ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Mathematical Psychology 47 (2003) 150–165

Journal of Mathematical Psychology

http://www.elsevier.com/locate/jmp

An extension of the exemplar-based random-walk model to separable-dimension stimuli

Andrew L. Cohen* and Robert M. Nosofsky

Department of Psychology, Indiana University, 1101 East 10th Street, Bloomington, IN 47405-7007, USA

Received 14 November 2001; accepted 6 November 2002

Abstract

An extension of Nosofsky and Palmeri's (Psychol. Rev. 104 (1997a) 266) exemplar-based random-walk (EBRW) model of categorization is presented as a model of the time course of categorization of separable-dimension stimuli. Nosofsky and Palmeri (1997a) assumed that the perceptual encoding of all stimuli was identical. However, in the current model, we assume as in Lamberts (J. Exp. Psychol.: General 124 (1995) 161) that the inclusion of individual stimulus dimensions into the similarity calculations is a stochastic process with the probability of inclusion based on the perceptual salience of the dimensions. Thus, the exemplars that enter into the random-walk change dynamically during the time course of processing. This model is implemented as a Markov chain. Its predictions are compared with alternative models in a speeded categorization experiment with separable-dimension stimuli.

Keywords: Classification; Perception; Reaction time; Markov chain; Cognitive models

A fundamental issue in cognitive psychology concerns the nature of perceptual categorization. Much of the past modeling work in this area has concentrated primarily on the prediction of choice probabilities. Recently, a number of researchers have focused on the time course of processing and the modeling of reaction times (Ashby, 2000; Ashby, Boynton, & Lee, 1994; Lamberts, 1995, 1998; Lamberts & Freeman, 1999a, b; Nosofsky & Alfonso Reese, 1999; Nosofsky & Palmeri, 1997a). One of the major contributions to the study of the time course of classification has been the exemplar-based random-walk (EBRW) model of Nosofsky and Palmeri (1997a). The general idea of the model is that when an object is presented to be classified, category exemplars stored in memory race to be retrieved (Logan, 1988; Marley & Colonius, 1992) with rates that depend on how similar the exemplars are to the object. These retrieved exemplars then enter into a random-walk process for making classification decisions (Busemeyer, 1985; Link, 1992; Ratcliff, 1978). The EBRW has been very successful at accounting for the time course of classification decision making and at generating predictions of classification response times.

An important limitation of the EBRW, however, is that in its current form it accounts for classification performance only in domains involving integral-dimension stimuli, which are stimuli in which the dimensions combine into relatively unanalyzable, integral wholes. By contrast, many stimuli are composed of separable dimensions, which are dimensions that remain psychologically distinct when in combination (Garner, 1974; Shepard, 1964). In applying the EBRW to account for response times involving integral-dimension stimuli, a reasonable simplifying assumption was that the stimuli were perceived and encoded holistically (Lockhead, 1972). To provide an account of response times involving separable-dimension stimuli, the EBRW needs to be extended by modeling the time course with which each of the individual dimensions is processed.

The key idea involves integrating the exemplar-retrieval and decision-making processes that form the core of the EBRW with the dimensional perceptual sampling ideas advanced by Lamberts (1995, 1998, 2000) in his alternative exemplar-based model of speeded classification. Extending the EBRW in this manner allows the model to account for classification response times in a far wider range of stimulus domains, and forces deeper considerations regarding the intricate relations between perceptual and decision-making

*Corresponding author. Fax: +812-855-4691.
E-mail address: alcohen@indiana.edu (A.L. Cohen).

0022-2496/03/$ - see front matter
doi:10.1016/S0022-2496(02)00031-7

processes in classification. The main theoretical and empirical goals of this paper are, first, to develop this extended version of the EBRW that will enable the model to account for the speeded classification of separable-dimension stimuli; and, second, to test this new model by comparing its predictions to the data from an individual-participant classification experiment.

The remainder of this article is organized as follows. First, we briefly describe Lambert's (2000) EGCM-RT which introduces the idea of stochastic sampling of stimulus dimensions. Second, we present an extension of the EBRW that incorporates perceptual encoding (EBRW-PE). The EBRW-PE is described both as a random-walk and as a Markov chain. Third, we present a speeded classification experiment involving separable-dimension stimuli and compare the modeling results of the EBRW-PE to those of alternative models. Finally, we look at the predictions of the EBRW-PE for the response-signal paradigm.

1. The extended generalized context model for response times (EGCM-RT)

Lamberts' (2000) EGCM-RT is a model of classification reaction time that builds on Nosofsky's (1986) generalized context model (GCM). The EGCM-RT borrows the equations for computing similarity from the GCM, but extends them by making them time-dependent. The EGCM-RT is a complex model and the reader is referred to Lamberts (2000) for details. Here we sketch only the key ideas behind the model.

According to the EGCM-RT, perceptual categorization involves the gradual construction of a stimulus representation through a process of information accumulation. In particular, there is a stochastic process that determines which of the M perceptual dimensions of an object have been encoded at any given point in time. The similarity between the object and the exemplars stored in memory changes systematically depending on this stochastic perceptual encoding process. Early in processing, few dimensions will have been encoded, so the object will be similar to numerous distinct exemplars stored in memory. With unlimited processing time, all dimensions will be encoded, and the definition of similarity is identical to the one found in the standard GCM (Nosofsky, 1986).

At every point in processing, an observer sums the current similarity of an item to the exemplars of the alternative categories. The probability that information accumulation stops after a certain period of time is a function of the evidence for category membership at that time. When processing does stop, the classification decision is based on the relative magnitude of the summed similarities to each alternative category.

2. The exemplar-based random-walk model with perceptual encoding (EBRW-PE)

The EBRW-PE combines the random-walk accumulation of category evidence of the EBRW (Nosofsky & Palmeri, 1997a) with the gradual accumulation of stimulus information of the EGCM-RT (Lamberts, 2000). As in the EBRW, the EBRW-PE assumes that people store category examples in memory along with their category labels. These exemplars are represented as points in a multidimensional psychological space. During a classification judgment, all of the exemplars race to be retrieved. How fast an exemplar races is determined by its similarity to the test item. The winning exemplar adds incremental evidence to a random-walk process. The race is then repeated and a response occurs when the random-walk counter reaches a category boundary.

The EBRW assumes that an entire exemplar is processed, stored, and retrieved at once, that is, the EBRW treats all stimuli as integral wholes. The key contribution of the EBRW-PE over the EBRW is that the similarity relations of the exemplars can change based on the stimulus information available at a particular time. As in the EGCM-RT, we assume that each stimulus dimension is stochastically included in the similarity computation. Specifically, at each step of the random-walk, there is some probability that a stimulus dimension will be encoded and thus enter into the similarity computation.

2.1. Formal properties

The individual item exemplars are represented as points in an M-dimensional psychological space. Let x_{im} denote the coordinate of exemplar i on dimension m. At time t, the distance between exemplars i and j depends on the subset of dimensions, ϕ_t, that have been processed by t. Formally,

$$d_{ij}(\phi_t) = \sum_{m \in M} w_m i_m(\phi_t)|x_{im} - x_{jm}|, \tag{1}$$

where w_m $(0 \leqslant w_m, \Sigma w_m = 1)$ is the attention weight given to dimension m (Nosofsky, 1986, 1987) and $i_m(\phi_t)$ is a binary-valued function of the dimensions encoded at time t (Lamberts, 1995, 1998) such that

$$i_m(\phi_t) = \begin{cases} 1 & \text{if } m \in \phi_t, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Lamberts (1995, 1998, 2000) assumed an exponential inclusion function, such that the probability that dimension m is encoded at time t and therefore included in the distance computation is given by $q_m \exp(-q_m t)$, where q_m is a free parameter that represents the rate of inclusion for stimulus dimension m. Because the time steps in a random-walk are discrete, in what follows, we assume a discrete analog of this function. In particular, we assume that, on every step of the random-walk, there is some fixed probability, r_m, that dimension m will be included in the distance computation (assuming that it was not included in the previous steps). Once a dimension is included in the distance calculation, it is never removed. Thus, the probability that dimension m is included in the distance calculation by step n of the random-walk is geometrically distributed, $1-(1-r_m)^n$. The geometric distribution is the discrete analog of the exponential.

