values as there are choice objects, in this case 160 scale values.

Tversky and Sattath (1979) investigate an intermediate level of complexity in the structure of aspect sets. They investigate tree structures known as preference trees. An example preference tree is shown in figure 2.

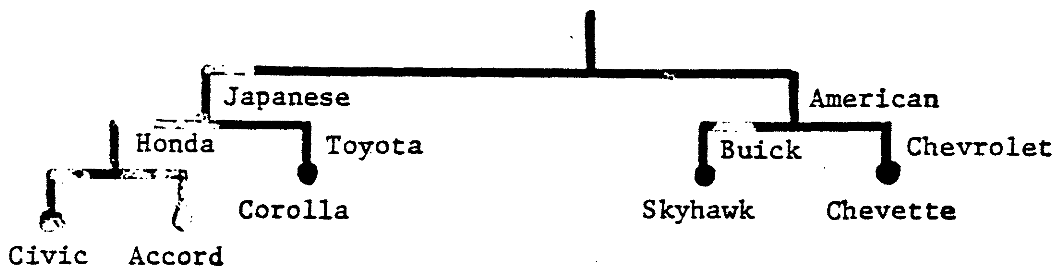


Figure 2: Example Preference Tree for Automobiles

A preference tree is an aspect structure which is a hierarchical aspect structure where there is no overlap among the branches. For example, in figure 2 there are Japanese and American automobiles. The Japanese cars are further subdivided into Hondas and Toyotas and the American cars are subdivided into Buicks and Chevrolets. The addition of an American Honda, say if consumers considered Hondas built in America as American, would introduce overlap and upset the tree structure. (See formal definition of preference trees in Tversky and Sattath 1979).

A preference tree structure greatly restricts the complexity of the inter-relationships among aspects. The 2^n-2 possible aspects are limited to at most $2n-2$ aspects corresponding to the maximal number of links in a tree with n terminal nodes. For example, in a choice set with 10 objects, the number of aspects possible are reduced from 1,022 to 18. Thus, preference trees represent a compromise between the generality of unrestricted EBA and the limitations of CRM. Nonetheless, preference trees can adequately represent many interesting choice situations. For example, the structured choice set, $\{x_1, x_2, y_1, y_2\}$, shown above is the preference tree in figure 3. (The vertical length of the branch indicates the importance of the feature.)

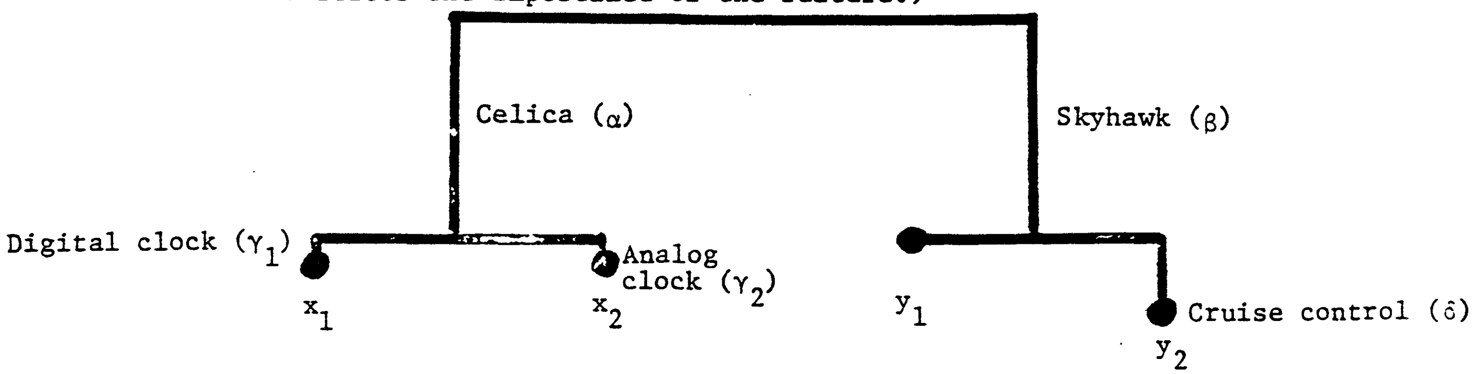


Figure 3: Preference Tree Representation of the Logical Counterexample to the Constant Ratio Model.

Hierarchical Elimination Models (HEM)

Neither CRM nor EBA are sequential decision rules. Preference trees are aspect structures, not decision rules. However, there is empirical evidence that individuals make explicit, hierarchical decisions. See among others Bettman (1979), Haines (1974), and Payne (1976).

In 1979, Tversky and Sattath introduced the hierarchical elimination model (HEM) to represent cognitive processing on a preference tree as a hierarchical series of choice points. The idea is that the individual sequentially compares sets of objects defined by branches in the preference tree. For example, in figure 2 he first decides among 'Japanese' and 'American' automobiles, then, if he selects 'Japanese', he chooses among 'Hondas' and 'Toyotas', and finally, if he selects 'Hondas', he chooses among 'Civics' and 'Accords'. The choice among branches is proportional to the measures of the aspects in that branch. For example, if α = 'Japanese', β = 'American'. γ = 'Honda', δ = 'Toyota',

μ = 'Buick', λ = 'Chevrolet', ξ = 'Civic', and η = 'Accord', then the probability he chooses the 'Japanese' branch is:

$$P(\text{Japanese}|T) = \frac{\alpha + \gamma + \delta + \xi + \eta}{(\alpha + \gamma + \delta + \xi + \eta) + (\beta + \mu + \lambda)} \tag{4a}$$

The probability he then chooses the Honda subbranch is:

$$P(\text{'Honda'}|\text{'Japanese'}) = \frac{\gamma + \xi + \eta}{(\gamma + \xi + \eta) + \delta} \tag{4b}$$

and

$$P(\text{'Civic'}|\text{'Honda'}) = \frac{\xi}{\xi + \eta} \tag{4c}$$

The overall probability is the product of the sequential conditional probabilities, i.e.,

$$P(\text{'Civic'}|T) = P(\text{'Civic'}|\text{'Honda'})P(\text{'Honda'}|\text{'Japanese'})P(\text{'Japanese'}|T) \tag{4d}$$

In other words, if A_1, A_2, ... A_ℓ, T is a sequence of choice sets such that A_i is contained in A_{i+1} and if the sequence corresponds to the branching in the preference tree, then a hierarchical probability model is given by:

$$P(x|\mathcal{A}) = P(x|A_1)P(A_1|A_2) \ldots P(A_{\ell-1}|A_\ell)P(A_\ell|T) \tag{5}$$

where $\mathcal{A}$ denotes the hierarchy.

For a preference tree, HEM is defined such that

$$P(x|A_1) = \frac{m(x)}{\sum_{y \in A_1} m(y)} \tag{6a}$$

$$P(A_i|A_{i+1}) = \frac{m(A_i)}{\sum_{B \in A_{i+1}} m(B)} \tag{6b}$$

where $m(A_i)$ denotes the measure of A_1 and $B \epsilon A_{i+1}$ indicates the sum is over all branches B of the set A_{i+1}. The measure of A_i is equal to the sum of all aspects in A_i except those shared by all other branches in A_{i+1}. The measure of x is the sum of all aspects of x except those shared by all objects in A_1. For example, in figure 2 and equation 4, m('Civic') = ξ, m('Accord') = η, m('Honda') = $\gamma + \xi + \eta$, m('Toyota') = δ, etc.

To see how HEM resolves the conceptual criticisms of CRM, we apply HEM to the preference tree in figure 3. Using equations 5 and 6 and the definitions of α, β, γ_1, and γ_2 for that figure, we obtain:

$$P(y_1|\{x_1, x_2, y_1\}) = \beta/(\alpha+\beta+\gamma_1+\gamma_2) \doteq \beta/(\alpha+\beta) = 1/2$$

$$P(x_1|\{x_1, x_2, y_1\}) = \left(\frac{\alpha+\gamma_1+\gamma_2}{\alpha+\gamma_1+\gamma_2+\beta}\right)\cdot\left(\frac{\gamma_1}{\gamma_1+\gamma_2}\right) \doteq \left(\frac{\alpha}{\alpha+\beta}\right)\left(\frac{\gamma_1}{\gamma_1+\gamma_2}\right) = 1/4$$

Clearly, HEM and EBA are very different cognitive decision rules. Intuitively, we would not expect them to yield the same choice probabilities. However, the reader may wish to verify that for the preference tree in figure 3, the forecast choice probabilities are the same for EBA and HEM. Tversky and Sattath (1979, p. 548) prove the surprising result that for any preference tree, EBA and HEM yield the same choice probabilities. Thus, on a preference tree, we could never distinguish the two alternative decision rules by simply observing choice probabilities. (However, we might distinguish them by other means such as verbal protocols or choice reaction time.)

To date, HEM has only been defined for preference trees. The preference tree defines a natural sequence of decisions. We now extend the preference tree/hierarchical elimination concepts to general aspect structures, general sequential decisions and to imposed sequential constraints on choice decisions.

3. AGENDAS

If the aspect structure forms a preference tree and the consumer follows HEM, then we can think of the choice process as following a sequence of constraints. That is, the consumer first chooses a set, A_ℓ, from among subsets of T, next he (or she) chooses a set, $A_{\ell-1}$, from among the subsets of A_ℓ, proceeding until he (or she) finally chooses an object, x, from the choice set A_1. The choice sequence is illustrated in figure 4 where the nodes represent choices. We have numbered the nodes to indicate the order of processing in HEM. In HEM the sequence is defined by the natural structure of the aspect tree.

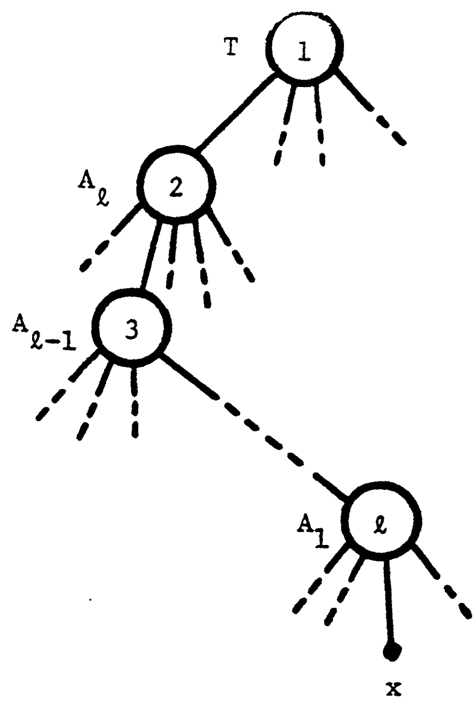


Figure 4: Schematic Representation of Sequential Choice Process in HEM on a Preference Tree.

There are, however, many situations in which the aspect structure is not a preference tree and/or the sequence of decisions are not determined by the aspects. For example, the aspect structure in figure 1 is not a preference tree because the aspects, 'color' and 'B & W' appear on both the 'consoles' branch and the 'portables' branch. In fact, figure 1 is more like a factorial structure than a tree structure. We expect such factorial structures to be common in mature markets like television sets where the market has evolved such that competitors have exploited all segments.

Even if the aspect structure is a tree, there may be external constraints which force a sequential process that does not follow the tree. Consider the faculty candidate who is constrained to a geographic area because of a spouse's opportunities or the automobile owner who is constrained in a choice of service stations because the automobile requires diesel fuel.

Even if no constraints are imposed, a consumer may wish to make choices sequentially. Consider restaurant choice in an unfamiliar city. We often simplify our decisions by first choosing price range, say high, medium, or cheap eats, and then style, say French, Italian, German, American, Chinese, Israeli, or Lithuanian. Such a choice process can be viewed as a set of (internally imposed) constraints.

We define an agenda as a sequence of constraints. In particular, an agenda is a tree of objects such that at any node, say A_ℓ in figure 4, the consumer must choose among those branches exiting that node. We label the nodes to indicate the order in which they are processed. For example, in figure 4, the agenda is processed from the top down as indicated.

However, an agenda need not be top down. It can also be bottom up as shown in figure 5.

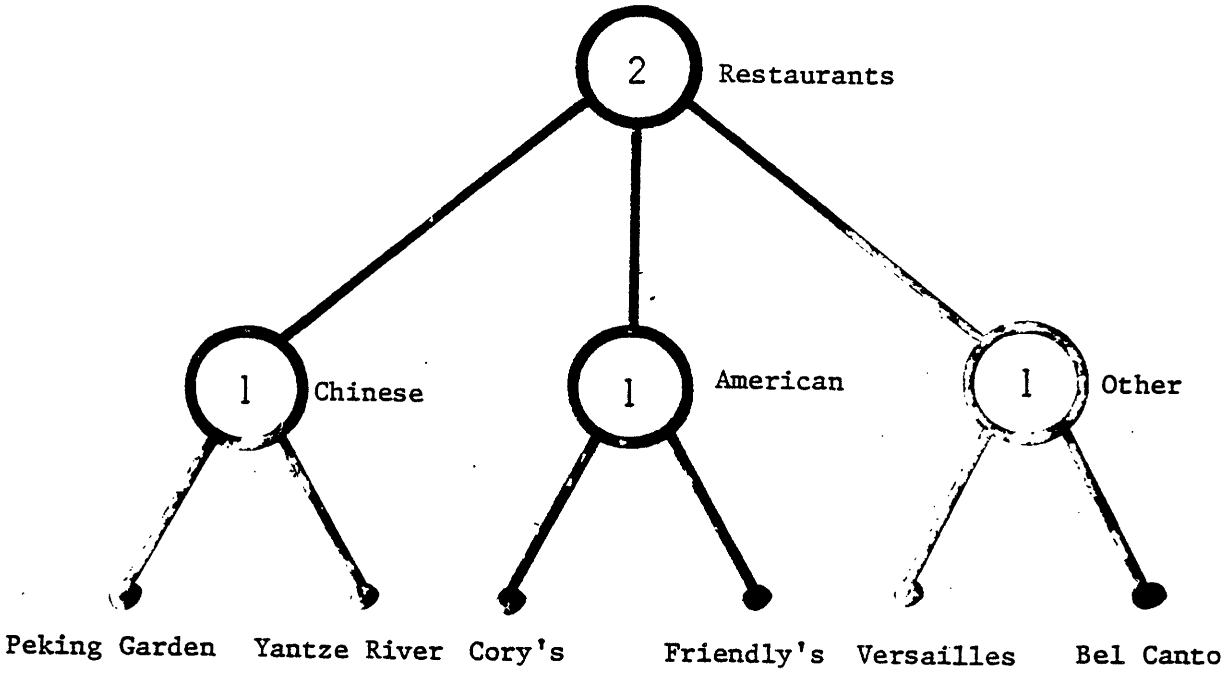


Figure 5: Bottom Up Agenda for Restaurant Choices in Lexington, MA.

In figure 5, we first choose the best restaurant in each class, say Yantze River for Chinese, Cory's for American, and Versailles for other. We then choose the restaurant for a night out by comparing the best with the best, in this case, Yantze River vs. Cory's vs. Versailles. Such an agenda might be used by a resident of Lexington, MA who is familiar with all six restaurants. On the other hand, a tourist unfamiliar with Lexington, may process this tree with a top down agenda. Notice that the processing sequence is a logical constraint not necessarily a temporal constraint in the sense that all '1' nodes do not need to be processed simultaneously. We only require that all '1' nodes be processed prior to any '2' nodes.

We also allow agendas to be mixed. For example, a consumer may process all bottom nodes, then the top node, and finally the intermediate nodes.

