JOURNAL OF MATHEMATICAL PSYCHOLOGY **41**, 260–274 (1997)
ARTICLE NO. MP971167

Dynamic Stochastic Models for Decision Making under Time Constraints

Adele Diederich

C.v. Ossietzky Universität, Oldenburg, Germany

This paper introduces the *multiattribute dynamic decision model* (MADD) to describe both the dynamic and the stochastic nature of decision making. MADD is based on information processing models developed by Diederich. It belongs to the class of sequential comparison models and generalizes and extends the so-called *decision field theory* (DFT) of Busemeyer and Townsend. Describing the decision maker's choice behavior for multiattribute choice alternatives in an uncertain environment by a stochastic process, the *Ornstein–Uhlenbeck process*, MADD predicts choice probabilities and mean choice response times when the decision maker has to respond within a given time limit. This paper outlines the prediction of two different versions of the model in detail.

1. INTRODUCTION

Investigating decision making under time constraints has become increasingly popular over the last decades. (Recently, an entire book was devoted to this subject, Svenson and Maule, 1993.) Two theoretical approaches can be distinguished among the predominantly experimentally guided research efforts:

(a) According to one approach, the decision maker possesses a collection of different decision strategies and chooses one depending on the choice situation. Theories differ with respect to situation/person determinants for selecting a strategy. The *cost/benefit* approach (Beach & Mitchell, 1978; Payne, Bettman, & Johnson, 1988) assumes that decision strategies can be distinguished in terms of their *accuracy* and *cognitive effort*. The selection of a decision strategy is seen as a compromise between maximizing accuracy and minimizing effort. Under time constraints, the decision maker is forced to select a strategy that is less time consuming but also less accurate. Janis and Mann (1977) assume that emotional factors and stress factors influence the selection of a decision strategy. Time constraints inducing stress lead to the adoption of less effective decision strategies. For a review of other studies and findings see Edland and Svenson (1993). I call these models *decision rule* (*strategy*) *selection models*. Test of the cost/benefit approaches usually involve choices among two or more alternatives.

(b) According to the other approach, the decision maker considers features of choice alternatives sequentially over time. Each feature comparison results in a value, which is integrated with the other values, producing a preference state at each moment in time. The preference moves in a random walk manner until a preset decision criterion (or decision boundary), characteristic for each choice alternative, is reached. At that point the process stops and a decision is made. These models are called *sequential comparison models* (Albert, Aschenbrenner, & Schmalhofer, 1988; Aschenbrenner, Albert, & Schmalhofer, 1984; Busemeyer, 1985; Busemeyer & Townsend, 1992, 1993; Diederich, 1995b; Wallsten & Barton, 1982, for probabilistic inference). The main difference among these models is how information about the choice alternatives is selected, evaluated, and integrated with the evaluations. Within these approaches, setting the decision maker under time pressure changes the decision maker's criterion rather than changing her or his strategies. The criterion is an increasing function of the time available for a decision. Tests of these models usually involve binary choices.

Sequential comparison models explicitly model the time it takes to reach a decision and, therefore, they naturally yield predictions for decision making under time constraints.

In the following, I present a dynamic stochastic decision model for binary multiattribute decision problems, called *multiattribute dynamic decision model* (MADD). MADD is a sequential comparison model and it extends and generalizes the so-called *decision field theory* (DFT) by Busemeyer and Townsend (1992, 1993) which applies to unidimensional choice alternatives.

First, I describe the model according to its assumptions about information selection, evaluation, and integration, as well as about how the process terminates. Then, I present a mathematical development of the model, followed by

Correspondence and reprint requests should be addressed to Adele Diederich, C.v. Ossietzky Universität, Institut für Kognitionsforschung, FB 5, Psychologie, A6, D-26111 Oldenburg, Germany. E-mail: diederich@psychologie.uni-oldenburg.de.

0022-2496/97 $25.00

specific assumptions about the composition of the deliberation process. In this paper I particularly focus on MADD's predictions with respect to decision making under time constraints.

2. MULTIATTRIBUTE DYNAMIC DECISION MODEL

Choice alternatives in multiattribute decision problems are usually characterized by attributes and by specific values for each attribute. These values are mapped on a scale specified in the particular research area (for example, utility, Keeny and Raiffa, 1976; attractiveness, Aschenbrenner, Albert, & Schmalhofer, 1984; Montgomery & Svenson, 1976; Svenson, 1979). Instead, MADD assumes that the decision maker uses attributes as memory retrieval cues for concrete ideas and aspects about the choice objects. The decision maker tries to anticipate and evaluate all possible consequences produced by choosing either of the alternatives. The anticipated consequences are retrieved from a complex associative memory process and it takes time to retrieve, compare and integrate the comparisons. The deliberation process might be accompanied by conflict and doubts. Obviously it is not possible to describe or predict this deliberation process in any detail. Rather, the goal of the model presented below is to understand the main cognitive mechanisms that guide the deliberation process that is involved in decisions made under uncertainty.

2.1. *Information Selection*

When confronted with a choice problem the decision maker draws the information about the alternatives from his or her memory. The possible consequences connected with choosing either of the alternatives are learned from experience and the decision maker remembers them more or less well. The decision maker considers attributes the alternatives may or may not have in common. Assume the attribute that comes first to the decision maker's mind will be processed first. This attribute might be the most important or the most salient one for the decision maker (see Aschenbrenner *et al.*, 1984; Wallsten & Barton, 1982). The decision maker may reconsider attributes several times.

2.2. *Evaluation and Integration*

A preference state is activated as soon as the decision maker is confronted with a decision problem and starts thinking about an attribute. The process is modelled as a continuous-time stochastic process. The variable, t, represents the amount of time that has elapsed since the decision maker has been confronted with the decision problem. At each point in time, the decision maker has a new *preference state*, denoted by X_t. Positive values, $X_t > 0$, indicate a momentary preference for alternative A, say,

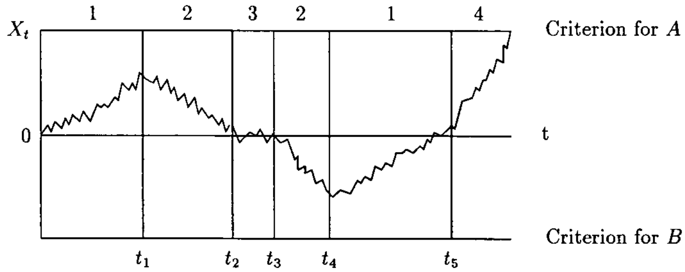

FIG. 1. A stochastic trajectory representing a hypothetical decision process for a binary choice situation. Each alternative is characterized by four attributes. Attributes may be reconsidered during the deliberation process.

negative values, $X_t < 0$, represent a momentary preference for alternative B, and $X_t = 0$ indicates indifference between the two alternatives at time t. If the initial preference state, that is the state at $t = 0$, is $X_0 > 0$ the decision maker is biased towards alternative A, whereas for $X_0 < 0$, the decision maker has an initial bias towards alternative B. For $X_0 = 0$, the decision maker has no *a priori* bias towards an alternative. Figure 1 illustrates one possible evolution of the process.

In this example the decision maker considers attribute 1 first. He or she has no a priori bias towards either alternative. Thinking about the associated aspects of the first attribute and retrieving the possible consequences for choosing either of the alternatives directs his or her preference towards alternative A, thus $X_t > 0$. After time t_1 the decision maker considers attribute 2 and the momentary preference for alternative A gets smaller. Then attribute 3 comes to the decision maker's mind. But thinking about this attribute does not help much to make a decision: X_t is close to 0 and the decision maker is indifferent between the two alternatives. Then he or she reconsiders attribute 2 and clearly the momentary preference is directed toward alternative B; thus $X_t < 0$. Then attribute 1 is reconsidered and the preference for alternative A grows as time goes by. Finally, a last attribute is considered and a choice is made in favor of alternative A.

Each comparison of a given attribute, k, results in a comparison value at any time t, called valence and denoted by $V_k(t)$.[1] $V_k(t)$ is assumed to be a random variable with mean δ_k and variance σ_k^2, representing the moment to moment changes that occur as the decision maker retrieves,

[1] The valence may be produced by a different at time t, i.e., $V(t) = V_A(t) - V_B(t)$, as Busemeyer and Townsend (1993) assume. Obviously, there is a degree of arbitrariness in choosing a *difference* value for presenting the decision maker's valence. The fraction $V_A(t)/V_B(t)$ leads to the same predictions. Both expressions are equivalent in the sense that they impose identical constraints on the valence $V(t)$ (see Falmagne's discussion, 1985, p. 117, in the context of Fechner's problem). Because of those identifiability problems, no a priori assumption on how the valence $V(t)$ is constituted will be made for MADD, but it will be specified when needed.

evaluates, and compares the values of the attributes, and considers possible consequences for choosing either of the alternatives. The valence $V_k(t)$ at time t is integrated with the previous preference state X_t, resulting in a new preference state $X_{t+\tau}$ at time $t+\tau$, where τ is a very small amount of time. The new preference state $X_{t+\tau}$ is defined as

$$X_{t+\tau} = (1-\gamma_k \tau) X_t + \tau V_k(t). \tag{1}$$

$(1-\gamma_k\tau)$ weighs the previous preference state with respect to attribute k and τ weighs the new valence value producing a new preference updated continuously over time.

2.3. *Choice*

It is assumed that the deliberation process continues until a preset decision criterion θ is reached and a choice is made. That is, the preference state has to reach a value that is equal to or more extreme than the value of θ before a decision can be made. In this example, the decision maker chooses alternative A as soon as $X_t > \theta$; alternative B is chosen as soon as $-X_t > \theta$.