The similarity of exemplar i to exemplar j at time t is an exponential decay function of the distance between the two objects at time t,

$$s_{ij}(\phi_t) = \exp(-cd_{ij}(\phi_t)), \tag{3}$$

where c is an overall sensitivity parameter used to scale distances in the space (Shepard, 1987).

When a test item is presented, all the exemplars in memory race to be retrieved with a rate based on their level of activation. The activation of exemplar j in memory at time t given the presentation of exemplar i is given by

$$a_{ij}(\phi_t) = M_j s_{ij}(\phi_t), \tag{4}$$

where M_j is the memory strength of exemplar j. The memory strength of an exemplar may vary because of factors such as presentation frequency, recency of presentation, and so forth. For simplicity, however, in the present article we set all memory strengths equal to one.

All exemplars in memory race exponentially to be retrieved with rate given by their activation at time t. Specifically, the probability density that exemplar j finishes the race at time τ is given by

$$f(\tau) = a_{ij}(\phi_t)\exp(-a_{ij}(\phi_t)\cdot\tau). \tag{5}$$

If exemplar j wins the race, then it is retrieved and enters the random-walk process. The random-walk is changed by unit value towards the category boundary associated with the winning exemplar. If the random-walk reaches one of the category boundaries, the appropriate response is given. Otherwise, a new race is initiated, another exemplar is retrieved (possibly the same exemplar as in the previous step), and the process continues.

Let $S_{iA}(\phi_t)$ denote the summed activation at time t of all exemplars from category A given exemplar i, that is,

$$S_{iA}(\phi_t) = \sum_{j\in A} a_{ij}(\phi_t). \tag{6}$$

Given the processing assumptions outlined above and the mathematical properties of the exponential distribution, it can be shown (Nosofsky & Palmeri, 1997a) that for two categories, A and B, the probability that the random-walk takes a step towards category A at processing time t is

$$p_i^A(\phi_t) = \frac{S_{iA}(\phi_t)}{S_{iA}(\phi_t)+S_{iB}(\phi_t)} \tag{7}$$

and the probability of moving towards the B boundary is given by

$$p_i^B(\phi_t) = 1 - p_i^A(\phi_t). \tag{8}$$

Following Nosofsky and Palmeri (1997a), it can also be shown that the expected time to take a step at processing time t given exemplar i is given by

$$\bar{T}_{\text{step}} = \alpha + \frac{1}{S_{iA}(\phi_t)+S_{iB}(\phi_t)}, \tag{9}$$

where α is a constant term associated with each step.

If all M dimensions are always included in the distance measure ($i_m(\phi_t)=1,\ \forall m,t$) this model reduces to the standard EBRW of Nosofsky and Palmeri (1997a). However, the analytic predictions are more complex when the time-dependent distance measure is involved. We now turn to the derivation of these predictions.

2.2. A Markov chain implementation

There is a straightforward route to deriving analytic predictions from the EBRW-PE, which is necessary if the model is to be studied with reasonable efficiency. The key idea is to conceive of the EBRW-PE as a Markov chain. The states in the model are the set of dimensions that have been encoded on each step of the random-walk together with the location of the random-walk counter. For example, if on step n of the walk, the observer has encoded dimensions 1 and 3 but not 2 and 4 of a four-dimensional stimulus, and the random-walk counter is now at $+3$ units away from the starting point of 0, then the process is in state $(+3, \{1, 3\})$. Given the assumptions in the model, the probability that on step $n+1$ the process makes a transition to state $(+4, \{1, 3, 4\})$ would be given by $(1-r_2)r_4p_i^A(\{1,3,4\})$, i.e., the joint probability that the participant fails to encode dimension 2, successfully encodes dimension 4, and that, given the new similarity relations defined on the present set of encoded dimensions, an exemplar from category A is retrieved. The absorbing states in the Markov chain are those states in which the random-walk counter reaches either one of the response criteria (A or $-B$).

Given such a Markov chain, there are well-known techniques for deriving the expected number of steps to reach an absorbing state, the probability that a given absorbing state is reached, and so forth (Diederich, 1997). Thus, it becomes straightforward to derive analytic predictions of mean response times and choice probabilities for the EBRW-PE. We now turn to a formalization of the EBRW-PE as a Markov chain and the derivation of these measures. Diederich (1997) (see also Busemeyer & Townsend, 1992, 1997) has proposed similar ideas about shifting attention to dimensions within a random-walk model of preferential choice; however, there are no notions involving stimulus similarity or retrieval of stored exemplars in her dynamic decision-making model.

Let Ω be the state space for the EBRW-PE random-walk. The EBRW-PE has $A+B+1$ random-walk positions, where A and $-B$ are the criteria for categories A and B, respectively. However, the state space must also take into account the set of dimensions, ϕ, included in the distance calculations at any time t.[1] For a stimulus set with M dimensions, ϕ is one of N sets, where

$$N = \sum_{k=0}^{M} \binom{M}{k} = 2^M \tag{10}$$

gives all possible combinations of M or fewer dimensions. Let Φ be the set of all possible ϕ. Then, Ω has $(A+B+1)N$ elements,

$$\Omega = \{-B, -B+1, \ldots, -1, 0, 1, \ldots, A-1, A\} \times \phi. \tag{11}$$

The set Ω contains all possible states of the Markov chain, where each state includes a random-walk position and a set of encoded dimensions. In what follows, we define (l, ϕ) as position l of the random-walk with the set ϕ of dimensions included in the distance computation.

$$\mathbf{Q}_{\phi',\phi} = \begin{array}{c|cccccccc|}
 & -B+1 & -B+2 & \cdots & -1 & 0 & 1 & \cdots & A-2 & A-1 \\
-B+1 & 0 & p_i^A(\phi) & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
-B+2 & p_i^B(\phi) & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
-1 & 0 & 0 & \cdots & 0 & p_i^A(\phi) & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & \cdots & p_i^B(\phi) & 0 & p_i^A(\phi) & \cdots & 0 & 0 \\
1 & 0 & 0 & \cdots & 0 & p_i^B(\phi) & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
A-2 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & p_i^A(\phi) \\
A-1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & p_i^B(\phi) & 0
\end{array} \cdot h(\phi', \phi) \tag{13}$$

State (l, ϕ) is *absorbing* if and only if there is probability zero that the random-walk leaves (l, ϕ) once it has been reached. The EBRW-PE has $2N$ absorbing states, $\{A, -B\} \times \Phi$. State (l, ϕ) is *transient* if and only if there is a non-zero probability that the random-walk may not return to (l, ϕ) once it has been reached. In the EBRW-PE, all $\Omega - (\{A, -B\} \times \Phi)$ non-absorbing states are transient.

State (l, ϕ) is a *starting* state if and only if at time $t = 0$, the random-walk is in state (l, ϕ). We assume here that the starting state of the EBRW-PE is $(0, \emptyset)$, where $\emptyset$ is the set of zero dimensions. Let $\mathbf{Z}$ be an $N(A+B-1)$ column vector that represents the starting state of the random-walk over the transient states, then $\mathbf{Z}$ contains a 1 on state $(0, \emptyset)$ and zeros elsewhere. $\mathbf{Z}$ can be thought of as a probability distribution over the states at $t = 0$.

Let r_m be the probability of adding dimension m to the distance calculation on any step of the random-walk given that it was not included on the previous step. Then

$$h(\phi', \phi) = \begin{cases} \prod\limits_{m \in \phi - \phi'} r_m \prod\limits_{m \in \phi - \phi} (1 - r_m) & \text{if } \phi' \subseteq \phi, \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

gives the probability of processing the dimensions in ϕ on the current step of the random-walk, given that the dimensions in ϕ' were processed on the previous step. For example, for a four-dimensional stimulus, the probability of adding dimensions 1 and 2 and not adding dimension 4 given that dimension 3 is already included in the similarity computation is $h(\{3\}, \{1, 2, 3\}) = r_1 \times r_2 \times (1 - r_4)$.

