PROOFS OF THEOREMS AND OTHER RESULTS

Appendix to "Agendas and Consumer Choice"

January 1986

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 $x \in T$, elimination by aspects yields:

$$P(x|T) = \prod_{\alpha_{ij} \in x'} \alpha_{ij} \Big/ \prod_{\ell=1}^{K} \left(\sum_{n=1}^{k_\ell} \alpha_{\ell n} \right) \qquad \text{(A1)}$$

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 $x \varepsilon A \subset T$ where $P(A|T) = \Sigma_{y \varepsilon A} P(y|T)$, then $P(x|A_1)P(A_1|A_2) \ldots P(A_n|T) = P(x|T)$ for $x \varepsilon A_1$ and $A_i \subset A_{i+1}$ where $P(A_i|A_{i+1}) = \Sigma_{y \varepsilon A_i} P(y|A_{i+1})$.

Proof. If $P(x|A)P(A|T) = P(x|T)$ for all x and A such that $x \varepsilon A \subseteq T$, then, specifically, $P(y|A_2)P(A_2|T) = P(y|T)$ for all y such that $y \varepsilon A_1 \subseteq A_2$. Then $\Sigma_{y \varepsilon A_1} P(y|A_2)P(A_2|T) = \Sigma_{y \varepsilon A_1} P(y|T)$. 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:

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

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

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

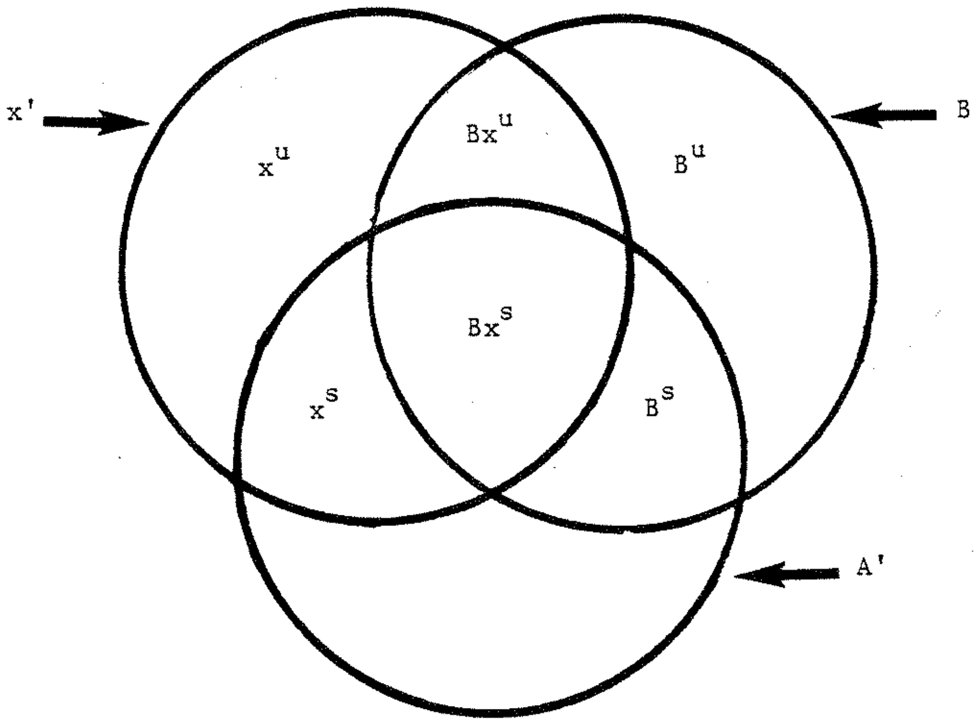
$$x^s = \{\gamma \mid \gamma \in x',\ \gamma \notin B'\ \gamma \in A'\}$$

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

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

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

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 $P(x \mid A^*) \{\gtreqless\} P(x \mid T)$ iff the following condition, A2, holds:

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

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

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

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

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 $\mu\epsilon Bx^u$ 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 $\lambda\epsilon Bx^s$.

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:

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

Lemma 5. Let T and A^* be defined such that $A^* = \{B^+, A\}$. Equality holds in A2 for all $x\epsilon B^+$ and $z\epsilon A$ 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 $y \varepsilon B^+$, condition A2 reduces to

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

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

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:

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

The above equation will be satisfied if $P(x|T_{\alpha_{\ell 1}}) = RP(x|B^+_{\alpha_{\ell 1}})$ and $\sum_{y \varepsilon B} P(y|T_{\alpha_{\ell 1}}) = R \sum_{y \varepsilon B} P(y|B^+_{\alpha_{\ell 1}})$ for ℓ = 1 to K for all $y \varepsilon B$, and if $\sum_{y \varepsilon B} P(y|T_{\alpha_{\ell n}}) = R$ for ℓ = 2 to K and n = 2 to k_ℓ for all $y \varepsilon B$, 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 $x \epsilon T$ and associated B. This implies a compatible preference tree.