The criterion is set by the decision maker when confronted with the choice task. It may depend on personal characteristics and/or on task and situation characteristics. For example, for an important decision, the decision maker sets a high θ, allowing him or her to accumulate more evidence prior to choice. The criterion may also depend on the available time for making a decision. That is, θ is assumed to be smaller with, than without, time constraints. The criterion is assumed to be fixed within a choice situation but may vary across choice situations.[2]

Note, that the preference state at any point in time is not observable and not necessarily open to introspection and, therefore, the hypothetical deliberation process is not directly observable. Nevertheless, some properties of the model seem obvious: the preference for alternative A and B may change during the deliberation process; the strength of preference may depend on the attribute; the order of attributes the decision maker is thinking of may be crucial for his or her decision; setting the decision maker under time pressure may reverse a decision. In order to elaborate these features of the model we now turn to its formal development.

3. MATHEMATICAL DEVELOPMENT OF MADD

The decision process representing the decision maker's preference for alternatives A or B at any time t is described by a continuous-time, continuous-state stochastic process, $\{X_t\colon t \geqslant 0\}$, X_t for short, with two absorbing boundaries, θ and $-\theta$. The space state, S, in which the values x of the random variable X_t lie, is assumed to consist of different degrees of preference, called the *preference states*. In particular, I assume such a stochastic process (diffusion process) for each attribute k considered by the decision maker.

Each process I consider here is characterized by two coefficients, a drift coefficient $\mu_k(x) = \delta_k - \gamma_k x$ and a diffusion coefficient $\sigma_k^2(x) = \sigma_k^2$.[3] The drift coefficient determines the direction and velocity of the process, the diffusion coefficient indicates the variance of the increments of the process.[4] Here, δ_k represents the mean valence resulting from the comparison between the alternatives A and B with respect to attribute k. Assuming the valence to be produced by the difference of valences for A and B, then, without loss of generality, $\delta_k > 0$ indicates a preference for alternative A with respect to attribute k, say, and $\delta_k < 0$ indicates a preference for alternative B with respect to attribute k. Assuming instead the valence to be produced by the ratio of the valences for A and B, then $\delta_k > 1$ indicates a preference for alternative A, and $\delta_k < 1$ indicates a preference for alternative B with respect to attribute k. The second part of the drift coefficient, $-\gamma_k x$, indicates that the drift varies proportionally to the value of the process.[5]

The parameter γ_k which I call the *conflict parameter* accounts for various conflict situations. $\gamma_k > 0$ indicates an avoidance–avoidance conflict situation. That is, the weight of the previous preference state becomes smaller (see Eq. 1) and the increments of the process get increasingly smaller as the distance from the starting position, i.e., from the initial preference state at time $t=0$, becomes larger. This implies that it takes longer to reach a preset criterion. $\gamma_k < 0$ indicates an approach–approach conflict situation. That is, the weight of the previous preference state becomes larger and the increments of the process get increasingly larger as the distance from the initial preference state increases and therefore the process more quickly reaches a preset decision criterion. $\gamma_k = 0$ indicates no conflict.[6] With γ_k refering to a conflict with respect to attribute k, three different conflict situations, i.e., approach–approach; avoidance–avoidance; and approach–avoidance, in the spirit of Lewin (1931, 1951) and Miller (1944), can be accounted for; σ^2 accounts for the moment-to-moment variability of evaluating the possible outcomes.

[2] Compare speed–accuracy trade-off.

[3] A diffusion process with these specific drift and diffusion coefficients is called an *Ornstein–Uhlenbeck process*.

[4] The *increment* of the process in the small time interval $(t, t+\tau]$ is $X_{t+\tau} - X_t$. The drift coefficient for a diffusion process, also called *infinitesimal first moment* or *infinitesimal mean*, is in general defined as $\mu(x, t) = \lim_{\tau\to 0} E[X(t+\tau) - X(t) \mid X(t) = x)]/\tau$. The diffusion coefficient, also called *infinitesimal second moment* or *infinitesimal variance*, is in general defined as $\sigma^2(x, t) = \lim_{\tau\to 0} E[X(t+\tau) - X(t) \mid X(t) = x)]/\tau$.

[5] In physics and biology this quantity reflects a restoring force directed to the origin, i.e., a force that pulls the process (particle) back to the origin. The force is larger as the distance from the origin increases.

[6] Note, that with $\gamma = 0$ the Ornstein–Uhlenbeck process reduces to a *Wiener process with drift*.

Note, while the diffusion coefficient is independent of the state of preference, the drift changes with location x, i.e., the process does not have independent increments. Further, the process is time-homogeneous with respect to each single attribute the decision maker considers.

The standard way to determine the choice probability and the mean choice response time for diffusion process models is to solve a partial differential equation, the Kolmogorov backward equation. Unfortunately, no closed form solution is known for the Ornstein–Uhlenbeck process with absorbing boundaries (Ricciardi, 1977, p. 128; Bhattacharya & Waymire, 1990, p. 389). Therefore, a birth–death chain, a discrete-time discrete-state stochastic process, for which solutions exist, is used to approximate the Ornstein–Uhlenbeck diffusion process, which is a continuous-time, continuous-state stochastic process. The derivations for this procedure can be found in Appendix A.

3.1. Choice Probability and Mean Choice Response Time

Next the probability to choose an alternative and the time it takes to decide, both interpreted as preference strength, have to be determined. That is, the probability that the process eventually reaches one of the boundaries of the process (the criterion to decide for one alternative) for the first time, the *first passage probability*, has to be calculated. Moreover, the time taken by the process to go to one of the boundaries for the first time, the *first passage time*, has to be determined. The decision criteria or boundaries of the process are *absorbing states*[7] of the process, since once the decision process reaches a criterion a decision is made and the process is absorbed. The transition probability density $p(t;\, x,\, y)$ of the diffusion process is approximated by transition probabilities $p_{ij}^{(n)}$ of the birth–death chain the transition probability matrix $\mathbf{P}$ for that chain has to be determined.

The state space of the process, S, is defined as the set of states of preference for the decision process. Assume that the preference changes by very small steps of size Δ (see also Appendix A). Then, the total number of preference states can be expressed as a function of the step size Δ and the criterion θ by setting $\theta = \ell\Delta$. Consequently S equals

$$S = \{0,\, \pm\Delta,\, \pm 2\Delta,\, \pm \cdots \pm (\ell - 1)\,\Delta,\, \pm \ell\Delta\}$$

and the cardinality of S is $2 \cdot \ell + 1 \equiv m$ or equivalently, $m = 2 \cdot \theta/\Delta + 1$. The process is bounded by $\ell\Delta$ and $-\ell\Delta$, two absorbing states for deciding either for alternative A for alternative B, respectively. For convenience these states are numbered from 1 to m, in order of magnitude.

First, the equations for one part of the decision process are developed, i.e., for thinking about just one attribute. This case is identical to the unidimensional decision problem, for which Busemeyer and Townsend (1992, 1993) developed their model. Then I will generalize the model for multiattribute decision problems.

With $\mu(x) = \delta - \gamma x$ and $\sigma^2(x) = \sigma^2$ the transition probabilities for the process with respect to one attribute take the following form (for convenience I drop the index k indicating attribute k for a while):

$$p_{i,j} = \begin{cases} \dfrac{1}{2}\left(1 - \dfrac{(\delta - \gamma \cdot i\Delta)}{\sigma}\sqrt{\tau}\right), & \text{for} \quad j - i = -1 \\ \dfrac{1}{2}\left(1 + \dfrac{(\delta - \gamma \cdot i\Delta)}{\sigma}\sqrt{\tau}\right), & \text{for} \quad j - i = +1 \\ 1 - p_{i,\,i-1} - p_{i,\,i+1}, & \text{for} \quad j = 1 \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

for $i,\ j = 1, 2, ..., m,$ and with $p_{11} = 1$ and $p_{mm} = 1$. $(\delta - \gamma i\Delta)/\sigma$ has to be within the limits $\pm 1/\sqrt{\tau}$ to produce a probability matrix. Note that the transition probabilities depend on the state in which the process is located.

Presenting the transition matrix $\mathbf{P} = \|p_{ij}\|$ for the two alternative choice process in its canonical form yields

$$\mathbf{P} = \left[\begin{array}{c|c} \mathbf{P}_1 & \mathbf{0} \\ \hline \mathbf{R} & \mathbf{Q} \end{array}\right]$$

$$= \begin{array}{c|cc|ccccc|} & 1 & m & 2 & 3 & \cdots & m-2 & m-1 \\ 1 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ m & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \hline 2 & p_{21} & 0 & p_{22} & p_{23} & \cdots & 0 & 0 \\ 3 & 0 & 0 & p_{32} & p_{33} & \cdots & 0 & 0 \\ 4 & 0 & 0 & 0 & p_{43} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ m-3 & 0 & 0 & 0 & 0 & \cdots & p_{m-3,\,m-2} & 0 \\ m-2 & 0 & 0 & 0 & 0 & \cdots & p_{m-2,\,m-2} & p_{m-2,\,m-1} \\ m-1 & 0 & p_{m-1,\,m} & 0 & 0 & \cdots & p_{m-1,\,m-2} & p_{m-1,\,m-1} \end{array}, \tag{3}$$

with $\mathbf{P_1}$ being a 2×2 matrix with two absorbing states, one for each choice alternative. $\mathbf{Q}$, an $(m-2) \times (m-2)$ matrix, contains the transition probabilities p_{ij} stated above. $\mathbf{R}$, an $(m-2) \times 2$ matrix, contains the transition probabilities from the transient[8] to the absorbing states.

Next consider the starting position of the process. At time $t = 0$ the process is set in motion either by starting it at a fixed state s_j, $s_j = (j - 1 - (m-1)/2)\,\Delta$, $j = 1, ..., m$, called

[7] A state is said to be an *absorbing* state if and only if $p_{ii} = 1$.