We denote agendas by upper case script letters, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, ..., and represent them by labeled diagrams such as figures 4 and 5. For top down agendas we use a superscripted star, i.e., $\mathcal{A}^*$, and for bottom up agendas we use a subscripted star, i.e., $\mathcal{A}_*$. When an agenda is either all top down or all bottom up we can represent the agenda more simply by nested sets. For example, $\{\{x, y\}, \{v, w\}\}^*$ indicates we first choose among $\{x, y\}$ and $\{v, w\}$ and then within either $\{x, y\}$ or $\{v, w\}$. Alternatively, $\{\{x, y\}, \{v, w\}\}_*$ indicates we first choose among x or y and among v or w, and then compare the best with the best, say x with v.

Relationships to Choice Probabilities.

We illustrate first the relationship of EBA and HEM to top down agendas. Next, we illustrate the effect of bottom up agendas and then we investigate which choice rules are affected by agendas.

Elimination-by-Aspects (EBA) and Agendas. Suppose that a consumer uses EBA whenever he is presented with a set of choice objects, but we constrain him to follow an agenda from the top down. For example suppose that T is partitioned into A and B and we force the consumer to choose first among A and B and then within A or B. Following Tversky and Sattath (1979) we assume that the probability of choosing A from T equals the sum of the probabilities of choosing any object in A from T. In our example, such a combination of constraints and EBA says simply:

$$P(x|\mathcal{A}) = P(x|A)P(A|T) \qquad (7)$$

$$\text{where } P(x|A) = \sum_{y \in A} P(y|T)$$

and $P(x|A)$ and $P(y|T)$ are given by EBA. (We can clearly generalize equation 7 to more than one intermediate level. For example, see equation 5.)

To see that such constraints can affect choice probabilities, let $T = \{x, y, v, w\}$ let $\mathcal{A}^* = \{\{x, w\}, \{y, v\}\}^*$, and let the aspect structure be given by figure 6. That is, $x' = \{\alpha_1, \beta_1\}$, $y' = \{\alpha_1, \beta_2\}$, $v' = \{\alpha_2, \beta_1\}$, and $w' = \{\alpha_2, \beta_2\}$. Notice that T is a 2 x 2 factorial structure. Then, after some algebra:

$$P(w|T) = \alpha_2\beta_2/[(\alpha_1 + \alpha_2)(\beta_1 + \beta_2)] \quad (8)$$

whereas $P(w|A^*) = (\alpha_2 + \beta_2)(\alpha_1\beta_1 + \alpha_2\beta_2)/[(\alpha_1 + \alpha_2)(\beta_1 + \beta_2)]$.

For example, suppose $\alpha_1 = .251$, $\alpha_2 = .249$, $\beta_1 = .499$, and $\beta_2 = .001$. Then $P(w|T) = .001$, but $P(w|\overset{*}{A}) = .125$. Clearly, an agenda can influence choice!

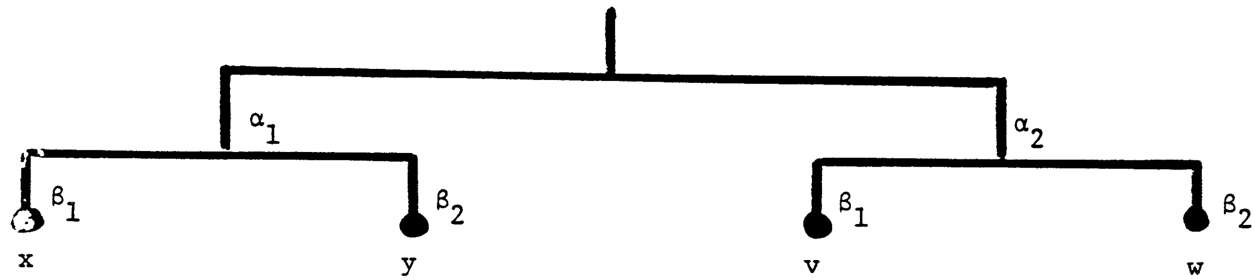


Figure 6: Example Factorial Aspect Structure

Generalized Elimination Model (GEM). Consider the factorial aspect structure in figure 6. (For example, α_1 may equal 'consoles', α_2 = 'portables', β_1 = 'color', and β_2 = 'B & W' as in figure 1.) If we apply the HEM equations, equations 5 and 6, following the aspect structure, i.e., $A^* = \{\{x, y\}, \{v, w\}\}^*$, and defining the measure of a branch as equal to the sum of all aspects that are unique to that branch, then we obtain:

$$P(x|T) = P(x|\{x, y\})P\{x, y\}|A) = \left(\frac{\beta_1}{\beta_1 + \beta_2}\right)\left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right) \quad (9)$$

Note that such a definition reduces to that used for preference trees because, by definition, there are no aspects shared among branches of a preference tree.

We define a more general model, GEM, by equations 5 and 6, but we allow the measure of a branch to be a general function of the unique and shared aspects of a branch. The above definition is but one example where the shared aspects, in this case, β_1 and β_2, have no effect on the probability of choosing a branch.

To illustrate that the choice of how one computes the measure of a branch can have an impact on choice probabilities, consider the aspect structure in figure 6 and the top down agenda, $\mathcal{B}^* = \{\{x, y\}, v\}^*$. Suppose $\alpha_1 = .03$, $\alpha_2 = .07$, $\beta_1 = .80$, and $\beta_2 = .10$. Apply GEM with no consideration of shared aspects when computing $m(\{x, y\})$ would imply:

$$P(\{x, y\}|\mathcal{B}^*) = \frac{\alpha_1 + \beta_2}{(\alpha_1 + \beta_2) + \alpha_2} \tag{10}$$

$P(\{x, y\}|\mathcal{B}^*) > P(\{x, y\}|A^*)$ as expected by the axiom of regularity and β_2 enters the calculation, but intuitively we would not expect the deletion of w to raise $P(x|\mathcal{B}^*)$ from .3 to .6 as equations 9 and 10 suggest. More likely, we would expect that β_1, which can be obtained on both branches, would also affect $P(\{x, y\}|\mathcal{B}^*)$.

Perhaps we should consider shared aspects in computing branching probabilities. For example, we might compute $P(\{x, y\}|\mathcal{B}^*)$ by including β_1 thus:

$$P(\{x, y\}|\mathcal{B}^*) = \frac{\alpha_1 + \beta_2 + \theta\beta_1}{(\alpha_1 + \beta_2 + \theta\beta_1) + (\alpha_2 + \theta\beta_1)} \tag{11}$$

where we have weighted the measure of β_1 by a scale factor of θ in the range $0 < \theta < 1$. If $\theta = 0$, equation 11 reduces to equation 10. However, if $\theta = 1$ using the scale above, we obtain $P(\{x,y\}|\mathcal{B}^*) = .46$ and $P(x|\mathcal{B}^*) = .4$, which is a more reasonable increase in the probability of choosing x.

It is an empirical question as to how shared aspects affect branching probabilities. Since equation 11 reduces to HEM on a preference tree for any value of θ, we define HEM(θ) as a special case of GEM when the measure of a branch equals the sum of all unique aspects plus θ times the sum of all shared aspects. In section 5, we investigate the strategic implications of shared aspects, that is, of θ.

Bottom Up Agendas. In a bottom up agenda, we represent a choice set by its best choice object, a processing rule often advocated by economists. Processing proceeds much as an athletic team proceeds through a tournament. In particular, for the aspect structure in figure 6 and the agenda, $\mathbb{A}_* = \{\{x, y\}, \{v, w\}\}_*$, we compute equation 12 if the consumer uses EBA for any unconstrained choice.

$$P(x|\mathbb{A}_*) = P(x|\{x, y\})[P(x|\{x, v\})P(v|\{v, w\}) + P(x|\{x, w\}) P(w|\{v, w\})] \quad (12)$$

$$= \left(\frac{\beta_1}{\beta_1 + \beta_2}\right)\left[\left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right)\left(\frac{\beta_1}{\beta_1 + \beta_2}\right) + \left(\frac{\alpha_1 + \beta_1}{\alpha_1 + \beta_1 + \alpha_2 + \beta_2}\right)\left(\frac{\beta_2}{\beta_1 + \beta_2}\right)\right]$$

Using $\alpha_1 = .051$, $\alpha_2 = .049$, $\beta_1 = .64$, and $\beta_2 = .26$, we obtain $P(x|\mathbb{A}_*) = .40$ compared to $P(x|T) = .36$ which we would obtain via unconstrained choice. In this case we see that a bottom up agenda enhances slightly the probability that the "best" choice object, x, is selected. ($x' = \{\alpha_1, \beta_1\}$ and $\alpha_1 > \alpha_2$, $\beta_1 > \beta_2$, hence, we call x "best".) We will see in section 5 that, in general, certain bottom up agendas enhance "good" objects while making "bad" objects less likely to be chosen.

For now, suffice it to say that bottom up agendas do affect choice and do so differently than top down agendas. For example, with EBA on the above aspect measures, $P(x|\mathbb{A}^*) = .36$, which turns out to be the same as the unconstrained choice, $P(x|T)$.

Invariance. We have already shown many examples where agendas affect choice probabilities. We might wonder whether agendas always affect choice probabilities or whether there is any choice rule for which choice probabilities are not affected by agendas. We call such unaffected choice rules invariant.

Consider the factorial aspect structure in figure 6. The example in equation 8 has shown that the agenda, $\mathbb{A}^* = \{\{x, w\}, \{y, v\}\}^*$ affects EBA choice probabilities. Similarly, if we compute HEM(0) for $\mathbb{A}^*$, we obtain:

$$P(w|\mathbb{A}^*) = 1/2\left(\frac{\alpha_2 + \beta_2}{\alpha_1 + \alpha_2 + \beta_1 + \beta_2}\right) \quad (16)$$

which is clearly different than the expression for $P(w|T)$ obtained in equation 8. This example illustrates that, in general, neither EBA nor HEM (0) are invariant with respect to top down agendas.

It is easy to see that for top down agendas, a cognitive processing model is invariant if and only if $P(x|T) = P(x|A)P(A|T)$ for any A whenever A is a subset of T and $P(x|T)$ is non-zero. However, this condition is Luce's (1959) choice axiom which is the defining property of CRM. Thus, for top down agendas, CRM is the only invariant decision model.

Consider the bottom up agenda, $\mathcal{B}_* = \{\{x, y\}, \{v, w\}\}_*$. Applying CRM to each pair gives us:

$$P(x|\mathcal{B}_*) = \left(\frac{u(x)}{u(x) + u(y)}\right)\cdot\left[\left(\frac{u(x)}{u(x) + u(v)}\right)\cdot\left(\frac{u(v)}{u(v) + u(w)}\right) + \left(\frac{u(x)}{u(x) + u(w)}\right)\cdot\left(\frac{u(w)}{u(v) + u(w)}\right)\right]$$

which does not reduce to CRM model of

$$P(x|T) = \frac{u(x)}{u(x) + u(y) + u(v) + u(w)}$$

Thus, not even CRM is invariant with respect to bottom up agendas. Since EBA and HEM(θ) are equivalent to CRM when there is no overlap among the aspect sets of x, y, v, and w, this example is sufficient to show that CRM, EBA, and HEM (θ) can be affected by bottom up agendas. To date, we know of no decision rules which are invariant with respect to bottom up agendas.

We state this result as a theorem since it is important conceptually even if it is easy to prove mathematically.

> *Theorem 1 (Invariance): The constant ratio model is the only decision rule invariant with respect to top down agendas. On the other hand, each of the decision rules, CRM, EBA, and HEM(Θ) can be affected by bottom up agendas.*

Theorem 1 is encouraging. Agendas do affect choice. We should be able to identify at least some agendas that affect choice in scientifically interesting and managerially useful ways.

Theorem 1 has an additional benefit because it provides a means to test whether CRM is a reasonable descriptive model of behavior. That is, if no top down agendas can be found to affect choice, then CRM is not eliminated as a decision rule. If any top down agenda affects choice, then CRM cannot be the decision rule.

Behavioral Hypotheses - Familiarity

The preceding analysis introduced agenda constraints and illustrated their effect on choice probabilities. In some cases the constraints will be imposed externally, but in other cases, they will be the result of self-imposed simplifications in cognitive processing. We close this section by setting forth initial hypotheses as to when self-imposed agendas are likely to be used by consumers. These hypotheses are empirically testable. While our analytic results in subsequent sections do not depend explicitly on the validity of these hypotheses, these hypotheses do serve to motivate and interpret self-imposed agendas.

Top Down vs. Bottom Up Agendas. In figure 5 we illustrated bottom up agendas with restaurant choice. Suppose you are at a conference in an unfamiliar city, say San Francisco, and you are faced with a restaurant choice. You have not experienced the various restaurants, but you have some idea as to what to expect. You may first decide on a seafood restaurant because San Francisco is known for seafood, then decide on one near your hotel, ask or read about nearby seafood restaurants and make a choice. When you arrive at the restaurant you select an item from the menu. This is clearly a top down hierarchical choice process.

Suppose instead you are at home and are faced with a restaurant choice. Since you eat out often, you know what you are likely to order at each restaurant, i.e., the restaurant is represented by its "best" item according to your tastes and taking prices into account. You compare these "best" items when selecting the "best" seafood restaurant, and you compare the "best" seafood restaurant to the "best" Chinese reseaurant to the "best" Italian restaurant, etc. when making your final choice. This is a bottom up hierarchical choice process.

Consider the hiring decision in the introduction to this paper. If there are many candidates and the search cost is high, we would likely narrow the field by deciding on an area of interest. On the other hand, if search cost is low as in the case of few candidates, we might use a bottom up, "best" versus "best" decision rule.

Similarly, if we are hiring at the "junior" level, we may be more likely to use a top down decision process because we are uncertain, even after a campus interview, how productive the candidate will be. At the 'senior' level we have much better information on research and teaching productivity and may be more likely to use a bottom up process where we compare, say, the best available model builder with the best available consumer behavioralist.

In each of these anecdotes, the key variable is familiarity. In general, we posit that:

> When one is very familiar with objects or search cost is low, and uncertainty is low, then a bottom up agenda will be favored. When one is unfamiliar with the choice objects, and search cost is high, or uncertainty is high, then a top down agenda will be favored.

One implication of the above familiarity hypothesis is that lack of information will lead an individual to a top down choice process and could conceivably lead him to eliminate an optimal choice object early in the hierarchy. For example, suppose the "best" Seattle restaurant is not a seafood restaurant or the "best" junior faculty candidate is not in the area we choose.

EBA vs. GEM. GEM computes representative measures to summarize the attractiveness of a branch. Constrained EBA sums the probabilities of the individual items. As with top down agendas, we expect representative measures to be used more often when detailed information about the objects is unknown, uncertain, or difficult to obtain. We posit that:

For constrained top down agendas, GEM will be the operant rule when detailed information is not available, while EBA will be favored when detailed information is available.

In analyzing the familiarity hypothesis, it is important to recognize that both GEM and EBA are paramorphic models of cognitive processing. GEM does not necessarily imply that an individual uses the postulated arithmetic rule to compute the measures, $m(A_\ell)$, of each branch, but rather that the arithmetic rule will provide a good estimate of the weight he (or she) will attach to each branch. Thus, according to the familiarity hypothesis, an individual will assign weights to branches based on limited familiarity. These weights will be similar to what we, as analysts, compute by the GEM rules.

