Psychological Review
2001, Vol. 108, No. 2, 370–392

0033-295X/01/$5.00 DOI: 10.1037//0033-295X.108.2.370

Multialternative Decision Field Theory: A Dynamic Connectionist Model of Decision Making

Robert M. Roe, Jerome R. Busemeyer, and James T. Townsend
Indiana University Bloomington

The authors interpret decision field theory (J. R. Busemeyer & J. T. Townsend, 1993) as a connectionist network and extend it to accommodate multialternative preferential choice situations. This article shows that the classic weighted additive utility model (see R. L. Keeney & H. Raiffa, 1976) and the classic Thurstone preferential choice model (see L. L. Thurstone, 1959) are special cases of this new multialternative decision field theory (MDFT), which also can emulate the search process of the popular elimination by aspects (EBA) model (see A. Tversky, 1969). The new theory is unique in its ability to explain several central empirical results found in the multialternative preference literature with a common set of principles. These empirical results include the similarity effect, the attraction effect, and the compromise effect, and the complex interactions among these three effects. The dynamic nature of the model also implies strong testable predictions concerning the moderating effect of time pressure on these three effects.

Preferential choice is a complex topic that requires examination from many different perspectives. Take, for example, the relatively simple task of buying a new car. From one point of view, this is a search problem in which a very large set of options is winnowed down to a much smaller set of satisfactory options (Simon, 1955; Tversky, 1972). From another point of view, this is an evaluation problem requiring trade-offs among multiple conflicting attributes such as safety, quality, performance, and cost (Keeney & Raiffa, 1976; Von Winterfeldt & Edwards, 1986). From a third point of view, this is a discrimination problem in which the strengths of competing candidates probabilistically compete for ultimate selection (De Soete, Feger, & Klauer, 1989; Thurstone, 1959).

This article presents a general decision theory that encompasses all of these points of view within a single theoretical framework. The present theory is an elaboration of an earlier theory known as decision field theory (Busemeyer & Townsend, 1993), which is based on the idea that information is sequentially sampled and accumulated over time to make a decision. This basic idea forms the foundation of a wide range of cognitive decision models, including sensory detection (Smith, 1995), perceptual discrimination (Link, 1992), memory recognition (Ratcliff, 1978), conceptual categorization (Ashby, 2000; Nosofsky & Palmeri, 1997), and multiattribute decision making (Aschenbrenner, Albert, & Schmalhofer, 1984). The present formulation also relates to previous artificial neural network models of decision processes (Grossberg & Gutowski, 1987; Leven & Levine, 1996; Usher & Zakay, 1993). In particular, the principle of *lateral inhibition* (Grossberg, 1982; McClelland & Rumelhart, 1981) is incorporated into the choice process, and this principle performs a crucial part in explaining paradoxical findings from the preferential choice literature.

Decision field theory was originally developed to explain choice behavior for decision making under uncertainty by Busemeyer and Townsend (1993). Later, Townsend and Busemeyer (1995) extended the theory to explain the relationships among choice, selling prices, and certainty equivalents. More recently, it was extended to account for multiattribute decision making by Diederich (1997). However, these previous applications were limited to binary choice situations; the present development extends the theory to multiple (more than two) choice problems and offers initial but critical probing of the theory's ability to predict several central findings in multialternative choice.

This article is organized as follows. The first section reviews the central or pivotal empirical findings from the multialternative choice literature. This includes (a) the *similarity effect* (Tversky, 1972) produced by adding a similar competing alternative to the choice set, (b) the *attraction effect* (Huber, Payne, & Puto, 1982) produced by adding a dominated alternative to the choice set, and (c) the *compromise effect* (Simonson, 1989) produced by adding an intermediate alternative to the choice set. Although specific alternative explanations have been proffered for each of these central qualitative findings, this is the first attempt to account for them all within a single unified theory. The next section presents the basic ideas of multialternative decision field theory (MDFT). We show how the classic multiattribute value model (Keeney & Raiffa, 1976; Von Winterfeldt & Edwards, 1986) and the classic preferential choice model (De Soete et al., 1989; Thurstone, 1959) can be derived as special cases from MDFT under certain ideal conditions and task constraints. The third section describes how MDFT ex-

Robert M. Roe, Jerome R. Busemeyer, and James T. Townsend, Department of Psychology, Indiana University Bloomington.

This research was supported by National Institute of Mental Health (NIMH) Grant F321MH11988-02; NIMH Perception and Cognition Grant R01 MH55680; National Science Foundation Decision Risk Management Science Grant SBR-9602102; and by the James McKeen Cattell Fellowship, which was awarded to Jerome R. Busemeyer. We thank Rob Goldstone, Roger Ratcliff, Barbara Mellers, and Greg Ashby for their comments on an earlier version of this work.

Correspondence concerning this article should be addressed to Jerome R. Busemeyer, Department of Psychology, Indiana University Bloomington, 1101 East 10th Street, Bloomington, Indiana 47405-7007. Electronic mail may be sent to jbusemey@indiana.edu.

plains the similarity, attraction, and compromise effects as well as their complex interactions, using a common set of principles and parameters. The fourth section shows how the winnowing search process of the elimination-by-aspects (EBA) model (Tversky, 1972) can be mimicked by alternative versions of MDFT. Finally, in the last section, we compare the explanatory power of MDFT with other preferential choice models.

Central Findings on Multialternative Preferential Choice

Basic Paradigm

Real life preferential choice problems normally entail a large number of alternatives and attributes. For example, when buying a new car, the buyer needs to compare a broad range of alternative manufacturers and models on a wide variety of performance and economic attributes. In contrast, laboratory experiments strive to test basic theoretical principles by examining a small number of choices that vary on only a few experimentally controlled attributes. In this article, we use a simple laboratory example to provide the background for clearly describing the central empirical results. However, the empirical principles are potentially applicable to the more complex real-life choices as well.

Consider the case of a new car purchase and suppose the choice set has been reduced to a few cars, which differ primarily on only two attributes, performance quality and driving economy. Figure 1 presents a graphical depiction of the problem and clearly illustrates the three major findings.

Similarity Effect

One of the first important results to arise from studies of preferential choice concerns the effect of adding a new competitive option to a choice set that already contains two dissimilar options

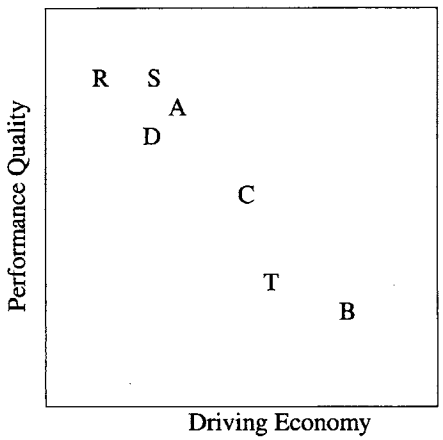

Figure 1. A graphical depiction of the problem of choosing between car options based on the two attributes of performance and economy. The horizontal axis represents the value of each car on the driving economy attribute, and the vertical axis represents the value of each car on the performance quality attribute. Each car is then represented as a point in this two-dimensional space. For example, car A is high on quality and low on economy, whereas car B is low on quality and high on economy. The similarity, attraction, and compromise effects can be illustrated clearly using this simple example.

(Sjoberg, 1977; Tversky, 1972). For example, suppose an industry considers the effect of introducing a new competitive product S on a market that already has two dissimilar competing products, A and B. Furthermore, suppose the new product is highly similar to product A and dissimilar to the other product B. The main finding is that the introduction of the new competitive product to the choice set reduces the probability of choosing similar products more than dissimilar products (Sjoberg, 1977; Tversky, 1972). In terms of market share, the new product steals more from similar products. This effect has significant practical implications for marketing and consumer research (Batsell & Polking, 1985; Bettman, Johnson, & Payne, 1991; Lehmann & Pan, 1994).

Our three-car example in Figure 1 shows a similar situation. The three options are represented in a two-dimensional space of quality and economy. Options A and B are located such that option A has better quality but poorer economy than B, and option B has worse quality but better economy than A. Suppose the high-quality car A is slightly more popular than the economical car B in a binary choice (say 55% favor A over B). If another high-quality car S is introduced, then it steals choices away from the original high-quality car A, evenly dividing the choices between A and S (27.5% each). However, the choices favoring the economical car B remain intact (45%), thus making it most popular in the trinary choice set. In general, a similarity effect occurs whenever the following reversal of choice probabilities is obtained (Sjoberg, 1977; Tversky, 1972): $\Pr[A \mid \{A, B\}] > \Pr[B \mid \{A, B\}]$ but $\Pr[A \mid \{A, B, S\}] < \Pr[B \mid \{A, B, S\}]$. (Note: $\Pr[A \mid \{A, B, S\}]$ denotes the probability of choosing option A from the set containing A, B, and S.)

The similarity effect produces violations of a preferential choice property called *independence from irrelevant alternatives.* According to this property, if x and y are both elements of a choice set T, which in turn is a subset of a larger choice set U, then $\Pr[x \mid T] > \Pr[y \mid T]$ implies $\Pr[x \mid U] > \Pr[y \mid U]$. A large class of probabilistic choice models, known as *simple scalable* utility models, must satisfy this property (Tversky, 1972). The simple scalable class includes all models that assume that each alternative can be assigned a utility scale value, independent of composition of the choice set, and choice probability is determined from the utilities by the general formula $\Pr[x \mid T] = F[u(x), u(y), \ldots, u(z)]$, where F is an increasing function of the first variable and a decreasing function of the all other variables. For example, Luce's (1959) well-known "ratio of strengths" choice model satisfies this property. The similarity effect violates the independence between irrelevant alternatives property and rules out the entire class of simple scalable choice models.

Attraction Effect

A second important finding from studies of preferential choice is the effect of adding a new option that is dominated by one of the other options in the original choice set (Huber, Payne, & Puto, 1982; Ratneshwar, Shocker, & Stewart, 1987; Simonson, 1989; Wedell, 1991). For example, suppose an industry considers introducing a new product D on a market that already has two highly dissimilar competitive products, A and B. Once again, the new product D is designed to be highly similar to an older product A, but in this case, A dominates D in the sense that A is better than D on all the primary attributes. At the same time, D does not

dominate B nor does B dominate D. In this case, D is called an *asymmetrically dominated decoy*. The *attraction effect* refers to the fact that the introduction of the new dominated product to a choice set increases the probability of choosing the dominant product. In terms of market share, the new product enhances the market share of the product that dominates it. (Note that this is the opposite of the similarity effect produced by a new competitive product.)

Figure 1 illustrates this choice situation for the car purchase example. As before, car A is superior on quality, whereas car B is superior on economy. Car D is now slightly inferior on both quality and economy as compared with car A. The attraction effect refers to the empirical finding that adding D to the choice set increases the probability that option A is chosen: $\Pr[A \mid \{A, B\}] < \Pr[A \mid \{A, B, D\}]$.

The attraction effect produces a violation of a general principle implied by a large class of random utility models called the *regularity principle* (cf. Colonius, 1984; MacKay & Zinnes, 1995; Marley, 1989). According to the regularity principle, for any option x that is an element of set W, which is in turn a subset of set U, $x \in W \subseteq U$, the probability of choosing x from W must be greater than or equal to choosing x from U, $\Pr[x; W] \geq \Pr[x; U]$. In other words, the addition of option D to the set already containing A and B should only decrease the probability that option A will be chosen, not increase it. For example, the classic Thurstone (1959) preferential choice model must satisfy regularity, and the attraction effect rules out this entire class of models.

The regularity effect is rather robust. For example, Huber et al. (1982) investigated a variety of different choice conditions producing a wide range of binary choice probabilities. Adding the dominated decoy (D) increased the probability of choosing the dominant alternative (A) under all of these conditions. This includes (a) when A was chosen less frequently than B in a binary choice, (b) when A and B were chosen equally often in a binary choice, and (c) when A was chosen more frequently than B in a binary choice.

Compromise Effect

A third important finding concerns the effect of adding a new option that lies between two competing extreme options in the original choice set (Simonson, 1989; Simonson & Tversky, 1992; Tversky & Simonson, 1993). Suppose there are three equally attractive products, A, B, and C, as indicated by their pairwise preferences, but suppose two of the products, say A and B, are extremely different, and the third product is a compromise that lies in between these two extremes. The compromise effect refers to the empirical finding that when all three options are available for choice, the compromise is chosen more frequently than either of the extremes. Unlike the previous decoy effect, the attractiveness of the compromise option is enhanced by introducing a new competitive (as opposed to dominated) option.

Figure 1 illustrates this choice situation for the car purchase example. As before, car A is superior on quality whereas car B is superior on economy. Car C is a compromise lying in between these two extremes—C is not as good as A on quality but better than A on economy, while C is not as good as B on economy but better than B on quality. The compromise effect refers to the empirical finding that $\Pr[A \mid \{A, B\}] = \Pr[A \mid \{A, C\}] = \Pr[B \mid \{B, C\}]$, but $\Pr[C \mid \{A, B, C\}] > \Pr[A \mid \{A, B, C\}]$ and $\Pr[C \mid \{A, B, C\}] > \Pr[B \mid \{A, B, C\}]$. In other words, the compromise is enhanced when viewed within the context of the two extremes. Furthermore, the effect is found even when the trinary choice set is presented before the three binary comparisons (and thus the result is not due to new information about options that changes the perception of attribute space used to describe the options).

Previous Explanations

Tversky (1972) developed the EBA model to explain the similarity effect. However, the EBA model also obeys the regularity principle (see Tversky, 1972), and therefore it is ruled out by the attraction effect. More recently, Tversky and Simonson (1993) proposed a context-dependent advantage model to account for the attraction and compromise effects. However, as proved in Appendix A, the context-dependent advantage model cannot account for the similarity effect. Taken together, these three central findings continue to remain a deep puzzle for decision theorists. To date, no single theoretical explanation has been brought forth to explain all three within a common theory.

Multialternative Decision Field Theory

The basic intuition underlying decision field theory is that a decision maker's preference for each option evolves during deliberation by integrating a stream of comparisons of evaluations among options on attributes over time. Consider, for example, the car purchase decision discussed earlier. Initially, the decision maker's attention may focus on the most important attribute (e.g., quality) and some specific aspects (e.g., initial acceleration, control on turns, stability at high speeds, stopping power) of this attribute are evaluated for a period of time. During this time period, the evaluation of each option is compared with others and these comparisons change the preferences up or down depending on whether an option has an advantage or disadvantage on the attended attribute. A few moments later, attention may switch to another less important attribute (economy), and comparisons of detailed aspects (e.g., price, gas efficiency, repair costs, reliability, durability) related to this second attribute are added to the previous preferences. Attention may then switch back to an earlier attribute for additional comparisons, and these comparisons continue to update the preferences for each option. Eventually a decision is reached either by an externally imposed time constraint (e.g., the car dealer presses for a final decision) or by a self-imposed criterion (e.g., preference exceeds a threshold and the buyer announces a decision).

Multialternative Dynamic Decision Process

The decision process described above is an example of a large class of decision models called *sequential sampling* models (Link & Heath, 1975; Ratcliff, 1978; Vickers, 1979). Decision field theory builds on this earlier theoretical work by extending the application of these models to multialternative preferential choice situations. The sequential sampling decision process described above can be stated more formally, as follows.

Valences. At any moment in time, each alternative in the choice set is associated with a valence value. The valence for option i at time t, denoted $v_i(t)$, represents the momentary advan-

tage or disadvantage of option i when compared with other options on some attribute under consideration. The ordered set of valences for all the options forms a valence vector, denoted $\mathbf{V}(t)$. For example, a choice among three alternatives {A, B, C} produces the three-dimensional valence vector $\mathbf{V}(t) = [v_A(t), v_B(t), v_C(t)]'$. This valence vector is determined by three different components.