[8] A state i is said to be *transient* if and only if, starting from state i, there is a positive probability that the process may not eventually return to this state.

the *initial state*, or by randomly locating it in the state space according to a probability distribution $\mathbf{Z}$ of S, called the *initial distribution*. In the former case, $\mathbf{Z}$ is the distribution concentrated at the state s_{j_0}, i.e., $z_j = 1$ if $j = j_0$, $z_j = 0$ if $j \neq j_0$. In the latter case, the probability is z_j that at time $t = 0$ the process will be found in state j, where $0 \leqslant z_j \leqslant 1$ and $\sum_j z_j = 1$. Therefore the initial distribution $\mathbf{Z}$ is an $(m - r)$ vector containing the probability distribution over the transient states. For convenience it is assumed here that the process starts at a fixed state. Here the initial distribution $\mathbf{Z}$ is an $(m - 2)$ vector containing the initial probability distribution over the transient states. Depending on whether the person's initial preference state is biased towards an alternative ($X_0 > 0$ or $X_0 < 0$) or is unbiased ($X_0 = 0$), the $\mathbf{Z}$ vector contains 1 on position $j > (m - 1)/2$ or on position $j < (m - 1)/2$ or on position $j = (m - 1)/2$, respectively, and zeros otherwise.

The equations for the mean choice time and choice probability are derived by using standard methods developed in Markov chain theory (e.g., Bhat, 1984; Bhattacharya & Waymire, 1990).

Let $\mathbf{T}_s$ be the set of transient states and $\mathbf{T}_s^c$ be the set of recurrent[9] states. Further, let $f_{ij}^{(n)}$ be the probability that, starting from transient state i, the process enters the absorbing state j in n steps. Since a transition from an absorbing state to a transient one is not possible, the number of steps for a first passage transition from i to j can be considered as proportional to the time needed. Defining T_{ij} to represent this random variable, we may write $f_{ij}^{(n)}$ as its distribution given by

$$\Pr(T_{ij} = n) = f_{ij}^{(n)}, \quad i \in \mathbf{T}_s, \quad j \in \mathbf{T}_s^c,$$

and let

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

be the probability of eventual passage to j. Let $\mathbf{F}^{(n)}$ be the matrix with elements $f_{ij}^{(n)}$ and let $\mathbf{F}$ be the matrix with elements f_{ij}. Then

$$\mathbf{F}^{(n)} = \mathbf{Q}^{n-1}\mathbf{R}, \quad n = 1, 2, ..., \infty \tag{4}$$

and

$$\mathbf{F} = \sum_{n=1}^{\infty} \mathbf{F}^{(n)} = (\mathbf{I} - \mathbf{Q})^{-1}\,\mathbf{R}, \tag{5}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are defined earlier and $\mathbf{I}$ is the identity matrix (see, e.g., Bhat, 1984, p. 79). The probability of choosing alternative A at time t, $t = n\tau$, starting from a initial preference state z, equals according to Eq. (4)

$$\Pr[T = t \cap \text{choose } A] = \mathbf{Z}'\mathbf{Q}^{n-1}\mathbf{R}_A, \quad n = 1, 2, ..., \infty, \tag{6}$$

where $\mathbf{Z}$ and $\mathbf{Q}$ are defined earlier. $\mathbf{R}_A$ is an $(m-2) \times 1$ vector of the matrix $\mathbf{R}_{(m-2)\times 2}$, containing the transient probabilities for alternative A, i.e.,

$$\mathbf{Z}'\mathbf{Q}\mathbf{R}_A = |0 \cdots z \cdots 0|$$

$$\times \begin{vmatrix} p_{22} & p_{23} & 0 & \cdots & 0 & 0 \\ p_{32} & p_{33} & p_{34} & \cdots & 0 & 0 \\ 0 & p_{43} & p_{44} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p_{m-3,\,m-2} & 0 \\ 0 & 0 & 0 & \cdots & p_{m-2,\,m-2} & p_{m-2,\,m-1} \\ 0 & 0 & 0 & \cdots & p_{m-1,\,m-2} & p_{m-1,\,m-1} \end{vmatrix} \begin{vmatrix} p_{21} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{vmatrix}.$$

The probability of choosing A is according to Eq. (5)

$$\Pr[\text{choose } A] = \mathbf{Z}' \sum_{n=1}^{\infty} \mathbf{Q}^{n-1}\,\mathbf{R}_{\mathbf{A}} = \mathbf{Z}'(\mathbf{I} - \mathbf{Q})^{-1}\,\mathbf{R}_A. \tag{7}$$

The rth moment for the distribution of times to choose alternative A is therefore

$$E[T^r \mid \text{choose } A] = \frac{\tau^r \mathbf{Z}' \sum_n n^r \mathbf{Q}^{n-1}\mathbf{R}_A}{\mathbf{Z}'(\mathbf{I} - \mathbf{Q})^{-1}\,\mathbf{R}_A}, \qquad n = 1, 2, ..., \infty. \tag{8}$$

In particular, the mean time equals

$$E[T \mid \text{choose } A] = \frac{\tau \mathbf{Z}'(\mathbf{I} - \mathbf{Q})^{-1}\,\mathbf{R}_A}{\mathbf{Z}'(\mathbf{I} - \mathbf{Q})^{-1}\,\mathbf{R}_A}. \tag{9}$$

The choice probability and the mean choice response time of choosing alternative B can be determined accordingly by replacing the vector $\mathbf{R}_A$ by the vector $\mathbf{R}_B$.

So far, I have developed the equations for thinking about just one attribute. Five parameters are required for this situation: δ, the mean valence for alternatives A and B with respect to a given attribute; γ, the conflict parameter, to account for various conflict situations; σ^2, the diffusion coefficient, that accounts for momentary fluctuation of evaluating possible consequences; θ, the criterion or absorbing boundary for the decision process; z, the starting position of the process, that accounts for a bias towards an alternative. Δ is chosen sufficiently small to approximate the diffusion process, and $\tau = \Delta^2/\sigma^2$ is given by determining Δ and σ.

In the multiattribute decision problem a stochastic difference equation is defined for each attribute k the decision

[9] A state i is said to be *recurrent* if and only if, starting from state i, eventual return to this state is certain.

maker considers. Therefore, for each attribute k a drift coefficient and a diffusion coefficient,

$$\mu_k(x) = \delta_k - \gamma_k \cdot x, \tag{10}$$

$$\sigma_k^2(x) = \sigma_k^2, \tag{11}$$

respectively, has to be determined. δ_k is the mean valence for alternative A and B with respect to attribute k; γ_k determines the conflict with respect to attribute k; σ_k^2 is the diffusion coefficients for each attribute k. That is, each additional attribute adds three new parameters to the five parameters mentioned above.

The question pursued in the next section is how the deliberation process across attributes can be modeled given the developments for a single attribute.

4. MODES OF ATTRIBUTE SELECTION

The general assumption is that the processing of attributes characterizing two alternatives in a choice situation can be described by a diffusion process with a specific drift rate and a diffusion coefficient. Given that, in general, the alternatives possess several attributes, the following questions arise: First, how are these attributes combined to determine the choice? Second, which additional assumptions about the order of processing of the attributes have to be made? I suggested earlier that the decision maker may consider order of salience, beginning with the most salient one. Figure 1 clearly presumes serial processing with a known order of attributes. This is similar to Wallsten and Barton's (1982) approach, which assumed that the order of the attributes the decision maker considers is naturally given, or to the Aschenbrenner *et al.* (1984) model assuming that the importance of an attribute can be determined by ranking them prior to the decision making process. In real-life situations, the order in which attributes are considered may be induced by emphasizing or naming certain attributes first, as is typically done in sales situation, advertisements, or political speeches, for example.

Figure 1 also presumes that the amount of time the decision maker devotes to each attribute is preset. This assumption may be reasonable if the decision maker is instructed (e.g., by a sales person or an experimenter) to think about the attributes in a certain order for given amounts of time. Alternatively, the amount of time spent on thinking about an attribute may be related to the salience of that attribute. In any case, when the order of the attributes the decision maker considers is preset *and* the time he or she spends thinking about each attribute is preset, the probability to choose an alternative and the expected time to choose that alternative can easily be determined. I have developed various *versions* of the MADD model that differ with respect to the assumptions about the order of attributes the decision maker considers and the amount of time the decision maker spends thinking about each (Diederich, 1996). Let T_f and T_u refer to the situation where the amount of time spent on an attribute is preset or **f**ixed and is not preset or is **u**nspecified, respectively. Further, let O_f and O_u refer to the situation where the order of considered attributes is preset or **f**ixed and not preset or **u**nspecified, respectively. Table 1 shows four versions of MADD, where, e.g., MADD/T_fO_f refers to the model where both the amount of time spent thinking about each attribute and the processing order of the attributes are assumed to be preset.

A fifth version, labeled MADD/pp, that does not fit exactly in the above scheme does not make explicit assumptions about processing order or processing time for each attribute but rather assumes probabilities for switching from one attribute to another.

I present two of the versions, MADD/T_fO_f and MADD/pp next. They take two extreme positions with respect to the assumptions made about processing order and processing time. The exposition is limited to these two versions because of space and, moreover, it is hoped that these two extreme versions articulate the range of possibilities most clearly.

For simplicity, I limit the number of attributes that characterize the alternatives to three. An extension to choice alternatives with more than three attributes is straightforward.

4.1. *MADD/$T_f O_f$: Preset Amount of Time, Preset Order of Attributes*

MADD/T_fO_f describes situations in which the decision maker considers attributes in preset order and for preset periods of time. That is, attribute order, e.g., $1 \to 2 \to 3$, is given, as well as the points in time, t_1 and t_2, when the next attribute is considered, e.g., prescribed by the experimenter, and, therefore, they are not free parameters of the model. No further assumptions are made about how this order and the switching times come about.