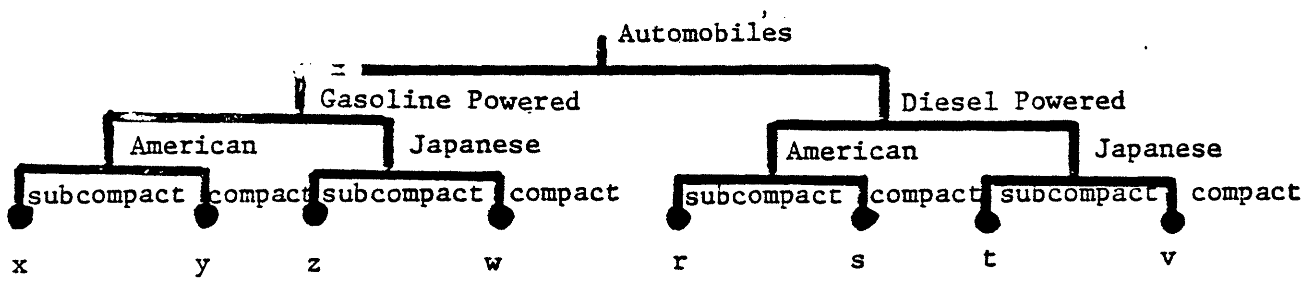
4. ASPECT STRUCTURES AND AGENDAS

We have already considered special aspect structures in the form of preference trees such as the automobile example in figure 2. There is another class of aspect structures that is important and which is also more parsimonious than a general structure.

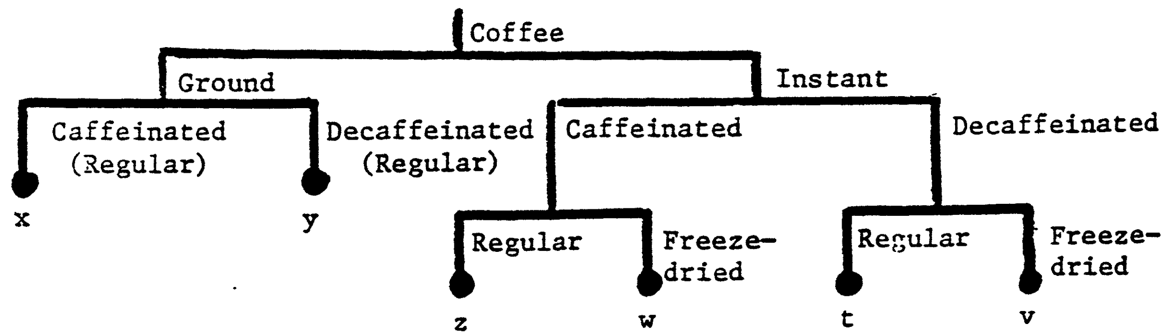
Consider controlled experiments where the researcher manipulates the choice objects, product categories where features dominate unique brand images, and product categories where manufacturers have exploited market segmentation by offering all possible combinations of features. In each of these instances we might expect to see factorial structures. For example, figure 7a is a factorial structure which represents an alternative hypothesis about the automobile market. In a factorial structure there is overlap among branches and that overlap is complete in the sense that aspects are partitioned into groups such that each object has exactly one aspect from each group and all possible combinations of these aspects are represented in the choice set.

Factorial structures (FS) occur often in marketing. For example, applications of conjoint analysis in marketing rely heavily on factorial designs (possibly fractional) for data collection. See review by Green and Srinivasan (1979). Similarly, factorial experimental designs are used to investigate consumer behavior theories. See Sternthal and Craig (1981). Some markets naturally evolve as factorial designs as line extensions are introduced to fill every market niche.

For example, one might describe instant coffees as 'caffeinated' versus 'decaffeinated' at one level, and 'regular' versus 'freeze dried' at another level. See empirical evidence in Urban, Johnson and Hauser (1984). However, as figure 7b illustrates, the entire coffee market may not be a factorial structure because there are no freeze-dried ground coffees. (On the other hand, the first two levels in figure 7b can be considered an FS for coffees.)



a) Factorial Aspect Structure for Automobiles



b) Instant coffee is a Factorial Aspect Structure, but coffee is not.

Figure 7: Aspect Structures

At this point it is worth digressing to recognize that the definition of aspects may be only an approximation. For example, in the instant coffee market there may be other differences between brands than those shown in figure 7b. Taster's Choice may have a different color label than Folger's and that color may matter in consumer choice. By analyzing the instant coffee market as an FS we are assuming that the specified features (aspects) dominate the features we do not model explicitly. The validity of such assumptions can only be answered empirically, but such assumptions do allow us to isolate and study agenda effects recognizing that, in application, we might have to include other effects in our analysis.

Compatibility

Theorem 1 shows that only CRM is unaffected by all top down agendas. But, this does not mean that all agendas affect choice outcomes. For example, it is clear that an agenda that matches, or at least does not disrupt, a GEM hierarchy will not affect GEM choice probabilities. But how about EBA? EBA is not a hierarchical processing rule but is instead a random access processing rule. Intuitively, we expect agendas to influence EBA probabilities.

But not all agendas do affect EBA probabilities. Consider the factorial structure in figure 6 and the "compatible" top down agenda structure in figure 8. First we compute an unconstrained EBA probability.

$$P(x|T) = \alpha_1 \left(\frac{\beta_1}{\beta_1 + \beta_2}\right) + \beta_1 \left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right) = \left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right)\left(\frac{\beta_1}{\beta_1 + \beta_2}\right) \tag{13}$$

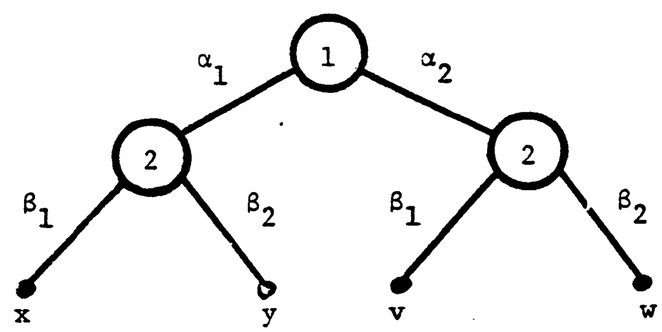


Figure 8: Agenda Structure "Compatible" with Factorial Aspect Structure in Figure 6.

where we have assumed $\alpha_1 + \alpha_2 + \beta_1 + \beta_2 = 1$ without loss of generality. Now we compute the EBA probability constrained by the agenda in figure 8.

$$P(x|A^*) = P(x|\{x, y\})P(\{x, y\}|T)$$

$$= \left(\frac{\beta_1}{\beta_1 + \beta_2}\right) \cdot \left[\alpha_1\left(\frac{\beta_1}{\beta_1 + \beta_2}\right) + \beta_1\left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right) + \alpha_1\left(\frac{\beta_2}{\beta_1 + \beta_2}\right) + \beta_2\left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right)\right]$$

$$= \left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right)\left(\frac{\beta_1}{\beta_1 + \beta_2}\right) \qquad (14)$$

They are the same! As it turns out, this result generalizes.

Following Tversky and Sattath (1979), we call two trees compatible if and only if there exists a third tree, defined on the same choice objects that is a refinement of both. For example, $\{\{\{x, y\}, z\}, \{t, v, w\}\}$ and $\{\{x, y, z\}, \{\{t, v\}, w\}\}$ are compatible since $\{\{\{x, y\}, z\}, \{\{t, v\}, w\}\}$ refines both agendas. On the other hand, $\{\{x, w\}, \{y, v\}\}$ is not compatible with $\{\{x, y\}, \{v, w\}\}$ since there is no tree which refines both these agendas. Note that the degenerate tree $\{x, y, v, w, \ldots\}$, implied by CRM is compatible with all trees on T. For factorial structures, an agenda is compatible if each branch of the agenda corresponds to dividing the factorial structure on the (factorial) groups of aspects.

Tversky and Sattath (1979) have shown that, in general, constrained top down agendas do not affect choice probabilities if the aspect structure is a preference tree and the agenda is compatible with the preference tree. We show in the appendix that this result also holds if the aspect structure is a factorial structure and the agenda is compatible with the factorial structure.

We also show the more surprising result that compatible preference trees and compatible factorial structures *are the only* agendas that do not affect choice probabilities. Except for specific choices of aspect measures[1], all

[1]We can always find degenerate cases in which $P(x|T) = P(x|A^*)$ for some choice of aspect measures. For compatible pretrees and FS's and only for compatible pretrees and FS's, $P(x|T) = P(x|A^*)$ for all choices of non-zero aspect measures. Define compatibility for non-tree/non-factorial when, for each split due to the agenda, there is a set of aspects contained in each branch that are contained in no other branch in the split.

other agendas will affect EBA choice probabilities. Even fractional factorial agendas and incomplete factorial agendas similar to figure 7b will affect choice. Formally,

> *Theorem 2 (Compatibility): For an arbitrarily chosen set of aspect measures, a constrained top down agenda,* $\mathbb{A}^*$*, has no effect on a family of EBA choice probabilities,* $P(x|T)$ *for all* $x \in T$*, if and only if either (1) the aspect structure forms a preference tree and* $\mathbb{A}^*$ *is compatible with the tree or (2) the aspect structure forms a factorial structure and* $\mathbb{A}^*$ *is compatible with the factorial structure.*

The detailed proof to theorem 2 is long and complex. Basically, we first show that EBA is algebraically equivalent to a hierarchical rule for preference trees and factorial structures and only for preference trees and factorial structures. We then show that a compatible agenda does not upset the hierarchical nature of the calculation.

Theorem 2 is both interesting and useful. Consider the television example in figure 1, which is a factorial structure with aspects 'consoles' versus 'portables' and 'color' versus 'B & W'. For EBA, theorem 2 implies that an advertising campaign will have no effect if it attempts to influence consumers to make decisons according to the hierarchy in figure 1. Nor will a 'color' versus 'B & W' campaign have an effect. However, a campaign encouraging the comparison 'color portables' to 'B & W consoles' will affect choice probabilities. Theorem 2 also cautions behavioral researchers to avoid agenda experiments on compatible factorial structures if they wish to idenfity agenda effects from observed behavior.

To date, we know of no aspect structures which are not affected by bottom up agendas.

Equivalence

Compare the calculation of $P(x|T)$ via HEM(0) in equation 9 to the calculation of $P(x|T)$ via EBA in equation 13. Although the procedure by which we calculate the choice probability differs dramatically, we obtain the same answer. This is true despite the fact that HEM and EBA are quite different hypotheses about how a consumer processes information to make a choice. For example, HEM is an explicit, sequential, top down decision process while EBA is a random access elimination process.

The calculation in equations 9 and 13 were based on the factorial structure in figure 5. It turns out that this result generalizes to all factorial structures and, as shown earlier by Tversky and Sattath (1979), to preference trees. But the result holds for no other structure.[2] Furthermore, the result does not hold when shared aspects are considered, i.e., when θ is not zero. Formally,

> *Theorem 3 (Equivalence): For an arbitrarily chosen set of aspect measures, the generalized hierarchial elimination model and elimination by aspects yield equivalent choice probabilities if and only if (1) the aspect structure is a preference tree or (2) the aspect structure is a factorial structure, $\theta=0$, and the hierarchy associated with HEM is compatible with the preference tree or factorial structure.*

An immediate corollary of Theorem 3 is that HEM(0) probabilities are independent of the order in which aspect partitions are processed. (For $\theta \neq 0$, the order can be shown to matter.) Thus, if a researcher uses a factorial structure experiment to investigate hierarchies and the subject(s) is using HEM(0) or EBA, the researcher will not be able to identify the order of aspect processing or the decision rule the subject is using by simply observing the choice outcomes. However, the researcher may be able to identify orderings or decision rules by other means such as verbal protocols or response time.

Furthermore, in a market that has evolved fully to a factorial structure, say some automobile submarkets or the instant coffee market, managerial actions to influence agendas will not depend upon the specific cognitive processing hierarchy as long as EBA or a compatible HEM(0) applies.

---

[2]Except, of course, for specific choices of aspect measures. We seek the result that holds for arbitrarily chosen aspect measures.

Summary

Theorems 2 and 3 illustrate the very special nature of preference trees and factorial structures. For such aspect structures and only for such aspect structures, compatible agendas do not affect EBA probabilities and these EBA probabilities are equivalent to hierarchical elimination probabilities. Theorems 2 and 3 do suggest that other agendas will affect choice probabilities.

We turn now to an analysis of which agendas do affect choice probabilities.

5. STRATEGIC IMPLICATIONS ON FACTORIAL STRUCTURES

In general, agendas affect choice, but section 4 suggests that the effect depends upon the type of agenda, the aspect structure, the decision rule, and the measures of the aspects. To a marketing manager seeking to improve the probability that his (or her) product is chosen, these are very important questions. For example, if the manager wants to design an advertising campaign to influence consumer agendas, that is, to influence to which competitive products his product is compared, he (or she) will want to evaluate the likely directional effect of his (her) campaign. If the directional effect depends only upon the aspect structure, not upon the specific measures of the aspects, so much the better.

In this section we illustrate the strategic implications of agendas on 2 x 2 factorial structures. Such factorial structures serve to illuminate more general results, but are sufficiently easy to visualize so as to not obscure the intuitive understanding of agenda effects.

We begin with dissimilar groupings, a class of top down agendas that enhance a lesser target object. We then illustrate bottom up agendas which enhance greater objects. Finally, we explore the comparative implications EBA and HEM and the effects due to shared aspects. Throughout this development, we assume the factorial aspect structure in figure 6, that is, $x' = \{\alpha_1, \beta_1\}$, $y' = \{\alpha_1, \beta_2\}$, $v' = \{\alpha_2, \beta_1\}$, and $w' = \{\alpha_2, \beta_2\}$.

Enhancement of a Lesser Object

Marketing folk wisdom suggests that perhaps it is always effective for a low market share product to force a comparison between itself and a high market share product. For example, there are many advertisements in which a 'cola' is compared to Coca-Cola. A similar phenomenon appears to be true for the Pitney Bowes/Savin/Xerox example. However, many automobile manufacturers go out of their way to compare themselves to the low share, but prestige, products of BMW and Mercedes.

Consider the agenda, A^*, in figure 9 and suppose we choose the aspects such that $\alpha_1 > \alpha_2$ and $\beta_1 > \beta_2$. That is, we choose the aspects such that the least prefered object, w, is compared to the most preferred[3] object, x. For example, if $\alpha_1 = .03$, $\alpha_2 = .01$, $\beta_1 = .81$, and $\beta_2 = .15$ then $P(w|A^*) = .08$ and $P(w|T) = .04$. The folk wisdom appears to be true for this example.

However, for the folk wisdom to be true for other choice objects in T, the condition that y has a lower unconstrained probability than v should be sufficient to assure that the probability of choosing y is enhanced by the agenda, A^*. In this example, $P(y|T) = .12$ which is less than $P(v|T) = .21$. But, the agenda A^* actually hurts y, i.e., $P(y|A^*) = .09$ which is less than the unconstrained probability, $P(y|T) = .12$. Thus, we have generated an example where it is not effective strategically to compare a low share product to a higher share product.

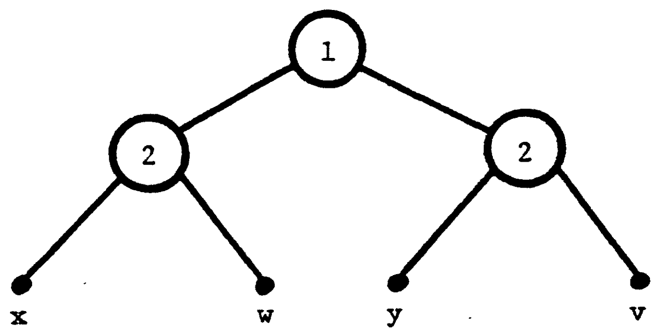


Figure 9: Dissimilar Grouping Top Down Agenda for Factorial Structure in Figure 6.