Suppose that exemplar i is presented and that the observer was in transient state (l, ϕ') on the preceding step. Then $\mathbf{Q}_{\phi',\phi}$, as developed in Eq. 13, is a matrix representation of the EBRW-PE random-walk transition probabilities to the states in which the set of dimensions ϕ is processed on the current step and the random-walk moves towards the A or B boundaries. The matrix entries give the probabilities of moving from the current position, the row labels, to the next position, the column labels. $\mathbf{Q}_{\phi',\phi}$ is an $(A+B-1) \times (A+B-1)$ matrix containing the transition probabilities from Eqs. (7) and (8). The probability of the random-walk remaining at the same position is zero. The elements of $\mathbf{Q}_{\phi',\phi}$ are multiplied by h to incorporate the probability

[1] Both ϕ and ϕ_t introduced in Eq. (1) represent a set of encoded dimensions. The t subscript was dropped to highlight that ϕ is now associated with a particular state of the Markov chain rather than a particular time.

of moving to the new set of encoded dimensions during that time step.

In a complete transition matrix, the row probabilities must sum to one. Even if the absorbing boundaries were included, the rows of $\mathbf{Q}_{\phi',\phi}$ do not sum to one because it is a submatrix of a larger matrix $\mathbf{Q}$ that holds the overall transition probabilities for all elements of Φ. The probability of moving from any transient state to any other transient state is given by the $N(A+B-1) \times N(A+B-1)$ matrix

$$\mathbf{Q} = \begin{vmatrix} \mathbf{Q}_{\phi,\phi} & \cdots & \mathbf{Q}_{\phi,\phi'} & \cdots & \mathbf{Q}_{\phi,\phi''} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \mathbf{Q}_{\phi',\phi} & \cdots & \mathbf{Q}_{\phi',\phi'} & \cdots & \mathbf{Q}_{\phi',\phi''} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \mathbf{Q}_{\phi'',\phi} & \cdots & \mathbf{Q}_{\phi'',\phi'} & \cdots & \mathbf{Q}_{\phi'',\phi''} \end{vmatrix}, \tag{14}$$

for all ϕ, ϕ', and ϕ'' where the submatrices are given as in Eq. (13).[2]

The transition probabilities of moving from a transient state (l, ϕ') to an absorbing state is given by $\mathbf{R}_{\phi'}$, an $(A+B-1) \times 2$ matrix such that

$$\mathbf{R}_{\phi'} = \begin{matrix} & -B & A \\ -B+1 & \sum_{\phi} h(\phi',\phi) \cdot p_i^B(\phi) & 0 \\ -B+2 & 0 & 0 \\ \vdots & \vdots & \vdots \\ -1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ A-2 & 0 & 0 \\ A-1 & 0 & \sum_{\phi} h(\phi',\phi) \cdot p_i^A(\phi) \end{matrix}. \tag{15}$$

The random-walk can move to an absorbing state only from an adjacent position. Given that the system starts with the set of dimensions ϕ', the two sums in Eq. (15) matrix give the probability that the system processes the dimensions in ϕ and then moves to an absorbing state. The first column of $\mathbf{R}_{\phi'}$ gives the probability of moving to the category B boundary. The second column gives the probability of moving to the category A boundary.

The probability of moving from any transient state to an absorbing state is given by the $N(A+B-1) \times 2$ matrix

$$\mathbf{R} = \begin{vmatrix} \mathbf{R}_{\phi} \\ \vdots \\ \mathbf{R}_{\phi'} \\ \vdots \\ \mathbf{R}_{\phi''} \end{vmatrix}, \tag{16}$$

where the submatrices are given as in Eq. (15)[3]. Recall that the first and second columns of Eq. (16), $\mathbf{R}_B$ and $\mathbf{R}_A$, give the probability of moving from a transition state to the absorbing state for categories B and A, respectively.

The benefit of using a Markov chain to implement the EBRW-PE is that analytic solutions for choice probabilities and reaction times are well known. The probability of choosing category A after n steps of the random-walk is given by

$$P(\text{number of steps} = n \text{ and choose } A) = \mathbf{Z}'\mathbf{Q}^{n-1}\mathbf{R}_A, \tag{17}$$

where $\mathbf{Z}$, $\mathbf{Q}$, and $\mathbf{R}_A$ are defined above. $\mathbf{Q}$ is multiplied $n-1$ times because the nth step occurs with the post-multiplication by $\mathbf{R}_A$. The overall probability of choosing category A is derived by letting n range over all possible values. Following Diederich (1997),

$$P(\text{choose } A) = \mathbf{Z}' \sum_{n=1}^{\infty} \mathbf{Q}^{n-1}\mathbf{R}_A = \mathbf{Z}'(\mathbf{I}-\mathbf{Q})^{-1}\mathbf{R}_A, \tag{18}$$

where $\mathbf{I}$ is the identity matrix. Given an outcome of category A, the mean time spent in the random walk is

$$E(T_{rw}|\text{choose } A) = \frac{\alpha \cdot \mathbf{Z}'(\mathbf{I}-\mathbf{Q})^{-2}\mathbf{R}_A}{\mathbf{Z}'(\mathbf{I}-\mathbf{Q})^{-1}\mathbf{R}_A} = \frac{\alpha \cdot \mathbf{Z}'(\mathbf{I}-\mathbf{Q})^{-2}\mathbf{R}_A}{P(\text{choose } A)}, \tag{19}$$

where α is a parameter that gives the average time for a random-walk step. Similar equations exist for category B by replacing $\mathbf{R}_A$ with $\mathbf{R}_B$ in Eqs. (18) and (19). The predicted overall choice time is given by

$$\begin{aligned} E(T) &= E(T_{rw}|\text{choose } A) \cdot P(\text{choose } A) \\ &\quad + E(T_{rw}|\text{choose } B) \cdot P(\text{choose } B) + \mu \\ &= \alpha(\mathbf{Z}'(\mathbf{I}-\mathbf{Q})^{-2}\mathbf{R}_A + \mathbf{Z}'(\mathbf{I}-\mathbf{Q})^{-2}\mathbf{R}_B) + \mu, \end{aligned} \tag{20}$$

where μ is a parameter that represents the mean of any residual response-execution and encoding times. In summary, Eqs. (18) and (20) give the analytic predictions of choice probability and mean response time from the EBRW-PE.

2.3. Submodel

To evaluate the utility of the additional encoding assumptions of the EBRW-PE, we also tested a model

[2] Note that ϕ, ϕ', and ϕ'' here are simply specific elements of Φ. Note also that the number of submatrices in $\mathbf{Q}$ is a function of the number of stimulus dimensions. For example, for a four-dimensional stimulus, $\mathbf{Q}$ will contain $2^4 \times 2^4$ submatrices, one row and one column for every possible subset of dimensions. The ellipses in Eq. (14) are fillers for additional submatrices. See Diederich (1997) for an explanation of why the row probabilities of $\mathbf{Q}$ sum to one.

[3] The sets ϕ, ϕ', and ϕ'' correspond to the sets of Eq. (14) and the ellipses serve the same purpose.

without these additional processes. Thus, in the restricted model, EBRW-R, at time $t = 0$, ϕ is the full set of stimulus dimensions. This restriction is equivalent to requiring all $r_m = 1$. Recall that in the standard EBRW, the time for each individual step in the random-walk is a constant α plus the time to retrieve the winning exemplar (see Eq. (9)). Thus, the standard EBRW is technically not a special case of the EBRW-PE because the latter model does not include the exemplar-retrieval component. However, because α is generally much larger than the expected retrieval time of the exemplars, this difference is a minor one. Thus the EBRW-R is essentially identical to the EBRW. In fact, although we do not report fits of the EBRW in the current article, the EBRW and EBRW-R make nearly identical predictions for the fits reported below.[4]

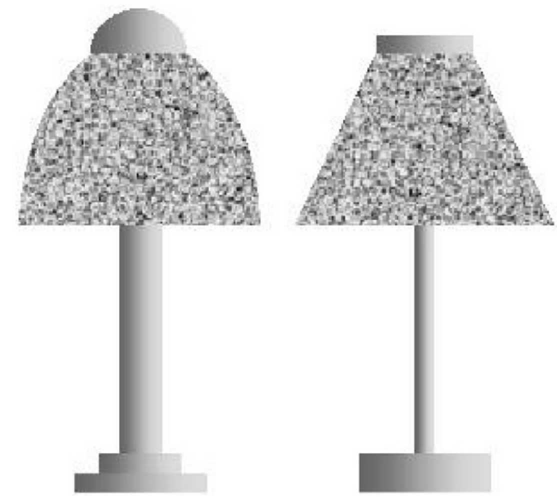

Fig. 1. Two of the stimuli from the experiment. The left stimulus has a value of 0 for all dimensions. The right stimulus has a value of 1 for all dimensions.