The first component of valence is the personal evaluation of each option on each attribute. In general, the value m_{ij} denotes the subjective value of option i on attribute j. For example, consumer-oriented magazines or Web pages provide the reader with large matrices indicating the objective values of each option on a wide variety of attributes. Using this objective information, the reader can assess his or her personal or subjective evaluations ($m'_{ij}s$). The general model can accommodate any number of evaluations. But for the simplified car decision problem presented above, there are only two primary attributes—economy and quality. The vector $\mathbf{M}_E = [m_{AE}, m_{BE}, m_{CE}]'$ represents the three evaluations for the three cars on economy: If car A gets lower economy than car B, then m_{AE} is assigned a lower scale value than m_{BE}, so that $m_{AE} < m_{BE}$. Similarly, define $\mathbf{M}_Q = [m_{AQ}, m_{BQ}, m_{CQ}]'$ as a vector of evaluations for the three cars on quality: If car A has higher quality than car B, then m_{AQ} is a assigned a higher scale value than m_{BQ}, so that $m_{AQ} > m_{BQ}$. Concatenation of these two vectors forms a 3×2 value matrix, $\mathbf{M} = [\mathbf{M}_E \mid \mathbf{M}_Q]$.

The second component of valence is the attention weight allocated to each attribute at a particular moment in time. The momentary attention to attribute j is represented by an attention weight, $W_j(t)$, at time t. For example, consumer-oriented magazines or Web pages facilitate evenhanded attention to a wide variety of attributes, allowing the reader to rely on his or her own judgments regarding the importance or relevance to the decision. Magazine, television, or Web advertisements attempt to manipulate or draw the viewer's attention to particular attributes that favor the sponsor's product. The general model can accommodate any number of attributes, but for the simplified car purchase example, there are only two attention weights, $W_E(t)$ for economy, and $W_Q(t)$ for quality. The attention weights vary from moment to moment due to changes and fluctuations in attention to the attributes over time. For example, at one moment in time, the decision maker may focus on one attribute (e.g., acceleration), but at later moment, attention may switch to another attribute (e.g., rising cost of gas). In general, the attention weights change and fluctuate across time according to a stationary process. For the present application, it is sufficient to assume that attention shifts in an all-or-none manner from one attribute at one moment, $W_Q(t) = 1$, $W_E(t) = 0$, to another attribute at another moment, $W_Q(t + 1) = 0$, $W_E(t + 1) = 1$, with some fixed probabilities. The probability of attending to the economic dimension is denoted w_E, and the probability of attending to the quality dimension is denoted w_Q. This implies an exponential waiting time for attention to shift from one given attribute to another.[1] The attention weights for all of the attributes form a weight vector $\mathbf{W}(t)$. For the car purchase example with only two attributes (quality and economy), the weight vector is a two-dimensional vector $\mathbf{W}(t) = [W_E(t)W_Q(t)]'$.

The matrix product of weights and values, $\mathbf{MW}(t)$, determines the weighted value of each alternative at each time point. For example, when choosing among three cars, the ith row of the product $\mathbf{MW}(t)$ equals the weighted value of ith option: $W_E(t)m_{iE} + W_Q(t)m_{iQ}$. At first glance, this looks like the classic weighted utility model. However, unlike the classic weighted utility model these weighted values are stochastic because of fluctuations in the attention weights $W_E(t)$ and $W_Q(t)$. A static version of this random weight utility model was successfully used by Fischer, Jia, and Luce (2000) to explain inconsistencies in ratings obtained from multiattribute judgment research.

The third and final component used to determine the valences of each option is the comparison process that contrasts the weighted evaluations of each option. This comparison process is needed to determine the relative advantage or disadvantage of each option on the attribute being considered at that moment. In general, the valence for each option is produced by contrasting the weighted value of one alternative against the average of all the others. For example, with three alternatives, {A, B, C}, the valence for option A is computed by the contrast $v_A(t) = W_E(t)m_{AE} + W_Q(t)m_{AQ} - [(W_E(t)m_{BE} + W_Q(t)m_{BQ}) + (W_E(t)m_{CE} + W_Q(t)m_{CQ})]/2$. This comparison process also can be represented by a matrix operation by defining a contrast matrix:

$$\mathbf{C} = \begin{bmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{bmatrix}.$$

Using this matrix definition, the valence vector is formed by the matrix product

$$\mathbf{V}(t) = \mathbf{CMW}(t). \tag{1a}$$

Each row of the matrix product produces a valence or comparison value similar to that shown for option A.

The car purchase example described above contains for simplicity, only three alternatives described by two primary attributes. However, most decisions involve a larger number of attributes, and Equation 1a is applicable with arbitrary numbers of alternatives and attributes. However, in practice it is useful to group the possibly large number of attributes into two subgroups, a relatively small subgroup of primary attributes of importance and a larger subgroup of irrelevant attributes. For example, the experiments discussed later usually design the options by manipulating a few primary dimensions, but these options may also differ on a number of other irrelevant attributes. Even in these more complex settings, a simplified analysis can be performed by partitioning a p-dimensional attention weight vector into two components $\mathbf{W}(t)' = [\mathbf{W}_1(t)', \mathbf{W}_2(t)']$, where $\mathbf{W}_1(t)$ is a q-dimensional component containing the primary dimensions and $\mathbf{W}_2(t)$ contains the remaining p–q irrelevant dimensions. Then Equation 1a can be rewritten in terms of the primary and irrelevant dimensions as follows:

$$\mathbf{V}(t) = \mathbf{CMW}(t) = \mathbf{CM}_1\mathbf{W}_1(t) + \boldsymbol{\epsilon}(t), \tag{1b}$$

where $\boldsymbol{\epsilon}(t) = \mathbf{CM}_2\mathbf{W}_2(t)$ can be treated as an stochastic error or residual term.

Preferences. At any moment in time, each alternative in the choice set is associated with a preference strength. The preference

[1] In this application we assume that the attention weights are identically and independently distributed over time according to a simple Bernoulli process. However, Diederich's (1997) multiattribute decision field model employs a more sophisticated Markov process for switching attention to attributes. The Bernoulli process is a simple special case of the more general Markov process, but it is sufficient for the present purposes.

strength for alternative i at time t, denoted $P_i(t)$, represents the integration of all the valences considered for alternative i up to that point in time. The preferences for all the alternatives form a preference state vector, denoted $\mathbf{P}(t)$. For example, a choice among three options {A, B, C} produces the three-dimensional preference state $\mathbf{P}(t) = [P_A(t), P_B(t), P_C(t)]'$.

A new state of preference $\mathbf{P}(t + 1)$ is formed at each moment from the previous preference state $\mathbf{P}(t)$ and the new input valence vector, $\mathbf{V}(t)$, according to the following linear stochastic difference equation:

$$\mathbf{P}(t + 1) = \mathbf{SP}(t) + \mathbf{V}(t + 1). \quad (2)$$

According to this simple updating equation, the new preference state is a weighted combination of the previous preference state and the new input valence. The dynamic behavior of this model is determined by two factors: the initial preference state $\mathbf{P}(0)$ at time $t = 0$, and the feedback matrix $\mathbf{S}$.

Initial preference state. In general, the initial preference state represents a residual bias left over from previous experience with choice problems. For example, status quo effects (Samuelson & Zeckhauser, 1988), previous habits, or experience and memory for the previous history of choices can be captured by the initial state. For novel choice problems, the initial state may be considered unbiased, in which case $\mathbf{P}(0) = 0$. The applications presented below are based on the latter assumption.

Feedback matrix. The feedback matrix $\mathbf{S}$ shown in Equation 2 contains the self-connections and interconnections among the choice alternatives. The diagonal elements, S_{ii}, determine the memory of the previous preference state for a given alternative. These self-feedback loops are needed to integrate the valences for a given option over time, and they allow the preference strength within an alternative to grow or decay over time. If the self-feedback loop is set to zero, then that option has no memory of its previous state. If the strength of the self-connection is set to one, then that option has perfect memory of its previous state. Intermediate strengths, between zero and one, provide partial memory and limited decay. For the present work, the self-feedback loops are identical for all options, that is, S_{ii} is the same for all i.

The off-diagonal elements, S_{ij} for $i \neq j$, determine the influence of one alternative on another. These interconnections are generally negative so that they produce competitive influences. If the interconnections are all zero, then the alternatives do not compete at all and instead they grow or decay independently and in parallel. If the interconnections are negative, then strong alternatives suppress weak alternatives. The strengths of the interconnections are determined by the concept of lateral inhibition (discussed again in the section on connectionist networks). The basic idea is that the strength of the lateral interconnection between a pair of options is a decreasing function of the distance between these two options in the multiattribute space.

More formally, it is assumed that each alternative is represented as a point in a multidimensional space with dimensions defined by the attributes used to characterize the choice alternatives. For example, Figure 1 illustrates a small set of cars placed within a two-dimensional space characterized by quality and economy. In this example, Options A and B are highly dissimilar, and two other options, S and D, are both highly similar to option A. Thus pairs (A, S) and (A, R) have much stronger inhibitory interconnections than pairs (A, B), (S, B), or (R, B). Define d_{ij} as the psychological distance between options i and j in this multiattribute preference space. Then the interconnection between options i and j is determined by $S_{ij} = F[d(A_i, A_j)]$, where F is a decreasing function. At this point we do not need to specify the form of F (e.g., exponential is one possibility; see Shepard, 1964). But this equation does place two important constraints on the lateral inhibition connections: One is symmetry, $S_{ij} = S_{ji}$, and the second is that the interconnection decreases with distance.

Conceptually, options that are unrelated to each other elicit little or no competition, whereas options that are closely related produce greater competition. These interconnections provide a dynamic mechanism for producing bolstering effects (Janis & Mann, 1977) and justification effects (Simonson, 1989) in preference. When a weak unattractive alternative is compared with a strong attractive alternative, the negative values from the weak option feed back through a negative connection and the product of these two factors produces a net positive effect that bolsters or justifies the strong option.

Multiattribute utility model. If the feedback matrix is set to zero ($\mathbf{S} = \mathbf{0}$), then according to Equation 2, the preference state equals the valence input, $\mathbf{P}(t) = \mathbf{V}(t)$. If it is also assumed that attention does not fluctuate across time, $\mathbf{W}(t) = \mathbf{w}$, and the residuals in Equation 1b are zero, $\boldsymbol{\epsilon}(t) = \mathbf{0}$, then valence $\mathbf{V} = \mathbf{CMw}$ produces exactly the same rank order as classic multiattribute weighted additive values (Keeney & Raiffa, 1976; Von Winterfeldt & Edwards, 1986). Applying the car purchase example to this special case, the valence for option A reduces to $v_A = (w_E m_{AE} + w_Q m_{AQ}) - [(w_E m_{BE} + w_Q m_{BQ}) + (w_E m_{CE} + w_Q m_{CQ})]/2$. Furthermore, $v_A > Max(v_B, v_C)$ implies that $(w_E m_{AE} + w_Q m_{AQ}) > Max[(w_E m_{BE} + w_Q m_{BQ}), (w_E m_{CE} + w_Q m_{CQ})]$. Thus, $\mathbf{V}$ produces exactly the same rank order over alternatives as the classic multiattribute model under these restrictive conditions.

Dynamic Thurstone model. On the basis of the multivariate central limit theorem, the distribution of the preference states $\mathbf{P}(t)$ converges to the multivariate normal distribution as the number of steps (t) in Equation 2 becomes large. If the time period between steps is small, then this convergence will occur very rapidly. If the feedback matrix $\mathbf{S}$ is set equal to an identity matrix (all cross-feedback set to zero), at any fixed time point t, Equation 2 reduces to a multivariate Thurstone preferential choice model (Bock & Jones, 1968; see Appendix B for details). Note that MDFT is a dynamic generalization of the classic Thurstone preference model, as it describes how the mean vector and variance–covariance matrix of the preference state evolve systematically over time.[2] The mean preferences for each alternative can change signs over time so that initially one alternative (say A) may have the largest mean preference, but later another alternative (say C) may come to dominate.

Summary of parameters. Altogether there are four sets of parameters that need to be specified to derive predictions from

[2] Previous theorists working in sensory, perception, and memory have shown that unidimensional sequential sampling models provide generalizations of the unidimensional signal detection models (Link and Heath, 1975; Ratcliff, 1978). The present work extends these ideas to preferential choice and provides a generalization of the multivariate Thurstone model including arbitrary variance–covariance structures.

MDFT. Two sets of parameters are contained in the mean weight and value matrices, **w** and **M**, which are also required by the classic multiattribute utility model. A third parameter is the residual variance contributed by the irrelevant attributes, which is also required by any Thurstone-type probabilistic choice model. The final set is contained in the symmetric feedback matrix, **S**, which is required by any dynamic connectionist model. The parameters in the **S** matrix are a function of the distance between alternatives in the attribute space. In the case of three alternatives, the feedback matrix produces at most three new parameters (self-feedback, A to B inhibition, and A to C inhibition).

Multialternative Choice Rules

Eventually, the evolving output preferences, $\mathbf{P}(t)$, determine the final choice. But the specific decision rule for determining this choice varies depending on whether the decision time is externally imposed versus subject controlled, as described next.

Externally controlled stopping time. Choice tasks are termed *externally controlled* when the decision is made at an appointed time or designated time point (Ratcliff, 1978; Vickers, Smith, & Brown, 1985). For example, a woman who has just received a proposal for marriage may be asked to announce her decision at breakfast the next morning. Alternatively, a woman who has just been offered a job may be asked to sign her contract by the end of the week. For the car choice example, the car dealer, on seeing other customers waiting for help, may lose his patience, interrupt the purchaser, and pressure her to make an immediate decision.

The vertical line in Figure 2 illustrates this type of stopping rule for the new car example. In this case, the option with the highest preference at the designated time point is chosen (option B in Figure 2).

Formally, the probability that one alternative (say A) is chosen from a set of three alternatives {A, B, C} at a fixed time t is determined by the relation:

$$\Pr[\text{A} \mid \{\text{A, B, C}\} \text{ at time } t] = \Pr[P_{\text{A}}(t) > P_{\text{B}}(t) \text{ and } P_{\text{A}}(t) > P_{\text{C}}(t)]. \quad (3)$$

Determination of choice probabilities for choice sets with N alternatives follows the same principle outlined in Equation 3, except that it requires the conjunction of $N - 1$ events of the form $P_{\text{A}}(t) > P_i(t)$, for $i \neq \text{A}$. Appendix B provides the mathematical formulas for calculating probabilities using Equation 3.

By stopping the deliberation process at various designated time points and estimating the choice probabilities as a function of deliberation time, it is possible to observe the evolution of preferences dynamically over time. Cognitive researchers have used this type of paradigm to study the dynamics of memory (Dosher, 1984; Gronlund & Ratcliff, 1989; Hintzman & Curran, 1997; Ratcliff, 1978; Ratcliff, 1980; Ratcliff & McKoon, 1982; Ratcliff & McKoon, 1989; Reed, 1973; Wickelgren, Corbett, & Dosher, 1980). Although this method has not been applied to preferential choice, the present theory provides simple predictions for this type of task. These predictions will be presented later during the review of the empirical results for preferential choice. The second type of choice task is presented next.

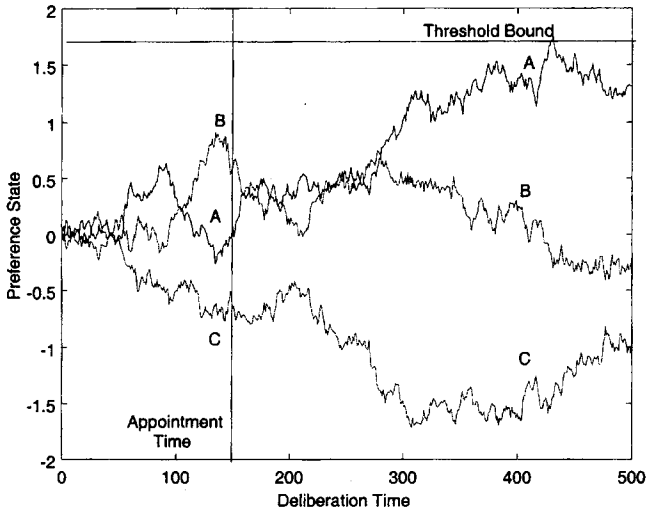

Figure 2. Illustration of the two stopping rules. The abscissa represents time, and the ordinate represents level of preference. The three trajectories (labeled A, B, and C) represent the preferences of each option as they evolve stochastically over time. The vertical line to the left represents the appointed time for the decision according to the externally controlled stopping rule. In this case, option B would be chosen at the designated time $t = 150$. The horizontal line on the top represents the inhibitory threshold that must be reached to make a choice for the subject-controlled stopping rule. In this case, option A would be chosen when it crosses the threshold at time $t = 430$.