3Here we loosely interpret "preference" as the EBA probability.

Fortunately, we can identify an interpretable condition where an agenda is effective independent of the specific aspect measures. In particular, if we group the object, w, in a top down agenda with that object, x, for which w is maximally dissimilar but inferior, then the agenda will enhance the probability that w will be chosen. This result holds independent of the specific aspect measures as long as $\alpha_1 > \alpha_2$ and $\beta_1 > \beta_2$. We state the result for a 2 x 2 factorial structure, but in the appendix we show it generalizes for an ℓ-level factorial and for more aspects.

> *Result 1 (Dissimilar Grouping): For the factorial structure in Figure 6, the top down agenda,* $\mathbb{A}^* = \{\{x, w\},\{y,v\}\}^*$, *enhances the EBA probability that the least preferred object, w, is chosen and hurts the EBA probability that the most preferred object, x, is chosen. That is,* $P(w|\mathbb{A}^*) > P(w|T)$ and $P(x|T) > P(x|\mathbb{A}^*)$.

For example, for the FS in figure 1, a campaign that encourages comparisons between B & W portable televisions and color console televisions will always be effective for B & W portables as long as '*color*' > '*B & W*' and '*console*' > '*portable*'.

We show in the appendix (Result 1.1) that the result holds even if there are some additional aspects that are common to the two dissimilar objects. Thus, a manager introducing a copier, w, which is "just as good as a Xerox" on some aspects and (probabilistically) weaker on all other aspects, say the image of the brand names, could increase w's market share if consumers could be encouraged to compare w to only a Xerox. We recognize, of course, that such a campaign might be effective for other reasons such as the comparison giving the lesser copier, w, a quality image. Result 1 suggests that the agenda effect reinforces such image effects.

Enhancement of a Greater Object

Result 1 indicates how a marketing manager might use an agenda to enhance the probability that a lesser object, w, is chosen. We now examine agendas that enhance the stronger object, x.

Anyone familiar with sports tournaments (tennis, basketball, soccer, etc.) knows that "seeding", i.e., the selection of the tournament agenda, affects the probabilities of the outcomes. Common belief is that the "right" seeding will lead to the best teams finishing well and worst teams being eliminated early. The "wrong" seeding leads to upset victories.

According to our familiarily hypothesis in section 3, we believe consumers tend to use bottom up agendas when they are familiar with the objects. Our intuitive belief might be that bottom up agendas are "good" decision rules. That is, they lead to "better" choices in the sense that choice objects with better aspects (higher aspect measures) are more likely to be chosen and "poorer" choice objects (lower aspect measures) more likely to be eliminated.

In this section we illustrate that such intuition is reasonable, some bottom up agendas alter choice probabilities to favor choice objects with better aspects. However, just as in sports tournaments, this phenomena is operant only with the "right" agenda. The "wrong" agenda can be counter productive and lead to "upsets".

To illustrate the effect of bottom-up agendas, consider the three possible pairing agendas on a 2 x 2 factorial structure. Following our convention and without loss of generality, we assume $\alpha_1 > \alpha_2$ and $\beta_1 > \beta_2$. The first agenda, $\mathfrak{A}_* = \{\{x, v\}, \{y, w\}\}_*$, makes the first comparisons with respect to products that differ only on α_1 and α_2. The second agenda, $\mathfrak{B}_* = \{\{x, y\}, \{v, w\}\}_*$, makes the first comparisons with respect to β_1 and β_2. Finally, the third agenda, $\mathfrak{C}_* = \{\{x, w\}, \{y, v\}\}_*$, is analogous to top down dissimilar groupings in that it "seeds" the best product against the worst and the two middle products against each other. Finally, without loss of generality, assume that $\{\alpha_1, \alpha_2\}$ is the more important aspect pair, i.e., $\alpha_1/(\alpha_1 + \alpha_2) > \beta_1/(\beta_1 + \beta_2)$. These assumptions assure that the EBA ordering is::

$$P(x|T) > P(y|T) > P(v|T) > P(w|T)$$

We first examine the agendas, $\mathbb{A}_*$ and $\mathbb{B}_*$, that are compatible with the factorial structure. Intuitively, we expect that the agenda, $\mathbb{A}_*$, which first encourages the "easy" comparison (α_1 vs. α_2), would enhance the "good" products, x and y. We expect $\mathbb{B}_*$ to do just the opposite and enhance "upsets", v and w. It turns out that this is indeed the case. In particular,

Result 2 (Bottom-Up Agendas): For compatible bottom-up agendas, $\mathbb{A}_*$ *and* $\mathbb{B}_*$*, on a 2 x 2 factorial structure where* $P(x|T) > P(y|T) > P(v|T) > P(w|T)$*, doing the easy comparison first* (α_1 *vs.* α_2) *enhances objects with already higher probability and doing the difficult comparison* (β_1 *vs.* β_2) *enhances objects with lower probabilities. That is:*

$$(1)\ P(x|\mathbb{A}_*) > P(x|T) > P(x|\mathbb{B}_*)$$

$$(2)\ P(y|\mathbb{A}_*) > P(y|T) > P(y|\mathbb{B}_*)$$

$$(3)\ P(v|\mathbb{B}_*) > P(v|T) > P(v|\mathbb{A}_*)$$

$$(4)\ P(w|\mathbb{B}_*) > P(w|T) > P(w|\mathbb{A}_*)$$

To illustrate this bottom up agenda effect in another way consider the concept of entropy, $H(\mathbb{A})$, for an agenda, $\mathbb{A}$.

$$H(\mathbb{A}) = -[P(x|\mathbb{A}) \ln P(x|\mathbb{A}) + P(y|\mathbb{A}) \ln P(y|\mathbb{A}) + P(v|\mathbb{A}) \ln P(v|\mathbb{A}) + P(w|\mathbb{A}) \ln P(w|\mathbb{A})]$$

Entropy measures the uncertainty of the system. High entropy means that the agenda tells us little about choice outcomes. (Entropy is maximized when all choice objects are equally likely to be chosen.) Reductions in entropy can be considered information and should be favored by consumers. See Gallagher (1968), Hauser (1978) and Jaynes (1957).

Since $\mathbb{A}_*$ makes probabilities more extreme, i.e., closer to 1.0 or 0.0, we expect it to decrease entropy relative to EBA. Similarly, we expect $\mathbb{B}_*$ to increase entropy. Formally.

Result 2.1 (Entropy): According to the conditions of Result 2, performing the easy comparisons first decreases entropy and performing the difficult comparisons first increases entropy. That is:

$$H(\mathbb{A}_*) < H(T) < H(\mathbb{B}_*)$$

By analogy to a sports tournament "seedings", we expect that a bottom up dissimilar grouping, i.e., first matching the "best" product, $x' = \{\alpha_1, \beta_1\}$, with the "worst" product, $w' = \{\alpha_2, \beta_2\}$, would maximize the probability of choosing the "best" object.

As it turns out, (see Appendix, result 2.2) the dissimilar grouping agenda, $\mathfrak{C}_*$, may or may not enhance x relative to unconstrained choice. However, $\mathfrak{C}_*$ never enhances x more than the best compatible agenda, $\mathfrak{A}_*$. While at first glance, this may seem counter intuitive, it does make good intuitive sense. The first comparison in the best compatible agenda, say x vs. v, is made with respect to the most favorable aspect, in this case α_1 vs. α_2. The first dissimilar grouping comparison, x vs. w., is made with respect to both aspect pairs, α_1 and α_2 and β_1 vs. β_2. Thus, although $\beta_1 > \beta_2$, the comparison with respect to β_1 vs. β_2 dilutes the strength of α_1 vs. α_2. Hence, according to EBA, x is more likely to be chosen over v than over w. It is not surprising, therefore, that the best compatible agenda is better than the dissimilar agenda. Furthermore, in the appendix we note that $\mathfrak{C}_*$ may actually do worse than unconstrained choice if $\beta_1 \beta_2 > \alpha_1 \alpha_2$.

In summary, our analyses of bottom up agendas suggest that marketing managers with superior quality products can enhance the probability that their product is chosen if they encourage consumers to use bottom up processing strategies and make the easy comparisons first.

Result 2.1 also suggests that the consumer who wants to improve his decision making with implicit agendas should use a bottom up agenda and make "easy" comparisons first.

Shared Aspects and Processing Rules.

Results 1 and 2 suggest when and how a marketing manager can use agenda effects to the advantage of a product. We can also imagine situations where an advertising or salesforce presentation can influence consumers to use one or another processing strategy. For example, Tversky (1972 a,b) suggests a number of advertising strategies consistent with EBA.

In this section we focus on the relative strategic effect of alternative processing rules, in particular, EBA and HEM(θ). For ease of exposition and without loss of generality, we label the first comparison in HEM(θ) with α_1 and α_2. That is, for figure 6, $A^* = \{\{x, y\}, \{v, w\}\}^*$.

For example, let $\alpha_1 = \beta_1 = .4$ and $\alpha_2 = \beta_2 = .1$, then by Theorem 3 we know that the EBA and HEM(0) probabilities are equal because the HEM hierarchy is compatible with the aspect structure. But what happens as θ increases? Calculating, we obtain:

$$P(x|A^*) = .64 \qquad \text{for } \theta = 0$$

$$P(x|A^*) = .52 \qquad \text{for } \theta = 1/2$$

$$P(x|A^*) = .48 \qquad \text{for } \theta = 1$$

Thus, it appears that the more the consumer considers the shared aspects, β_1 and β_2, the lower the likelihood that the "best" product, x, is chosen. As it turns out, for all 2 x 2 factorial structures, shared aspects hurt the products that are better on the first comparison, x and y, and enhance the products that do worse on the first comparison, v and w. Formally,

> *Result 3 (Shared Objects): For a 2 x 2 factorial structure with the first comparison made with respect to α_1 and α_2, and for $\alpha_1 > \alpha_2$, shared aspects in hierarchial processing, i.e., HEM(θ), enhance those objects which contain α_2 and hurt those objects which contain α_1. The effect increases as the importance of the shared objects increases, i.e., as θ increases.*

This result is best visualized with the graph in figure 10. By the equivalence theorem, the HEM and EBA probabilities start out the same at $\theta = 0$, but as θ increases, the choice probability increases (decreases)[4] whenever the choice object contains α_2 (α_1). (Assuming of course $\alpha_1 > \alpha_2$.)

---

[4]The curvature is as drawn. Second order conditions are negative for objects containing α_2 and positive for objects containing α_1.

Thus, the manager of product x, where $x' = \{\alpha_1, \beta_1\}$, should discourage the consideration of shared aspects while the manager of product w, $w' = \{\alpha_2, \beta_2\}$, will wish to encourage the consideration of shared aspects. Managers of "mixed" products, $y' = \{\alpha_1, \beta_2\}$ and $v' = \{\alpha_2, \beta_2\}$, may or may not wish to encourage shared aspects depending upon whether the α's or the β's are considered first.

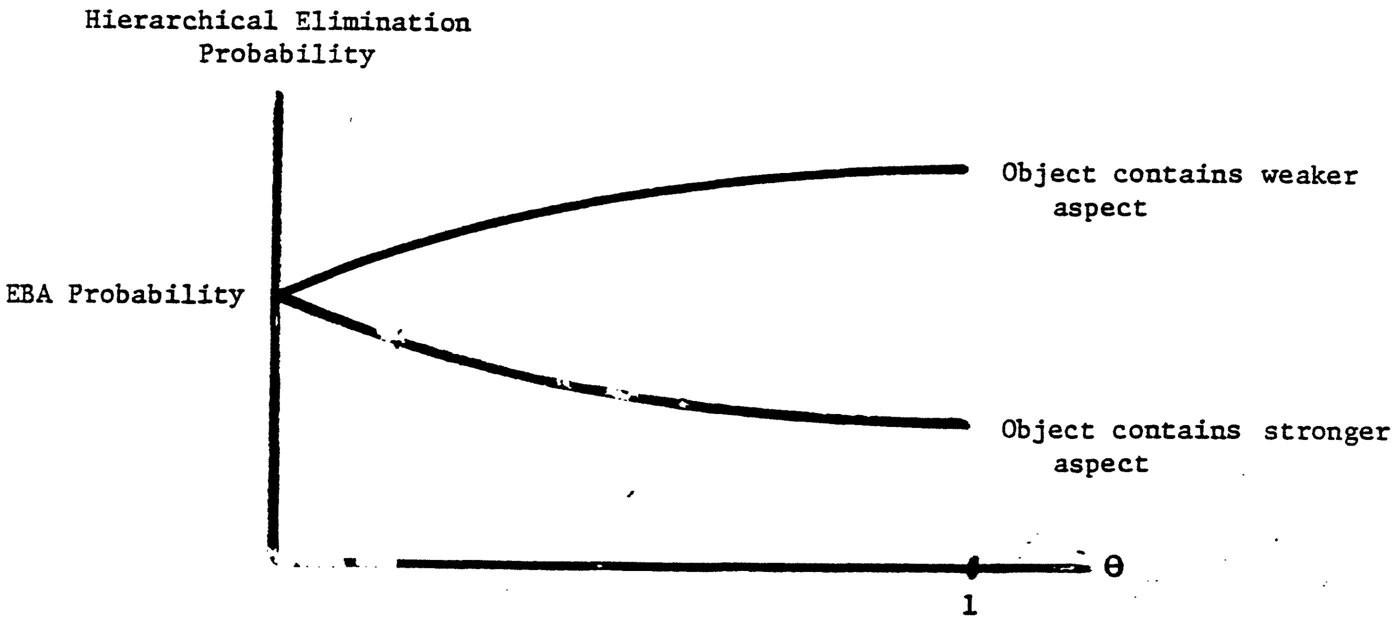


Figure 10: The Effect of Shared Aspects on Choice Probabilities. θ measures the importance of shared objects.

Summary

Agendas and processing rules do affect choice probabilities. By understanding their effects, the marketing manager can begin to generate strategies with which to increase market share. In this section we have explored some of these effects illustrating our results on 2 x 2 factorial structures.

In particular, our results suggest:

- managers with a lesser product should encourage consumers to use a top down agenda grouping together "dissimilar" products,
- managers with a greater product should encourage consumers to use a bottom up agenda making the "easy" comparisons first,
- managers with a strong aspect, say α_1, should encourage consumers to use a random access rule (EBA),
- if consumers do use a hierarchical rule, considering α_1 vs. α_2 first, managers with a strong aspect should encourage consumers <u>not</u> to consider shared aspects
- managers with a weak aspect, say α_2, should encourage consumers to use a hierarchical rule, consider α_1 vs. α_2 first and place high weight on the shared aspects β_1 and β_2. Finally,

- consumers who wish to increase the likelihood that they will choose the "best" product should use a bottom up agenda making the "easy" comparison first.

These results are illustrative of agenda effects. We leave extensions and generalizations to future research.

6. DOMINANCE AND REGULARITY

The previous section considered strategic managerial implications. In this section we explore more theoretical implications of agendas. In particular, we show that agendas can cause two commonly assumed axioms of probabilistic choice theory to be violated. Furthermore, we illustrate that such violations are quite reasonable and should be expected.