3. Experiment

In several studies, Nosofsky and colleagues (e.g., Nosofsky & Alfonso Reese, 1999; Nosofsky & Palmeri, 1997a, b) have demonstrated that the standard EBRW successfully predicts mean choice probabilities and reaction times for integral-dimension stimuli. Recently, Lamberts (2000) successfully applied the EGCM-RT to predict classification response times for separable-dimension stimuli. The purpose of this experiment was to replicate Lamberts (2000) and to obtain a rich set of RT data suitable for quantitative fitting by the EBRW-PE, EBRW-R, and EGCM-RT. We hoped to gain evidence for the utility of the added perceptual encoding assumptions of the EBRW-PE by comparing its fits to those of its special case, the EBRW-R. We were also interested in comparing the ability of the EBRW-PE and EGCM-RT to account for the data.

During a training phase, participants first learned to classify nine stimuli into two categories. In a 9-day transfer phase, participants then made speeded classifications of the nine training stimuli and seven transfer stimuli. The stimuli were composed of four, separable dimensions.

3.1. Method

3.1.1. Participants

Eight Indiana University graduate and undergraduate students participated in this 10-day study. Each was paid $7 a day, plus a bonus for good performance. None of the participants was familiar with the issues under investigation.

3.1.2. Stimuli

Following Lamberts (2000), the stimuli were computer-generated images of lamps displayed on a black background (see Fig. 1). The lamps varied along four binary-valued, separable dimensions: top (curved or flat); shade (curved or angular); upright (fat or thin); and base (two or one sections). For coding purposes, the first value for each dimension is labeled 0 and the second value is labeled 1. The stimuli were approximately 11.5 cm high and 5.5 cm wide and were viewed from a distance of about 60 cm.

3.1.3. Design and procedure

The category structure is shown in Table 1. This category structure was taken from Medin and Schaffer (1978) and was used by Lamberts (2000). To minimize any stimulus-specific effects, the assignment of each physical stimulus dimension (top, shade, base, and upright) to the logical stimulus dimensions (D1–D4 of Table 1) was different for each participant. The assignments are given in Table 9 of Appendix A.

On day 1, participants learned the category structure by induction over exemplars. On each trial, a training stimulus (stimuli 1–9 in Table 1) was shown and the participant judged its category assignment. Feedback was given after each trial. There were 35 training blocks, but participants could leave early if they achieved perfect performance for eight blocks in a row. During each block, each training stimulus was seen once. The order of presentation was randomized within blocks. Learning the category structure was emphasized over fast responding.

[4] Subsequent to completing the present work, we learned of earlier research reported by Shiffrin and Thompson (1988) who developed techniques for implementing nonconstant step times within a Markov-chain framework. These techniques could be used to implement a version of the EBRW-PE that is a true generalization of the EBRW.

Table 1
Category structure

Category and stimulus	Dimension D1	D2	D3	D4
Category A				
1	1	1	1	0
2	1	0	1	0
3	1	0	1	1
4	1	1	0	1
5	0	1	1	1
Category B				
6	1	1	0	0
7	0	1	1	0
8	0	0	0	1
9	0	0	0	0
Transfer				
10	0	0	1	0
11	0	0	1	1
12	0	1	0	0
13	0	1	0	1
14	1	0	0	0
15	1	0	0	1
16	1	1	1	1

Note. The assignment of each physical dimension (top, shade, base, and upright) to each logical stimulus dimension (D1–D4) was different for each participant. See the text for details.

Table 2
Observed and predicted choice proportions and mean response times (in ms) for participant 1

Stimulus	Observed	EBRW-PE	EBRW-R	EGCM-RT
1	0.97	0.98	0.99	0.99
	453	450	454	465
2	0.96	0.98	1.00	0.99
	490	447	452	465
3	1.00	0.99	1.00	0.99
	385	400	402	387
4	0.93	0.98	1.00	0.96
	461	475	485	482
5	0.92	0.96	1.00	0.97
	478	470	474	476
6	0.09	0.02	0.00	0.00
	441	440	440	436
7	0.03	0.01	0.00	0.00
	462	454	461	455
8	0.03	0.01	0.00	0.00
	410	428	423	422
9	0.00	0.00	0.00	0.00
	371	355	337	361
10	0.04	0.01	0.00	0.00
	461	460	465	455
11	0.92	0.96	1.00	0.97
	479	472	478	476
12	0.00	0.00	0.00	0.00
	362	356	338	361
13	0.02	0.01	0.00	0.00
	426	437	432	422
14	0.04	0.03	0.00	0.00
	432	441	443	436
15	0.94	0.97	1.00	0.96
	487	482	490	482
16	1.00	0.99	1.00	0.99
	380	400	402	387

Note: For each stimulus, the top row holds choice proportions for category A and the bottom row holds reaction times. EBRW-PE is the exemplar-based random walk with perceptual encoding, EBRW-R the restricted exemplar-based random walk, and EGCM-RT the extended generalized context model for response times.

Days 2–10 were identical to one another. The experiment proceeded much as on day 1; however, all 16 training and transfer stimuli were included in each block. Order of stimulus presentation was randomized within each block. Participants still received feedback after each training trial and a message of "Thank You" was given after each transfer trial. There were 29 blocks each day. Participants were urged to respond as quickly as possible without sacrificing accuracy. On each day, twice during the experiment and once afterwards, participants were apprised of their average percent correct and mean reaction time for the training stimuli.

3.2. Results

Trials with an RT of less than 150 ms or greater than 4000 ms were excluded from further analysis (less than 1% of the data). Both accuracy and reaction times had generally leveled off by day 3. Because we were interested in asymptotic classification performance, our subsequent analyses are based only on the data from day 3 to day 10. Furthermore, because we are interested in modeling the performance of only those participants who learned the category structure, participants who achieved less than 60% correct on any of the nine training stimuli are not reported. Three of the eight participants did not meet this criterion. The individual participant data for the remaining five participants are reported in Tables 2–6.

3.3. Modeling results

We used the logical dimension coding of Table 1 to fit the EBRW-PE, EBRW-R, and EGCM-RT to each individual participant's mean RT and accuracy data for each stimulus. The free parameters in the Markov implementation of the EBRW-PE are the dimensional attention weights, w_1, w_2, and w_3 (with $w_4 = 1 - w_1 - w_2 - w_3$; Eq. (1)); the dimension encoding probabilities, r_1, r_2, r_3, and r_4 (Eq. (12)); the scaling parameter, c (Eq. (3)); the step time, α (Eqs. (9), (19) and (20)); the residual time, μ (Eq. (20)); and the category criteria, A and B (Eqs. (11), (13) and (15)). The EBRW-R adds the restriction that $r_m = 1$, $m = 1$–4. The EGCM-RT uses four utility values, u_1, u_2, u_3 (with $u_4 = 1 - u_1 - u_2 - u_3$); four-dimensional inclusion rates, q_1, q_2, q_3, and q_4, a generalization parameter, c^*; a residual time, t_{res}; a bias parameter, b; and θ, a response-scaling parameter.