Internally controlled stopping time. Choice tasks are called *internally controlled decisions* when the decision maker is free to decide how long to deliberate before finally announcing or committing to a particular choice (Ratcliff, 1978; Vickers et al., 1985). In the car purchase example, the buyer may inform the car dealer that she wishes to go home and think about the purchase and that she will call back as soon as she makes up her mind. This is a common type of choice task used in laboratory experiments on decision making.

The horizontal line in Figure 2 illustrates this type of stopping rule for the new car example. In this case, a choice is made as soon as the strength of preference for an option crosses a threshold, and the first option to exceed the threshold is then chosen (option A in the figure).

The choice probabilities and decision times for the internally controlled task are determined by the first passage time distribution for a sample path to cross the threshold boundary (see Bhattacharya & Waymire, 1990; Cox and Miller, 1965; Smith, 2000). Formulas for computing the first passage time distribution (as well as the moments including choice probabilities and mean decision times) have been derived by Busemeyer and Townsend (1992) and by Diederich (1997) for the binary choice case and by Busemeyer and Diederich (2000) for more than two options. Alternatively, the first passage time distribution for the multichoice (more than two) case can be obtained through computer simulation (see Appendix C for details).

Connectionist Interpretation

Figure 3 provides an interpretation of the multialternative dynamic model as a connectionist network.[3] The first layer is a connectionist feed-forward network (Rumelhart & McClelland, 1986). The evaluation nodes (labeled **M**) to the left of the figure represent the evaluations of each alternative on each attribute. For example, if the decision is to choose among three cars on the basis of quality and economy attributes, then there are six possible evaluations. (Of course, a real car purchase entails many more attributes and alternatives.) These evaluations are filtered by momentary attention weights (labeled **W**) linked to the attributes. Subsequently, the weighted evaluations are transformed by contrast coefficients (labeled **C**) to produce comparisons among the weighted evaluations. The outputs of the first layer (labeled **V**) in Figure 3 are the valences, which represent the advantage or disadvantage being considered for each alternative at a particular time point. These valences change stochastically over time as the decision maker's attention shifts unpredictably from one attribute to another.

The second layer in Figure 3 is a competitive recursive network (Grossberg, 1982; Grossberg, 1988; Rumelhart & McClelland, 1986). The three nodes (labeled A, B, and C) represent the three choice alternatives (e.g., three cars). More generally, one decision node in the recursive network corresponds to each choice alternative. The input into each decision node is the valence from the first layer. The output activation from each decision node represents the strength of preference for the corresponding alternative at a particular point in time. The activation level increases with positive valences (advantages) and decreases with negative valences (disadvantages). Thus, the activation of a choice node at any point in time represents the strength of the evolving preference formed by the temporal integration of the stream of input valences.

Each decision node in the network is connected with every other node and each decision node also has a self-feedback loop. The self-feedback loop integrates the valences for a given option over time, allowing activation within a node to grow or decay. The interconnections among nodes represent a competitive system so that activation of one node inhibits the other nodes. The interconnection strengths are assumed to be a decreasing function of the perceived dissimilarity between alternatives within the multiattribute space. This corresponds to the principle of lateral inhibition used in competitive neural networks. Lateral inhibition is a key principle for producing edge enhancement effects and Mach band effects in perception (see Cornsweet, 1970). In this application, the lateral inhibition effects are used to produce bolstering or justification effects observed in decision making (Janis & Mann, 1976; Simonson, 1989). This completes the presentation of the basic theory, and now we turn to several important empirical applications for multialternative preferential choice.

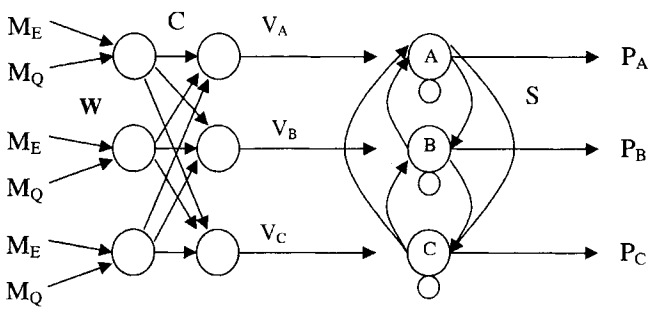

Figure 3. Connectionist interpretation of multialternative decision field theory consisting of two layers. The first layer is a connectionist feed-forward network. The evaluations (labeled M) to the left of the figure represent the evaluations of each alternative on each attribute. These evaluations are filtered by momentary attention weights (labeled W) linked to the attributes. Subsequently, the weighted evaluations are transformed by contrast coefficients (labeled C) to produce comparisons among the weighted evaluations. The outputs of the first layer (labeled V) are the valences. The second layer is a competitive recursive network. The three nodes (labeled A, B, and C) represent the three choice alternatives (e.g., three cars). Each node is connected to every other node and has a self-feedback loop. The values of these connections are given in **S**, the feedback matrix. The outputs are preferences (labeled P).

Applications of MDFT to Central Empirical Findings

Predictions for the Similarity Effect

According to MDFT, the following explanation causes the similarity effect. Consider the choice among options A, B, and S in Figure 1. Whenever attention happens to focus on the quality attribute, then both options A and S gain advantages while option B gets a disadvantage. Likewise, whenever attention happens to focus on the economy attribute, then both options A and S get disadvantages while option B gains an advantage. Thus, the valences of A and S are positively correlated with each other and negatively correlated with B. Participants who tend to focus more on economy will choose option B whereas participants who tend to focus more on quality will choose either A or S. Thus S only hurts or takes away choices from option A and not option B. A more formal explanation is presented below, after a discussion of the details about the assignment of parameters of Equation 2 to the conditions indicated in Figure 1.

According to the value structure shown in Figure 1, car A is high on quality and low on economy whereas car B is just the opposite. Car S is similar to A but slightly lower on economy and slightly better on quality. This pattern of evaluations can be represented in a simple form by the value matrix:

$$\mathbf{M}_1 = \begin{array}{c} \begin{matrix} \text{E} & \text{Q} \end{matrix} \\ \begin{bmatrix} 1 & 3 \\ .85 & 3.2 \\ 3 & 1 \end{bmatrix} \end{array} \begin{matrix} \text{A} \\ \text{S} \\ \text{B} \end{matrix} .$$

The precise numerical values are not critical to produce the predicted pattern from the model, as long as S is a competitive alternative with values close to option A.

The attention weights, $[W_E(t), W_Q(t)]$, were assumed to fluctuate over time steps according to a simple Bernoulli process. Specifically, the probability of attending to the economy attribute was assigned a probability w_E, and the probability of attending to the

[3] The connectionist interpretation presented here uses a local rather than a distributed representation of the input features (see Rumelhart & McClelland, 1986). The local representation was chosen because it provides a more direct relation to previous theories in decision making. It is possible to employ a distributed representation, but this was not required for the phenomena considered here.

quality attribute was assigned a probability w_Q, with independent sampling across time steps (but see Footnote 1). The probability of attending to the quality attribute was set to a slightly higher level (w_Q = .45) than the probability of attending to the economy attribute (w_E = .43), and there was some small residual probability allowed for attention to irrelevant attributes. The reason for the slight difference in attention to each attribute was to produce a slight preference in favor of car A over car B for the binary choice condition. This was needed to satisfy the antecedent condition for the test of independence from irrelevant alternatives. These same parameters were then used to make new predictions for the trinary choice set.

The parameters for the feedback matrix were chosen as follows. First, the self-connections were set to a high value (S_{ii} = .94) to produce slow decay of memory. The inhibitory connections between distant alternatives were set to very low values ($S_{AB} = S_{BA} = S_{SB} = S_{BS} = -.001$). The inhibitory connections between the similar alternatives were set to relatively greater magnitudes ($S_{AS} = S_{SA} = -.025$). These parameter values satisfy a stability requirement (i.e., the eigenvalues of the feedback matrix **S** are all less than 1 in magnitude). The predicted pattern does not change much as long as the inhibitory connections are not too large.

The predictions computed from Equation 3 are shown in Figure 4, which plots the probability of choice as a function of deliberation time separately for different choice alternatives. The upper and lower curves, indicated with "+" and "□" symbols, respectively, represent the choice probabilities for cars A and B from the binary choice comparison. As required for the antecedent condition of the independence test, the probability of choosing A grows over time to become larger than B for the binary choice. The critical feature is revealed by the upper and lower curves indicated with "○" and "x" symbols, respectively, representing the choice probabilities for B and A from the trinary choice comparison. Consistent with previous research, the model predicts that the probability of choosing B is higher than A for the trinary choice set, thus violating the independence from irrelevant alternatives property.

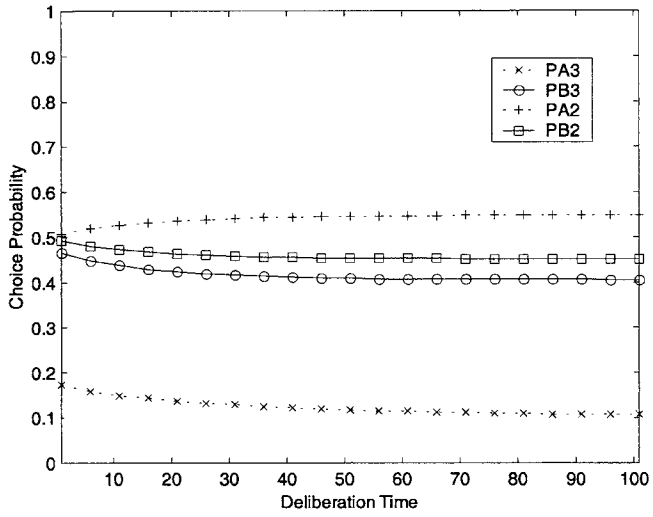

Figure 4. Multialternative decision field theory predictions for the similarity effect. The probability of choice is plotted as a function of deliberation time separately for different choice alternatives. The upper and lower curves indicated with "+" and "□" symbols, respectively, represent the choice probabilities for cars A and B (PA and PB, respectively) from the binary choice comparison. The upper and lower curves indicated with an "○" and "x" symbols, respectively, represent the choice probabilities for B and A from the trinary choice comparison.

Note that similarity also could be manipulated by placing the new alternative S lower and to the right of A (by slightly increasing the economy and slightly decreasing the quality of S compared with A) to produce the following value matrix:

$$\mathbf{M}_1 = \begin{matrix} \\ A \\ S \\ B \end{matrix} \begin{matrix} E & Q \\ \left[\begin{matrix} 1 & 3 \\ 1.25 & 2.8 \\ 3 & 1 \end{matrix}\right. \end{matrix}$$

Although the dynamics are slightly different, the model still correctly predicts the same pattern for the similarity effect. The binary choice probability for alternative A over B asymptotes at Pr[A | A, B] = .55, but the trinary choice probability for option B, Pr[B | A, B, S] = .4, exceeds that for option A, Pr[A | A, B, S] = .3 at this same time point.

According to MDFT, similarity effects are caused by the correlations among valences produced by the primary attributes. When attention is focused on the economy attribute, then the valence for option B will be greater than the valences for options A and S. Alternatively, whenever attention is focused on the quality attribute, then the valence for option B will be less than the valences for options A and S. This causes the differences between B and A to be positively correlated with the differences between B and S. Figure 5 illustrates the effects of this correlation on choice probability. The left and right panels show the equal density contours from the multivariate normal densities used to compute the choice probabilities for the trinary choice set. The left panel shows the contour for option A, and the right panel shows the contour for option B. Choice probability is related to the area above and to the right of the zero preference state on the vertical and horizontal axes. Note that the strong positive correlation for option B rotates the elliptical contour into the upper right corner, thus increasing the choice probability for option B relative to A.

If the effect of the correlations among the primary attributes is diminished and instead the residual variance due to irrelevant attributes is amplified (by increasing the variance of the noise term ϵ), then the similarity effect disappears.[4] For example, if the residual variance is increased tenfold (see Appendix B for details),

[4] If the covariance matrix for the valence is constrained to satisfy sphericity (zero correlations), then the multivariate Thurstone model reduces to the Case V version, and the latter is a special case of the simple scalable class of models (cf. Bockenholt, 1992). However, according to MDFT, the covariance matrix for the valence reflects the similarity structure among options in the choice set, and this implies that the covariance matrix does not satisfy sphericity. Decision field theory is not the first to make important use of the covariance matrix for modeling stimuli within a multidimensional space. Previous applications of the multiattribute Thurstone model for binary choices (see De Soete et al., 1989) used this concept to explain violations of strong stochastic transitivity. The covariance matrix also plays a crucial role in Ashby and Townsend's (1986) general recognition theory, and in Ashby & Maddox's (1992) decision-bound theory of categorization.

then the asymptotic binary choice probability of A over B remains at Pr[A | {A, B}] = .55, but the trinary choice probability of choosing A, Pr[A | {A, B, S}] = .36, remains higher than that for alternative B, Pr[B | {A, B, S}] = .30.

It is interesting that the similarity effect is strongest when all the inhibitory interconnections are set to zero. Thus, lateral inhibition impedes the similarity effect. However, inhibitory interconnections are needed for the other effects discussed later. To maintain consistency in our assumptions, we used the same lateral inhibition principle across all three of the empirical applications.

Figure 6 shows a broader examination of the predictions for the similarity effect derived from MDFT. This figure shows the difference between the probability of choosing option A and the probability of choosing option B from the trinary choice set, plotted as a function of the two key theoretical parameters. One parameter is the strength of the inhibitory connection between the nodes for options A and S, S_{AS}, and the second is the standard deviation of the residuals from the irrelevant attributes (*SD* of ϵ). The probabilities of attending to each dimension were equated, $w_E = w_Q$ so that the binary choices between A and B are predicted to be equal. The similarity effect occurs when the difference Pr[A | {A, B, S}] − Pr[B | {A, B, S}] is negative (below the line).

A close look at Figure 6 shows that the similarity effect is strongest when both the standard deviation (*SD*) of the residuals and amount of lateral inhibition are low. Consider, for example, the curve produced by the low residual *SD*: As the lateral inhibition increases the effect diminishes, but one can see that even with relatively high lateral inhibition the effect is still present. Slightly increasing the residual *SD* when inhibition is high, however, causes the effect to disappear. Now consider the curve produced when lateral inhibition is low: As the residual *SD* increases the effect diminishes, but even at a relatively high *SD* one can see that the effect is still present. Under relatively high residual *SD*, lateral inhibition still must be increased a sizable amount before the effect disappears. In sum, the similarity effect occurs whenever the

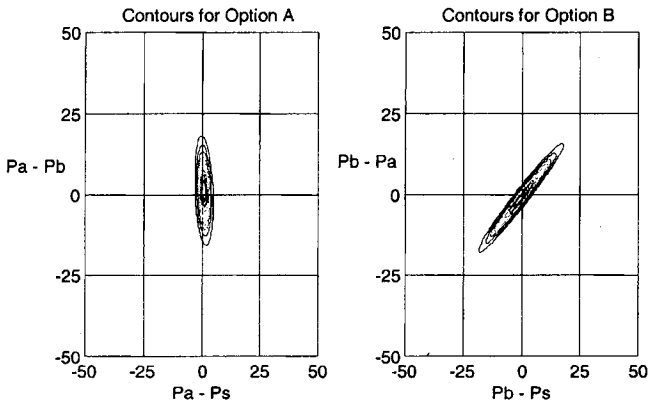

Figure 5. A contour plot showing the effects on choice probability of correlation among valences produced by the primary dimensions. The left and right panels show the equal density contours from the multivariate normal densities used in Equation 3 to compute the choice probabilities for the trinary choice set. The left panel shows the contour for option A, and the right panel shows the contour for option B. Choice probability is related to the area above and to the right of the zero preference state on the vertical and horizontal axes.