Consider the following four vacations:

x: Japan via Japanese Airlines (JAL) with free drinks on the plane.

y: Japan via Northeast Orient (NW).

v: Japan via JAL, and

w: Hong Kong via NW.

If a person were choosing from the set, $\{x, y, v, w\}$, vacation x dominates vacation v and, hence, we would expect that no rational consumer would choose v from the set, $\{x, y, v, w\}$. Indeed, with a random access rule such as EBA, the probability that v is chosen is zero. Structurally, with EBA, dominated choice objects cannot be chosen because either the dominating aspect will be chosen, or at some point in the decision process all other aspects will have been considered and the choice will be x vs. v.

Such is not the case with agendas. Consumers may pay a price for simplifying their decision rules with agendas.

Consider the top-down agenda, $\{\{x, y\}, \{v, w\}\}^*$, shown in figure 11. As we might expect, the right-hand branch $\{v, w\}$, has a non-zero probability of being chosen and v has a non-zero probability of being selected from the branch. Furthermore, such a choice makes good intuitive sense if the consumer is using a two-step decision process. Suppose that, in January, a consumer must choose a tour company and, in June, he must choose which of two tours to take. A rational consumer might choose the $\{v, w\}$ tour company to keep his options open and maintain maximum variety. He would then choose vacation v if, in June, he decides he would really rather go to Japan than to Hong Kong. He may even feel that the sacrifice of free drinks is well worth the chance to delay his decision on which country to visit. (Such a decision is analogous to the decisions made by business travelers who avoid "supersaver" fares in favor of more flexible regular fares.)

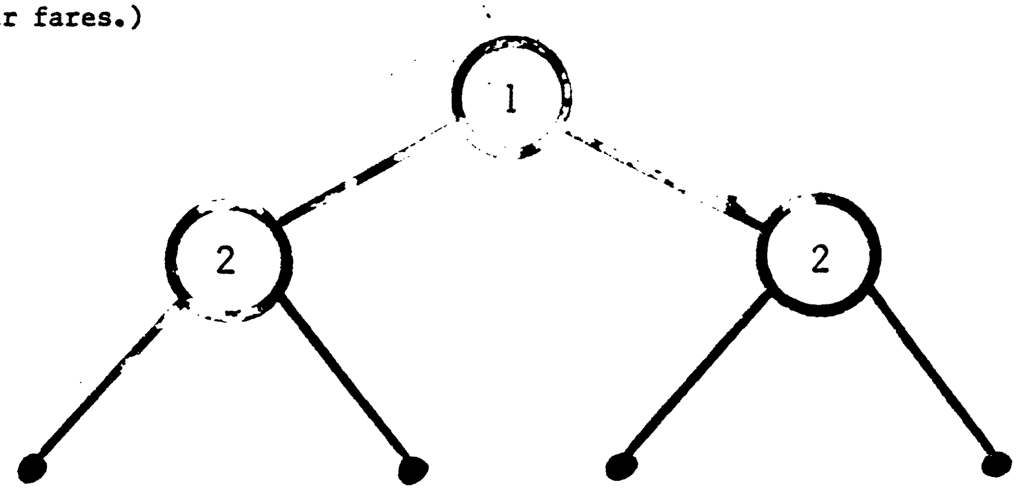


Figure 11: Illustrative Agenda in which a Dominated Choice Object Might be Chosen.

Depending on the aspect measures, we can make this effect as strong as we like. For example, suppose α_1 = 'Japan', α_2 = 'Hong Kong', β_1 = 'JAL', β_2 = 'NW', γ = 'free drinks', $\alpha_1 = .94$, $\alpha_2 = .05$, $\beta_1 = .008$, $\beta_2 = .001$, and $\gamma = .001$. If the consumer uses EBA, we get:

$$P(x|T) = .76$$

$$P(v|T) = .00$$

which is much as we would expect if, in June, he can still choose from the entire choice set. However, if the consumer must use a hierarchical decision rule, HEM(0), with the agenda in figure 11, we get:

$$P(x|A^*) = .02$$

$$P(v|A^*) = .93$$

which is again as we would expect for the "story" cited above.[5]

If agendas can cause consumers to violate dominance, we should not be surprised to see agendas cause consumers to violate regularity. That is, we should expect that with an agenda-based processing rule, the probability of choosing an object, say v, can be enhanced by adding another object, say w, to the choice set.

Consider again the two tour companies. Suppose that the first tour company still offers the two Japan vacations, $\{x, y\}$, but the second tour company only offers the one Japan vacation, v. Now both variety and the fact that x dominates v favor the first tour company. Thus, as expected, even with HEM(0), $P(v|\{\{x,y\}, v\}^*) = 0$. However, as shown above, if the second tour company adds the option of the Hong Kong flight, the second tour company is favored and $P(v|\{\{x, y\}, \{v, w\}\}^*) = .93$. Furthermore, this violation of regularity makes good intuitive sense.

Notice also the extent that the second tour company benefits by adding the Hong Kong trip even though there is only a 5% chance that it is chosen, i.e., $P(w|A^*) = .05$. Their "market share" goes from 0% to 98% with the addition of Hong Kong trip.

For empirical evidence that regularity can be violated see Huber, Payne, and Puto (1982). However, we note that their experiments appear to exploit perceptual effects, not agenda effects.

In summary, agendas can cause dominance and regularity to be violated, but such violations are intuitive and help us better understand the phenomena of agendas.

---

[5] We get the same effect with HEM(1), but the effect is less severe, i.e., $P(x|A^*) = .44$, and $P(v|A^*) = .49$.

7. DISCUSSION, SUGGESTED EXPERIMENTS, AND EXTENSIONS

Consumer choice behavior is by its very nature complex. Any attempt to study it requires that we make tradeoffs between full complexity, i.e., modeling all phenomena we can postulate, and parsimony, i.e., focusing on specific phenomena to understand their role in complex decision making. In this paper we focus on specific phenomena, agendas, and how they influence choice probabilities.

To study agendas we draw upon, but generalize, well known, empirically documented probabilistic models of consumer choice behavior. From this theoretical base, we examine top down and bottom up constraints on consumer choice, constraints that can be external or, perhaps, self-imposed. We show that only one choice rule is not affected by top down agendas and that all of the rules we review are affected by bottom up agendas.

Aspect structures interact with agendas. We show that EBA is unaffected by compatible factorial structures and preference trees and only by those aspect structures. Furthermore, EBA and HEM(0) yield equivalent predictions if and only if the aspect structure is factorial or a preference tree.

Agendas can have strategic managerial implications. We demonstrate top down agendas that enhance lesser objects and bottom up agendas that enhance greater objects. Furthermore, depending upon the aspect structure and the strength of the aspects, we illustrate that a manager may or may not wish to encourage consumers to consider shared aspects.

Finally, we demonstrate that two commonly assumed axioms of choice, dominance and regularity, can be violated due to agenda effects. Furthermore, such violations are plausible and suggest viable managerial strategies.

We feel our results are scientifically interesting. By isolating agenda effects we are better able to understand their implications. Our mathematical results provide a framework with which to study these effects and suggest means by which to test our hypotheses and models. For example, results 1, 2, and 3 suggest that certain agendas will influence choice in predictable directions. The clever experimentor can use these results to test hypotheses about decision rules and/or aspect strucutres or to infer indirectly the operant decision rule.

We feel that our results are useful managerially. Section 5 illustrates selected agenda effects with 2 x 2 factorial structures, many of these results are generalizable,and more results are obtainable. We have chosen those results which illustrate the potential of agendas and which lead to a better intuitive understanding of agendas.

We suggest below some potential experiments and extensions.

Experiments

We have developed our theory from a few simple hypotheses. Nonetheless, our results make strong predictions about behavior. Most of these predictions can be tested experimentally. For example, Tversky and Sattath (1979) have already demonstrated agenda effects on preference trees. Their experiments had 100 subjects choose among triples of gambles. In one choice setting, subjects were constrained by a compatible preference tree and, in another choice setting, by a grouping of dissimilar objects. The results were as predicted and were replicated with a second choice problem.

One interesting experiment to test our results would be to manipulate top down vs. bottom up agendas. To induce a top down agenda, the experimenter can ask subjects to first choose groups of objects and then choose within the chosen group. To induce a bottom up agenda, the experimentor can ask subjects to first choose objects from pairs of objects and them compare the "best" with the "best". The experimentor would also collect data on unconstrained choice. By result 2.1, the best bottom up agenda should increase entropy relative to the random access rule while the compatible top down agenda should induce no change if the subjects are using EBA and HEM(0). The directional predictions of other theorems and results can also be tested by such experiments. Alternatively, one can manipulate the importance of shared aspects by emphasizing (deemphasizing) shared aspects in the statement to the subjects of the choice problem.

Perhaps the most interesting set of experiments for future study are those that test our familiarity hypotheses. Such experiments are more difficult because familiarity must be carefully manipulated. With the right inductions and manipulation checks, the experimentor can vary familiarity and attempt to measure the choice rule and agenda (bottom up, top down, or random access) that the subjects use. The experimentor might wish to use protocol analysis to measure the dependent variable. Alternatively, he can use our theorems and results to infer the operant choice rule and agenda through their effect on the directional change in choice probabilities.

Extensions

It is possible to generalize some of our analysis. We suggest some directions and leave details to future study.

Aggregation. Tversky and Sattath (1979, pp. 552-554) develop an aggregate interpretation of EBA where $P(x|A)$ is the proportion of consumers who select an object, x, from the choice set, A. They show in their Aggregation Theorem (p. 553) that their aggregate interpretation is compatible with EBA.

Since our choice rules are generalizations of their choice rules, we can consider the same interpretations for analyzing agendas. That is, we can interpret choice probabilities as population proportions where the mechanism is that consumers are heterogeneous, but that the aggregate effect of individual elimination processes can be summarized with an EBA-like or GEM-like rule.

Generalized Agendas. We have defined agendas with respect to hierarchical partitionings of choice sets, e.g., $\{\{x, y\}, \{v, w\}\}$. We can generalize this definition by dropping the requirement that each level be a partition. For example, we can allow the following generalized agenda: $\{\{x, y, v\}, \{y, v, w\}, \{\{x, y\}, w\}\}^*$. This generalization allows us to visualize an agenda as a directed graph with the universe, T, as its source and single objects, say x, as sinks.

Alternative Measures. We analyzed the effect of agendas on consumer choice probabilities. The definitions of CRM, EBA, and GEM can be modified for alternative measures that have related properties, i.e., that are greater than zero and sum to one across choice objects. For example, CRM, EBA, and HEM can be defined for any intensity measure such as constant sum preference comparisons. See Torgerson (1958) and Hauser and Shugan (1980) for relevant axioms.

REFERENCES

Aaker, D. and J. Myers, Advertising Management, (Englewood Cliffs, NJ: Prentice-Hall), 1975.

Bettman, J. R., An Information Processing Theory of Consumer Choice, (Reading, MA: Addison-Wesley Publishing Company), 1979.

Day, G. S., A. D. Shocker, and K. Srivastava, "Consumer-Oriented Approaches to Identifying Product Markets," Journal of Marketing, Vol. 43, No. 4, (Fall 1979), pp. 8-20.

Debreu, G., "Review of R. D. Luce, Individual Choice Behavior: A Theoretical Analysis," American Economic Review, Vol. 50, (1960), pp. 186-188.

Farquharson, R., Theory of Voting, (New Haven, CT: Yale University Press), 1969.

Gallagher, Robert, Information Theory and Reliable Communication, (New York: John Wiley & Sons, Inc.) 1968.

Green, P. E. and V. Srinivasan, "Conjoint Analysis in Consumer Research: Issues and Outlook," Journal of Consumer Research, Vol. 5, No. 2, (September 1978) pp. 103-123.

Haines, G. H., "Process Models of Consumer Decision Making," in G. D. Hughes and M. L. Ray (eds.) Buyer/Consumer Information Processing, (Chapel Hill, NC: University of North Carolina Press), 1974.

Hauser, J. R., "Testing the Accuracy, Usefulness, and Significance of Probabilistic Models: An Information Theoretic Approach," Operations Research, Vol. 26, No. 3, (May-June 1978), pp. 406-421.

Hauser, J. R. and S. M. Shugan, "Intensity Measures of Consumer Preference," Operations Research, Vol. 28, No. 2, (March-April 1980), pp. 278-320.

Hauser, J. R. and G. L. Urban, "A Normative Methodology for Modeling Consumer Response to Innovation," Operations Research, Vol. 25, No. 4, (July-August 1975), pp. 579-619.

Huber, J., J. W. Payne, and C. Puto, "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and Similarity Hypothesis," Journal of Consumer Research, Vol. 9, (June 1982), pp. 90-98.

Jaynes, E. T., "Information Theory and Statistical Mechanics," Physics Review, Vol. 106, (1957), p. 620.

Jeuland, A. P., F. M. Bass, and G. P. Wright, "A Multibrand Stochastic Model Compounding Heterogeneous Erlang Timing and Multinomial Choice Processes," Operations Research, Vol. 28, No. 2, (March-April 1980), pp. 255-277.

Johnson, R. M., "Beyond Conjoint Analysis: A Method of Pairwise Trade-Off Analysis," Advances in Consumer Research, (October 1975).

Kalwani, M. and D. Morrison, "A Parsimonious Description of the Hendry System," Management Science, Vol. 23, No. 5, (January 1977), pp. 467-477.

Lancaster, K., Consumer Demand: A New Approach, (New York, NY: Columbia University Press), 1971.

Luce, R. D. Individual Choice Behavior, (New York: John Wiley & Sons, Inc.), 1959.

Luce, R. D., "The Choice Axiom After Twenty Years," Journal of Mathematical Psychology, Vol. 15, (1977), pp. 215-233.

Luce, R. D., and P. Suppes, "Preference, Utility and Subjective Probability," in R. D. Luce, R. R. Bush, and E. Galenter (eds.), Handbook of Mathematical Psychology, Vol. 3, (New York: Wiley & Sons, Inc.), 1965, pp. 249-410.

Mandell, M. I., Advertising, 2nd ed., (Englewood Cliffs, NJ: Prentice-Hall, Inc.) 1974.

McFadden, D., "Econometric Models for Probabilistic Choice Among Products," Journal of Business, Vol. 53, No. 3, Part 2, (July 1980), pp. 513-529.

McKelvey, R. D., "A Theory of Optimal Agenda Design," Management Science, Vol. 27, No. 3, (March 1981), pp. 303-321.

McKelvey, R. D., and R. G. Niemi, "A Multistage Game Representation of Sophisticated Voting for Binary Procedures," Journal of Economic Theory, Vol. 18, No. 1, (June 1978), p. 1-22.

Payne, J. W., "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," Organization Behavior and Human Performance, Vol. 16, (June 1976), pp. 355-387.

Pessimier, E. A., Product Management: Strategy and Organization, (New York: Wiley/Hamilton), 1977.

Plott, C. and M. Levine, "A Model of Agenda Influence on Committee Decisions," American Economic Review, Vol. 68, (1978), pp. 146-159.

Punj, G. N. and R. Staelin, "The Choice Process for Graduate Business Schools," Journal of Marketing Research, Vol. XV, (November 1978), pp. 588-598.

Rao, V. R. and Sabavala, D. J., "Inference of Hierarchical Choice Processes from Panel Data," Journal of Consumer Research, Vol. 8, (June 1981), pp. 85-96.

Restle, F., Psychology of Judgment and Choice, (New York: Wiley & Sons, Inc.), 1961.