Table 3
Observed and predicted choice proportions and mean response times (in ms) for participant 2

Stimulus	Observed	EBRW-PE	EBRW-R	EGCM-RT
1	0.97	0.99	1.00	0.99
	558	544	546	556
2	0.97	0.99	1.00	0.99
	567	544	544	556
3	1.00	0.99	1.00	0.99
	528	543	526	556
4	0.97	0.96	0.96	0.99
	759	714	731	736
5	0.97	0.97	0.96	0.99
	650	701	734	696
6	0.07	0.09	0.08	0.03
	816	788	786	816
7	0.08	0.11	0.07	0.24
	752	793	781	759
8	0.01	0.00	0.00	0.00
	559	582	568	556
9	0.00	0.00	0.00	0.00
	566	579	530	556
10	0.08	0.11	0.08	0.24
	753	793	786	759
11	0.97	0.97	0.95	0.99
	681	701	740	696
12	0.00	0.00	0.00	0.00
	570	579	533	556
13	0.00	0.00	0.00	0.00
	556	582	577	556
14	0.02	0.09	0.08	0.03
	828	788	791	816
15	0.95	0.96	0.95	0.99
	759	714	736	736
16	1.00	0.99	1.00	0.99
	534	542	527	556

Note: For each stimulus, the top row holds choice proportions for category *A* and the bottom row holds reaction times. EBRW-PE is the exemplar-based random walk with perceptual encoding, EBRW-R the restricted exemplar-based random walk, and EGCM-RT the extended generalized context model for response times.

Table 4
Observed and predicted choice proportions and mean response times (in ms) for participant 3

Stimulus	Observed	EBRW-PE	EBRW-R	EGCM-RT
1	0.94	0.99	1.00	1.00
	849	799	704	733
2	0.99	0.99	1.00	1.00
	781	736	704	731
3	1.00	0.99	1.00	1.00
	655	717	704	729
4	0.99	0.99	1.00	0.99
	898	848	818	1022
5	0.93	0.99	1.00	0.99
	864	806	818	971
6	0.15	0.17	0.11	0.06
	1108	1062	1081	1081
7	0.07	0.30	0.11	0.22
	936	1042	1081	1018
8	0.15	0.07	0.00	0.00
	967	940	879	977
9	0.01	0.03	0.00	0.00
	911	934	879	919
10	0.23	0.49	0.52	0.26
	1035	1079	1214	1018
11	0.74	0.72	0.82	0.99
	1051	1033	1113	971
12	0.02	0.09	0.06	0.00
	1097	1038	1029	1007
13	0.89	0.90	0.91	0.89
	993	1007	1029	1065
14	0.51	0.66	0.59	0.51
	1192	1222	1206	1081
15	0.84	0.84	0.87	0.99
	1085	1037	1073	1022
16	1.00	0.99	1.00	1.00
	614	728	704	732

Notes: For each stimulus, the top row holds choice proportions for category *A* and the bottom row holds reaction times. EBRW-PE is the exemplar-based random walk with perceptual encoding, EBRW-R the restricted exemplar-based random walk, and EGCM-RT the extended generalized context model for response times.

Some of the EGCM-RT parameters serve functions similar to the EBRW-PE parameters. Both w and u are dimension weights as in Eq. (1). As stated above, the r and q parameters determine the probability of including a dimension in the similarity computation. Unequal category criteria, A and B, and a b different from 0.5 result in a bias for one of the categories. Increasing A and B or θ will produce more deterministic responses. Both c and c^* scale the similarity space as in Eq. (3) and μ and t_{res} are both residual times. See Lamberts (2000) for a more thorough discussion of the parameters. The EBRW-PE has 12 free parameters, the EBRW-R has 8 free parameters, and the EGCM-RT has 11 free parameters. A single set of parameters was used to simultaneously fit both the RT and accuracy data. A different set was used for each participant.

There is currently no ideal method for combining goodness of fit measures for reaction times and choice proportions. Following Lamberts (2000), the models were fitted to the data by maximizing the sum of the proportions of variance accounted for in the reaction time and choice probability data (32 data points). Tables 2–6 show the model-fitting results for each participant. The summary measures of fit are reported in Table 7. The best-fitting parameters for the EBRW-PE, EBRW-R, and EGCM-RT are reported in Tables 10–12 respectively of Appendix A.

In general, all three models did a reasonably good job of describing the data. In particular, the EBRW-PE provided a good account of both the choice probabilities and reaction time data for all of the individual participants. Note that most of the fit variability resides in the reaction times. Fig. 2 provides an illustration of the reaction time fits for each model and participant. Direct comparisons between models are difficult because of the difference in the number and nature of the free

Table 5
Observed and predicted choice proportions and mean response times (in ms) for participant 4

Stimulus	Observed	EBRW-PE	EBRW-R	EGCM-RT
1	0.94	0.96	0.97	0.99
	978	922	852	978
2	0.92	0.96	0.97	0.99
	975	922	852	978
3	1.00	0.99	1.00	0.99
	567	593	568	585
4	0.91	0.93	0.96	0.99
	1025	984	895	979
5	0.96	0.99	1.00	0.99
	589	621	626	585
6	0.01	0.00	0.00	0.00
	591	573	584	585
7	0.09	0.18	0.15	0.00
	981	963	1135	980
8	0.07	0.09	0.10	0.00
	976	1017	1056	979
9	0.00	0.00	0.00	0.00
	564	537	482	585
10	0.11	0.18	0.15	0.10
	957	963	1135	980
11	0.97	0.99	1.00	0.99
	572	621	626	585
12	0.00	0.00	0.00	0.00
	557	537	584	585
13	0.11	0.09	0.10	0.00
	969	1017	1056	979
14	0.02	0.00	0.00	0.00
	557	573	482	585
15	0.94	0.93	0.96	0.93
	983	984	895	979
16	1.00	0.99	1.00	0.99
	558	593	568	585

Note: For each stimulus, the top row holds choice proportions for category A and the bottom row holds reaction times. EBRW-PE is the exemplar-based random walk with perceptual encoding, EBRW-R the restricted exemplar-based random walk, and EGCM-RT the extended generalized context model for response times.

Table 6
Observed and predicted choice proportions and mean response times (in ms) for participant 5

Stimulus	Observed	EBRW-PE	EBRW-R	EGCM-RT
1	0.92	0.98	1.00	0.99
	661	647	618	658
2	0.96	0.98	1.00	0.99
	670	647	618	658
3	1.00	0.99	1.00	0.99
	472	482	478	481
4	0.96	0.95	1.00	0.99
	668	682	643	691
5	0.95	0.99	1.00	0.99
	500	492	526	481
6	0.04	0.00	0.00	0.00
	466	487	508	482
7	0.15	0.19	0.01	0.14
	698	704	777	703
8	0.11	0.09	0.01	0.02
	727	720	713	708
9	0.00	0.00	0.00	0.00
	465	472	421	482
10	0.11	0.19	0.01	0.14
	718	704	777	703
11	0.97	0.99	1.00	0.99
	502	492	526	481
12	0.00	0.00	0.00	0.00
	495	472	421	482
13	0.11	0.09	0.01	0.02
	705	720	713	708
14	0.02	0.00	0.00	0.00
	466	487	508	482
15	0.92	0.95	1.00	0.99
	681	682	643	691
16	0.99	0.99	1.00	0.99
	477	482	478	481

Note: For each stimulus, the top row holds choice proportions for category A and the bottom row holds reaction times. EBRW-PE is the exemplar-based random walk with perceptual encoding, EBRW-R the restricted exemplar-based random walk, and EGCM-RT the extended generalized context model for response times.

parameters, but a few modeling trends deserve notice. First, the fits of the EGCM-RT and the EBRW-PE are very similar. Although the EGCM-RT accounts for more of the data variability for participants 1 and 2, the EBRW-PE still captures the main trends in these reaction time data. The EBRW-PE not only accounts for more of the data variability for participant 3, but, as can be seen in Fig. 2, it also does a better job of accounting for the fine-grain differences in the reaction times. Both models do an excellent job for the relatively less variable data from participants 4 and 5. It appears that more diagnostic experiments will be required to sharply distinguish these two models.