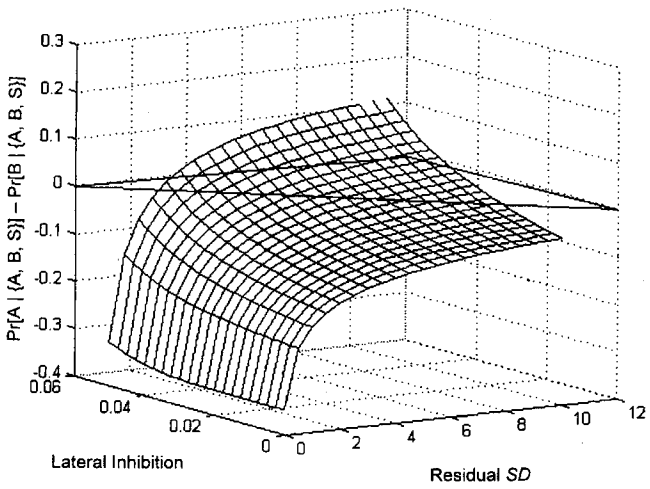

Figure 6. A response surface plot showing the difference between the probability of choosing option A and the probability of choosing option B from the trinary choice set plotted as a function of two key theoretical parameters. One parameter is the strength of the inhibitory connection between the nodes for options A and S, S_{AS}, and the second is the standard deviation of the residuals for the irrelevant attributes, ϵ. The similarity effect occurs when the difference Pr[A|{A, B, S}] − Pr[B|{A, B, S}] in the figure is negative (below the line).

covariance matrix for the primary dimensions is relatively more important than the residual variance. Note that experiments on the similarity effect used carefully designed stimuli that minimized the influence of irrelevant attributes to maximize the similarity effect.[5]

Similarity effects are not restricted to multialternative choice contexts, and they also have been observed using only binary choices (Mellers and Biagini, 1994). Once again, refer to Figure 1 and first consider two pairwise choices: One choice between option A and option R, and a second choice between option B and option R. In this paradigm, option R is used as a standard for comparing the preference strengths of options A and B. Note that option A is more similar to option R than option B: Option R has almost the same quality as A but differs more on economy. Under this condition, Mellers and Biagini (1994) found that Pr[A | {A, R}] > Pr[B | {B, R}]. In other words, when option R is used as a standard for comparison, option A appears to produce a stronger preference than option B.

Now consider two other pairwise choices: One choice between option A and option T, and a second between option B and option T. In this case, option T is used as a standard for comparing the preference strengths of options A and B. Note that option B is more similar to option T than option A: Option T has almost the

[5] To further check the robustness of the similarity effect across different stopping rules, the choice probabilities were also computed using the internally controlled stopping rule (see Appendix C for details). The results from the internally controlled stopping rule replicate those shown in Figure 4. The probability that A is chosen over B in the binary choice is Pr[A|{A, B}] = .55, as before, but the trinary probability of choosing B, Pr[B|{A, B, S}] = .35, is higher than the trinary probability of choosing A, Pr[A|{A, B, S}] = .25, again violating the independence from irrelevant alternatives property.

same quality as B but differs more on economy. Under this condition, Mellers and Biagini (1994) found Pr[B | {B, T}] > Pr[A | {A, T}]. In other words, when option T is used as a standard for comparison, option B appears to produce a stronger preference than option A. This reversal in binary choice probabilities constitutes another type of violation of independence between alternatives. This violation rules out all simple scalable utility models for binary choices (Tversky, 1972).

MDFT predicts these results using exactly the same parameters as used in Figure 4, except for changes in the value matrices:

$$\mathbf{M}_1 = \begin{matrix} \\ A \\ R \\ B \end{matrix}\!\!\begin{matrix} E & Q \\ \end{matrix}\begin{bmatrix} 1 & 3 \\ .5 & 3.1 \\ 3 & 1 \end{bmatrix} \qquad \mathbf{M}_1 = \begin{matrix} \\ A \\ T \\ B \end{matrix}\!\!\begin{matrix} E & Q \\ \end{matrix}\begin{bmatrix} 1 & 3 \\ 2.5 & 1.1 \\ 3 & 1 \end{bmatrix}$$

The asymptotic predictions derived from Equation 3 reproduce the violation of independence observed by Mellers & Biagini (1994): Pr[A | {A, R}] = .99 > Pr[B | {B, R}] = .63; Pr[B | {B, T}] = .99 > Pr[A | {A, T}] = .77. These predictions are robust with respect to lateral inhibition: The same result is obtained with $S_{AR} = S_{BT} = -.025$ and with $S_{AR} = S_{BT} = 0$.

The intuitive reason for this predicted pattern is simple to explain. First consider the choice between options A and R: When attention is focused on quality, little difference in valence occurs, but when attention is focused on economy, a large positive valence favoring option A occurs. Next consider the choice between options B and R: When attention is focused on quality, large valences favoring option R occur, but when attention is focused on economy, large valences favoring B occur. Analogously, when choosing between B and T, the valence almost always favors B, but when choosing between A and T, valences oscillate back and forth, one moment favoring A, the next moment favoring T. Technically, binary choice probability is determined by the ratio of the mean difference and the standard deviation of the difference (see Appendix B). The variance of the preference difference is larger for the choice between dissimilar (e.g., B and R) as compared to similar (e.g., A and R) options.

So far we have shown that MDFT can reproduce the well-known similarity effect under a wide variety of conditions when the valences on the primary attributes are correlated in a manner that reflects the similarity structure in the choice set. But several earlier choice models were developed to explain these results (EBA model of Tversky, 1972, and the multivariate Thurstone model of Edgell & Geisler, 1980, for multiple choice; Mellers & Biagini, 1994, for binary choice). The next challenge is faced by simultaneously explaining the similarity effect as well as the attraction effect. All of the above mentioned multiple-choice models satisfy a general property known as the *regularity principle.* The attraction effect described next produces empirical violations of the regularity principle.

Predictions for the Attraction Effect

According to MDFT, the following explanation causes the attraction effect. Consider a choice among options A, B, and D in Figure 1. Comparisons of the dominated decoy with the average of the other two options eventually produces a negative preference state for the dominated decoy, D. Then this negative preference state from the dominated decoy feeds through a negative inhibitory link to the closely positioned dominant option, A. The two negatives cancel to produce a net positive bolstering effect of the dominated decoy on the dominant option. Thus the decoy makes the dominant option "appear" stronger, similar to an edge enhancement effect in perception. Option B does not experience any bolstering effect because it is too dissimilar to D, and the inhibitory connection is too weak to produce the effect. A more formal analysis is provided below.

The only change in assignment of parameters that needs to be made for computing the predictions from Equation 3 for the attraction effect is the value matrix **M** (which is necessary to represent the new location of dominated option D in Figure 1). All of the remaining parameters were kept constant across the two different applications.

According to the value structure shown in Figure 1, car A is high on quality and low on economy whereas car B is just the opposite. Car D is similar to A but slightly inferior on economy and quality. This pattern of evaluations can be represented in a simple form by the value matrix

$$\mathbf{M}_1 = \begin{matrix} \\ A \\ D \\ B \end{matrix}\!\!\begin{matrix} E & Q \\ \end{matrix}\begin{bmatrix} 1 & 3 \\ .5 & 2.5 \\ 3 & 1 \end{bmatrix}$$

Once again, the precise numerical values are not critical to produce the predicted pattern from the model. For example, very similar results are obtained when the values of D are set to (.75, 2.75) instead of (.5, 2.5).

Figure 7 shows the predictions of MDFT for the attraction effect. The figure plots choice probability as a function of deliberation time separately for different choice alternatives. The upper

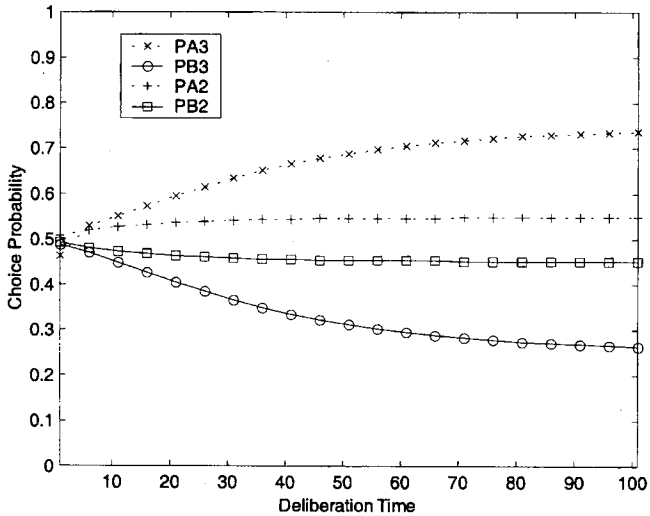

Figure 7. A plot of the multialternative decision field theory predictions for the attraction effect. Choice probability is plotted as a function of deliberation time separately for different choice alternatives. The upper and lower curves indicated with "+" and "□" symbols, respectively, represent the choice probabilities for cars A and B (PA and PB, respectively) from the binary choice comparison. The lower and upper curves indicated with an "○" and "x" symbols, respectively, represent the choice probabilities for B and A from the trinary choice comparison.

and lower curves indicated with "+" and "□" symbols, respectively, represent the choice probabilities for cars A and B from the binary choice set. Similar to Figure 4, the probability of choosing A grows over time to become larger than B for the binary choice. The upper and lower curves, indicated with "x" and "○" symbols, respectively, represent the choice probabilities for A and B from the trinary choice set. Consistent with previous research, the model correctly predicts that the probability of choosing A is higher from the trinary choice set (the "x" curve) as compared to the probability of choosing A from the binary choice set (the "+" curve), thus violating the regularity property.

To check the robustness of this prediction, it was recomputed using Equation 3 after changing the probabilities attending to each of the primary attributes. In one case, the probability of attending to each attribute was equated ($w_E = w_Q = .45$), producing binary choice probabilities equal to .50 for A and B, but the probability of choosing A from the trinary set rose to .69. In another case, they were reversed ($w_E = .43$, $w_Q = .45$), producing a binary choice probability equal to .45 for option A, but once again the probability of choosing A from the trinary set rose to .65. In both cases, the same pattern of predictions was obtained for the dominating alternative—the decoy increased the predicted probability of choosing the dominating alternative (i.e., adding D increased the probability of choosing A) in agreement with results from Huber et al. (1982).

Note that Figure 7 was generated using exactly the same model parameters as Figure 4, except for the change in the **M** matrix to reflect the change of option D from a competitive to a dominated alternative close to option A. But the predicted effect of adding the dominated option (Figure 7) was just the opposite of the predicted effect of adding the competitive option (Figure 4). The theoretical reason for this dramatic change in predictions needs to be understood. If the lateral inhibitory connections are set to zero so that the feedback matrix **S** is set equal to a diagonal matrix, then the attraction effect disappears. In this case, the predictions computed from Equation 3 produced the following results: The decoy D is virtually ignored, the probability of choosing A from the binary and trinary choice sets remain identical and both equaled .55, and the decoy has no effect on the dominating alternative A.

It turns out that the correlations among the primary attributes do not play a crucial role for the attraction effect. Figure 8 shows the predictions when the lateral inhibition is reset to its original value (as used in Figure 7), but the effect of the residual variances (*SD* of ϵ) is increased tenfold. As can be seen in Figure 8, the attraction effect still occurs, although now it takes some time to build up. In sum, lateral inhibitory connections are crucial and covariance structure is less important for MDFT to produce the attraction effect.

Figure 9 shows a more detailed examination of the predictions for the attraction effect derived from MDFT. This figure shows the difference between the probability of choosing option A from the trinary set and the probability of choosing A from the binary set, plotted as a function of the two key theoretical parameters: the strength of the inhibitory connection between the nodes for options A and D, S_{AD}, and the standard deviation of the residuals for the irrelevant attributes, *SD* of ϵ. The probabilities of attending to each dimension were equated, $w_E = w_Q$ so that the binary choices between A and B are predicted to be equal. The attraction effect occurs when the difference $\Pr[A \mid \{A, B, D\}] - \Pr[A \mid \{A, B\}]$ is positive (above the line).

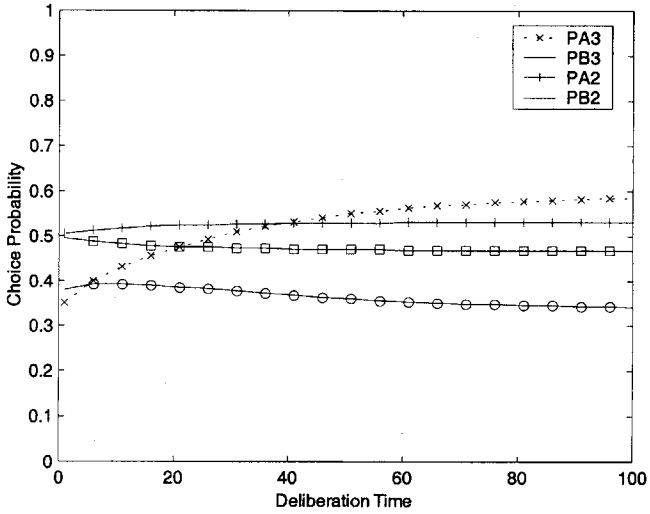

Figure 8. A plot of the multialternative decision field theory predictions for the attraction effect when the influence of the diagonal matrix of residual variances is increased tenfold. The upper and lower curves indicated with "+" and "□" symbols, respectively, represent the choice probabilities for cars A and B (PA and PB, respectively) from the binary choice comparison. The lower and upper curves indicated with "○" and "x" symbols, respectively, represent the choice probabilities for B and A from the trinary choice comparison.

A close look at Figure 9 shows that the attraction effect is strongest when the *SD* of the residuals is low and amount of lateral inhibition is high. Consider, for example, the curve produced by the low residual *SD*: As the lateral inhibition increases, the effect quickly becomes present and increases dramatically with increasing amounts of lateral inhibition. Now consider the curve produced when lateral inhibition is low: When the residual *SD* is low, the effect is not present, and increasing residual *SD* does not lead to the occurrence of the effect. In sum, lateral inhibition is much more critical to explaining the attraction effect than is amount of residual *SD*.

A complete explanation for similarity and attraction effects must include a formal explanation for their intricate interactions. The next section considers the effects of systematically varying the position of the decoy on similarity and attraction effects.

Similarity and Attraction Interactions

Distance effects. A fundamental implication of the lateral inhibitory explanation for the attraction effect is that it should diminish with psychological distance between the decoy and the dominant option. Refer to Figure 10, and consider moving the decoy from position R through C and straight down the quality attribute to position D. According to MDFT, this movement increases the distance between the decoy and dominant option, causing the lateral inhibition between the decoy and the dominant option to decrease. MDFT predicts that the bolstering effect of the decoy should eventually be eliminated after moving from positions R through C to D.