Three drift coefficients and three diffusion coefficients have to be determined according to Eq. (10) and Eq. (11). With these coefficients the transition probabilities given each attribute can be calculated according to Eq. (2). The

TABLE 1

Four Versions of MADD

	Preset order	
Preset time	Yes	No
Yes	MADD/T_fO_f	MADD/T_fO_u
No	MADD/T_uO_f	MADD/T_uO_u

transition probabilities can be expressed as transition matrices (Eq. (3)) for each attribute by labeling the submatrices $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{R}_3$ and $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$ for attributes 1, 2, and 3, respectively. The starting position for the preference process is determined by the vector $\mathbf{Z}$. The deliberation process regarding the first attribute evolves until time t_1 when the next attribute comes into consideration. That means, the probability of choosing alternative A, say, before time t_1, starting from the initial preference state z is, based on Eq. (6),

$$\Pr[T < t_1 \cap \text{choose } A] = \mathbf{Z}' \sum_{i=1}^{n_1} \mathbf{Q}_A^{i-1} \mathbf{R}_{A_1}, \tag{12}$$

with $n_1 = t_1/\tau$ (n_1 is restricted to an integer). Then the decision maker considers the next attribute, 2. The deliberation process for that attribute starts from the preference state vector where the prior process ended, i.e., from $\mathbf{Z}'\mathbf{Q}_1^{n_1}$. The decision maker thinks about attribute 2 and its attached consequences until time t_2. The probability of choosing alternative A after t_1 and before time t_2 is, therefore,

$$\Pr[t_1 \leqslant T < t_1 \cap \text{choose } A] = \mathbf{Z}'\mathbf{Q}_1^{n_1} \sum_{i=n_1+1}^{n_2} \mathbf{Q}_2^{i-(n_1+1)} \mathbf{R}_{A_2}, \tag{13}$$

with $n_2 = t_2/\tau$ (n_2 is restricted to an integer). The deliberation process for the last attribute considered starts from the preference state vector in which the second process ended, i.e., from $\mathbf{Z}'\mathbf{Q}_1^{n_1}\mathbf{Q}_2^{n_2-n_1}$ and ends when a criterion for choosing either of the alternatives is reached. The probability of choosing alternative A when starting to think about attribute 3 at time t_2 and continuing until the criterion is reached is, therefore,

$$\Pr[t_2 \leqslant T \cap \text{choose } A] = \mathbf{Z}'\mathbf{Q}_1^{n_1}\mathbf{Q}_2^{n_2-n_1} \sum_{i=n_2+1}^{\infty} \mathbf{Q}_3^{i-(n_2+1)} \mathbf{R}_{A_3}. \tag{14}$$

The probability of choosing alternative A is determined by adding the single parts of the process. Therefore, when considering three attributes, the choice probability for A is

$$\begin{aligned}\Pr[\text{choose } A] &= \mathbf{Z}' \sum_{i=1}^{n_1} \mathbf{Q}_1^{i-1} \mathbf{R}_{A_1} + \mathbf{Z}'\mathbf{Q}_1^{n_1} \sum_{i=n_1+1}^{n_2} \mathbf{Q}_2^{i-(n_1+1)} \mathbf{R}_{A_2} \\ &\quad + \mathbf{Z}'\mathbf{Q}_1^{n_1}\mathbf{Q}_2^{n_2-n_1} \sum_{i=n_2+1}^{\infty} \mathbf{Q}_3^{i-(n_2+1)} \mathbf{R}_{A_3}.\end{aligned} \tag{15}$$

A term must be added for each additional attribute the decision maker takes into consideration. Successive stages of the process start with the final preference state of the previous stage. Note that $\mathbf{Z}'\mathbf{Q}_1^{n_1}$ and $\mathbf{Z}'\mathbf{Q}_1^{n_1}\mathbf{Q}_2^{n_2-n_1}$ are defective initial distributions. Further note that the stochastic process is time homogeneous within each time interval $[0, T_1)$, $[t_1, t_2)$, and $[t_2, \infty)$ but nonhomogeneous across $[0, \infty)$ (see Diederich, 1992, 1995). Assuming the same drift and diffusion parameters for each attribute results in the same transition matrix for all attributes. This equivalence can be expressed by dropping the indices specific for each attribute in Eq. (15), reducing it to Eq. (7). In other words, considering several attributes all with the same parameters is equivalent to considering any one of them alone.

The mean response time for choosing alternative A is, according to Eq. (8) and Eq. (15),

$$\begin{aligned}E[T \mid \text{choose } A] &= \tau \Bigg[\mathbf{Z}' \sum_{i=1}^{n_1} i\mathbf{Q}_1^{i-1}\mathbf{R}_{A_1} + \mathbf{Z}'\mathbf{Q}_1^{n_1} \sum_{i=n_1+1}^{n_2} i\mathbf{Q}_2^{i-(n_1+1)}\mathbf{R}_{A_2} \\ &\quad + \mathbf{Z}'\mathbf{Q}_1^{n_1}\mathbf{Q}_2^{n_2-n_1} \sum_{i=n_2+1}^{\infty} i\mathbf{Q}_3^{i-(n_2+1)}\mathbf{R}_{A_3} \Bigg] \Bigg/ \Pr[\text{choose } A].\end{aligned} \tag{16}$$

The probability and the mean response time for choosing alternative B can be determined accordingly.

MADD/$\mathrm{T_fO_f}$ has the following parameters: z, the starting position for the process; δ_1, δ_2, δ_3, the mean valences for alternative A and B with respect to the given attributes 1, 2, 3; γ_1, γ_2, and γ_3, the conflict parameters for each attribute; σ_1^2, σ_2^2, σ_3^2, the diffusion coefficient for each attribute; and the criterion θ to decide for one alternative. Each additional attribute adds three new parameters. However, for many decision situations, the conflict parameters may be assumed to be the same for all attributes, in particular when the conflict is an approach–approach or avoidance–avoidance situation rather than an approach–avoidance situation. Moreover, the diffusion coefficients may be assumed to be the same for all attributes when the uncertainty about the possible consequences is assumed to be the same for all considered attributes. Thus, for K attributes, the minimum number of attributes is $K+4$.

4.2. *MADD/pp*

Now I consider the version of the model where neither the specific order nor the amount of time the decision maker thinks about each attribute are fixed, i.e., MADD/pp. A major distinction between this version and MADD/$\mathrm{T_fO_f}$ is that now temporal additivity over the processing of the attributes is not assumed. Further, the strict assumption of a single-pass serial system is given up.

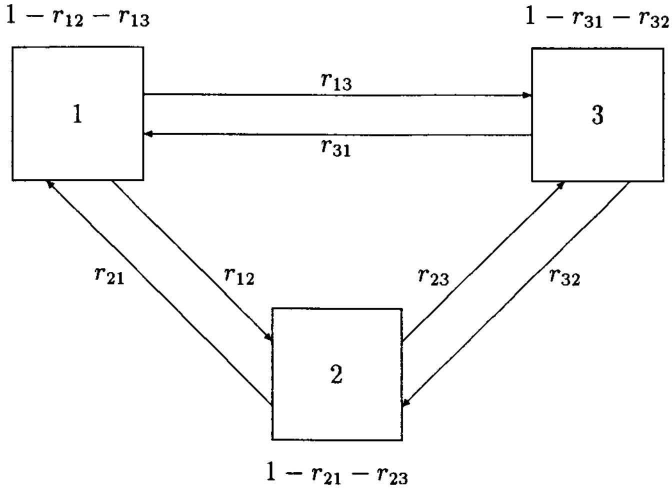

FIG. 2. A transition state diagram for three attributes.

The decision maker is allowed to consider the attributes in any order and to reconsider them many times. Reconsideration was not excluded in MADD/$T_f O_f$. However, each time another attribute, new or old, came into consideration a new starting preference state vector for that attribute had to be determined. The processes were connected in series and each starting preference state vector depended on *all* the preceding attributes already considered.

In MADD/pp, the amount of time spent thinking about an attribute and the frequency of considering it may be a function of its salience. One way to incorporate these assumptions into the model is to assign constant probabilities (weights) for staying with one attribute and for switching to another attribute.[10] Note, that the salience may be defined in the model by the weights r_{ij}, and in some cases can be measured by the estimates of the weights from the data. Denote the probability of remaining with attribute i as r_{ii}, $i = 1, 2, 3$, and the probability of switching from attribute i to attribute j as r_{ij}, $j = 1, 2, 3$, $i \neq j$. With $r_{ii} = 1 - r_{ij} - r_{ik}$, $k = 1, 2, 3$, $k \neq i, j$, increasing the probability of staying within an attribute simultaneously tends to decrease the probabilities of switching to the others. Furthermore, note that increasing r_{ij}, the frequency of reconsidering attributes, i.e., increasing the frequency of switching back and forth between attributes, may turn a strictly serial information process into a pseudo-parallel information process. Figure 2 presents a state transition diagram with three attributes, 1, 2, 3.