Recall that the EBRW-R is a submodel of the EBRW-PE with all inclusion rates set to one. To the extent that the additional perceptual encoding assumptions of the EBRW-PE are useful, they will manifest themselves in a better fit of the EBRW-PE over the EBRW-R. For participants 1 and 2, these two models provide almost identical fits. However, the EBRW-PE does provide a better account of the data for participants 3, 4, and 5. For these three participants, it appears that there is indeed a benefit of the extra processing assumptions.

A cross-validation analysis (Browne, 2000) was also performed to determine the extent to which the models are capturing random noise in the data. In a cross-validation analysis, a model is fit to a subset of the data, the calibration sample, and then tested on the remaining data, the validation sample. If a model is fitting noise in the calibration sample, it will perform relatively poorly on the validation sample. For example, according to the fits reported in Table 7, the EGCM-RT is clearly fitting the data of participant 1 better than the EBRW-PE. If the EGCM-RT is simply fitting the noise in the data, then the EGCM-RT will fit the calibration sample better than the EBRW-PE, but it will fit the validation sample

Table 7
Fit summaries for the EBRW-PE, EBRW-R, and EGCM-RT as measured by choice probability and reaction time proportion of variance accounted for (PVAF) for each participant

	Model		
Statistic	EBRW-PE	EBRW-R	EGCM-RT
Participant 1			
Choice	1.00	0.99	0.99
RT	0.87	0.85	0.94
Participant 2			
Choice	1.00	1.00	0.98
RT	0.91	0.90	0.97
Participant 3			
Choice	0.93	0.95	0.94
RT	0.87	0.77	0.73
Participant 4			
Choice	0.99	0.99	0.99
RT	0.97	0.83	0.99
Participant 5			
Choice	0.99	0.98	0.99
RT	0.98	0.86	0.98

Note: EBRW-PE is the exemplar-based random walk with perceptual encoding, EBRW-R the restricted exemplar-based random walk, and EGCM-RT the extended generalized context model for response times.

worse. All three models were fitted to the data obtained by collapsing across all the even-numbered trials. Then, holding the best-fitting parameters fixed, the models were evaluated on the data from the odd-numbered trials. The results are given in Table 13 of Appendix A. For all three models, all five participants, and both the choice probabilities and reaction times, the pattern of results mimic those of Table 7. That is, the relative performance across the calibration and validation samples was nearly identical. Thus, the relative performance of the models does not appear to be attributable to their ability to fit random noise in the data.

4. Modeling signal-to-respond data with the EBRW-PE

Both the EBRW-PE and the EGCM-RT appear to do a good job of fitting both the choice probabilities and reaction times in a speeded classification task involving separable-dimension stimuli. In numerous experiments, Lamberts has demonstrated that his dynamic model of similarity also successfully predicts classification choice probabilities that are observed under conditions of time pressure. The design of one particularly diagnostic experiment (Lamberts & Freeman, 1999a) is presented in Table 8. As illustrated in the table, the stimuli were composed of four binary-valued dimensions. Stimuli with values of 0 along each dimension tend to belong to category A, whereas stimuli with values of 1 along each dimension tend to belong to category B. However, Stimulus 5 (1111) is an exception stimulus that belongs to category A. Following an initial learning phase, participants engaged in a response-signal paradigm, in which a signal was presented at either 150 or 300 ms following the stimulus presentation. Participants were instructed to respond immediately to the response signal. Trials on which no signal was presented were also included in the design. The key result of interest is displayed in Fig. 3. For the 300 ms signal, participants classified the exception stimulus into the wrong category; whereas on the no-signal trials, there was a crossover in which participants classified the exception stimulus into the correct category.

Importantly, as pointed out by Lamberts and Freeman (1999a), the standard EBRW is unable to predict this crossover in classification that is observed with increased processing time. The reason is that the EBRW assumes a fixed probability that the counter moves in the direction of category A on each individual step of the random-walk. Therefore, increasing the amount of processing time can only increase the expected *distance* of the random-walk counter from the starting point of zero, it cannot change the expected *direction* of the random-walk counter. However, with the addition of perceptual encoding, the EBRW-PE can predict this crossover effect. Both the EBRW-PE and the EGCM-RT correctly predict this pattern of results for similar reasons. For the very early response signal, performance is near chance because no perceptual dimensions have been encoded. For the 300 ms signal, only a subset of the perceptual dimensions has been encoded, and so the exception stimulus will have a greater summed similarity to category B than to category A. Finally, on no-signal trials, all dimensions are encoded, and the exception stimulus will have greater summed similarity to the exemplars of its own category, in particular to itself, than to the exemplars of the contrast category.

Here we derive the analytic predictions of the EBRW-PE for the response-signal paradigm. Let $\mathbf{R}_+$ be an $N(A+B-1)$ column vector such that each element of $\mathbf{R}_+$ represents a transition state (l,ϕ) in the same order as the elements of $\mathbf{R}_A$. Then,

$$\text{the } (l,\phi)\text{th element of } \mathbf{R}_+ = \begin{cases} 1 & \text{if } l>0, \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

and similarly for $\mathbf{R}_-$ and $\mathbf{R}_0$ with the condition on l appropriately changed. The probability that the EBRW-PE will respond A after n steps is given by

$$P(A|n \text{ steps}) = \mathbf{Z}'\sum_{i=0}^{n-1}\mathbf{Q}^i\mathbf{R}_A + \mathbf{Z}'\mathbf{Q}^n\mathbf{R}_+ + \rho\cdot\mathbf{Z}'\mathbf{Q}^n\mathbf{R}_0, \tag{22}$$

where ρ is the base-rate of category A. The first term gives the probability of hitting the A absorbing boundary at or before n steps. We assume that if the random-walk hits a boundary before the signal, then the participant waits until the signal and then gives that