Table 1 illustrates this effect, where the results are computed from Equation 3 using the same parameters as in Figure 7, except

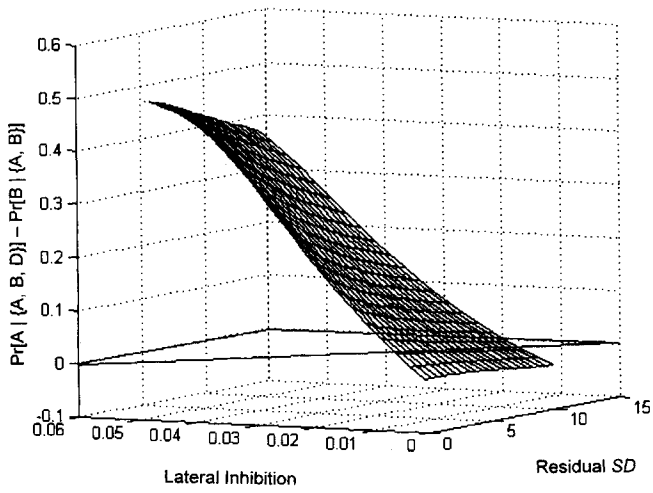

Figure 9. A response surface plot showing the difference between the probability of choosing option A from trinary as compared with binary choice sets, plotted as a function of two key theoretical parameters. One parameter is the strength of the inhibitory connection between the nodes for options A and D, S_{AD}, and the second is the standard deviation of the residuals for the irrelevant attributes, ϵ. The attraction effect occurs when the difference Pr[A|{A, B, D}] − Pr[A|{A, B}] is positive (above the line).

for the changes indicated in the table. The first column represents the value of the decoy option on the quality attribute, and the second column represents the lateral inhibition between option A and the decoy. The last two columns show the probabilities of choosing option A from the binary and trinary choice sets at asymptote. The first and last rows represent decoy positions R and D, respectively, in Figure 10, and the intermediate rows represent intermediate positions for option C. As can be seen in the table, the attraction effect gradually disappears with distance. In agreement with this prediction, the attraction effect has been empirically

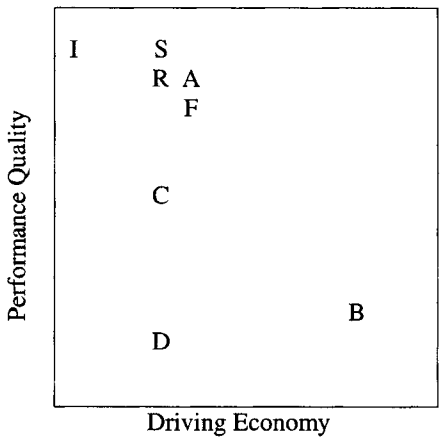

Figure 10. A graphical depiction of possible placement of decoys in the attribute space. The horizontal axis represents the value of each car on the driving economy attribute, and the vertical axis represents the value of each car on the performance quality attribute. Each car is then represented as a point in this two-dimensional space.

Table 1
Effect on the Size of the Attraction Effect of Increasing the Distance Between Decoy Option and Dominant Option

Quality	A to decoy inhibition	Pr(A|A, B)	Pr[A|{A, B, C}]
3.00 (R)	−.025	.55	.65
1.50 (C)	−.008	.55	.63
.75 (C)	−.003	.55	.58
.50 (D)	−.001	.55	.55

found to decrease and eventually disappear with increasing distance (T. B. Heath & Chatterjee, 1991; Wedell, 1991).

Range versus frequency decoys. A more refined prediction can be tested by considering the differential effects of decoys R versus F shown in Figure 10. The range decoy, R, is dominated by A because it has the same quality but worse economy than A. It is called a *range decoy* because adding it to the choice set increases the range on the economy dimension. The frequency decoy, F, is dominated by A in that it has the same economy as A but worse quality. It is called a *frequency decoy* because adding it to the choice set increases the frequency of items below A. Using range and frequency decoys, Huber et al. (1982) found that the decoy effect was stronger with range decoys as compared with frequency decoys.

According to MDFT, changing the position of the decoy changes the similarity of the decoy to option B. Notice in Figure 10 that the range decoy is further away from B than the frequency decoy. The principle of lateral inhibition states that inhibition decreases as a function of distance. This implies that the inhibitory connection between nodes for options F and B is stronger than the inhibitory connections between nodes for options R and B.

The above explanation for the differences between range and frequency decoys was verified by computing the predictions of MDFT using different inhibitory connections in the feedback matrix **S** for range and frequency decoys. Furthermore, to check the robustness of the theory's predictions concerning the attraction effect, the internally controlled stopping rule was used to compute the probabilities for this application (but note that similar results are obtained using Equation 3). For the range decoy, the following **S** matrix shown was used:

$$\mathbf{S} = \begin{array}{c} \\ \text{A} \\ \text{R} \\ \text{B} \end{array} \begin{array}{c} \begin{array}{ccc} \text{A} & \text{R} & \text{B} \end{array} \\ \begin{bmatrix} .95 & -.09 & -.001 \\ -.09 & .95 & -.003 \\ -.001 & -.003 & .95 \end{bmatrix} \end{array}$$

For the frequency decoy, a slightly different **S** matrix was used:

$$\mathbf{S} = \begin{array}{c} \\ \text{A} \\ \text{F} \\ \text{B} \end{array} \begin{array}{c} \begin{array}{ccc} \text{A} & \text{F} & \text{B} \end{array} \\ \begin{bmatrix} .95 & -.09 & -.001 \\ -.09 & .95 & -.02 \\ -.001 & -.02 & .95 \end{bmatrix} \end{array}$$

Notice that the only difference between these two matrices is the inhibition between the decoy and B. For the range decoy the value was −.003, and for the frequency decoy it was −.02. Once again, referring to Figure 10, these numbers reflect the fact that F is closer to B than is R. Thus, the only change needed to capture the different effects of range versus frequency decoys was a change in

lateral inhibition that reflects the distances of the objects in attribute space.

Figure 11 shows the predictions of MDFT under various choice conditions for range and frequency decoys. Each point in the figure represents a pair of probabilities, Pr[A | {A, B}], Pr[A | {A, B, C}], obtained from the same choice condition. In Figure 11, the regularity principle requires all of the points to lie on or below the identity line. Violations of regularity occur when any of the points appear above the identity line. As can be seen in Figure 11, the model predicts that adding either a frequency or a range decoy produces violations of regularity across the entire range of binary choice probabilities. Furthermore, the range decoy has a larger effect as compared with the frequency decoy, consistent with the findings of Huber et al. (1982).

Inferior decoys. A final test MDFT is obtained by considering the effect of adding what is called an *inferior decoy* to the choice set. For example, consider the decoy labeled "I" in Figure 10. Technically, this is a competitive option because it is superior to option A on quality. Practically, however, it is inferior to option A because the small advantage in terms of quality is offset by a large disadvantage in terms of economy. Inferior decoys are like dominated decoys in the sense that they are rarely ever chosen. Huber and Puto (1983) examined the effects of inferior decoys and found that they produced attraction effects similar to dominated decoys.

Note that if option I shown in Figure 10 is shifted horizontally over to the right toward the position of option S in Figure 10, then the inferior option changes into a highly competitive option. Huber and Puto (1983) also examined the effect of gradually changing the inferior option into a competitive option in this manner. They found that the proportion of choices for option A decreased, the proportion of choices for option I increased, but the proportion of choices for option B changed very little. In other words, both attraction and similarity effects were demonstrated within the same study.

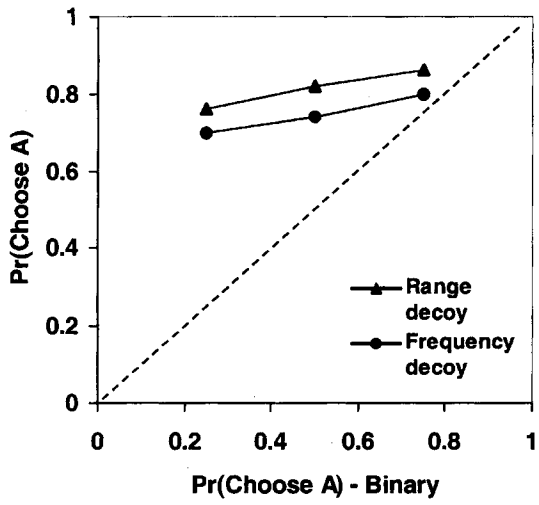

Figure 11. A plot showing results from multialternative decision field theory under various choice conditions for range and frequency decoys. The abscissa represents the probability that the dominating option (A) is chosen from a binary choice set. The ordinate represents the probability that A is chosen out of a trinary choice set that includes a decoy. Each point in the figure represents a pair of probabilities (Pr), Pr[A|{A, B}], Pr[A|{A, B, C}], obtained from the same choice condition. The line connected by solid circles represents the probability that A is chosen when the frequency decoy was presented for various choice conditions. The line connected by the solid triangles represents the probability that A is chosen when the range decoy was presented for various choice conditions. The diagonal identity line represents the separation of where the attraction effect does and does not occur. Any points above the line represent areas where the effect does occur and below where it does not.

These interactions are precisely the effects expected from MDFT. Table 2 shows the results at asymptote computed from Equation 3 using the same parameters as used to produce Figures 4 and 7, except for changes in the value of economy for the inferior option. The first column shows the gradual shifts in the value of the economy attribute from .85 (i.e., the value used to define option S in Figure 10) to .55 (i.e., the value used to define option I in Figure 10). The second and third rows show the probabilities of choosing options A and B from the trinary choice set {A, B, I}. As can be seen in the table, decreasing the economy drastically increased the probability of choosing option A, while the probability for B changed very little. Also note that an attraction effect is produced by the inferior option when the economy was very low in value. These predictions are in accord with the findings by Huber and Puto (1983).

Although earlier choice models have been proposed specifically to account for the attraction effect (e.g., Ariely & Wallsten, 1995; Dhar & Glazer, 1996), only MDFT has been simultaneously applied to both the similarity and attraction effects and their interactions. Furthermore, these earlier models never attempted to formally explain another finding known as the compromise effect described next.

Predictions for the Compromise Effect

The compromise effect presents a real challenge for MDFT because no special mechanisms were built into the theory to produce this effect. Nevertheless, the same mechanisms that MDFT used to explain the similarity and attraction effect work together to produce the compromise effect.

The experimental conditions used to produce the compromise effect required a few changes in model parameters. First, the values in the matrix **M** were changed to represent the new location of compromise option C in Figure 1 as follows. According to the value structure shown in Figure 1, car A is high on quality and low on economy whereas car B is just the opposite. Car C is in between these two, being inferior to B on economy and inferior to A on quality. This pattern of evaluations can be represented in a simple form by the value matrix:

$$\mathbf{M} = \begin{matrix} \\ \\ \\ \end{matrix} \begin{matrix} \text{E} & \text{Q} \\ \end{matrix} \begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{bmatrix} \begin{matrix} \text{A} \\ \text{C} \\ \text{B} \end{matrix}$$

Once again, the precise numerical values are not critical to produce the predicted pattern from the model as long as the compromise position indicated in the matrix is satisfied.

The inhibitory connections in the feedback matrix also need to be changed to reflect the equal distances between the compromise option and the two extreme options. The self-connections were set to $S_{ii} = .94$ (as before), the inhibitory connections between the two

extreme options were set to $S_{AB} = S_{BA} = -.001$ (as before), and the inhibitory connections between the compromise and each extreme were set to $S_{AC} = S_{CA} = S_{CB} = S_{BC} = -.025$. Note that all these parameters are exactly the same as those used to produce the attraction effect in Figure 8. The only difference is that the inhibitory connections between A and C are now set to the same values as the inhibitory connections between C and B, reflecting the fact that the compromise is placed in between the two extremes.

Finally, the probability of attending to the quality attribute was set equal to the probability of attending to the economy dimension ($w_E = w_Q = .45$). This was necessary to meet the antecedent conditions for the compromise effect. Using these parameters forced all of the binary choice probabilities equal to .50. The remaining model parameters used to compute the predictions from Equation 3 were assigned exactly the same values as those used to produce the attraction effect shown in Figure 8.

Figure 12 shows the predictions of MDFT for the compromise effect. The figure plots choice probability as a function of deliberation time separately for different choice alternatives. The upper curve indicated with a "+" symbol represents the choice probability for the compromise car C out of the trinary choice set. The overlapping two lower curves represent the choice probabilities for the two extreme options, A and B, from the trinary choice set. The choice probabilities from the binary choice set are not shown because all three are exactly equal (.50). Figure 12 shows that the probability of choosing the complement alternative, C, grows over time to become larger than both A and B. Consistent with previous research, the model correctly predicts that the probability of choosing the compromise from the trinary set is higher than the extremes, despite the fact that the binary choice probabilities are all equal.[6]

Once again, lateral inhibition is crucial for producing this effect. If the lateral inhibitory interconnections are eliminated, then the predicted compromise effect disappears. But the explanation is quite different from that used to explain the attraction effect. Lateral inhibition does not bolster the mean preference of the compromise alternative because all three options are assigned equally attractive weighted values. The mean valence input is zero for all, and multiplication of the feedback matrix, **S**, by the zero mean input produces zero mean output, and thus the mean preference state remains at zero over time.

Table 2
Effect of Inferior Option on Trinary Choice Probabilities

Economy for I	Pr[A|{A, B, I}]	Pr[B|{A, B, I}]
.85	.33	.43
.75	.50	.42
.65	.57	.40
.55	.61	.38

Note. The first column refers to the value of the **M** matrix corresponding to the first attribute, economy, for the inferior option I. The first row corresponds to the position of option S in Figure 10, and the last row corresponds to the position of option I in Figure 10. The second and third rows correspond to positions of inferior alternatives shifted horizontally along the line from I to S. The binary choice probability of choosing A over B was uniformly equal to .55 for all rows.

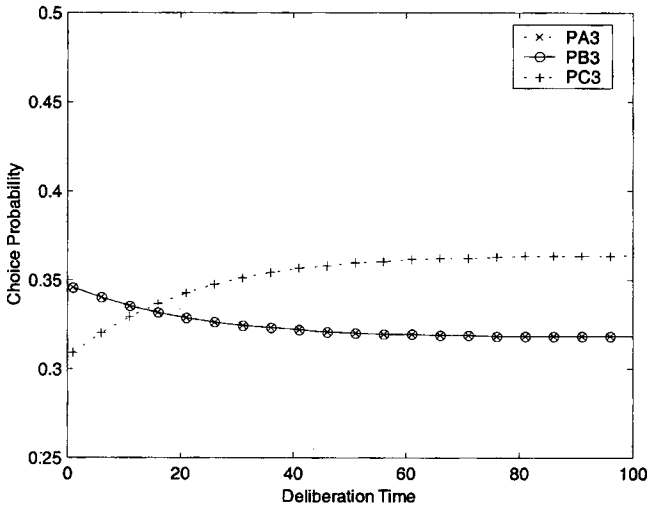

Figure 12. Predictions of multialternative decision field theory for the compromise effect. Choice probability is plotted as a function of deliberation time separately for different choice alternatives. The upper curve indicated with a "+" symbol represents the choice probability for the compromise car C (PC3) out of the trinary choice set. The overlapping two lower curves represent the choice probabilities for the two extreme options, car A (PA3) and car B (PB3), from the trinary choice set.

The source of the effect lies in the operation of lateral inhibition on the momentary fluctuations in valence. The negative inhibitory connections between the nodes for the compromise and extreme options cause the compromise to be negatively correlated with both of the extremes. This in turn causes the differences between the compromise and extreme option A to be positively correlated with the difference between the compromise and extreme option B. Finally, the positive correlation for the compromise provides a probabilistic advantage over the extreme option in Equation 3.

The probabilistic advantage produced by the positive correlation can be seen in Figure 13. The left and right panels show the equal density contours from the multivariate normal densities used in Equation 3 to compute the choice probabilities for the trinary choice set. The left panel shows the contour for an extreme option (e.g., A), and the right panel shows the contour for the compromise option C. Choice probability is related to the area above and to the right of the zero preference state on the vertical and horizontal axes. Note that the relatively stronger positive correlation for the compromise rotates the elliptical contour further into the upper right corner, thus increasing the choice probability for the compromise option.

At this point, we can examine one of the most important issues concerning the explanation of the similarity, attraction, and com-

[6] To check the robustness of the predictions for MDFT, the predictions were computed again using the internally controlled stopping rule and slightly different parameters (see Appendix C for details). The parameters were assigned in a manner to guarantee that all three binary choice probabilities were equal to .50. The results for the trinary choice set produced the same pattern of results as shown in Figure 12: The probability of choosing the compromise (.48) exceeded the probability of choosing either the upper extreme option A (.27) or the lower extreme option B (.25).