The next step is to determine the choice probability and choice response time for this version of the model. As before, for each attribute a matrix with transition probabilities according to Eq. (2) has to be created. The transition probabilities for the transient preference states must be weighted by the parameters r_{ii}. As a consequence, the resulting matrices are no longer transition matrices since the rows do not add up to 1. The problem is remedied by considering the probabilities of switching attention from one attribute to another without changing preference states. Define $\mathbf{P}^*_{(1)}$, $\mathbf{P}^*_{(2)}$, and $\mathbf{P}^*_{(3)}$, as the defective preference state transition matrices for the attributes 1, 2, and 3, respectively. That is, for m preference states

$$\mathbf{P}^*_{(i)} =$$

$$\begin{array}{c|ccccccc|}
 & 1 & 2 & 3 & \cdots & m-2 & m-1 & m \\
1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
2 & r_{ii}p_{21} & r_{ii}p_{22} & r_{ii}p_{23} & \cdots & 0 & 0 & 0 \\
3 & 0 & r_{ii}p_{32} & r_{ii}p_{33} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
m-2 & 0 & 0 & 0 & \cdots & r_{ii}p_{m-2,m-2} & r_{ii}p_{m-2,m-1} & 0 \\
m-1 & 0 & 0 & 0 & \cdots & r_{ii}p_{m-1,m-2} & r_{ii}p_{m-1,m-1} & r_{ii}p_{m-1,m} \\
m & 0 & 0 & 0 & \cdots & 0 & 0 & 1
\end{array} \tag{17}$$

with $i = 1, 2, 3$ for attributes 1, 2, 3. Next the matrix containing the probabilities r_{ij} for switching from attribute i to attribute j as

$$\mathbf{r}_{ij} = \begin{array}{c|ccccccc|}
 & 1 & 2 & 3 & \cdots & m-2 & m-1 & m \\
1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
2 & 0 & r_{ij} & 0 & \cdots & 0 & 0 & 0 \\
3 & 0 & 0 & r_{ij} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
m-2 & 0 & 0 & 0 & \cdots & r_{ij} & 0 & 0 \\
m-1 & 0 & 0 & 0 & \cdots & 0 & r_{ij} & 0 \\
m & 0 & 0 & 0 & \cdots & 0 & 0 & 0
\end{array}. \tag{18}$$

With these matrices the matrix

$$\begin{vmatrix}
\mathbf{P}^*_{(1)} & \mathbf{r}_{12} & \mathbf{r}_{13} \\
\mathbf{r}_{21} & \mathbf{P}^*_{(2)} & \mathbf{r}_{23} \\
\mathbf{r}_{31} & \mathbf{r}_{32} & \mathbf{P}^*_{(3)}
\end{vmatrix} \tag{19}$$

is constructed which is again a transition matrix of size $(3 \cdot m) \times (3 \cdot m)$. The matrix above can be represented in the canonical form according to Eq. (3). Now $\mathbf{P}_1$ is a $(3 \cdot 2) \times (3 \cdot 2)$ submatrix with ones in the diagonal, the transition probabilities for the absorbing states; $\mathbf{R}$ is a $(3 \cdot m - 3 \cdot 2) \times (3 \cdot 2)$ matrix with the weighted transition probabilities from the transient to the absorbing states.

[10] Versions MADD/$T_u O_f$ and MADD/$T_u O_u$ assume that the longer an attribute is considered the more likely the process is to switch to the next one to account for the decrease in the rate of information extracted from one attribute (Diederich, 1996).

$\mathbf{Q}$ is a $(3 \cdot m - 3 \cdot 2) \times (3 \cdot m - 3 \cdot 2)$ matrix containing the weighted transition probabilities for the transient states and the switching probabilities.

The probability of choosing alternative A is according to Eq. (7),

$$\Pr[\text{choose } A] = \mathbf{Z}'(\mathbf{I} - \mathbf{Q})^{-1}\, \mathbf{R}_A,$$

and the mean time for choosing A is according to Eq. (8),

$$E[T \mid \text{choose } A] = \frac{\tau \mathbf{Z}'(\mathbf{I} - \mathbf{Q})^{-2}\, \mathbf{R}_A}{\Pr[\text{choose } A]}.$$

$\mathbf{Z}$ is a $3(m-2)$ vector containing the starting position for the deliberation process. For example, assuming the same probability of being considered first for all three attributes, $\mathbf{Z}$ contains $\frac{1}{3}$ on three positions according to the initial state for the three attributes with respect to both alternatives.

MADD/pp has the following parameters: $\mathbf{Z}$, the vector with the starting positions for the process; δ_1, δ_2, δ_3, the mean valences for alternative A and B with respect to the given attributes 1, 2, 3; the conflict parameters γ_1, γ_2, and γ_3 for each attribute; σ_1^2, σ_2^2, σ_3^2, the diffusion coefficients for each attribute; the criterion θ to decide for one alternative. These parameters are exactly the same as for MADD/$\mathrm{T_fO_f}$. Additional parameters for this version are the switching probabilities (or weights), r. In case of three attributes, six new parameters r_{12}, r_{13}, r_{21}, r_{23}, r_{31}, r_{32} have to be estimated. With a number K of attributes there will be $K(K-1)$ additional r-parameters. To reduce the number of parameters the conflict parameters, as well as the diffusion coefficients, may be assumed to be the same for all attributes for many choice situations (see above). Moreover, the switching probabilities r may be functionally related.

5. PREDICTIONS OF MADD: CHANGE OF PREFERENCE AS A FUNCTION OF TIME LIMITS

MADD makes several parameter-free predictions about the relationship between choice probabilities and choice response times for choosing among two alternatives. In particular, it predicts an inverse relation between choice probabilities and choice response times. That is, if the probability for choosing A over B, $\Pr(AB)$, is greater or equal to 0.5, then as $\Pr(AB)$ increases the mean choice response time for choosing A over B decreases. Moreover, MADD predicts increasing choice response times with an increasing number of attributes. Further, it predicts longer mean response times for adverse choice situations compared to desirous ones (avoidance–avoidance conflicts compared to approach–approach conflicts; see Diederich, 1996). This paper focuses on another aspect of the model, the change of preference as a function of time limits. The parametric predictions of these aspects are outlined for the two versions presented in the previous section.

Two alternatives A and B with three attributes, 1, 2, 3, are used throughout the examples described below. The parameters belonging to an attribute are indicated by subscripts 1, 2, and 3, respectively. For simplicity, several parameters are set to simple values. In particular, the conflict parameters are set to $\gamma_1 = \gamma_2 = \gamma_3 = 0$, i.e., no conflict, the diffusion coefficients are assumed to be $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1$, and the stepsize is $\Delta = 1$.[11] The criterion is assumed to be an increasing function of the time available for a decision (see above). It is determined in terms of the matrix size m, i.e., $\theta = (m-1)/2$. That is, decreasing the available time to make a decision decreases the decision maker's criterion and, therefore, in terms of the model, decreases the size of the matrix. For the predictions below, the size of the matrix was always an odd number to avoid an a priori bias towards either of the alternatives. The probability is plotted as a function of the matrix size. I present several examples for MADD/pp and MADD/$\mathrm{T_fO_f}$ that differ systematically in the remaining parameters. All computations for the predictions were carried out using the matrix language GAUSS.

5.1. *MADD/pp: Same Salience Parameters r_{ij} for all Attributes*

Assume that two of the three attributes favor alternative A, say, and one attribute favors alternative B, i.e., of the three mean valences δ_1, δ_2, and δ_3, two of them have signs different from the third one. Without loss of generality, let us assume that δ_1 and δ_2 are positive and favor alternative A, and δ_3 is negative and favors alternative B. Specially, let $\delta_1 = 0.3$, $\delta_2 = 0.1$, and $\delta_3 = -0.2$. Moreover, let the parameters r_{ij} be the same for all attributes, i.e., none of the attributes gets a special weight. Now we assume that we know that the decision maker considers the attribute that favors alternative B first, i.e., with probability 1 the process starts with a drift directed to the criterion for alternative B. The left panel of Fig. 3 shows this result for different r_{ij}.

At the beginning of the decision task, the decision maker focuses attention on the attribute that favors alternative B ($\delta_3 < 0$). Therefore, his or her preference is directed toward that alternative: Then he or she considers those attributes that are in favor of alternative A ($\delta_1, \delta_2 > 0$) and preference gradually drifts towards this alternative. As time on the decision problem increases the decision maker reverses his or her initial preference for alternative B to a preference for alternative A since overall there is more evidence to decide for alternative A.[12]

[11] Note that for approximating a diffusion process by a birth–death chain Δ is chosen sufficiently small ($\ll 1$). However, for the following predictions it is not important to carry them out in *real* time.

[12] By *preference reversal* is meant here a crossing of the $P = 0.5$ line.

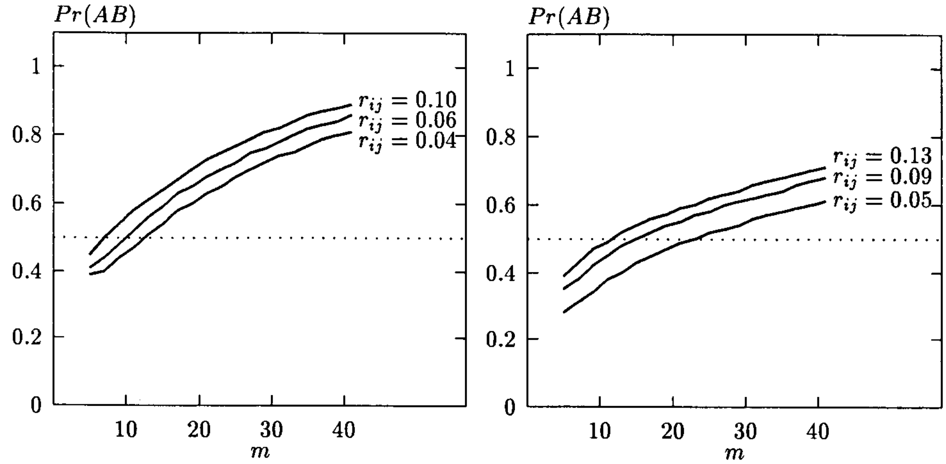

FIG. 3. Change of preference over time when the decision maker focuses her attention at the beginning of the decision problem on the attribute that favors alternative B, ($\delta_3 < 0$). Pr(AB) states the probability for choosing A over B. m indicates the size of the transition matrix for each attribute. The criterion, θ, $\theta = (m-1)/2$ is assumed to be an increasing function of the time limit. Different lines indicate different r_{ij}. For the left panel the valence parameters are $\delta_1 = 0.3$, $\delta_2 = 0.1$, $\delta_3 = -0.2$. For the right panel the parameters are $\delta_1 = 0.3$, $\delta_2 = 0.2$, $\delta = -0.4$.