Case (2) implies that $P(x|T_\gamma) = 1$ for $\gamma \epsilon x^s$, $\Sigma_B P(y|T_\delta) = 1$ for $\delta \epsilon B^s$, and $P(y|T_\lambda) = P(y|B_\lambda^+)$ for all $\lambda \epsilon Bx^s$ and $y \epsilon B^+$. 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 $y \epsilon B^+$. 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, $z \epsilon A$, 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 $\delta \epsilon B^s$. Case (4) implies $B^s = Bx^s = \emptyset$ and $P(x|T_\gamma) = 1$ for $\gamma \epsilon x^s$. 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 $\gamma \epsilon x^s$ 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 $\gamma \epsilon x^s$ and $\lambda \epsilon Bx^s$. (Note $Bx^s = \emptyset$ for a 2x2 factorial and $x^s = \emptyset$ for a 2^n factorial where $n > 2$.) By similar arguments, $\sum_{y \epsilon B} P(y|T_\delta) = R$ and $\sum_{y \epsilon B} P(y|T_\lambda) = R \sum_{y \epsilon B} P(y|B_\lambda^+)$ for $\delta \epsilon B^s$ and $\lambda \epsilon Bx^s$. By hypothesis, equality holds in (A2), hence we can write $R = P(B_\xi^+|T)$ for $\xi \epsilon x^s \cup B^s \cup Bx^s$. 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

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

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

Putting together the condition that $R=P(B^+_\xi|T)$ for all $\xi\varepsilon\ x^s \cup Bx^s$ and that $\mu\varepsilon x$, we can write x' as $\{\underline{\mu}, \xi_1, \xi_2, \ldots\}$. We then limit successively on $\xi_i \varepsilon x^s \cup Bx^s$ 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 $y\varepsilon B$ there must exist a matching A'_y in A'. Since the conditions must hold for all x, y, $z\varepsilon T$ 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 $x\varepsilon T$ 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, 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 A^* is compatible with the tree or (2) the aspect structure forms a factorial structure and A^* is compatible with the factorial structure.

Proof. By lemmas 4 and 5, $P(x|A^*) = P(x|T)$ iff the aspect structure is a pretree or FS and compatible with $A^* = \{\{x, B\}, A\}^*$. 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:

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

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

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

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

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

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

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

$m(B^+)/[m(B^+) + m(T-B^+)] = P_h(B^+|T)$ 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):

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

Applying the above equation iteratively yields

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

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

<u>Proof</u>. 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, $P(x|\mathcal{A}_*) = q[pq + r(1-q)]$, $P(x|\mathcal{B}_*) = p[pq + r(1-p)]$, and, by Lemma 1, $P(x|T) = pq$. Rearranging terms yields $P(x|\mathcal{A}_*) = pq + q(1-q)(r-p)$ and $P(x|\mathcal{B}_*) = pq + p(1-p)(r-q)$. Thus, since $(r-p) > 0$, $P(x|\mathcal{A}_*) > pq = P(x|T)$ and since $(r-q) < 0$, $P(x|\mathcal{B}_*) < pq = P(x|T)$. This completes the proof of part (1). We show the other conditions similarly. For example, $P(y|\mathcal{A}_*) = q(1-p) + q(1-q)[t - (1-p)] > q(1-p) = P(y|T)$ 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: $H(\mathcal{A}_*) < H(T) < H(\mathcal{B}_*)$.

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

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

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

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

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

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

Result 2.2 (Bottom-up Dissimilar Agendas): For the dissimilar grouping agenda, $\mathcal{C}_*$, 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) $P(x|\mathcal{C}_*) > P(x|T)$ iff $\alpha_1\alpha_2 > \beta_1\beta_2$

(b) $P(x|\mathcal{A}_*) > P(x|\mathcal{C}_*) > P(x|\mathcal{B}_*)$.

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,

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

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

Thus, $P(x|\mathfrak{C}_*) > P(x|T)$ 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, $P(x|\mathfrak{C}_*) > P(x|T)$ 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 $P(x|\mathfrak{A}_*) > P(x|\mathfrak{C}_*)$ 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 $P(x|\mathfrak{C}_*) > P(x|\mathfrak{B}_*)$ 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

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

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

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

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

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.