Rumelhart, D. L. and J. G. Greeno, "Similarity Between Stimuli: An Experimental Test of the Luce and Restle Choice Models," Journal of Mathematical Psychology, Vol. 8, (1971), pp. 370-381.

Shugan, S. M., "The Cost of Thinking." Journal of Consumer Research, Vol. 7, (September 1980), pp. 99-111.

Silk, A. J. and G. L. Urban, "Pretest Market Evaluation of New Package Goods: A Model and Measurement Methodology," Journal of Marketing Research, Vol. 15, (May 1978), pp. 171-191.

Sternthal, B. and C. S. Craig, Consumer Behavior Theory and Strategy, (Englewood Cliffs, NJ: Prentice-Hall, Inc.), 1982.

Srivastava, R. K., A. D. Shocker, and R. P. Leone, "Market Structure Analysis: Hierarchical Clustering of Products Based on Substitition in Use," Journal of Marketing, Vol. 45, No. 3, (September 1981), pp. 38-48.

Torgerson, W. S., Methods of Scaling, (New York: John Wiley & Sons, Inc.), 1958, pp. 105-112.

Tversky, A., "Choice by Elimination," Journal of Mathematical Psychology, Vol. 9, (1972), pp 341-367.

Tversky, A , "Elimination by Aspects: A Theory of Choice," Psychological Review, Vol. 79, (1972), pp. 281-299.

Tversky, A. and S. Sattath, "Preference Trees," Psychological Review, Vol. 86, No. 6, (1979), pp. 542-573.

Tybout, A. M. and J. R. Hauser, "A Marketing Audit Using a Conceptual Model of Consumer Behavior: Applications and Evaluation," Journal of Marketing, Vol. 45, No. 3, (Summer 1981).

Urban, G. L., P. Johnson, and J. R. Hauser, "Testing Competitive Market Structures." Marketing Science, Vol. 3, No. 2, (Spring 1984), pp. 83-112.

Urban, G. L., and G. Katz, "Pre-Test Market Models: Validation and Managerial Implications," Journal of Marketing Research, Vol. 20, No. 3, (August 1983), pp. 221-234.

Wilkie, W. L., and E. A. Pessemier, "Issues in Marketing's Use of Multiattribute Models," Journal of Marketing Research, Vol. 10, (November 1973), pp. 428-441.

APPENDIX

PROOFS OF THEOREMS AND OTHER RESULTS

Throughout this appendix, we use Greek letters to denote both aspects and their measures whenever this can be done without ambiguity. We begin with five lemmas that simplify our proofs.

Lemma 1: Let T be a factorial structure (FS) with matched aspects, $\{\alpha_{11}, \alpha_{12}, \cdots \alpha_{1k_1}\} \times \{\alpha_{21}, \alpha_{22}, \cdots \alpha_{2k_2}\} \times \cdots \times \{\alpha_{K1}, \alpha_{K2}, \cdots \alpha_{Kk_K}\}$. Then, for all $x \epsilon T$, elimination by aspects yields:

$$P(x|T) = \prod_{\alpha_{ij} \epsilon x'} \alpha_{ij} \Big/ \prod_{\ell=1}^{K} \left(\sum_{n=1}^{k_\ell} \alpha_{\ell n} \right)$$

where the aspects are scaled to sum to 1.0.

Proof. We proceed by induction on K, the number of levels on the FS. Without loss of generality (wlog) let $x' = \{\alpha_{11}, \alpha_{21}, \cdots \alpha_{K1}\}$. Lemma 1 is clearly true for K = 1 since $P(x|T) = \alpha_{11}/\sum_{n=1}^{k_1} \alpha_{1n}$ since all elements of T have disjoint aspect sets for K = 1.

Assume equation A1 holds for (K - 1) and note that $T_{\alpha_{\ell 1}}$ is a (K - 1) level FS. Then:

$$P(x|T) = \sum_{i=1}^{K} \alpha_{i1} P(x|T_{\alpha_{i1}}) = \sum_{i=1}^{K} \alpha_{i1} \left[\frac{\prod_{\ell \neq i} \alpha_{\ell 1}}{\prod_{\ell \neq i} \sum_{n=1} \alpha_{\ell n}} \right]$$

$$P(x|T) = \sum_{i=1}^{K} \left[\left(\prod_{\ell=1}^{K} \alpha_{\ell 1} \right) \left(\sum_{j=1}^{k_i} \alpha_{ij} \right) \right] \Big/ \left[\prod_{\ell=1}^{K} \left(\sum_{n=1}^{k_\ell} \alpha_{\ell n} \right) \right]$$

$$= \left[\prod_{\ell=1}^{K} \alpha_{\ell 1} \Big/ \prod_{\ell=1}^{K} \left(\sum_{n=1}^{k_\ell} \alpha_{\ell n} \right) \right] \cdot \left[\sum_{i=1}^{K} \sum_{j=1}^{k_i} \alpha_{ij} \right]$$

which completes the induction since $\sum_{i=1}^{K} \sum_{j=1}^{k_i} \alpha_{ij} = 1$ by the scaling convention.

Lemma 2. Suppose that $P(x|A)P(A|T) = P(x|T)$ for all x and A such that $x \in A \subset T$ where $P(A|T) = \Sigma_{y \in A} P(y|T)$, then $P(x|A_1)P(A_1|A_2) \ldots P(A_n|T) = P(x|T)$ for $x \in A_1$ and $A_i \subset A_{i+1}$ where $P(A_i|A_{i+1}) = \Sigma_{y \in A_i} P(y|A_{i+1})$.

Proof. If $P(x|A)P(A|T) = P(x|T)$ for all x and A such that $x \in A \subseteq T$, then, specifically, $P(y|A_2)P(A_2|T) = P(y|T)$ for all y such that $y \in A_1 \subseteq A_2$. Then $\Sigma_{y \in A_1} P(y|A_2)P(A_2|T) = \Sigma_{y \in A_1} P(y|T)$. Multiply both sides by $P(x|A_1)$ yields $P(x|A_1)P(A_1|A_2)P(A_2|T) = P(x|A_1)P(A_1|T) = P(x|T)$ where the last step uses the hypothesis of the lemma. Finally, we proceed by induction to the result.

Lemma 3. Suppose that $P(x|T) = P(x|A_1)P(A_1|A_2) \ldots P(A_{n-1}|A_n)$ for some sequence $A_1, \ldots A_n$, such that $A_n = T$, $A_i \subset A_{i+1}$ and the cardinality of A_i equals $i + 1$. Consider another sequence, $B_1, \ldots B_m$, such that $B_j = A_i$ and $B_{j+1} = A_{i+t}$ for some t. Then $P(x|T) = P(x|B_1)P(B_1|B_2) \ldots P(B_{m-1}|T)$.

Proof. See Tversky and Sattath (1979), Appendix E, pp. 572-3. An alternative proof can be constructed similar to that in Lemma 2.

Before proceeding to Lemma 4, we consider a partition of the choice set such that $T = A \cup B \cup \{x\}$. We define aspect sets from x and B's perspective. Here, x^u and B^u are the unique aspects of x and B, respectively; Bx^u are the aspects that are shared by x and B but not by any element of A; x^s are the aspects x shares with at least one object in A but not with B'; B^s are aspects B' shares with A' but not with x'; Bx^s are aspects B' and x' share with A'; T^c is the set of all common aspects. In set notation:

$$x^u = \{\alpha \mid \alpha \in x', \alpha \notin B', \alpha \notin A'\}$$

$$B^u = \{\beta \mid \beta \notin x', \beta \in B', \beta \notin A'\}$$

$$Bx^u = \{\mu \mid \mu \in x' \cap B', \mu \notin A'\}$$

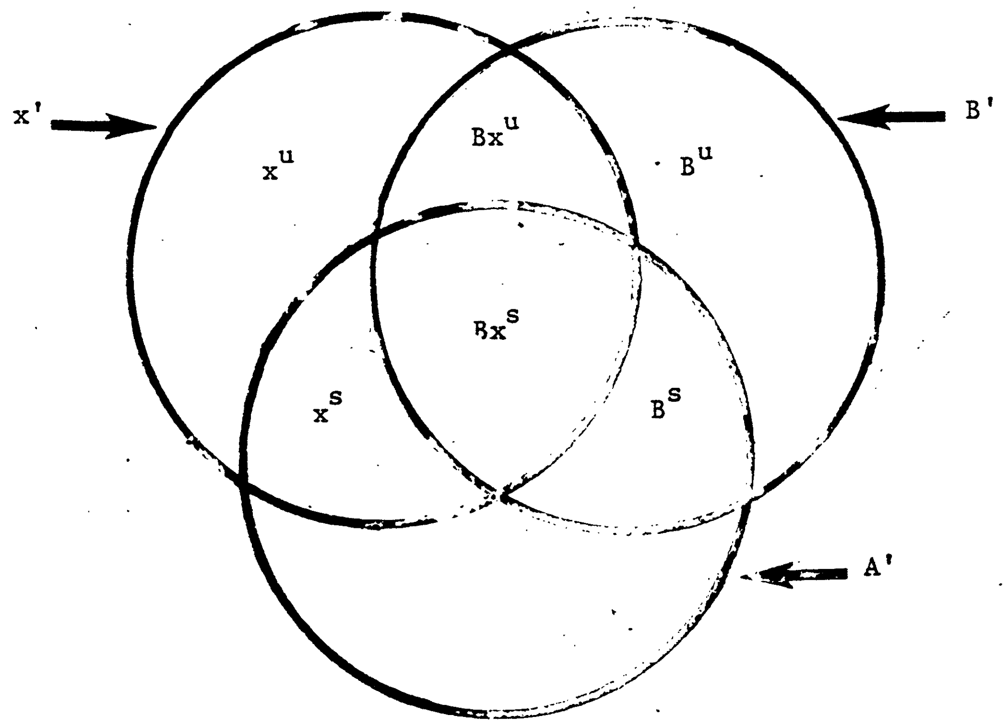
$$x^s = \{\gamma \mid \gamma \in x', \gamma \notin B' \; \gamma \in A'\}$$

$$B^s = \{\delta \mid \delta \notin x', \delta \in B', \delta \in A'\}$$

$$Bx^s = \{\lambda \mid \lambda \in x' \cap B' \cap A', \lambda \notin T^c\}$$

$$T^c = \{\eta \mid \eta \in z' \text{ for all } z \in T\}$$

These definitions can best be visualized by the following Venn diagram:



Lemma 4. Let x be a choice object and let A and B be sets of choice objects such that $T = A \cup B \cup \{x\}$. Let $B^+ = B \cup \{x\}$ and let A^* be a constrained agenda for elimination-by-aspects with hierarchy $\{\{x, B\}, A\}$. Then $P(x \mid A^*) \left\{\substack{\geq\\=\\\leq}\right\} P(x \mid T)$ iff the following condition, A2, holds:

$$\frac{\sum_{\alpha \in x^u} \alpha + \sum_{\gamma \in x^s} \gamma + \sum_{\lambda \in Bx^s} \lambda P(x \mid B_\lambda^+) + \sum_{\mu \in Bx^u} \mu P(x \mid B_\mu^+)}{\sum_{\beta \in B^u} \beta + \sum_{\delta \in B^s} \delta + \sum_{\lambda \in Bx^s} \left[\lambda \sum_{y \in B} P(y \mid B_\lambda^+)\right] + \sum_{\mu \in Bx^u} \left[\mu \sum_{y \in B} P(y \mid B_\mu^+)\right]} \left\{\substack{\geq\\=\\\leq}\right\}$$

$$\frac{\sum_{\alpha\varepsilon x^u}\alpha + \sum_{\gamma\varepsilon x^s}\gamma P(x|T_\gamma) + \sum_{\lambda\varepsilon Bx^s}\lambda P(x|T_\lambda) + \sum_{\mu\varepsilon Bx^u}\mu P(x|B_\mu^+)}{\sum_{\beta\varepsilon B^u}\beta + \sum_{\delta\varepsilon B^s}\left[\delta\sum_{y\varepsilon B}P(y|T_\delta)\right] + \sum_{\lambda\varepsilon Bx^s}\left[\lambda\sum_{y\varepsilon B}P(y|T_\lambda)\right] + \sum_{\mu\varepsilon Bx^u}\left[\mu\sum_{y\varepsilon B}P(y|B_\mu^+)\right]} \tag{A2}$$

Proof. By definition, $P(x|A^*) = P(x|B^+)P(B^+|T) = P(x|B^+) \cdot [P(x|T) + \sum_{y\varepsilon B} P(y|T)]$. Thus, by rearranging terms and recognizing $P(x|B^+) + \sum_{y\varepsilon B} P(y|B^+) = 1$, we get $P(x|A^*) \{\gtrless\} P(x|T)$ if and only if

$$\frac{P(x|B^+)}{\sum_{y\varepsilon B} P(y|B^+)} \{\gtrless\} \frac{P(x|T)}{\sum_{y\varepsilon B} P(y|T)}$$

Finally, applying EBA to each term and using the aspect set definitions for x^u, B^u, Bx^u, x^s, B^s, and Bx^s we obtain condition A2. Note that $T_\mu = B_\mu^+$ by definition and that $\mu\varepsilon Bx^u$ can affect the ratios in condition A2 because the selection of μ as the elimination aspect eliminates some alternatives in B but not all alternatives in B. Similarly for $\lambda\varepsilon Bx^s$.

If $B = \{y\}$, a singleton, then Bx^u and Bx^s will not affect the left side of condition A2 and Bx^u will not affect the right side of condition A2. The resulting condition simplifies to condition A3 for $A^* = \{\{x, y\}, A\}$ where we have written yx^s for Bx^s, y^u for B^u, and y^s for B^s:

$$\frac{\sum_{\alpha\varepsilon x^u}\alpha + \sum_{\gamma\varepsilon x^s}\gamma}{\sum_{\beta\varepsilon y^u}\beta + \sum_{\delta\varepsilon y^s}\delta} \{\gtrless\} \frac{\sum_{\alpha\varepsilon x^u}\alpha + \sum_{\gamma\varepsilon x^s}\gamma P(x|T_\gamma) + \sum_{\lambda\varepsilon yx^s}\lambda P(x|T_\lambda)}{\sum_{\beta\varepsilon y^u}\beta + \sum_{\delta\varepsilon y^s}\delta P(y|T_\delta) + \sum_{\lambda\varepsilon yx^s}\gamma P(y|T_\lambda)} \tag{A3}$$

Lemma 5. Let T and A^* be defined such that $A^* = \{B^+, A\}$. Equality holds in A2 for all $x\varepsilon B^+$ and $z\varepsilon A$ and for all possible values of non-zero aspect measures if and only if the aspect structure is (1) a preference tree or (2) a factorial structure compatible with the hierarchy associated with A^*.

Proof. (If preference tree equality holds.) If T' is compatible with {{x, B}, A} then $x^s = B^s = Bx^s = \emptyset$. Substituting these relationships reduces condition A2 to an identity.

(If FS equality holds.) By the definition of compatibility for FS's, there must exist some aspect (or aspect set) which is contained in all objects in B^+ but not in any objects in A. Wlog, let this aspect (or aspect set) be α_{11} and let $x' = \{\alpha_{11}, \alpha_{21}, \ldots, \alpha_{K1}\}$. For K = 1, $B = \emptyset$, $P(x|B^+) = 1$ and equality holds.