Increasing r_{ij} increases the frequency with which the decision maker switches from one attribute to the next. Comparing, e.g., the line with $r_{ij} = 0.04$ and the line with $r_{ij} = 0.10$ in Fig. 3, left panel, shows that the decision maker reverses his or her preference sooner when r_{ij} is larger because he or she spends less time considering the attribute that favors alternative B before tending to switch to those attributes that favor alternative A. Thus, decision makers under time pressure may choose differently than those not under pressure since they have not enough time to consider all relevant attributes. Moreover, setting a time limit may even cause the decision maker to concentrate on one attribute (or fewer attributes than available in real life situations). The right panel of Fig. 3 shows the predictions of MADD/pp for a different set of parameters, i.e., $\delta_1 = 0.3$, $\delta_2 = 2$, and $\delta_3 = -0.4$. The r_{ij} are indicated in the figure. Note that the lines of the right panel of Fig. 3 are flatter than those of the left one due to different overall evidence for alternative A. That is, the sum of the valence values is smaller for the example shown in the right panel (0.1) than for that one shown in the left panel (0.2).

5.2. *MADD/pp: Different Salience Parameters r_{ij} for the Attributes*

For this example let the mean valence parameters be $\delta_1 = 0.5$, $\delta_2 = 0.4$, favoring alternative A, and $\delta_3 = -0.3$, favoring alternative B. Now assume that the attribute in favor of alternative B gets a higher weight by increasing the probability to switch to that attribute, i.e., r_{13}, $r_{23} > r_{12}$, r_{21}, r_{31}, r_{32}. Moreover, assume that the decision maker considers any of the attributes first; i.e., he or she starts with probability $\frac{1}{3}$ of reflecting on any of the three attributes. Figure 4 shows the predictions for this situation with $r_{12} = r_{21} = r_{31} = r_{32} = 0.02$ and $r_{13} = r_{23}$, ranging from 0.06 to 0.09.

At the beginning of the decision situation, the decision maker's preference tends to be directed towards alternative A since the values of two attributes speak for choosing A. But these attributes are less salient than the attribute that favors alternative B; i.e., as time goes by, the decision maker has switched more often to the more salient attribute and has spent more time reflecting on it than on the remaining attributes. Increasing the parameters for switching back to a more salient attribute, in this example r_{13} and r_{23}, reverses the preference sooner.

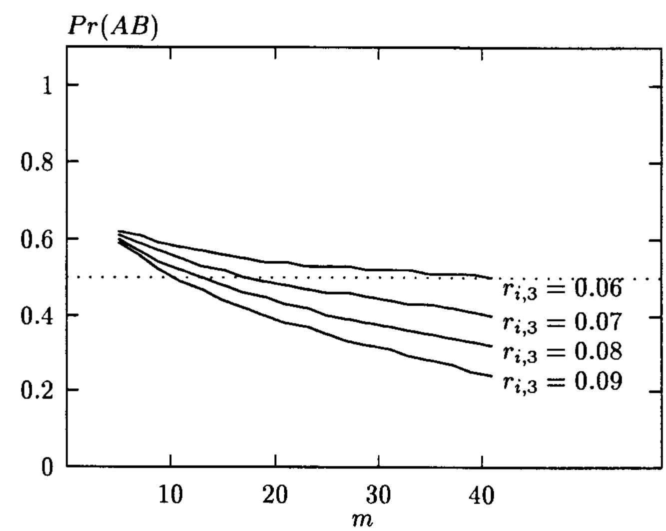

FIG. 4. Change in preference over time when the attribute in favor of alternative B gets higher salience. Different curves indicate different $r_{i,3}$, $i = 1.2$. For the uppermost curve $r_{13} = r_{23} = 0.06$, for the lowest curve $r_{13} = r_{23} = 0.09$, r_{12}, r_{21}, r_{31}, r_{32} equal 0.02 in all cases.

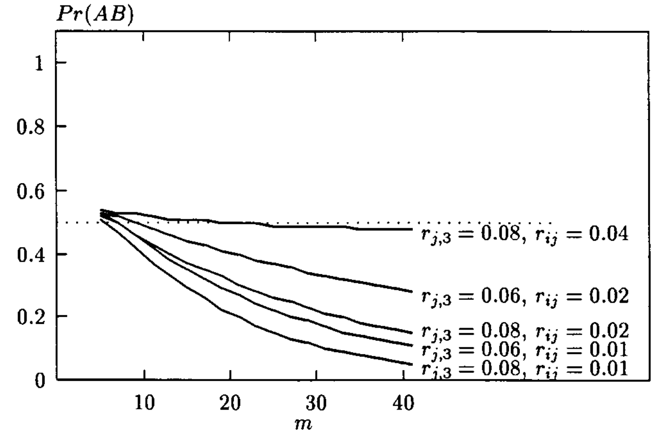

FIG. 5. Change in preference over time when the attribute in favor of alternative *B* gets higher salience. Different curves indicates different *r* with $i = 1, 2, 3$ and $j = 1, 2$.

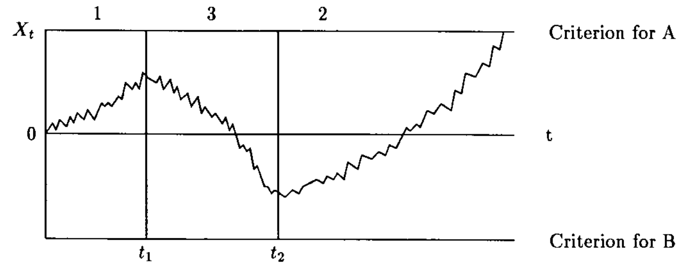

FIG. 6. A stochastic trajectory representing a hypothetical decision process with two alternatives characterized by three attributes. The decision maker presumably starts thinking about an attribute favoring alternative *A*, then considers an attribute that favors *B*, and finally thinks about an attribute that favors *A* again.

Figure 5 shows the predictions for a different set of parameters, i.e., $\delta_1 = 0.3$, $\delta_2 = 0.1$, and $\delta_3 = -0.2$. Again, the decision maker starts with probability $\frac{1}{3}$ of reflecting on any of the three attributes first. Attribute 3 gets a higher weight than attributes 1 and 2. Different lines result from different assumed weights as indicated in the figure.

5.3. *MADD*/$T_f O_f$, *Example* 1

Now assume that the attributes are processed according to MADD/$T_f O_f$; i.e., the processing order and processing time for each attribute are preset. Let the mean valence values be $\delta_1 = 0.5$, $\delta_2 = 0.4$, and $\delta_3 = -0.3$. The decision maker first considers an attribute that is in favor of alternative *A*; then he or she considers the attribute favoring *B*; finally he or she considers the remaining attribute, favoring *A*. Therefore, the order of attributes could, for example, be $1 \to 3 \to 2$. Consider, first, the situation where no time constraints are imposed: At the beginning of the decision process, the decision maker tends to prefer alternative *A* since δ is larger than 0 for the first attribute. After a certain amount of time he or she switches to the next attribute. This attribute favors alternative *B* and the decision maker's preference state is directed towards that alternative. Finally, he or she contemplates the third attribute, favoring alternative *A* and decides for this alternative. Figure 6 illustrates this process.

Now assume that the decision maker is set under time pressure. Remember that the criterion is assumed to be an increasing function of the time limit. With a short time limit the decision maker has a small criterion and the decision is more likely to be reached from considering the first attribute only and, therefore, he or she will prefer alternative *A*. With increased time, the decision criterion is raised and the decision maker likely will consider also the second attribute and the preference will tend to alternative *B*. The left panel of Fig. 7 shows the possible change of preference as a function of time for the parameters indicated above. Different curves represent different time limits to consider attribute 3 with

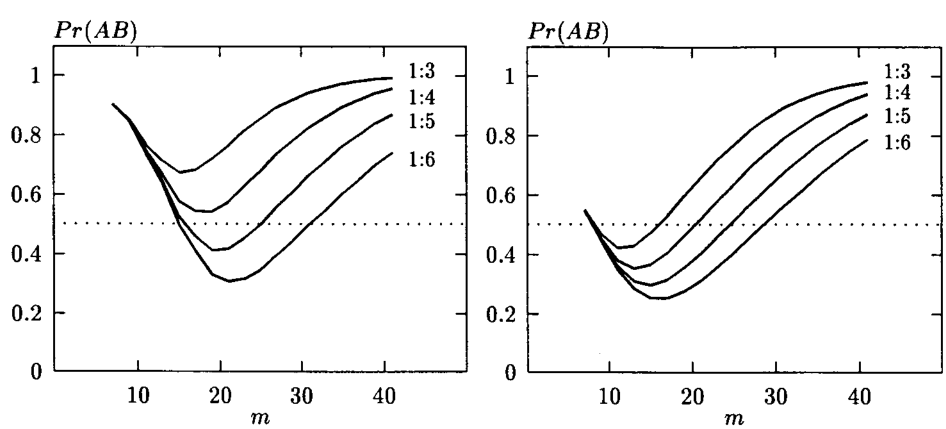

FIG. 7. Change of preference over time depending on the processing order of attributes and the amount of time spent thinking of each attribute. Different curves represent different amounts of time spent thinking about the first and second attribute (see text). Pr(*AB*) states the probability for choosing *A* over *B*; *m* indicates the size of the transition matrix for each attribute. The criterion, θ, $\theta = (m-1)/2$ is assumed to be an increasing function of the time limit.