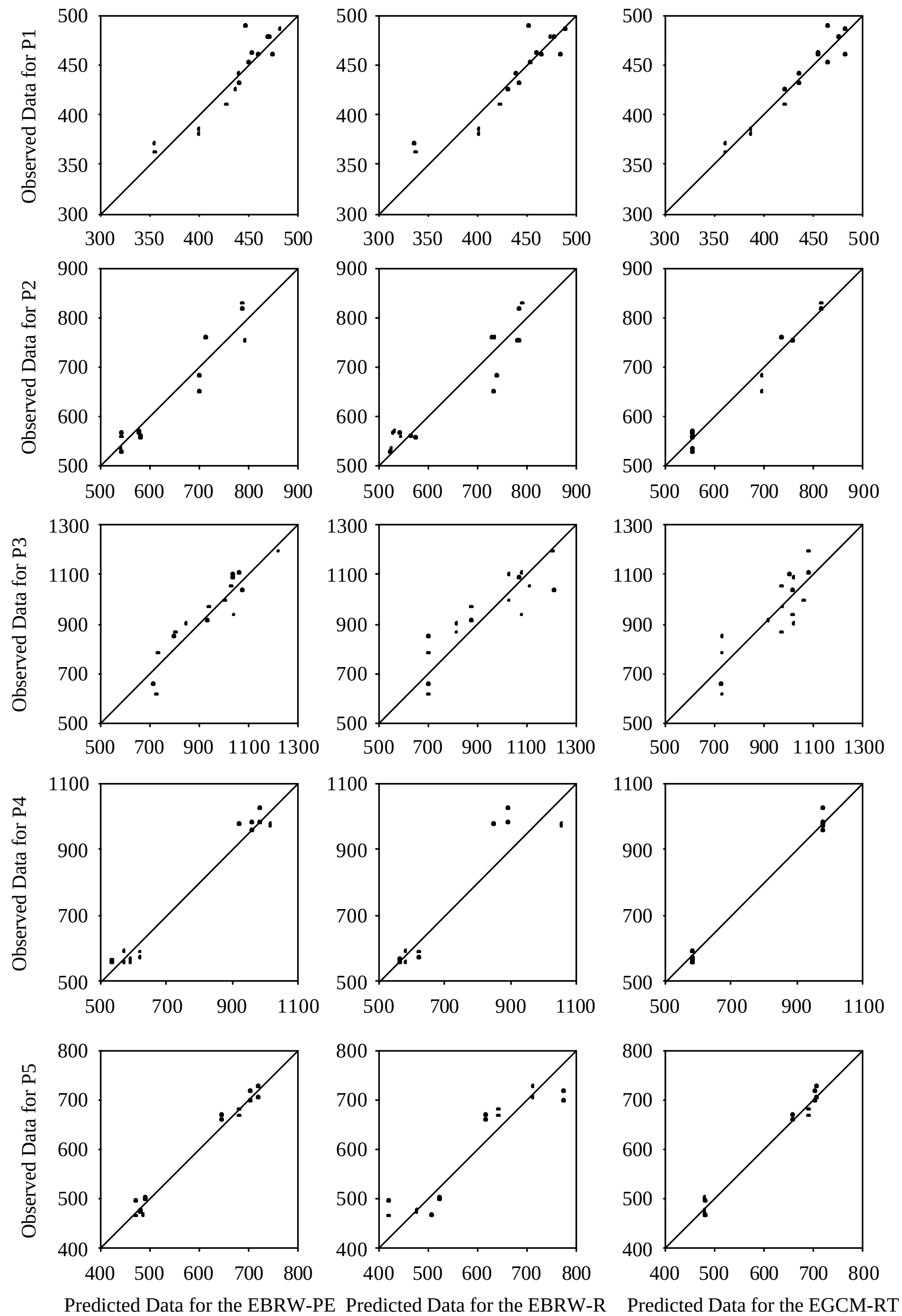

Fig. 2. Scatterplots of observed mean response time (ms) for each stimulus and each participant against the predicted mean response time for the EBRW-PE, EBRW-R, and EGCM-RT. P1–P5 are participants 1–5, respectively; EBRW-PE the exemplar-based random walk model with perceptual encoding; EBRW-R the restricted exemplar-based random walk model; and EGCM-RT the extended generalized context model for reaction times.

response. The second term gives the probability of being in a positive, transient random-walk position after n steps. The final term gives the probability that the random-walk is in position 0 after n steps, where we assume the participant guesses according to category base-rates, ρ.

We applied Eq. (22) of the EBRW-PE to the response-signal paradigm of Lamberts and Freeman (1999a). In the present example, we set the encoding probabilities, r_m, of all dimensions to 0.20. The attention weights, w_m, were set to 0.25, and the overall sensitive parameter was set at $c = 10.0$. The results for the critical exception stimulus and two of the non-exception stimuli are shown in the top panel of Fig. 4. Here we focus on the exception stimulus, i.e., Stimulus 5. As can be seen from the figure, in the very initial steps of the random-walk the counter is near the starting point of zero, so responding would be near chance. As processing continues, the counter moves in the direction of the category *B* response criterion, i.e., towards the wrong category. With still more processing time, however, the counter begins to switch directions and move towards category *A* boundary, i.e., the correct category choice for the exception. Thus, the EBRW-PE captures the key qualitative pattern of results reported by Lamberts and Freeman (1999a). The model behaves in this manner because early in processing, when relatively few dimensions have been encoded, the exception stimulus has greater summed similarity to the exemplars of category *B* than to category *A*, so the exemplars of category *B* are the ones that tend to be retrieved. As processing continues, all dimensions are eventually encoded, and the exception stimulus starts to retrieve its own memory trace, driving the random-walk in the direction of the correct category.

Table 8
Category structure from Lamberts and Freeman (1999a)

Category and Stimulus	Dimension D1	D2	D3	D4
Category A				
1	0	0	0	1
2	0	0	1	0
3	0	1	0	0
4	1	0	0	0
5	1	1	1	1
Category B				
6	1	1	1	0
7	1	1	0	1
8	1	0	1	1
9	0	1	1	1

Note: D1–D4 are stimulus dimensions 1–4, respectively.

The bottom panel of Fig. 4 shows the same fits with the encoding probabilities, r_m, set at 1.0, i.e., the EBRW-R model. With these encoding probabilities, all

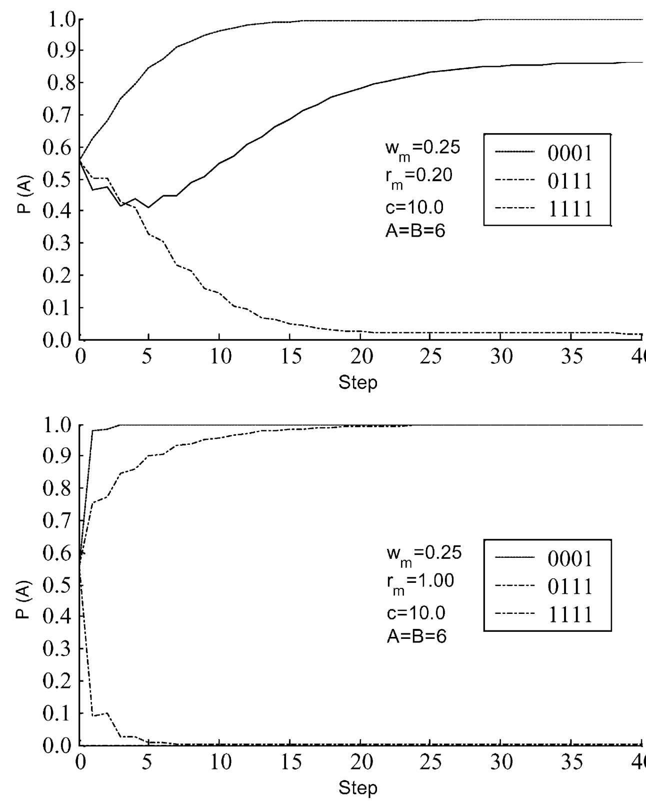

Fig. 4. Probability that the EBRW-PE classifies a stimulus into category *A* when applied to the response-signal paradigm of Lamberts and Freeman (1999a). w_m is the dimension attention weights, r_m the inclusion rates, *c* the sensitivity parameter, and *A* and *B* are categories *A* and *B* criterion, respectively. The fits in the upper panel had inclusion rates less than one. The inclusion rates in the lower panel were fixed at one, which is equivalent to the EBRW-R. Stimulus 1111 is the exception stimulus.

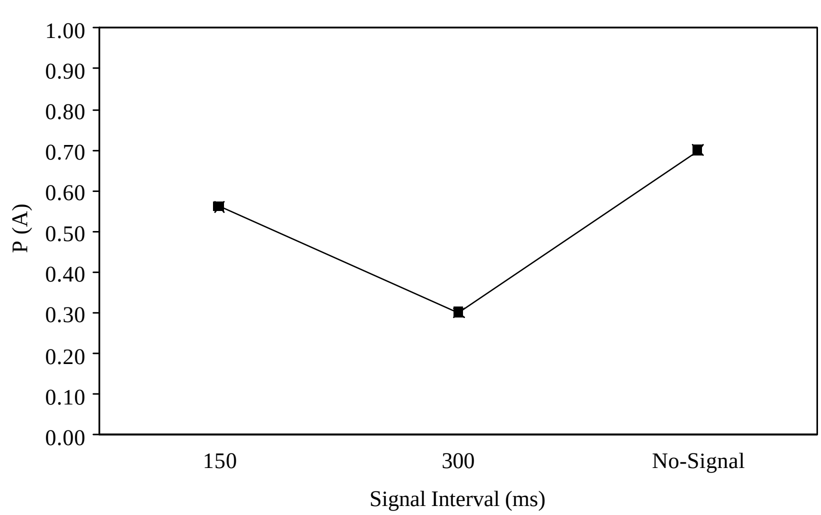

Fig. 3. Proportion of category *A* responses as a function of response signal for the exception stimulus (stimulus 5) in Experiment 3 of Lamberts and Freeman (1999a).

dimensions are encoded on the very first step of the random-walk. The important aspect of this figure to note is that the choice probability of the exception stimulus moves consistently towards category A, the correct category, without the dip towards category B as in the EBRW-PE. Thus, although the EBRW-R did a good job of fitting the data from the speeded-classification experiment, it cannot predict the qualitative pattern of results demonstrated in the response-signal paradigm of Lamberts and Freeman (1999a). The additional perceptual encoding assumptions of the EBRW-PE allow that model to account for such data.