Suppose K = 2. Then $x^u = B^u = Bx^s = \emptyset$, $Bx^u = \{\alpha_{11}\}$ $x^s = \{\alpha_{21}\}$, and $B^s = \{\alpha_{22}, \alpha_{23}, \ldots \alpha_{2k_2}\}$. Since α_{11} is common to all $y \in B^+$, condition A2 reduces to

$$\frac{\alpha_{21}}{\sum_{n=2}^{k_2} \alpha_{2n}} = \frac{\alpha_{21} P(x|T_{\alpha_{21}})}{\sum_{n=2}^{k_2} \alpha_{2n} \sum_{y \in B} P(y|T_{\alpha_{2n}})}$$

Finally, $P(x|T_{\alpha_{21}}) = P(y|T_{\alpha_{2n}}) = \alpha_{11}/(\Sigma_n \alpha_{1n})$ for $\alpha_{2n} \in y'$ for K = 2. Thus, these terms cancel and the equality holds since $\alpha_{2n} \in y'$ for exactly one $y \in B$.

Suppose K > 2. Then $x^u = B^u = x^s = \emptyset$, $Bx^u = \alpha_{11}$, $Bx^s = \{\alpha_{21}, \alpha_{31}, \ldots, \alpha_{K1}\}$, and $B^s = \{\alpha_{22}, \alpha_{23}, \ldots, \alpha_{2k_2}, \alpha_{32}, \alpha_{33}, \ldots \alpha_{3k_3}, \ldots, \alpha_{Kk_K}\}$. Substituting these terms in condition A2 yields:

$$\frac{\sum_{\ell=1}^{K} \alpha_{\ell 1} P(x|B^+_{\alpha_{\ell 1}})}{\sum_{\ell=2}^{K} \sum_{n=2}^{k_\ell} \alpha_{\ell n} + \sum_{\ell=1}^{K} \left[\alpha_{\ell 1} \sum_{y \in B} P(y|B^+_{\alpha_{\ell 1}})\right]} = \frac{\sum_{\ell=1}^{K} \alpha_{\ell 1} P(x|T_{\alpha_{\ell 1}})}{\sum_{\ell=2}^{K} \sum_{n=2}^{k_\ell} \left[\alpha_{\ell n} \sum_{y \in B} P(y|T_{\alpha_{\ell n}})\right] + \sum_{\ell=1}^{K} \alpha_{\ell 1} \sum_{y \in B} P(y|T_{\alpha_{\ell 1}})}$$

The above equation will be satisfied if $P(x|T_{\alpha_{\ell 1}}) = RP(x|B^+_{\alpha_{\ell 1}})$ and $\sum_{y \in B} P(y|T_{\alpha_{\ell 1}}) = R \sum_{y \in B} P(y|B^+_{\alpha_{\ell 1}})$ for ℓ = 1 to K for all $y \in B$, and if $\sum_{y \in B} P(y|T_{\alpha_{\ell n}}) = R$ for ℓ = 2 to K and n = 2 to k_ℓ for all $y \in B$, where R is some non-zero constant.

Recognizing that $B^+_{\alpha_{\ell 1}}$ and $B^+_{\alpha_{\ell n}}$ are (K-2)-level FS's whereas $T_{\alpha_{\ell 1}}$ and $T_{\alpha_{\ell n}}$ are (K-1)-level FS's we can apply lemma 1 yielding:

$$P(x|B^+_{\alpha_{\ell 1}}) = \frac{\prod_{j\neq 1,\ell} \alpha_{j1}}{\prod_{j\neq 1,\ell} \sum_{n=1}^{k_j} \alpha_{jn}} \quad ; \quad P(x|T_{\alpha_{\ell 1}}) = \frac{\prod_{j\neq \ell} \alpha_{j1}}{\prod_{j\neq \ell} \sum_{n=1}^{k_j} \alpha_{jn}}$$

Hence, for $\ell \neq 1$, $P(x|T_{\alpha_{\ell 1}}) = (\alpha_{11} / \sum_{n=1}^{k_1} \alpha_{1n}) \cdot P(x|B^+_{\alpha_{\ell 1}})$ which satisfies the above condition with $R = (\alpha_{11} / \sum_{n=1}^{k_1} \alpha_{1n})$. For $\ell=1$, $B^+ = T_{\alpha_{11}}$ and α_{11} is common across x and B. Thus, α_{11} can be shown to cancel from the EBA formulae. See discussion in Tversky (1972). We show the other terms similarly. (Recognize that the selection of $\alpha_{\ell n}$, $n \neq 1$ conditions out x. Because all remaining aspects are shared with T - B, the set B will be chosen with probability R.) Thus, condition A2 holds for a compatible FS.

(Equality requires Pretree or FS). We rule out the trivial case where A is empty or $P(A|T) = 0$. Wlog, assume the aspect measures on T' sum to 1.0. Then the condition for equality in A2 can be written in the form:

$$\frac{a + c + g}{b + d + h} = \frac{a + ce + g}{b + df + h}$$

where $a = \sum_{\alpha \varepsilon x^u} \alpha$, $b = \sum_{\beta \varepsilon B^u} \beta$, $g = \Sigma_\mu \mu P(x|B^+_\mu)$, $h = \Sigma_\mu [\mu \Sigma_B P(y|B^+_\mu)$,

$c = \Sigma_\gamma \gamma + \Sigma_\lambda \lambda P(x|B^+_\lambda)$, $ce = \Sigma_\gamma \gamma P(x|T_\gamma) + \Sigma_\lambda \lambda P(x|T_\lambda)$, $e = ce/c$,

and d, f are defined accordingly. Note that a, b, c, d, g, $h \varepsilon [0, 1]$ by the scaling assumption. e, $f \varepsilon [0, 1]$ since $P(y|B^+_\xi) \geq P(y|T_\xi)$ for all ξ because $B^+ \subset T$ and EBA satisfies regularity. The above relationship is equivalent to

$$(a+g)d(f-1) + (b+h)c(1-e) + cd(f-e) = 0.$$

Since the aspect measures can be chosen arbitrarily on the interval, [0, 1], subject to scaling restrictions and since the above equation must hold for any choice of aspect measures, equality cannot depend on a specific relationship between non-zero a, b, c, and d. Thus, the above relationship will only be satisfied by an aspect structure which implies either (1) $c = d = 0$, (2) $f = e = 1$, (3) $c = 0$, $f = 1$, (4) $d = 0$, $e = 1$, or (5) $a = b = 0$, $f = e \neq 1$, and (i) $g/h = c/d$ or (ii) $g = h = 0$. (Note that the cases (5)-(iii) $g = h = 0$, $a/b = c/d$ and (5)-(iv) $(a+g)/(b+h) = c/d$ for non-zero a, b, g, and h would require special relationships among arbitrarily chosen aspects and could not be satisfied by structure alone.)

Case (1) implies $x^s = B^s = Bx^s = \emptyset$ for all $x \in T$ and associated B. This implies a compatible preference tree.

Case (2) implies that $P(x|T_\gamma) = 1$ for $\gamma \varepsilon x^s$, $\Sigma_B P(y|T_\delta) = 1$ for $\delta \varepsilon B^s$, and $P(y|T_\lambda) = P(y|B_\lambda^+)$ for all $\lambda \varepsilon Bx^s$ and $y \varepsilon B^+$. But this implies $P(A|T_\gamma) = 0$, $P(A|T_\delta) = 0$, and $P(A|T_\lambda) = 0$ and at least one of x^s, B^s, or Bx^s is non-empty (else $f = e = 0$). Thus, all elements in A that share any common objects with B^+ are dominated by some $y \varepsilon B^+$. Finally; we rule out the case where objects in all A are identical to at least one object in B^+ by the $P(A|T_\xi) = 0$ conditions. Thus $P(z|T) = 0$ for objects, $z \varepsilon A$, that share some common aspects with objects in B^+. All other objects in A have aspects sets which are disjoint from x' and B'. Thus case (2) is a preference tree.

Case (3) implies $x^s = Bx^s = \emptyset$ and $\Sigma_B P(y|T_\delta) = 1$ for $\delta \varepsilon B^s$. Case (4) implies $B^s = Bx^s = \emptyset$ and $P(x|T_\gamma) = 1$ for $\gamma \varepsilon x^s$. These are special cases of case (2).

Case (5), $f = e = 1$. Let $f = e = R$. This condition must hold for arbitrary selection of aspect measures. By successively varying $\gamma \varepsilon x^s$ we can show that it must be true that $P(x|T_\gamma) = RP(x|B_\gamma^+) = R$ and $P(x|T_\lambda) = RP(x|B_\lambda^+)$ for all $\gamma \varepsilon x^s$ and $\lambda \varepsilon Bx^s$. (Note $Bx^s = \emptyset$ for a 2x2 factorial and $x^s = \emptyset$ for a 2^n factorial where $n > 2$.) By similar arguments, $\sum_{y \varepsilon B} P(y|T_\delta) = R$ and $\sum_{y \varepsilon B} P(y|T_\lambda) = R \sum_{y \varepsilon B} P(y|B_\lambda^+)$ for $\delta \varepsilon B^s$ and $\lambda \varepsilon Bx^s$. By hypothesis, equality holds in (A2), hence we can write $R = P(B_\xi^+|T)$ for $\xi \varepsilon x^s \cup B^s \cup Bx^s$. Finally, since $a=b=0$, $x^u = B^u = \emptyset$ and $x^s \cup B^s \cup Bx^s \cup Bx^u = (B+)'$.

Consider subcase (i), g/h = c/d. By definition, $(a+c+g)/(b+d+h) \equiv P(x|B^+)/\sum_{y\in B} P(y|B^+)$. By g/h = c/d and a=b=0, $(a+c+g)/(b+d+h) = g/h$, hence

$$\frac{g}{h} = \frac{\sum_{\mu\in Bx^u} \mu P(x|B_\mu^+)}{\sum_{\mu\in Bx^u} \mu \sum_{y\in B} P(y|B_\mu^+)} = \frac{P(x|B^+)}{\sum_{y\in B} P(y|B^+)}$$

By hypothesis this condition must hold for arbitrary choice of the measures of μ, hence it must be true that $P(x|B_\mu^+)/\sum_{y\in B} P(y|B_\mu^+) = P(x|B^+)/\sum_{y\in B} P(y|B^+)$ for all $\mu\in Bx^u$. Since μ does not affect these probability ratios, by the properties of EBA, it must be the case that $\mu\in y'$ for all $y\in B^+$. Since $\mu\in A'$ by definition, B^+ must differ from A by $\underline{\mu}\equiv Bx^u$. Since these conditions must hold for all $y\in B^+$ and for all $z\in A$, there must exist a complementary $\bar{\underline{\mu}}\in A'$.

Putting together the condition that $R=P(B_\xi^+|T)$ for all $\xi\in x^s \cup Bx^s$ and that $\mu\in x$, we can write x' as $\{\underline{\mu}, \xi_1, \xi_2, \ldots\}$. We then limit successively on $\xi_i \in x^s \cup Bx^s$ until only $\underline{\mu}$ is left to consider. It must then be the case that there exists $A'_x \subseteq A'$ such that $A'_x = \{\bar{\underline{\mu}}, \xi_1, \xi_2, \ldots\}$. Similarly for all $y\in B$ there must exist a matching A'_y in A'. Since the conditions must hold for all x, y, $z\in T$ we can find a factorial match for every choice object. Thus, the factorial is complete and compatible and not fractional.

Finally, subcase (ii), g=h=0, is a degenerate case where $Bx^u=\emptyset$ and the "factorial structure" splits on identical aspects.

Thus, the only cases where condition A2 can hold for all $x\in T$ is if (1) T' is a preference tree and A^* is compatible or (2) T' is factorial structure and A^* is compatible. This completes the proof of lemma 5.

Theorem 1 (Invariance): The constant ratio model is the only decision rule invariant with respect to top down agendas. On the other hand, each of our decision rules, CRM, EBA, and HEM(Θ), can be affected by bottom up agendas.

Proof. Both parts have been proven in the text. The first part is true by the definition of CRM, $P(x|A)P(A|T) = P(x|T)$ for all A. Counterexamples have been supplied for the second part.

Theorem 2 (Compatibility): For an arbitrarily chosen set of aspect measures, a constrained top down agenda, $\mathcal{A}^*$, has no effect on a family of EBA choice probabilities, $P(x|T)$ for all $x \in T$, if and only if either (1) the aspect structure forms a preference tree and $\mathcal{A}^*$ is compatible with the tree or (2) the aspect structure forms a factorial structure and $\mathcal{A}^*$ is compatible with the factorial structure.

__Proof.__ By lemmas 4 and 5, $P(x| \mathcal{A}^*) = P(x|T)$ iff the aspect structure is a pre-tree or FS and compatible with $\mathcal{A}^* = \{\{x, B\}, A\}^*$. Lemmas 2 and 3 extend this result to arbitrary compatible agendas. If a single-level agenda affects a family of choice probabilities, then a multi-level agenda must also since we cannot quarantee any cancellation of effects except by fortuitous choice of the aspect measures.

Theorem 3 (Equivalence). For an arbitrarily chosen set of aspect measures, the hierarchical elimination model (HEM) and elimination by aspects (EBA) yield equivalent choice probabilities if and only if (1) the aspect structure is a preference tree or (2) the aspect structure is a factorial structure, $\theta = 0$, and the hierarchy associated with HEM is compatible with the preference tree or factorial structure.

__Proof.__ (Compatible pretree implies equivalence.) Consider a hierarchy compatible with a pretree. Then by lemma 5, equality holds in condition A2. Hence, by lemma 4, we can write EBA as a hierarchical rule, i.e., $P(x|B^+)\ P(B^+|T) = P(x|T)$. HEM(0) is defined such that $P(x|B^+)\ P(B^+|T) = P(x|T)$. Lemmas 2 and 3 assure that this can be extended to multiple levels. Thus we need only show that $P_h(x|B^+) = P_e(x|B^+)$ and $P_h(B^+|T) = P_e(B^+|T)$ where $P_h(x|B^+)$ is computed by HEM(0) and $P_e(x|B^+)$ is computed by EBA. Define $P_h(B^+|T)$, $P_e(B^+|T)$ accordingly.

Applying EBA yields:

$$P_e(x|B^+) = \left[\sum_{\alpha\varepsilon x} u^\alpha + \sum_{\gamma\varepsilon x^s} \gamma + \sum_{\mu\varepsilon Bx^u} u^\mu P(x|B_\mu^+) + \sum_{\lambda\varepsilon Bx^s} \lambda P(x|B_\lambda^+)\right] \Big/ \sum_{\sigma\varepsilon B^{+\prime}} \sigma$$

Now on a compatible pretree, $Bx^s = x^s = B^s = \emptyset$ and for all $\mu\varepsilon Bx^u$, $\mu\varepsilon y'$ for all $y\varepsilon B^+$. Thus $\mu\varepsilon Bx^u$ does not affect $P_e(x|B^+)$, hence $P_e(x|T^+)$ reduces to

$$P_e(x|B^+) = \left(\sum_{\alpha\varepsilon x} u^\alpha\right) \Big/ \left[\left(\sum_{\alpha\varepsilon x} u^\alpha\right) + \left(\sum_{\beta\varepsilon B} u^\beta\right)\right]$$

Finally, according to the definition of HEM(0), see equations 5, 6, 9, 10, and 11, we have $m(x) = \sum_{\alpha\varepsilon x} u^\alpha$, $m(B) = \sum_{\beta\varepsilon B} u^\beta$, and

$$P_h(x|B^+) = \left(\sum_{\alpha\varepsilon x} u^\alpha\right) \Big/ \left[\left(\sum_{\alpha\varepsilon x} u^\alpha\right) + \left(\sum_{\beta\varepsilon B} u^\beta\right)\right]$$

Thus, $P_e(x|B^+) = P_h(x|B^+)$. Finally,

$$P_e(B^+|T) = \sum_{y\varepsilon B^+} P_e(y|T) = \left[\sum_{y\varepsilon B^+} \sum_{\alpha\varepsilon y} u^\alpha + \sum_{\mu\varepsilon Bx^u} u^\mu\right] \Big/ \sum_{\sigma\varepsilon T'} \sigma =$$

$m(B^+)/[m(B^+) + m(T-B^+)] = P_h(B^+|T)$ paralleling the arguments used to show $P_e(x|B^+) = P_h(x|B^+)$. For an alternative proof see Tversky and Sattath, 1979, Appendix B, pp. 568-570).