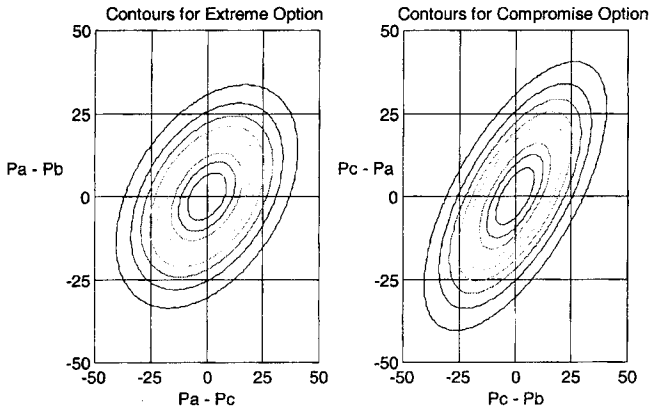

Figure 13. A contour plot showing the effects on choice probability of correlation among valences produced by the primary dimensions. The left and right panels show the equal density contours from the multivariate normal densities used in Equation 3 to compute the choice probabilities for the trinary choice set. The left panel shows the contour for an extreme option (e.g., A), and the right panel shows the contour for the compromise option C. Choice probability is related to the area above and to the right of the zero preference state on the vertical and horizontal axes.

promise effects: How robust are the predictions for all three effects? To answer this question, we first computed the effects predicted by MDFT for each pair of values for the two key parameters across the entire range of the parameter space. Then for each pair of parameter values, we determined whether all three effects were simultaneously present in the appropriate direction. The results of this response surface analysis are shown in Figure 14, where the vertical axis of Figure 14 represents 22 equally spaced levels of residual standard deviation, the horizontal axis represents 22 equally spaced levels of lateral inhibition, and the 22 by 22 matrix covers the entire range of parameter values from Figures 6 and 9. Each solid point in this figure indicates a pair of parameter values that produces all three effects.

From Figure 14, it is apparent that massive regions of the parameter space (i.e., combinations of lateral inhibition and residual *SD*) exist where all these effects are expected.

The compromise effect occurs when the compromise option hurts or takes away shares equally or symmetrically from both extreme alternatives (Simonson & Tversky, 1992; Tversky & Simonson, 1993). However, not all stimuli were precisely designed to produce this symmetry. Simonson and Tversky (1992) also reported unequal or asymmetric effects, called *polarization effects,* in which the compromise hurt one extreme more than the other extreme. The polarization effect also can be accommodated within MDFT by relaxing the assumption that the psychological distances between the compromise option and each extreme option are exactly equal. To demonstrate a polarization effect, the trinary choice probabilities are recomputed from Equation 3 using exactly the same parameters as used in Figure 12 with the following single exception: The inhibitory connection between option C and B was changed to $S_{BC} = -.01$ (slightly less than the inhibitory connection $S_{AC} = -.025$ between options C and A). The results for the binary choices were unchanged, and the results at asymptote for the trinary set produced a polarization effect. The compromise C is chosen with probability .35, the extreme option A is chosen with the same probability .35, but the probability of choosing the other extreme B was reduced to .30. Thus, allowing asymmetry in the lateral inhibition connections reproduces the polarization effect reported by Simonson and Tversky (1992).

Tversky and Simonson (1993) developed the context-dependent advantage model to explain the compromise effect and attraction effects. However, there are three advantages to be claimed for MDFT over this earlier theory. First, as shown in Appendix A, the context-dependent advantage model fails to account for the similarity effect. Second, this earlier theoretical treatment never attempted to formally address the interactions between similarity and attraction (see Tables 1 and 2). Third, the context-dependent advantage model provides only an algebraic description of the evaluation of multialternative choice options, and it does not have any mechanism for describing the predecisional search measures, described next.

Predecisional Search

It is commonly accepted among decision researchers that different decision strategies are used depending on the number of alternatives presented in the choice set (cf. Payne, Bettman, & Johnson, 1993; Wright & Barbour, 1977). At the beginning, when a large number of alternatives are being considered, a simple elimination process is used to quickly screen out options that are unacceptable on the first couple of attributes. Later, when only a few competitive options remain after this initial screening, a more deliberative compensatory process is used to make the final selection. For example, consumers may start by using an EBA-type

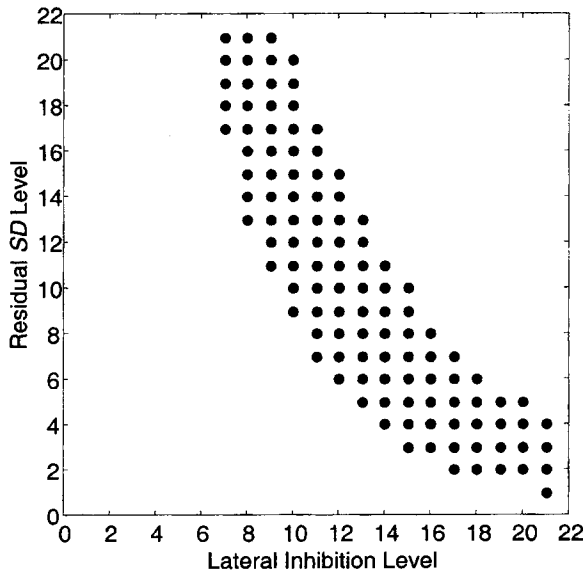

Figure 14. A plot of all the parameter values that simultaneously reproduce all three effects. The vertical axis represents 22 equally spaced levels of residual standard deviation, the horizontal axis represents 22 equally spaced levels of lateral inhibition, and the 22 by 22 matrix covers the entire range of parameters examined in Figures 6, 9, and 15. Each solid point in this figure indicates a pair of parameter values that produces all three effects.

choice rule (Tversky, 1972) and then later switch to a weighted additive difference rule (Tversky, 1969).

A large amount of evidence has accumulated over the past 20 years for strategy-switching dependent on choice set size (see Payne et al., 1993). One index used to monitor strategies is the proportion of information searched, defined as the ratio of the number of cells considered out of the total the number of cells available within an alternative by attribute value matrix. This index has been found to decrease as the choice set size increases. Another important index is the direction of search, defined as the ratio (between-attribute comparisons − within-attribute comparisons)/(within-attribute comparisons + between-attribute comparisons). This index has been found to decrease in the negative direction favoring more within-attribute comparisons as the choice set size increases. Both of these findings are consistent with the idea of strategy-switching from a more thorough compensatory process to a quick elimination process as the choice set size increases. Other process measures based on eye movements and memory recall of values also support this general idea (see I. P. Levin & Jasper, 1995; Russo & Rosen, 1975).

In this section, we show how MDFT mimics strategy-switching simply by adding a second lower elimination boundary to the internally controlled stopping rule (an idea first proposed by Ratcliff, 1978). Previous applications of the internally controlled stopping rule considered only the case where there was an upper acceptance boundary. Now consider the case where there is only a lower rejection boundary and no upper acceptance boundary. Figure 15 illustrates an example for a hypothetical deliberation process involving five new cars that are evaluated along three attributes. The abscissa represents time and the ordinate represents preference. The lower reject boundary is placed at −50 in the figure. Each line represents the evolution of preferences computed from Equation 2 for a particular car option over time. The vertical lines represent shifts of attention from one attribute to another during the deliberation process. In this case, the first attribute is economy, the second comfort, and the third quality. As can be seen, while focusing on the first attribute, economy, two of the options (B, E) were eliminated from the choice set. With the shift to the second attribute, comfort, no items were eliminated. However, with the shift to the third attribute, quality, two more options (C, D) were eliminated. Finally, the choice is based on the last remaining option (A).

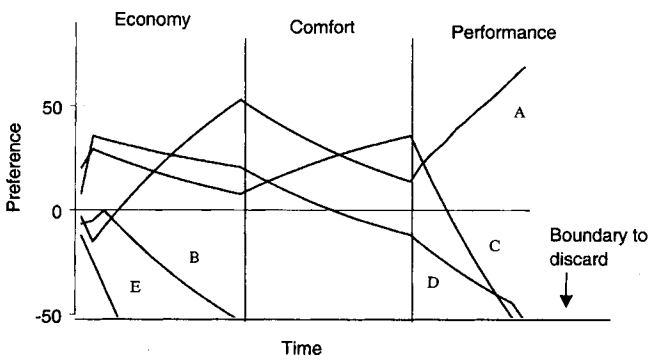

Figure 15. An illustration of the essential ideas for a hypothetical deliberation process involving five new cars that are evaluated along three attributes. The abscissa represents time and the ordinate represents preference. Each line represents the evolution of preferences for a particular car option over time. The vertical lines represent shifts of attention from one attribute to another during the deliberation process. The first attribute is economy, the second comfort, and the third quality.

The complete version of this decision process uses both an upper acceptance boundary and a lower rejection boundary. The lower boundary represents the criterion level of preference used to reject options from consideration. The upper boundary represents the criterion level of preference used to select an option as the final choice. But in this case, there are two ways an option can be chosen: One is to be the first to cross the upper boundary (stopping the search), and if that fails for all options then the other is to be the last to survive rejection.

The double-boundary version of MDFT can mimic strategy-switching by allowing the lower reject boundary to change depending on the number of options initially presented to the decision maker. When the choice set is large, the lower boundary is set equal to a small distance below the neutral point (zero preference state), thus allowing quick rejection of inferior alternatives and only allowing superior alternatives to survive further consideration. When the choice set is small, then the lower boundary is set farther below the neutral point to avoid eliminating any option too quickly and every alternative receives thorough consideration. In general, the criterion for rejecting becomes more lenient (closer to neutral) as the choice set size increases.

To show how this simple idea mimics strategy-switching phenomena, a simulation was performed using the choice sets presented in one of the original predecisional processing studies by Payne, Braunstein, and Carroll (1978). In this study, participants were presented 2, 7, or 12 alternatives described by 12 attributes, with one of three (low, medium, or high) evaluations assigned to each alternative–attribute cell. Furthermore, the alternatives were constructed so that each alternative was high on four attributes, medium on four, and low on four, and the distribution varied across alternatives so that no alternative dominated another.

One thousand simulated decision makers were run on the above multiple-choice task using Equation 2 to generate the preferences and double boundaries to make the final decision (see Appendix C for details). The criterion for rejection was set to high, medium, and low magnitudes for the 2, 7, and 12 alternative choice sets, respectively. The main results for this simulation are shown in Figure 16. The steeply declining curve shows the results produced by the 12-alternative choice set, the flat curve shows the results for the 2-alternative choice set, and the intermediate curve shows the results for the 7-alternative choice set.

For the 12-alternative set, the majority of options are rejected very quickly on the basis of comparisons within only the first or second attributes. For example, the probability of rejecting any one of the 12 options on the basis of the first attribute is .45, and the probability of rejecting an option on second attribute (given that it survived the first attribute) is .40. On the average, 8 out of 12 alternatives are eliminated on the basis of the first two attributes alone. But as the number of attributes processed in increases, the probability of rejection rapidly declines, so that the probability of rejecting any one of the superior options that survives until the seventh attribute is only .10 and this rejection probability gradually approaches zero. Thus the double-boundary model behaves like an elimination process when a large number of alternatives are available in the choice set.

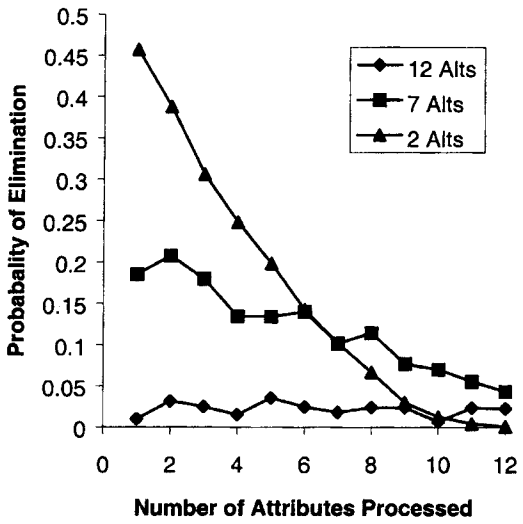

Figure 16. Strategy switching as a function of choice set size. The ordinate represents the probability that an option is rejected, conditioned on the event that it is not already rejected (i.e., number of options rejected divided by the number of options still remaining). The abscissa represents the number of attributes that have been processed during deliberation on a choice set. The steeply declining curve shows the results produced by the 12-alternative choice set, the flat curve shows the results for the 2-alternative choice set, and the intermediate curve shows the results for the 7-alternative choice set. Alts = alternatives.

For the 2-alternative set, the likelihood of eliminating an option is very small. In this case, the probability of rejection remains approximately constant at a very low average level equal to .01 across the 12 attributes. The probability that each option survives past the first 11 attributes is approximately .90. This allows time to consider all of the trade-offs between the two alternatives very carefully and thoroughly on the basis of a very large number of attributes before reaching a decision.

Using these simulation results, we also can compute the indices for the two main predecisional search measures mentioned earlier. The proportion of information search is a decreasing function of the rejection rate, and the rejection rate increases with choice set size. For this simulation, the proportion of information searched decreased from a high level of .89 for the 2-alternative set to .45 for the 7-alternative set, down to a low level of .25 for the 12-alternative set. The basic trend for the direction of search is to increase the proportion of within-attribute comparisons as the choice set size increases. For this simulation, the direction of search index changed from −.38 for the 2-alternative set, to −.79 for the 7-alternative set, and to −.89 for the 12-alternative choice set. This agrees qualitatively with the basic pattern of findings from previous research.

Summary of Empirical Applications

Thus far we have shown that MDFT provides a uniform and comprehensive account of the critical empirical findings from multialternative preferential choice studies. This includes explanations for similarity effects, attraction effects, compromise effects, and their subtle interactions. Most important, as shown in Figure 14, all three of these basic effects can be explained by using a common set of parameters. A closer examination of Figure 14 yields some interesting relations between the theoretical parameters and the occurrence of the three basic effects. First, we find that relatively low lateral inhibition and residual *SD* facilitate the similarity effect, whereas relatively high lateral inhibition and residual *SD* facilitate the compromise effect. High inhibition also facilitates the attraction effect, but the residual *SD* plays a minor role. This leads to an interesting prediction concerning individual differences: Individuals showing the strongest similarity effect may evince weaker compromise effects and vice versa. Thus, there should be a negative correlation between attraction effects and similarity effects as opposed to a positive correlation between the attraction and compromise effects. Confirmatory results supporting these individual difference predictions have been reported by Wedell (1993).

MDFT also makes interesting predictions regarding the dynamical aspects of the decision process. As can be seen in Figures 8 and 12, a directly testable prediction derived from the present theory is that both the attraction effect and the compromise effect should be attenuated by time pressure. In contrast, the similarity effect is not expected to decrease under time pressure, and instead this effect should remain fairly constant across time. Thus MDFT predicts qualitatively different effects of time pressure on choice probability for the three major findings. Although a formal test of this hypothesis remains to be performed, two separate lines of preliminary evidence converge to support the general idea that attraction and compromise effects increase with longer deliberation times. First, Wedell (1993) found attraction effects to be positively correlated with choice response time. Second, Simonson (1989) reported that attraction effects are enhanced when participants are motivated to carefully deliberate about their decisions.

Finally, we showed how MDFT assimilates findings regarding the predecisional search processes used to filter a large set of options down to a smaller competitive set. By using both an upper acceptance boundary and a lower rejection boundary, MDFT can emulate findings regarding predecisional search processes reported in the literature. Specifically, at the beginning of the decision-making process, when a large number of alternatives are being considered, a simple elimination process quickly screens out options that are unacceptable on the first couple of attributes. Later, when only a few competitive options remain after this initial screening, a more deliberative compensatory process makes the final selection. The last section compares MDFT with earlier theories of multialternative preferential choice.

Comparisons With Other Models

Decision Theories

During the past 20 years, decision theorists have made significant theoretical progress toward the development of theories for preferential choice. Some models were successful for explaining the similarity effect (Candel, 1997; Edgell & Geisler, 1980; Mellers & Biagini, 1994; Tversky, 1972; see also De Soete et al., 1989); others were successful for explaining the attraction effect (Ariely & Wallsten, 1995; Dhar & Glazer, 1996) or the compromise effect (Tversky & Simonson, 1993). However, a common formal theoretical explanation for all three pivotal effects has

eluded past researchers. Table 3 provides a comparison among models.