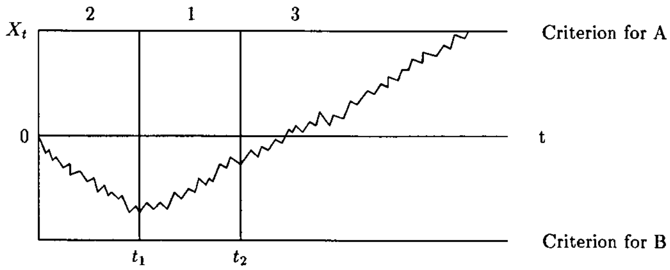

FIG. 8. A stochastic trajectory representing a hypothetical decision process with two alternatives characterized by three attributes. The decision maker starts thinking about an attribute favoring alternative *B* and the two different attributes that both favor alternative *A*.

$\delta_3 = -0.3$. For the lowest curve the time spent thinking about attribute 3 is six times as long as that spent on thinking about the first attribute. For the uppermost curve the deliberation time for attribute 3 is three times as long as that for the first attribute.

The right panel of Fig. 7 represents the predictions for a different set of valence parameters $\delta_1 = 0.1$, $\delta_2 = 0.3$, and $\delta_3 = -0.2$. The assumed processing order is $1 \to 3 \to 2$.

The depth and shift of the dips of the curves depend on the amount of time spent on attribute 3 ($\delta_3 < 0$). The longer the decision maker reflects on it the more likely it is to decide for alternative *B*.

5.4. *MADD/$T_f O_f$, Example* 2

Now consider a different processing order of attributes. Let the decision maker start thinking about attribute 3 with $\delta_3 = -0.3$. She tends to prefer alternative *B*. Then the remaining attributes are considered, both favoring alternative *A*. The preference for alternative *B* gets smaller, increasing the preference for alternative *A*. Figure 8 shows this hypothesized process.

As before, assume that the decision maker has to decide within a time limit. With a short deadline he or she has only enough time to think about the first attribute, here 3, and the preference tends towards alternative *B*. With more time available the next attribute can be considered and the preference is gradually shifted to alternative *A*. The left panel of Fig. 9 demonstrates the predictions of the model for the parameters $\delta_1 = 0.5$, $\delta_2 = 0.4$, and $\delta_3 = -0.3$. Different curves indicate different amounts of time the decision maker devotes to considering attribute 3. The time spent thinking about attribute 3 is four times as long for the upper one as for the lower one.

The right panel of Fig. 9 represents the predictions for a different set of valence parameters, i.e., $\delta_1 = 0.2$, $\delta_2 = 0.3$, and $\delta_3 = -0.1$. The assumed processing order is $3 \to 2 \to 1$.

5.5. *Discussion*

Note an important difference between Busemeyer and Townsend's decision field theory and the multiattribute extension developed here. In DFT a change in preference as a function of time can only occur by letting the decision maker have an initial preference for one alternative ($X_0 \neq 0$) and, simultaneously, a higher valence for the other alternative. For example, let the initial state of preference be directed to alternative *B* ($X_0 < 0$), and the mean valence be in favor of alternative *A* ($\delta > A$). Under a short time limit, the decision maker will more frequently decide for the initial preferred alternative *B*, but the probability to choose alternative *A* grows as a function of time. For MADD no bias towards an alternative has to be assumed to predict this effect.

Note that MADD/pp and MADD/$T_f O_f$ differ with respect to the predicted relation between preference and time. MADD/pp always predicts a monotonic relationship between these quantities, whereas MADD/$T_f O_f$ does

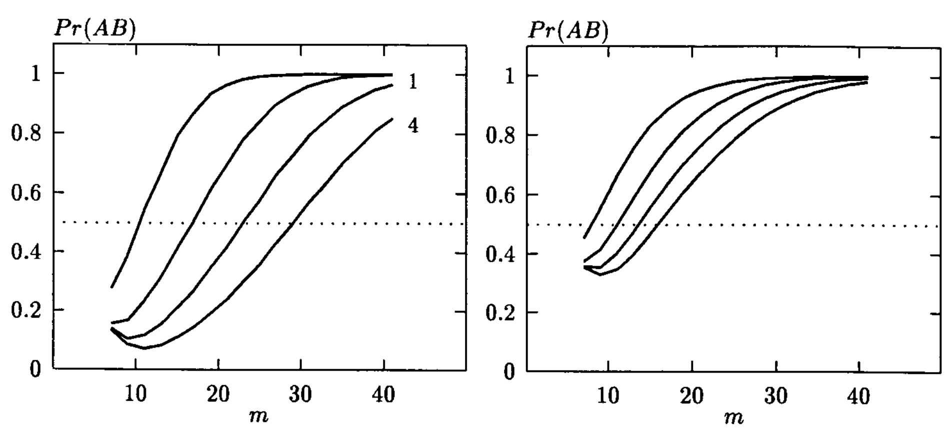

FIG. 9. Change of preference over time depending on the processing order of attributes and the amount of time spent thinking of each attribute. Different curves represent different amounts of time spent thinking about the first and second attribute (see text); m indicates the size of the transition matrix for each attribute. The criterion $\theta = (m-1)/2$ is assumed to be an increasing function of the time limit.

not. The reason for this is as follows. Switching back and forth, as assumed in MADD/pp, *averages* the valences, so to speak. It seems that the sign of the sum of all valences adjusted by the weights determines the direction of the function. Depending on the assumed processing order and sign of the mean valence of the attributes MADD/$\mathrm{T_f O_f}$ is also able to predict a nonmonotonic relationship between preference and time.

If one starts with an attribute that has an advantage for alternative B, say and all the remaining subsequently considered attributes favor alternative A, the relationship between choice probability and choice time is monotonic as seen in Fig. 9. On the other hand, if the sign of the mean valence alternates (e.g., $+-+$ or $-+-$) the predicted relationship is nonmonotonic as in Fig. 7.

Both versions, however, predict that with increasing time available to make a decision, the probability of choosing a particular alternative converges to one.

6. SUMMARY AND DISCUSSION

A dynamic stochastic decision model for multiattribute alternatives, presented in a binary choice task, was described and developed mathematically. It is called *multiattribute dynamic decision model* (MADD) and it belongs to the class of cognitive processing models for decision making, in particular, to the sequential comparison approach. The deliberation process to decide for either of the two alternatives is modelled as an *Ornstein–Uhlenbeck process* with two absorbing boundaries, approximated by a birth–death chain. MADD makes quantitative predictions for both the choice probability and the distribution of choice response times. In particular, the choice probability and the mean choice response time for choosing an alternative are specified by determining the first passage probability and the first passage time of the process. The capability of predicting decision time is an important feature that many other decision models do not possess. Some models, e.g., decision rule selection models consider time pressure as an important variable but do not have a mechanism to predict real time performance.

Two different versions of the model, MADD/pp and MADD/$\mathrm{T_f O_f}$ with specific assumptions on how to combine several attributes under consideration into a whole deliberation process, were proposed and their predictions with respect to decision making under time constraints were outlined in several examples. In MADD/$\mathrm{T_f O_f}$ a preset amount of time and a preset order for considering attributes is given, while MADD/pp allows switching between attributes with a certain probability. Time limits were assumed to be functionally related to the decision criterion. Both versions differ with respect to the functional relation between choice probability and time limit. MADD/pp always predicts a monotonic relationship between these quantities whereas for MADD/$\mathrm{T_f O_f}$ a nonmonotonicity can be obtained depending on the order and sign of the mean valence of the attributes. Unlike decision field theory (DFT) (Busemeyer & Townsend, 1993) MADD does not need to assume an a priori bias towards one alternative to predict a change of preference (from above $p=0.5$ to below $p=0.5$ for any of the alternatives) as a function of time limits.

Several experiments testing the predictions of the model are under way (see also Diederich, 1996).

7. APPENDIX

7.1. *Approximation of a Discussion Process by a Birth–Death Chain*

Some of the following derivations are close to the presentation in Bhattacharya and Waymire (1990, p. 233f, pp. 386–389). First, I describe the discrete stochastic process that is used to approximate the continuous process in detail.

Imagine a simple random walk that never skips states in its evolution, i.e., a Markov chain with state space $S=\{0, 1, ..., \}$ for which transitions from state n can only move to state $n+1$ or to state $n-1$ or stay in the same state. Markov chains with this property are called *birth–death chains*. The state space can be seen as the size of a population, the increase of the size by one as birth and the decrease by one as death. The transition probabilities p_{ij} of going from i to j for the birth–death chain are

$$p_{ij}=\begin{cases}\alpha_i, & \text{if } j=i-1\\ \beta_i, & \text{if } j=i+1\\ 1-\alpha_i-\beta_i, & \text{if } j=i\\ 0, & \text{otherwise.}\end{cases}$$

The transition probabilities may depend on the state in which the process is located.

Now we decrease the step size and increase the frequency for the discrete-time, discrete-state birth–death chain to approximate a diffusion process X_t with drift coefficient $\mu(x)$ and diffusion coefficient $\sigma^2(x)$.

Suppose there are two real-valued functions $\mu(x)$ and $\sigma^2(x)$ on $(-\infty, \infty)$ that are continuously differentiable and that σ'' exists and is continuous, and that $\sigma^2(x)>0$ for all x. Also assume that $\mu(x)$ and $\sigma^2(x)$ are bounded. Consider now a birth–death chain with state space $S=\{0, \pm\Delta, \pm 2\Delta, ...\}$ with step size $\Delta>0$ and having transition probabilities p_{ij} of going from $i\Delta$ to $j\Delta$ in one step given by

$$p_{i,i-1}=\alpha_i^{(\Delta)}:=\frac{\sigma^2(i\Delta)\,\tau}{2\Delta^2}-\frac{\mu(i\Delta)\,\tau}{2\Delta}$$

$$p_{i,i+1}=\beta_i^{(\Delta)}:=\frac{\sigma^2(i\Delta)\,\tau}{2\Delta^2}+\frac{\mu(i\Delta)\,\tau}{2\Delta}$$

$$p_{ii}=1-\frac{\sigma^2(i\Delta)\,\tau}{\Delta^2}=1-\alpha_i^{(\Delta)}-\beta_i^{(\Delta)},$$

where τ is given by

$$\tau = \frac{\Delta^2}{\sup_x \sigma^2(x)}$$

and is the actual time in between two successive transitions. With the conditions imposed on $\mu(x)$ and $\sigma^2(x)$ a sufficiently small Δ guarantees the nonnegativity of the transition probabilities p_{ii}, $p_{i,i-1}$, $p_{i,i+1}$. (See Bhattacharya & Waymire, 1990, p. 386.)