(Compatible FS implies equivalence.) At any level ℓ a compatible FS splits such that $\alpha_{\ell j}\varepsilon(B^+)'$ and $\{\alpha_{\ell n}|n \neq j\} \varepsilon (T-B^+)'$ and $\{\alpha_{mn}|m \neq \ell, \alpha_{mn}\varepsilon(B^+)'\}$ is contained in both $(B^+)'$ and $(T-B^+)'$. Thus, for HEM(0):

$$P_h(B^+|T) = \alpha_{\ell j} \Big/ \sum_{n=1}^{k_\ell} \alpha_{\ell n}$$

Applying the above equation iteratively yields

$$P(x|T) = \prod_{\alpha_{ij}\varepsilon x'} \alpha_{ij} \Big/ \prod_{\ell=1}^{K} \left(\sum_{n=1}^{k_\ell} \alpha_{\ell n}\right)$$

which is equivalent to $P_e(x|T)$ as shown by equation (A1) in lemma 1.

(Equivalence implies compatible pretree or FS.) By lemma 5, $P(x|B^+)\, P(B^+|T) = P(x|T)$ only if $A^* = \{\{x, B\}, A\}$ is a compatible pretree or FS. Thus, EBA will not become a hierarchical rule unless the aspect structure is a pretree or FS. Thus, except for fortuitous choices of aspect measures, HEM and EBA will not be equivalant unless the aspect structure is a compatible pretree or FS.

Result 1.1 (Dissimilar grouping on a 2^k factorial structure): Let $T = B \cup \{x, w\}$ be a factorial structure. Suppose for every $\alpha_{i1} \in x' - x' \cap w'$, $\alpha_{i1} > \alpha_{i2}$ where α_{i2} is the aspect matched to α_1 and $x' - x' \cap w'$ contains at least two elements. Then, the constrained agenda, $A^* = \{\{x, w\}, B\}$, is an effective EBA agenda for w and a counterproductive EBA agenda for x. I.e., $P(w|A^*) > P(w|T)$ and $P(x|T) > P(x|A^*)$.

__Proof__. According to EBA $P(x|A^*) = P(x|\{x, w\})[P(x|T) + P(w|T)]$. Thus, $P(x|A^*) < P(x|T)$ iff $P(x|\{x, w\})/P(w|\{x, w\}) < P(x|T)/P(w|T)$. [Use $P(w|\{x, w\}) = 1 - P(x|\{x, w\})$.] Let $I = \{i|\alpha_{i1} \in x' - x' \cap w'\}$ and let $J = \{j|\alpha_{j2} \in w' - w' \cap x'\}$. Note that I = J since $\alpha_{i1} \in x'$ is matched to $\alpha_{j2} \in w'$. Thus,

$$\frac{P(x|\{x, w\})}{P(w|\{x, w\})} = \frac{\sum_{i \in I} \alpha_{i1}}{\sum_{i \in I} \alpha_{i2}}.$$

Now, by Lemma 1,

$$\frac{P(x|T)}{P(w|T)} = \frac{\prod_{i \in I} \alpha_{i1} \cdot \prod_{x', k \notin I} \alpha_{kj}}{\prod_{i \in I} \alpha_{i2} \cdot \prod_{w', k \notin I} \alpha_{kj}}.$$

The second terms in the numerator and denominator cancel since $\{\alpha_{kj}|\alpha_{kj} \in x', k \notin I\} = \{\alpha_{kj}|\alpha_{kj} \in w', k \notin I\}$ by the hypotheses of the theorem. Thus we must show that

$$\frac{\sum_i \alpha_{i1}}{\sum_i \alpha_{i2}} < \frac{\prod_\ell \alpha_{\ell 1}}{\prod_\ell \alpha_{\ell 2}}$$

where $\alpha_{i1} > \alpha_{i2} > 0$. (All sums and products are over the set, I.) Because all terms are positive, this condition reduces to $\Sigma_i(\alpha_{i2}\Pi_\ell \alpha_{\ell 1} - \alpha_{i1}\Pi_\ell \alpha_{\ell 2}) > 0$. Rearranging terms yields $\Sigma_i \alpha_{i1}\alpha_{i2}(\Pi_{\ell \neq i}\alpha_{\ell 1} - \Pi_{\ell \neq i}\alpha_{\ell 2}) > 0$. This condition holds whenever $\{\ell|\ell \in I, \ell \neq i\} \neq \emptyset$ which is true whenever I contains at least two elements. Since the condition holds, we have shown $P(x|A^*) < P(x|T)$. The proof for $P(w|A^*) > P(w|T)$ is symmetric.

Result 1 (Dissimilar Grouping): For the factorial structure in figure 6, the top down agenda, A^*, = $\{\{x, w\}, \{y, v\}\}^*$, enhances the EBA probability that the least preferred object, w, is chosen and hurts the EBA probability that the most preferred object, x, is chosen. That is, $P(w|A^*) > P(w|T)$ and $P(x|T) > P(x|A^*)$.

Proof: Result 1 is a special case of Result 1.1 with k = 2.

Result 2 (Bottom-Up Agendas): For compatible bottom-up agendas, A_* and B_* on a 2 x 2 factorial structure where $P(x|T) > P(y|T) > P(v|T) > P(w|T)$,

doing the easy comparison first (α_1 vs. α_2) enhances objects with already higher probability and doing the difficult comparison first (β_1 vs. β_2) enhances objects with lower probabilities. That is,

(1) $P(x|A_*) > P(x|T) > p(x|B_*)$

(2) $P(y|A_*) > P(y|T) > P(y|B_*)$

(3) $P(v|B_*) > P(v|T) > P(v|A_*)$

(4) $P(w|B_*) > P(w|T) > P(w|A_*)$

Proof. For $x' = \{\alpha_1, \beta_1\}$, $y' = \{\alpha_1, \beta_2\}$, $v' = \{\alpha_2, \beta_1\}$, $w' = \{\alpha_2, \beta_2\}$, the conditions of the result translate to $\alpha_1 > \alpha_2$, $\beta_1 > \beta_2$, and $\alpha_1/(\alpha_1+\alpha_2) > \beta_1/(\beta_1+\beta_2)$. According to the definition of bottom-up agendas,

$$P(x|B_*) = P(x|xy)[P(x|xv)P(v|vw) + P(x|xw)P(w|vw)]$$

$$P(x|A_*) = P(x|xv)[P(x|xy)P(y|yw) + P(x|xw)P(w|yw)]$$

where we have used the shorthand notation $P(x|xy)$ for $P(x|\{x, y\})$, etc. Introduce the notation, $p = P(x|xy) = P(v|vw) = 1 - P(w|vw) = \beta_1/(\beta_1 + \beta_2)$; $q = P(x|xv) = P(y|yw) = 1 - P(w|yw) = \alpha_1/(\alpha_1+\alpha_2)$; and $r = P(x|xw) = (\alpha_1+\beta_1)/(\alpha_1+\beta_1+\alpha_2+\beta_2)$. By the conditions of the theorem, $q > p$. Furthermore, it is easy to show $q > r > p$.

Using our notation, $P(x|A_*) = q[pq + r(1-q)]$, $P(x|B_*) = p[pq + r(1-p)]$, and, by Lemma 1, $P(x|T) = pq$. Rearranging terms yields $P(x|A_*) = pq + q(1-q)(r-p)$ and $P(x|B_*) = pq + p(1-p)(r-q)$. Thus, since $(r-p) > 0$, $P(x|A_*) > pq = P(x|T)$ and since $(r-q) < 0$, $P(x|B_*) < pq = P(x|T)$. This completes the proof of part (1). We show the other conditions similarly. For example, $P(y|A_*) = q(1-p) + q(1-q)[t - (1-p)] > q(1-p) = P(y|T)$ where $t = (\alpha_1+\beta_2)/(\alpha_1+\alpha_2+\beta_1+\beta_2)$.

Result 2.1 (Entropy): According to the conditions of Result 2, performing the the easy comparisons first decreases entropy and performing the difficult comparisons first increases entropy. That is: $H(A_*) < H(T) < H(B_*)$.

Proof. It is sufficient to show that $(\partial H/\partial A) < 0$ where A is some function with the properties $\partial p_x/\partial A > 0$, $\partial p_y/\partial A > 0$, $\partial p_v/\partial A < 0$, $\partial p_w/\partial A < 0$, and $p_x > p_y > p_v > p_w$, where $p_x = P(x|A_*)$, etc. Using this notation, $H(A) = -p_x \ln p_x - p_y \ln p_y - p_v \ln p_v - p_w \ln p_w$. Using the chain rule for differentiation yields

$$\frac{\partial H}{\partial A} = -\ln p_x \frac{\partial p_x}{\partial A} - \ln p_y \frac{\partial p_y}{\partial A} - \ln p_v \frac{\partial p_v}{\partial A} - \ln p_w \frac{\partial p_w}{\partial A}$$

Where we have used $\partial (p_x + p_y + p_v + p_w)/\partial A = 0$.

Recognizing $-\partial p_v/\partial A = \partial (p_x + p_y + p_w)/\partial A$ and substituting yields:

$$\frac{\partial H}{\partial A} = -\ln\left(\frac{p_x}{p_v}\right)\frac{\partial p_x}{\partial A} - \ln\left(\frac{p_y}{p_v}\right)\frac{\partial p_y}{\partial A} - \ln\left(\frac{p_w}{p_v}\right)\frac{\partial p_w}{\partial A}$$

Finally, by inspection we see that all terms are negative.

Result 2.2 (Bottom-up Dissimilar Agendas): For the dissimilar grouping agenda, C_*, on a 2 x 2 factorial structure with $P(x|T) > P(y|T) > p(v|T) > P(w|T)$, and $\alpha_1/(\alpha_1+\alpha_2) > \beta_1/(\beta_1+\beta_2)$,

(a) $P(x|C_*) > P(x|T)$ iff $\alpha_1\alpha_2 > \beta_1\beta_2$

(b) $P(x|A_*) > P(x|C_*) > P(x|B_*)$.

Proof. We continue with the notation of result 2. Let $s = P(y|vy) = (\alpha_1+\beta_2)/(\alpha_1+\beta_1+\alpha_2+\beta_2) = \alpha_1+\beta_2$ where wlog $\alpha_1+\beta_1+\alpha_2+\beta_2 = 1$. Then according to the definition of bottom-up agendas,

$$P(x|\mathcal{C}_*) = P(x|xw)[(P(x|xy)P(y|yv) + P(x|xv)P(v|yv)]$$

$$\text{or } P(x|\mathcal{C}_*) = r[ps + q(1-s)]$$

Thus, $P(x|\mathcal{C}_*) > P(x|T)$ if

$$(\alpha_1+\beta_1)\left[(\alpha_1+\beta_2)\left(\frac{\beta_1}{\beta_1+\beta_2}\right)+(\alpha_2+\beta_1)\left(\frac{\alpha_1}{\alpha_1+\alpha_2}\right)\right] > \left[\frac{\alpha_1}{\alpha_1+\alpha_2}\right]\cdot\left[\frac{\beta_1}{\beta_1+\beta_2}\right]$$

which after much algebra reduces to,

$$\alpha_1\alpha_2(\alpha_1\beta_2 - \alpha_2\beta_1) > \beta_1\beta_2(\alpha_1\beta_2 - \alpha_2\beta_1)$$

Finally, since $\alpha_1/(\alpha_1+\alpha_2) > \beta_1/(\beta_1+\beta_2)$, we have $\alpha_1\beta_2 - \alpha_2\beta_1 > 0$, hence under the conditions of the theorem, $P(x|\mathcal{C}_*) > P(x|T)$ iff $\alpha_1\alpha_2 > \beta_1\beta_2$. Note that if $\beta_1/(\beta_1+\beta_2) > a_1/(\alpha_1+\alpha_2)$, the appropriate condition is $\beta_1\beta_2 > \alpha_1\alpha_2$.

(Part b.) To show $P(x|\mathcal{A}_*) > P(x|\mathcal{C}_*)$ we must show $q[pq + r(1-q)] > r[ps + q(1-s)]$ where p, q, r, and s are defined above. After much algebra, this condition reduces to $\alpha_1\beta_2 > \alpha_2\beta_1$ which is true since $\alpha_1/(\alpha_1+\alpha_2) > \beta_1/(\beta_1+\beta_2)$. We show $P(x|\mathcal{C}_*) > P(x|\mathcal{B}_*)$ by symmetry.

Result 3 (Shared objects): For a 2 x 2 factorial structure with the first comparison made with respect to α_1 and α_2, and for $\alpha_1 > \alpha_2$, shared aspects in hierarchical processing, i.e., HEM(Θ), enhance those objects which contain α_2 and hurt those objects which contain α_1. The effect increases as the importance of the shared objects increases, i.e., as Θ increases.

Proof. By definition

$$P(x|T) = \alpha_1\left(\frac{\beta_1}{\beta_1+\beta_2}\right)+\beta_1\cdot\left(\frac{\alpha_1}{\alpha_1+\alpha_2}\right)=\frac{\alpha_1\beta_1}{(\alpha_1+\alpha_2)(\beta_1+\beta_2)}$$

$$P(x|\mathcal{A}^*) = \left(\frac{\beta_1}{\beta_1+\beta_2}\right)\cdot\left(\frac{\alpha_1+\Theta(\beta_1+\beta_2)}{\alpha_1+\alpha_2+2\Theta(\beta_1+\beta_2)}\right)$$

First we recognize that $P(x|T) = P(x|A_*)$ for $\Theta = 0$. Next, taking derivatives of $P(x|A_*)$ yields:

$$\frac{\partial P}{\partial \Theta} = \frac{\beta_1}{\beta} \cdot \left[\frac{(\alpha_1+\alpha_2+2\Theta\beta)\ \beta - 2\beta(\alpha_1+\Theta\beta)}{(\alpha_1+\alpha_2+2\Theta\beta)^2} \right]$$

where $\beta = (\beta_1+\beta_2)$. After some algebra, this condition reduces to

$$\frac{\partial P}{\partial \Theta} = (\text{constant}) \cdot (\alpha_2-\alpha_1) \tag{A4}$$

where the constant is positive. Thus for $\alpha_2 < \alpha_1$, $\frac{\partial P}{\partial \Theta} < 0$, hence $P(x|A^*, \Theta>0) < P(x|A^*, \Theta = 0) = P(x|T)$.

If we reverse β_1 and β_2 we see the same result holds for $y' = \{\alpha_1, \beta_2\}$. If we reverse α_1 and α_2 we have the results for $v' = \{\alpha_2, \beta_1\}$ and $w' = \{\alpha_2, \beta_2\}$ because the derivative is now positive. Because A4 is negative for $\Theta > 0$, we have shown also the last statement that the effect increases as Θ increases.