On the one hand, models specifically designed to explain similarity effects cannot explain violations of regularity resulting from the attraction effect. On the other hand, models specifically designed to explain attraction effects cannot explain violations of independence from irrelevant alternatives produced by the similarity effect. As we showed in the previous sections, MDFT succeeds in providing a coherent explanation for all three classes of phenomena as well as the subtle interactions among these effects. For example, MDFT not only provides an explanation of the main effect for the dominated decoy, but also accounts for differences between range and frequency decoys and the positions of inferior decoys on the attraction effect.

Previous Artificial Neural Networks

Several artificial neural network models of decision making have recently been proposed that are also dynamic in nature. Leven and Levine (1996) proposed a multiattribute neural network decision model to explain consumer choice behavior; however, it was not applied to any of the major empirical findings on multialternative preferential choice discussed here. Another dynamic model proposed by Usher and Zakay (1993) was designed to implement an EBA choice process using a neural network framework. It provides an explanation for the similarity effect and for some search process results. However, like the EBA model, no explanations for the attraction effect and the compromise effect follow from this theory.

Sequential Sampling Models

Decision field theory follows a growing movement in cognitive psychology to model decision processes according to a sequential sampling process (Ashby, 2000; R. Heath, 1984; Link, 1992; Nosofksy & Palmeri, 1997; Ratcliff, 1978; Ratliff & Rouder, 1998; Ratcliff, Van Zandt, & McKoon, 1999; Smith, 1995). However, the present formulation provides a couple of innovations to this earlier work. One innovation is that decision field theory accumulates *valences* (comparative affective evaluations) whereas the earlier models accumulated *evidence* (likelihoods of competing hypotheses). A second major innovation is the incorporation of the principle of lateral inhibition to define the feedback matrix **S** in the linear stochastic dynamic system (see Equation 2).[7]

Table 3
Comparison of Models and Major Phenomena

	Phenomena			
Theory	Similarity	Attraction	Compromise	Search process
Context-dependent advantage model	No	Yes	Yes	No
Elimination-by-aspects model	Yes	No	No	Yes
Thurstone model	Yes	No	No	No
Earlier neural network	Yes	No	No	Yes
Decision field	Yes	Yes	Yes	Yes

Implications for Future Research

MDFT not only provides a more complete explanation of the major findings from past research on multialternative choice than previous theories, but also generates new testable predictions and provides new directions for research. For example, MDFT predicts that the compromise effect should gradually turn into a similarity effect as the compromise option is moved along the diagonal toward one of the extreme options, as it is in Figure 10 (analogous to the similarity–attraction interactions shown in Table 2). Another interesting prediction is that the attraction effect results from a buildup of lateral inhibition and therefore should take time to build up. Although preliminary evidence supports the latter hypothesis (Simonson, 1989; Wedell, 1993), more rigorous tests are required to rule out possible explanations based on random guessing under time pressure. We are currently beginning a new program of research to examine these and other predictions from MDFT. Several experimental tests of predictions derived from the original version of decision field theory have recently appeared, and the results were found to be in agreement with theory (see Diederich & Busemeyer, 1999; Dror, Busemeyer, & Basola, 1999). Our goal for the present article is to lay down the basic theoretical foundation for future model tests and model comparisons.[8]

Conclusion

We have recast decision field theory (Busemeyer & Townsend, 1993) into a connectionist framework and extended it to encompass multialternative decision making. Two earlier important theories (weighted additive utility; Thurstone choice theory) were shown to be derivable as special cases of our theory. The central thrust of the study then probed MDFT relative to three essential experimental findings concerning multichoice situations: the similarity effect, the attraction effect, and the compromise effect. MDFT is the only formal and quantitatively specified theory that has successfully explained all three effects; other future testable predictions were considered. In addition, we demonstrated that MDFT can simulate behavior associated with Tversky's (1972) EBA process and we discussed the critical time-oriented facets and potential predictions of MDFT. Finally, we compared MDFT with other multialternative choice models, in particular with those possessing artificial neural network interpretations. Decision field theory is based on a simple set of theoretical assumptions, all of them founded on long-standing psychological principles in motivation, decision making, and information processing. Over the past several years, this theory has begun to demonstrate itself capable of explaining a broad spectrum of classical results and making new testable predictions, many of which have been investigated and

[7] Ashby (1989) also mentioned the use of lateral inhibition in his stochastic version of general recognition theory.

[8] Unfortunately most of the data from previous experiments are not suitable for quantitative tests because of the small number of experimental conditions and the small sample sizes within each condition. For this reason we have considered only the qualitative predictions of MDFT. Quantitative tests of decision field theory have been performed in other applications (e.g., Busemeyer & Townsend, 1993; Dror et al., 1999), and new experiments are underway to provide both qualitative as well as quantitative tests of time pressure and attraction effects.

confirmed. The present work suggests that major multialternative choice phenomena naturally ensue from the extended theory.

References

Ariely, D., & Wallsten, T. S. (1995). Seeking subjective dominance in multidimensional space: An exploration of the asymmetric dominance effect. *Organizational Behavior and Human Decision Processes, 63,* 223–232.

Aschenbrenner, K. M., Albert, D., & Schmalhofer, F. (1984). Stochastic choice heuristics. *Acta Psychologica, 56(1–3),* 153–166.

Ashby, F. G. (1989). Stochastic general recognition theory. In D. Vickers & P. L. Smith (Eds.), *Human information processing: Measures, mechanisms and models* (pp. 435–457). Amsterdam: Elsevier.

Ashby, F. G. (1992). *Multidimensional models of perception and cognition.* Hillsdale, NJ: Erlbaum.

Ashby, F. G. (2000). A stochastic version of general recognition theory. *Journal of Mathematical Psychology, 44,* 310–329.

Ashby, F. G., & Maddox, T. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 50–71.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93,* 154–179.

Batsell, R. R., & Polking, J. C. (1985). A new class of market share models. *Marketing Science, 4,* 177–198.

Bettman, J. R., Johnson, E. J., & Payne, J. W. (1991). Consumer decision making. In T. S. Robertson & H. H. Kassarjian (Eds.), *Handbook of consumer behavior* (pp. 50–84). Englewood Cliffs, NJ: Prentice Hall.

Bhattacharya, R. N., & Waymire, E. C. (1990). *Stochastic processes with applications.* New York: Wiley.

Bock, R. D., & Jones, L. V. (1968). *The measurement and prediction of judgment and choice.* San Francisco: Holden-Day.

Bockenholt, U. (1992). Multivariate models of preference and choice. In F. G. Ashby (Ed.), *Multidimensional Models of Perception and Cognition,* pp. 89–111. Hillsdale, NJ: Erlbaum.

Busemeyer, J. R., & Diederich, A. (2000). *Survey of decision field theory.* Manuscript submitted for publication.

Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences, 23,* 255–282.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic cognition approach to decision making. *Psychological Review, 100,* 432–459.

Candel, M. J. J. M. (1997). A probabilistic feature model for unfolding tested for perfect and imperfect nestings. *Journal of Mathematical Psychology, 41,* 414–430.

Colonius, H. (1984). *Stochastische theorien individuellen wahlverhastens* [Stochastic theories of individual choice behavior]. Berlin, Germany: Springer-Verlag.

Cornsweet, T. (1970). *Visual perception.* Academic Press.

Cox, D. R., & Miller, H. D. (1965). *The theory of stochastic processes.* London: Chapman & Hall.

De Soete, G., Feger, H., & Klauer, K. C. (Eds.). (1989). *New developments in probabilistic choice modeling.* Amsterdam: North-Holland.

Dhar, R., & Glazer, R. (1996). Similarity in context: Cognitive representations and violation of preference and perceptual invariance in consumer choice. *Organizational Behavior and Human Decision Processes, 67,* 280–293.

Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology, 41,* 260–274.

Diederich, A., & Busemeyer, J. R. (1999). Conflict and the stochastic dominance principle of decision making. *Psychological Science, 10,* 353–359.

Dosher, B. A. (1984). Discriminating pre-experimental (semantic) from learned (episodic) associations: A speed accuracy study. *Cognitive Psychology, 16,* 519–555.

Dror, I. E., Busemeyer, J. R., & Basola, B. (1999). Decision making under time pressure: An independent test of sequential sampling models. *Memory & Cognition, 27,* 713–725.

Edgell, S. E., & Geisler, W. S. (1980). A set-theoretic random utility model of choice behavior. *Journal of Mathematical Psychology, 21(3),* 265–278.

Fischer, G. W., Jia, J., & Luce, M. F. (2000). Attribute conflict and preference uncertainty: The RandMAU model. *Management Science, 46,* 669–684.

Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 846–858.

Grossberg, S. (1982). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control.* Hingham, MA: Reidel.

Grossberg, S. (1988). *Neural networks and natural intelligence.* Cambridge, MA: MIT Press.

Grossberg, S., & Gutowski, W. E. (1987). Neural dynamics of decision making under risk: Affective balance and cognitive-emotional interactions. *Psychological Review, 94,* 300–318.

Heath, R. (1984). Random-walk and accumulator models of psychophysical discrimination: A critical evaluation. *Perception, 13,* 57–65.

Heath, T. B., & Chatterjee, S. (1991). How entrants affect multiple brands: A dual attraction mechanism. *Advances in Consumer Research, 18,* 768–771.

Hintzman, D. L., & Curran, T. (1997). Comparing retrieval dynamics in recognition memory and lexical decision. *Journal of Experimental Psychology: General, 126,* 228–247.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research, 9,* 90–98.

Huber, J., & Puto, C. (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of Consumer Research, 10,* 31–44.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment.* New York: Free Press.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preference and value tradeoffs.* New York: Wiley.

Lehmann, D. R., & Pan, Y. (1994). Context effects, new brand entry, and consideration sets. *Journal of Marketing Research, 31,* 364–374.

Leven, S. J., & Levine, D. S. (1996). Multiattribute decision making in context: A dynamic neural network methodology. *Cognitive Science, 20,* 271–299.

Levin, I. P., & Jasper, J. D. (1995). Phased narrowing: A new process tracing method for decision making. *Organizational Behavior and Human Decision Processes, 64,* 1–8.

Link, S. W. (1992). *The wave theory of difference and similarity.* Hillsdale, NJ: Erlbaum.

Link, S. W., & Heath, R. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40,* 77–111.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis.* New York: Wiley.

MacKay, D. B., & Zinnes, J. L. (1995). Probabilistic multidimensional unfolding: An anisotropic model for preference ratio judgments. *Journal of Mathematical Psychology, 39,* 99–111.

Marley, A. A. J. (1989). Random utility family that includes many of the "classical" models and has closed form choice probabilities and choice reaction times. *British Journal of Mathematical and Statistical Psychology, 42,* 13–36.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation

model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88,* 375–407.

Mellers, B. A., & Biagini, K. (1994). Similarity and choice. *Psychological Review, 101,* 505–518.

Nosofsky, R. M., & Palmeri, T. J. (1997). Comparing exemplar-retrieval and decision-bound models of speeded perceptual classification. *Perception & Psychophysics, 59,* 1027–1048.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* New York: Cambridge University Press.

Payne, J. W., Braunstein, M. L., & Carroll, J. S. (1978). Exploring predecisional behavior: An alternative approach to decision research. *Organizational Behavior and Human Decision Processes, 22,* 17–44.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R. (1980). A note on modeling accumulation of information when the rate of accumulation changes over time. *Journal of Mathematical Psychology, 21,* 178–184.

Ratcliff, R., & McKoon, G. (1982). Speed and accuracy in the processing of false statements about semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 16–36.

Ratcliff, R., & McKoon, G. (1989). Similarity information versus relational information: Differences in the time course of retrieval. *Cognitive Psychology, 21,* 139–155.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9,* 347–356.

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review, 106,* 261–300.

Ratneshwar, S., Shocker, A. D., & Stewart, D. W. (1987). Toward understanding the attraction effect: The implications of product stimulus meaningfulness and familiarity. *Journal of Consumer Research, 13,* 520–527.

Reed, A. V. (1973). Speed–accuracy tradeoff in recognition memory. *Science, 181,* 574–576.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations.* Cambridge, MA: MIT Press.

Russo, J. E., & Rosen, L. D. (1975). An eye fixation analysis of multialternative choice. *Memory & Cognition, 3,* 267–276.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty, 1,* 7–59.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology, 1,* 54–87.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69,* 99–118.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16,* 158–174.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 24,* 281–295.

Sjoberg, L. (1977). Choice frequency and similarity. *Scandinavian Journal of Psychology, 18,* 103–115.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review, 102,* 567–593.

Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology, 44,* 408–463.

Takane, Y. (1989). Analysis of covariance structures and probabilistic binary choice data. In G. De Soete, H. Feger, & K. C. Klauer (Eds.), *New developments in probabilistic choice modeling* (pp. 139–160). Amsterdam: North-Holland.

Thurstone, L. L. (1959). *The measurement of values.* Chicago: University of Chicago Press.

Townsend, J. T., & Busemeyer, J. R. (1995). Dynamic representation of decision-making. In R. F. Port & T. van Gelder (Eds.), *Mind as motion* (pp. 101–120). Cambridge, MA: MIT Press.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76,* 31–48.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 281–299.

Tversky, A., & Simonson, I. (1993). Context dependent preferences. *Management Science, 39,* 1179–1189.

Usher, M., & Zakay, D. (1993). A neural network model for attribute-based decision processes. *Cognitive Science, 17,* 349–396.

Vickers, D. (1979). *Decision processes in visual perception.* New York: Academic Press.

Vickers, D. B. J., Smith, P., & Brown, M. (1985). Experimental paradigms emphasizing state or process limitations: I. Effects of speed–accuracy tradeoffs. *Acta Psychologica, 59,* 129–161.

Von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research.* Cambridge, England: Cambridge University Press.

Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology, 17,* 767–778.

Wedell, D. H. (1993, November). *Studying effects of different decoys on choice in a within-subject design.* Paper presented at the 34th Annual Meeting of the Psychonomic Society, St. Louis, MO.

Wickelgren, W., Corbett, A., & Dosher, B. (1980). Priming and retrieval from short-term memory: A speed accuracy trade-off analysis. *Journal of Verbal Learning and Verbal Behavior, 19,* 387–404.

Wright, P., & Barbour, F. (1977). Phased decision strategies: Sequels to an initial screening. In M. Starr & M. Zeleny (Eds.), *Multiple criteria decision making: TIMS studies in the management sciences,* (Chap. 6, pp. 1–10). Amsterdam: North-Holland.

(Appendixes follow)

Appendix A

Derivations for Content-Dependent Advantage Model

This appendix proves that the context-dependent advantage model developed by Tversky and Simonson (1993) cannot simultaneously explain the similarity effect and the compromise effect. Tversky and Simonson never attempted to apply the model to the similarity effect, although it clearly can be applied in the same manner as the compromise effect.

Their model requires the definition of four parameters: α = the unique advantage of option A over B; β = the unique advantage of option B over A; γ = unique advantage of option A over S; and λ = the unique advantage of option S over A. Tversky and Simonson (1993) based all their derivations on the case where $\alpha = \beta$, so that according to their model, the options A and B are chosen equally often in the binary case.