Given that the process is at $x = i\Delta$, the mean displacement in a single step of size Δ in time τ is

$$\Delta\beta_i^{(\Delta)} + (-\Delta)\,\alpha_i^{(\Delta)} = \mu(i\Delta)\,\tau = \mu(x)\,\tau. \tag{20}$$

Therefore, the instantaneous rate of mean displacement per unit time, when the process is at x, is $\mu(x)$. The mean squared displacement in a single step of size Δ in time τ is

$$\Delta^2\beta_i^{(\Delta)} + (-\Delta)^2\,\alpha_i^{(\Delta)} = \sigma^2(i\Delta)\,\tau = \sigma^2(x)\,\tau. \tag{21}$$

That is, the instantaneous rate of mean squared displacement per unit time, when the process is at x, is $\sigma^2(x)$.

Hence, the transition probabilities $p_{i,i-1}$ and $p_{i,i+1}$ can be written as functions of the drift and diffusion coefficient $\mu(x)$ and $\sigma^2(x)$, respectively. With $\Delta^2 = \tau\sigma^2(x)$

$$p_{i,i-1} = \frac{1}{2}\left(1 - \frac{\mu(x)}{\sigma(x)}\sqrt{\tau}\right), \tag{22}$$

$$p_{i,i+1} = \frac{1}{2}\left(1 + \frac{\mu(x)}{\sigma(x)}\sqrt{\tau}\right). \tag{23}$$

Thus, the birth–death chain with transition probabilities p_{ij} approximates the diffusion process with drift coefficient $\mu(x)$ and diffusion coefficient $\sigma^2(x)$ as $\Delta \to 0$. (See the convergence theorem, e.g., Bhattacharya & Waymire, 1990, p. 387; Karlin & Taylor, 1981, p. 169). To see how a differential equation can be expressed by a difference equation let $p_{ij}^{(n)}$ be the n-step transition probabilities going from i to j. By definition $p_{ij}^{(n)} = \sum_{k \in S} p_{ik} p_{ik}^{n-1}$, $n = 2, 3, \ldots$ and S is the state space. Therefore, with $p_{ii} = 1 - \beta_i^{(\Delta)} - \alpha_i^{(\Delta)}$, we get for the birth–death process,

$$p_{ij}^{(n+1)} = p_{ii}p_{ij}^{(n)} + p_{i,i+1}p_{i+1,j}^{(n)} + p_{i,i-1}p_{i-1,j}^{(n)}$$
$$p_{ij}^{(n+1)} = (1 - \beta_i^{(\Delta)} - \alpha_i^{(\Delta)})\,p_{ij}^{(n)} + \beta_i^{(\Delta)}p_{i+1,i}^{(n)} + \alpha_i^{(\Delta)}p_{i-1,i}^{(n)}$$

or

$$\begin{aligned} p_{ij}^{(n+1)} - p_{ij}^{(n)} &= \beta_i^{(\Delta)}(p_{i+1,j}^{(n)} - p_{ij}^{(n)}) - \alpha_i^{(\Delta)}(p_{ij}^{(n)} - p_{i-1,j}^{(n)}) \\ &= \frac{\mu(i\Delta)\,\tau}{2\Delta}\left((p_{i+1,j}^{(n)} - p_{ij}^{(n)}) + (p_{ij}^{(n)} - p_{i-1,j}^{(n)})\right) \\ &\quad + \frac{1}{2}\frac{\sigma^2(i\Delta)\,\tau}{\Delta^2}\left(p_{i+1,j}^{(n)} - 2p_{ij}^{(n)} + p_{i-1,j}^{(n)}\right), \end{aligned}$$

With $t = n\tau$ and states $x = i\Delta$ and $y = j\Delta$ and $\Delta \to 0$, defining as approximate density

$$p^{(\Delta)}(n\tau;\, i\Delta,\, j\Delta) := \frac{p_{ij}^{(n)}}{\Delta},$$

we get

$$\begin{aligned} &p^{(\Delta)}((n+1)\,\tau;\, i\Delta,\, j\Delta) - p^{(\Delta)}(n\tau;\, i\Delta,\, j\Delta) \\ &\quad = \mu(i\Delta)(p^{(\Delta)}(n\tau;\,(i+1)\,\Delta,\, j\Delta) \\ &\qquad - p^{(\Delta)}(n\tau;\,(i-1)\,\Delta,\, j\Delta))\,\tau/2\Delta \\ &\qquad + \tfrac{1}{2}\sigma^2(i\Delta)(p^{(\Delta)}(n\tau;\,(i+1)\,\Delta,\, j\Delta) - 2p^{(\Delta)}(n\tau;\, i\Delta,\, j\Delta) \\ &\qquad + p^{(\Delta)}(n\tau;\,(i-1)\,\Delta,\, j\Delta))\,\tau/\Delta^2 \end{aligned}$$

which is the difference equation version of the partial differential equation

$$\frac{\partial}{\partial t}\,p(t;\,x,\,y) = \mu(x)\frac{\partial}{\partial x}\,p(t;\,x,\,y) + \frac{1}{2}\,\sigma^2(x)\frac{\partial^2}{\partial x^2}\,p(t;\,x,\,y),$$

for $t > 0$, $-\infty < x, y < \infty$. That is, the transition probability density $p(t;\,x,\,y)$ can be computed approximately by determining the n-step transition probabilities $p_{ij}^{(n)}$. This can be done by constructing a transition probability matrix and raising it to the nth power. A *transition probability matrix* is a square matrix $\mathbf{P} = \|p_{ij}\|$ with $p_{ij} \geqslant 0$ for all i and j and $\sum_{j \in S} p_{ij} = 1$ for all i.

For the Ornstein–Uhlenbeck diffusion process the characteristic coefficients are the drift coefficient $\mu(x) = \delta - \gamma x$ and the diffusion coefficient $\sigma^2(x) = \sigma^2$ which can be inserted into Eq. (22) and Eq. (23).

ACKNOWLEDGMENTS

This research was supported by Deutsche Forschungsgemeinschaft Grant No. Di 506/2-1 and Grant No. Di 506/2-2. I thank Bill Batchelder and Tom Wallsten for very helpful comments on an earlier version of this paper.

REFERENCES

Albert, D., Aschenbrenner, K. M., & Schmalhofer, F. (1988). Cognitive choice processes and the attitude-behavior relation. In A. Upmeyer (Ed.), *Attitudes and behavioral decisions*, pp. 61–99. Berlin: Springer-Verlag.

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

Bhat, U. N. (1984). *Elements of applied stochastic processes*. New York: Wiley.

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

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for selection of decision strategies. *Academy of Management Review*, **3**, 439–449.

Bettman, J. R., Johnson, E. J., & Payne, J. W. (1990). A comparison analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes*, **45**, 111–139.

Busemeyer, J. R. (1985). Decision making under uncertainty: A comparison of simple scalability, fixed sample, and sequential sampling models. *Journal of Experimental Psychology*, **11**, 538–564.

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-cognitive approach to decision-making in an uncertain environment. *Psychological Review*, **100**, 432–459.

Diederich, A. (1992). *Intersensory facilitation: Race, superposition, and diffusion models for reaction time to multiple stimuli*. Frankfurt: Peter Lang.

Diederich, A. (1994). A diffusion model for intersensory facilitation of reaction time. In G. H. Fischer & D. Laming (Eds.), *Contributions to mathematical psychology, psychometrics, and methodology*, pp. 207–220. New York, Springer-Verlag.

Diederich, A. (1995a). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology*, **39**, 197–215.

Diederich, A. (1995b). A dynamic model for multi-attributive decision problems. In J.-P. Caverni, M. Bar-Hillel, F. H. Barron, & H. Jungermann (Eds.), *Contributions to decision making I*, pp. 175–191. Amsterdam: Elsevier Science.

Diederich, A. (1996). Multiattribute dynamic decision model: A dynamic stochastic approach for multiattribute binary choice tasks. Habilitationsschrift.

Edland, A., & Svenson, O. (1993). Judgment and decision making under time pressure. Studies and findings. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making*, pp. 27–39. New York: Plenum.

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

Lewin, K. (1931). *Die psychologische Situation bei Lohn und Strafe*. Leipzig. Reprinted, Darmstadt: Wissenschaftliche Buchgesellschaft, 1964.

Lewin, K. (1951). Selected theoretical papers. In D. Cartwright (Ed.), *Field theory in social science*. New York: Harper.

Miller, N. E. (1944). Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*, Vol. 1, pp. 431–465. New York: Ronald Press.

Montgomery, H., & Svenson, O. (1976). On decision rules and information processing strategies for choices among multiattribute alternatives. *Scandinavian Journal of Psychology*, **17**, 283–291.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 534–552.

Ricciardi, L. M. (1977). *Lecture notes in biomathematics*. Berlin: Springer-Verlag.

Svenson, O. (1979). Process descriptions of decision making. *Organizational Behavior and Human Performance*, **23**, 86–112.

Svenson, O., & Maule, A. J. (Eds.) (1993). *Time pressure and stress in human judgment and decision making*. New York: Plenum Press.

Wallsten, T. S., & Barton, C. (1982). Processing probabilistic multidimensional information for decisions. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, **8**, 361–384.

Received March 6, 1997