Following Tversky and Simonson's (1993) derivation for the compromise effect, the similar option S, as shown in Figure 1, has a unique advantage, λ, over A, which is equal to the unique advantage, γ, of A over S, so that $\gamma = \lambda = \epsilon$. The main difference between the compromise and similarity effect is that for the latter, ϵ is an arbitrarily small magnitude. Applying their formulas to the similarity case yields

$$V(\mathrm{A}) = \alpha + \theta\left[\frac{\alpha}{\alpha + \delta(\beta)} + \frac{\gamma}{\gamma + \delta(\lambda)}\right]$$

$$= \alpha + \theta\left[\frac{\alpha}{\alpha + \delta(\alpha)} + \frac{\epsilon}{\epsilon + \delta(\epsilon)}\right],$$

$$V(\mathrm{B}) = \beta + \theta\left[\frac{\beta}{\beta + \delta(\alpha)} + \frac{\beta + \gamma}{\beta + \gamma + \delta(\alpha + \lambda)}\right]$$

$$= \alpha + \theta\left[\frac{\alpha}{\alpha + \delta(\alpha)} + \frac{\alpha + \epsilon}{\alpha + \epsilon + \delta(\alpha + \epsilon)}\right].$$

In the above formulas, θ is an unknown context weight parameter, and $\delta(x) > x$ is assumed to be a convex function. Strict convexity was required by Tversky and Simonson (1993, p. 1187) to explain the compromise effect.

To explain the similarity effect, it is necessary to have $V(\mathrm{B}) > V(\mathrm{A})$. But this is impossible if $\delta(x)$ is strictly convex. To see this, note that if $\delta(x)$ is strictly convex, then for any $\alpha > 0$, $\delta(\alpha) > \alpha\delta'(0)$, where $\delta'(0)$ is the derivative of δ evaluated at zero. This implies that as $\epsilon \to 0$, $\delta(\alpha + \epsilon)/(\alpha + \epsilon) > \delta(\epsilon)/\epsilon$. But the last inequality implies that

$$\frac{\alpha + \epsilon}{\alpha + \epsilon + \delta(\alpha + \epsilon)} < \frac{\epsilon}{\epsilon + \delta(\epsilon)}, \tag{A1}$$

and finally, this inequality (Equation A1) implies $V(\mathrm{A}) > V(\mathrm{B})$, contrary to the similarity effect. For example, consider the commonly used case where $\delta(x) = x^{\delta}$ is a power function and define $\eta = 1 + (\epsilon/\alpha)$.

$$(\alpha + \epsilon) > \epsilon \to (\eta\alpha)^{\delta - 1} > \epsilon^{\delta - 1} \to \frac{(\eta\alpha)^{\delta}}{\eta\alpha} > \frac{\epsilon^{\delta}}{\epsilon} \to \frac{(\alpha + \epsilon)^{\delta}}{\alpha + \epsilon}$$

$$> \frac{\epsilon^{\delta}}{\epsilon} \to V(\mathrm{A}) > V(\mathrm{B}).$$

For the linear case, it is easy to show that $\delta(x) = \delta x \to V(\mathrm{A}) = V(\mathrm{B})$, again contrary to the desired empirical finding for the similarity effect, $V(\mathrm{B}) > V(\mathrm{A})$. In fact, the inequality (Equation A1) is exactly the same inequality that must be satisfied to produce the compromise effect (see Tversky and Simonson, 1993, p. 1187), hence their need for imposing the convexity property. But this same property rules out the possibility of explaining the similarity effect.

Appendix B

Derivations for Decision Field Theory

This appendix derives the mathematical formula used to compute the probabilities for Equation 3. The valence vector defined by Equation 1 is a linear transformation of the stochastic weight vector, $\mathbf{W}(t)$. This weight vector is assumed to change across time according to a stationary stochastic process. The stationarity assumption for the weights implies that the valence vector ($\mathbf{V}$) is also a stationary stochastic process with mean $E[\mathbf{V}(t)] = E[\mathbf{C}\mathbf{M}_1\mathbf{W}_1(t) + \boldsymbol{\epsilon}(t)] = \mathbf{C}\mathbf{M}_1E[\mathbf{W}_1(t)] + E[\boldsymbol{\epsilon}(t)] = \mathbf{C}\mathbf{M}_1\mathbf{w}_1 + 0 = \boldsymbol{\mu}$ and variance–covariance matrix at each time point given by $Cov[\mathbf{V}(t)] = \mathrm{Cov}[\mathbf{C}\mathbf{M}_1\mathbf{W}_1(t) + \boldsymbol{\epsilon}(t)] = \mathbf{C}\mathbf{M}_1\mathrm{Cov}[\mathbf{W}_1(t)]\mathbf{M}_1'\mathbf{C}' + Cov[\boldsymbol{\epsilon}(t)] = \mathbf{C}\mathbf{M}_1\boldsymbol{\Psi}\mathbf{M}_1'\mathbf{C}' + \boldsymbol{\varsigma} = \boldsymbol{\Phi}$, where $\boldsymbol{\Psi} = \mathrm{Cov}[\mathbf{W}_1(t)] = E[(\mathbf{W}_1(t) - \mathbf{w}_1)(\mathbf{W}_1(t) - \mathbf{w}_1)']$ is the variance–covariance matrix for the primary weights, and $\boldsymbol{\varsigma} = Cov[\boldsymbol{\epsilon}(t)] = E[\boldsymbol{\epsilon}(t)\boldsymbol{\epsilon}(t)']$ is the variance–covariance matrix of the residuals. Analogous to a factor analytic model (cf. Takane, 1989), the residuals are assumed to be uncorrelated with the primary dimensions and uncorrelated with each other. This implies that $\boldsymbol{\varsigma}$ is a diagonal matrix.

The effects of the feedback matrix on the evolution of preference over time can be seen more clearly by expanding Equation 2:

$$\mathbf{P}(t) = \Sigma_{j=0,t-1}\mathbf{S}^j\mathbf{V}(t - j) + \mathbf{S}^t\mathbf{P}(0). \tag{B1}$$

This equation shows that the current preference state can be viewed as a weighted sum of the previous input valences. The weight placed on each previous input is determined by the feedback matrix raised to a power, where the power equals the lag between the current state and the previous input. For this system to be stable, the eigenvalues of the feedback matrix must be less than one in magnitude. In this case, the effect of the feedback matrix decays toward zero as the lag increases in value. Taking expectations and simplifying produces the mean preference over time:

$$\boldsymbol{\xi}(t) = E[\mathbf{P}(t)] = (\mathbf{I} - \mathbf{S})^{-1}(\mathbf{I} - \mathbf{S}^t)\boldsymbol{\mu} + \mathbf{S}^t\mathbf{P}(0). \tag{B2}$$

For stable systems, as $t \to \infty$, then $\boldsymbol{\xi}(t) \to (\mathbf{I} - \mathbf{S})^{-1}\boldsymbol{\mu}$, which is a simple formula for analyzing the effects of the feedback matrix on the asymptotic mean preference state. If weights are identically and independently distributed over time (see Footnote 1), then the variance–covariance matrix of the preference state evolves over time according to $\boldsymbol{\Omega}(t) = \mathrm{cov}[\mathbf{P}(t)] = E\{(\mathbf{P}(t) - E[\mathbf{P}(t)])(\mathbf{P}(t) - E[\mathbf{P}(t)])'\} = \Sigma_{j=0,t-1}\mathbf{S}^j\boldsymbol{\Phi}\mathbf{S}^{j'}$.

On the basis of the multivariate central limit theorem, the distribution of the preference states $\mathbf{P}(t)$ converges to the multivariate normal distribution as the number of steps (t) in Equation B1 becomes large. If the time period between steps is small, then this convergence will occur very rapidly. Given that $\mathbf{P}(t)$ is distributed according to the multivariate normal distribution with mean $\boldsymbol{\xi}(t)$ and variance–covariance matrix $\boldsymbol{\Omega}(t)$, then the probability of choosing A from the set {A, B, C} at fixed time point t is:

$$\Pr[P_A(t) - P_B(t) > 0 \text{ and } P_A(t) - P_C(t) > 0]$$

$$= \int_{\mathbf{X}>0} exp[-(\mathbf{X} - \mathbf{\Gamma})'\mathbf{\Lambda}^{-1}(\mathbf{X} - \mathbf{\Gamma})/2]/(2\pi|\mathbf{\Lambda}|^{.5})d\mathbf{X}, \quad \text{(B3)}$$

where $\mathbf{X} = [P_A(t) - P_B(t), P_A(t) - P_C(t)]'$, $\mathbf{\Gamma} = \mathbf{L}\xi(t)$, $\mathbf{\Lambda} = \mathbf{L}\Omega(t)\mathbf{L}'$ and

$$\mathbf{L} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}.$$

Techniques for integrating Equation B3 are described in Ashby (1992, p. 26). Equation B3 is the mathematical formula used to compute the probabilities in Equation 3.

For binary choices, $\mathbf{L} = [1 \ -1]$, and $\mathbf{\Gamma} = \mathbf{L}\xi(t)$, $\Lambda = \mathbf{L}\Omega(t)\mathbf{L}'$ are both scalars: Γ represents the mean difference in preference, and Λ represents the variance of this difference. For binary choices, Equation B3 reduces to a simple integration of the unidimensional normal distribution. The binary choice probability of choosing A over B at time t is $\Pr[P_A(t) - P_B(t) > 0] = F(\Gamma/\sqrt{\Lambda})$, where $F(x)$ is the standard cumulative normal distribution function.

Parameters used for Figures 4, 7, and 8: The attention weights for the two primary attributes were assumed to fluctuate according to a simple Bernoulli process with probabilities $\pi_1 = .43$ and $\pi_2 = .45$. Based on this assumption, we derive $\mathbf{w}_1 = E[\mathbf{W}_1(t)] = [w_E \ w_Q]'$, and $\mathbf{\Psi} = \text{diag}(\mathbf{w}_1) - \mathbf{w}_1\mathbf{w}_1'$. The values for the $\mathbf{M}_1$ matrix described in the text were then inserted to compute $\boldsymbol{\mu} = \mathbf{C}\mathbf{M}_1\mathbf{w}_1$ and $\mathbf{\Phi} = \mathbf{C}\mathbf{M}_1\mathbf{\Psi}\mathbf{M}_1'\mathbf{C}' + \mathbf{s}$, where $\mathbf{s}$ was initially set to $\mathbf{s} = \text{diag}(.1, .1)$ for Figures 4 and 7. But it was changed to $\mathbf{s} = \text{diag}(10, 10)$ for Figure 8 to examine the robustness of the attraction effect with respect to the residual variance. These matrices were then inserted into Equation B3 to produce the predictions shown in Figures 4, 7, and 8.

Appendix C

Binary Choice Probabilities

This appendix describes how model simulations were performed to compute predictions for the internally controlled stopping rule. At each time step the preference vector for each alternative was updated using the **C** matrix for binary choice given in the text and the following **S** matrix in Equation 2:

$$\mathbf{S} = \begin{bmatrix} .95 & 0 \\ 0 & .95 \end{bmatrix}.$$

Attention shifted between dimensions in that at each time step, with a probability of .5, a column of the **M** matrix was chosen for use in updating the preference vector. Once a column was chosen, an error vector distributed $N \sim (0, 7)$ was added to represent momentary fluctuations in value. The simulation continued until the preference for one option reached an upper boundary of 60. The option that reached the boundary first was the option chosen for that simulated subject. One thousand subjects were simulated for each effect. These same procedures, with the same parameters, were used in all trinary simulations with the exception of the **M, C,** and **S** matrices.

Similarity

Simulations were run exactly as with the binary case discussed above except with different **M, C,** and **S** matrices. The following **M** matrix was used:

$$\mathbf{M}_1 = \begin{array}{c} \\ \text{A} \\ \text{Similar} \\ \text{B} \end{array} \begin{array}{c} \begin{array}{cc} \text{E} & \text{Q} \end{array} \\ \begin{bmatrix} 28.2 & 12.2 \\ 29 & 11 \\ 10 & 30 \end{bmatrix} \end{array}$$

This matrix was chosen such that the values for A and B led to a slight advantage for choosing A over B in the binary case. The **C** matrix was the same as presented in the text and the following **S** matrix was used:

$$\mathbf{S} = \begin{array}{c} \begin{array}{ccc} \text{A} & \text{S} & \text{B} \end{array} \\ \begin{bmatrix} .95 & -.09 & -.001 \\ -.09 & .95 & -.003 \\ -.001 & -.003 & .95 \end{bmatrix} \end{array} \begin{array}{c} \\ \text{A} \\ \text{S} \\ \text{B} \end{array}$$

Values in this matrix were chosen to take into account the relative positions of the options in the multiattribute evaluation space. For example, the lateral inhibition between A and S is greater than A and B because it is closer in space (see Figure 1).

Attraction Effect

Simulations for the attraction effect were exactly the same as for the similarity effect except for the differences in the **M** and **S** matrices. The **M** matrices were chosen such that there was a range of probabilities of choosing A over B in the binary case.

$$\Pr[\text{A}|\{\text{A, B}\}] = .25 \quad \Pr[\text{A}|\{\text{A, B}\}] = .5 \quad \Pr[\text{A}|\{\text{A, B}\}] = .75$$

$$\mathbf{M}_1 = \begin{array}{c} \begin{array}{cc} \text{E} & \text{Q} \end{array} \\ \begin{bmatrix} 22 & 18 \\ 17.6 & 18 \\ 31 & 10 \end{bmatrix} \end{array} \quad \mathbf{M}_1 = \begin{array}{c} \begin{array}{cc} \text{E} & \text{Q} \end{array} \\ \begin{bmatrix} 22 & 18 \\ 17.6 & 18 \\ 30 & 10 \end{bmatrix} \end{array} \quad \mathbf{M}_1 = \begin{array}{c} \begin{array}{cc} \text{E} & \text{Q} \end{array} \\ \begin{bmatrix} 23 & 18 \\ 18.4 & 18 \\ 30 & 10 \end{bmatrix} \end{array}.$$

The **S** matrices for the range and frequency decoys were presented in the text. The rationale for the values in these matrices is given in the text.

Compromise Effect

Simulations for the compromise effect were exactly the same as the decoy effect with different **M** and **S** matrices.

$$\mathbf{M} = \begin{array}{c} \begin{array}{cc} \text{E} & \text{Q} \end{array} \\ \begin{bmatrix} 15 & 25 \\ 20 & 20 \\ 25 & 15 \end{bmatrix} \end{array} \begin{array}{c} \\ \text{A} \\ \text{C} \\ \text{B} \end{array} \quad \mathbf{S} = \begin{array}{c} \begin{array}{ccc} \text{A} & \text{C} & \text{B} \end{array} \\ \begin{bmatrix} .95 & -.05 & -.001 \\ -.05 & .95 & -.05 \\ -.001 & -.05 & .95 \end{bmatrix} \end{array} \begin{array}{c} \\ \text{A} \\ \text{C} \\ \text{B} \end{array}.$$

Values in the **M** matrix were chosen to coincide with their positions in the multiattribute evaluation space. The **S** matrix was chosen to have equal lateral inhibition between A and C and B and C also due to their positions in the multiattribute evaluation space.

(Appendix continues)

Search Variability

Values for the **M** matrix were chosen such that each option had 12 attributes, 4 of which had a high value, 4 medium, and 4 low (1, 0, −1, respectively). Attribute values were ordered such that no option dominated another. Simulations were run separately for one of three choice set sizes: 2, 7, or 12. Hence, for the 12-option set, the original **M** matrix was a 12 (number of options) × 12 (number of attributes) matrix with each row representing values across dimensions for each alternative. For the 7- and 2-option sets the **M** matrix was constructed similarly and were of size 7 × 12 and 2 × 12, respectively.

Because some options are eliminated from the choice set during deliberation, the size of all of the matrices changed. The **S** matrix consisted of a number of options remaining × number of options remaining diagonal matrix with .95 on the diagonal and zeros everywhere else. The **C** matrix was constructed to compare each alternative with the average of all other alternatives.

The upper preference boundary to choose an option was set at 50. The lower boundary to discard an option was set at −30, −20, −10 for the 2-, 7-, and 12-option sets, respectively. Attention shifted from one dimension to another in that on each trial there was a 95% probability that the same dimension would be used at the next time step and a 5% probability of moving to the next dimension. An error component of $N \sim (0, 7)$ was added to the column of the **M** matrix chosen on each trial. Simulations were run for each subject until an option crossed the upper boundary, all but one option was discarded, or all 12 dimensions were cycled through.

Received July 2, 1999
Revision received July 19, 2000
Accepted July 22, 2000 ■