5. Discussion

5.1. Summary

In this article, we proposed a new model of speeded classification that integrates the random-walk process of Nosofsky and Palmeri's (1997a) EBRW and the stochastic sampling of stimulus dimensions of Lamberts' (2000) EGCM-RT. As in the EBRW, the EBRW-PE represents exemplars as points in a multidimensional psychological space. When a test object is presented, all exemplars in memory race to be retrieved. How fast an exemplar races is based on its similarity to the test object. The winning exemplar adds incremental evidence to a random-walk process. A response occurs when the random-walk counter reaches a criterion. The main advance of the EBRW-PE over the standard EBRW is the inclusion of stochastic perceptual encoding. As in the EGCM-RT, on each step of the random-walk, there is a chance that a particular stimulus dimension will be encoded and the similarity of two objects depends on the set of encoded stimulus dimensions at the time of comparison. To derive analytic predictions, the EBRW-PE was implemented as a Markov chain.

In a speeded classification experiment involving separable-dimension stimuli, we demonstrated that the EBRW-PE is capable of producing accurate quantitative predictions of classification choice probabilities and mean response times for individual participants. Although further tests are needed to clearly demarcate the need for the perceptual encoding assumption in this task, there was some evidence that the perceptual encoding assumptions do indeed benefit the model. Also, because the standard EBRW does not involve the time-extended accumulation of stimulus information, it is not able to predict the crossover effect observed in the Lamberts and Freeman (1999a) response-signal paradigm. With the added assumption of stochastic perceptual encoding, the EBRW-PE can predict this crossover. The idea is that early in the random-walk, few dimensions have been included in the similarity computation, so the summed similarity of the exception stimulus is actually higher to the incorrect category and so the random-walk will move in the wrong direction. However, as time progresses, more dimensions will be included and the summed similarity to the correct category will eventually become greater, allowing the random-walk to move in the correct direction.

5.2. Benefits of the EBRW-PE

Based on the present results, we conclude that the EBRW-PE provides another viable account of the time course of categorization of separable-dimension stimuli. However, the data of the current experiment do not allow us to sharply discriminate between the EBRW-PE and the EGCM-RT. The main potential advantage of the EBRW-PE is that it maintains a traditional random-walk formalism whereas the EGCM-RT uses a less-studied stopping-rule decision mechanism. A vast amount of research suggests that random-walk models are consistent with detailed aspects of reaction time data such as the form of the reaction time distributions and both error and correct reaction times (Ratcliff, 1978). However, future work is needed to contrast the EBRW-PE and EGCM-RT.

5.3. Variable criteria

A limitation of the version of the EBRW-PE developed in this article is that it assumed fixed and integer-valued response criteria. The discrete values of the response criteria are required in the Markov-chain implementation of the random-walk and undoubtedly place limits on the degree of quantitative accuracy that the model can achieve in its predictions. A straightforward approach to extending the model is to make allowance for variable criterion settings across different trials of an experimental session (Laming, 1968; Ratcliff & Rouder, 1998; Rouder, 1996; Van Zandt & Ratcliff, 1995). For any given setting of the response criteria, the Markov-chain model would be used to generate predictions of choice probability and mean response time. The overall predictions would be obtained by averaging across a probability distribution of the individual Markov-chain model predictions. An additional virtue of this approach is that it would allow the model to account for speed-accuracy tradeoffs in which error responses are made more rapidly than are correct responses (see Van Zandt & Ratcliff, 1995, for a more detailed discussion).

5.4. Relations to Ratcliff's (1978) diffusion model

Ratcliff's diffusion model is arguably the most successful of all response time models, accounting for choice probabilities, means and distributions of reaction times, and error latencies that well describe experimental data across a large range of experimental paradigms. A diffusion process is basically a random-walk in which the information that drives a decision process is accumulated continuously over time instead of in discrete steps. In a diffusion process, the rate at which information accumulates, that is, the mean amount of information accumulated per unit of time, is called the drift rate, v, analogous to p_i^A or p_i^B of Eqs. (7) and (8). In Ratcliff's (1978) diffusion model, the drift rate is not constant, but rather varies within a particular trial. The variability is assumed to be normally distributed with standard deviation σ.

The relationship between v and σ can account for differences in stimulus reaction times and error rates. Each stimulus or stimulus type can have its own drift rate. For example, if a particular stimulus has a large, positive drift rate, then, on average, the diffusion process will move consistently towards the positive boundary. It will reach the positive boundary rapidly and rarely reach the negative boundary regardless of the amount of noise in the drift. However, noise will have a much larger effect on a stimulus with a lower positive drift rate. In this case, the diffusion process will waver back and forth, resulting in a much longer time to reach the correct boundary and a much higher probability of hitting the incorrect boundary. Thus, "easy" stimuli can be modeled as having extreme mean drift rates, they move quickly and accurately towards the correct boundary. "Difficult" stimuli can be modeled as having drift rates closer to zero, they move more slowly towards the correct boundary and there is a much higher chance of an incorrect response (Ratcliff & Rouder, 1998).

Table 9
Physical dimensions assigned to each logical dimension for each participant

	Logical dimensions			
Participant	D1	D2	D3	D4
1	Top	Base	Shade	Upright
2	Shade	Upright	Top	Base
3	Upright	Shade	Base	Top
4	Shade	Top	Upright	Base
5	Base	Upright	Top	Shade
6	Base	Top	Upright	Shade
7	Top	Shade	Base	Upright
8	Upright	Base	Shade	Top

Note: Only the data from participants 1–5 were analyzed.

Table 10
Best fitting parameters for the EBRW-PE for each participant when simultaneously fit to the response time and choice probability data

	Participants				
Parameter	P1	P2	P3	P4	P5
w_1	0.011	0.006	0.306	0.317	0.134
w_2	0.257	0.077	0.319	0.259	0.017
w_3	0.270	0.459	0.035	0.346	0.185
w_4	0.463	0.459	0.339	0.078	0.664
r_1	0.463	1.000	0.519	1.000	0.082
r_2	0.781	0.000	0.308	0.241	0.000
r_3	0.625	0.849	0.544	0.000	0.388
r_4	0.427	0.700	0.356	0.070	1.000
c	9.305	100.000	43.653	34.688	16.875
α	50.938	18.048	75.894	44.609	24.000
μ	174.219	466.797	523.438	182.031	293.750
A	4.000	4.000	2.000	8.000	7.000
B	3.000	6.000	5.000	6.000	6.000

Note: P1–P5 the participants 1–5, respectively; EBRW-PE the exemplar-based random walk with perceptual encoding; w_1–w_4 the attention weights; r_1–r_4 the encoding probabilities; c the sensitivity parameter; α the step-time constant; μ the reaction time intercept; A the category A criterion; and B the category B criterion.

Table 11
Best-fitting parameters for the EBRW-R for each participant when simultaneously fit to the response time and choice probability data

	Participants				
Parameter	P1	P2	P3	P4	P5
w_1	0.005	0.181	0.315	0.698	0.100
w_2	0.379	0.012	0.317	0.205	0.000
w_3	0.298	0.401	0.033	0.000	0.158
w_4	0.318	0.407	0.335	0.097	0.743
c	8.252	6.500	33.195	9.488	14.000
α	66.208	79.063	58.379	103.438	64.922
μ	64.063	275.156	587.500	121.875	140.000
A	5.000	3.000	2.000	4.000	5.000
B	4.000	3.000	5.000	3.000	4.000

Note: P1–P5 are the participants 1–5, respectively; EBRW-R the restricted exemplar-based random walk; w_1–w_4 attention weights; c the sensitivity parameter; α the step-time constant; μ the reaction time intercept; A the category A criterion; and B the Category B criterion.

Table 12
Best-fitting parameters for the EGCM-RT for each participant when simultaneously fit to the response time and choice probability data

	Participants				
Parameter	P1	P2	P3	P4	P5
u_1	0.000	0.090	0.293	0.499	0.085
u_2	0.487	0.004	0.307	0.424	0.000
u_3	0.184	0.420	0.246	0.030	0.464
u_4	0.329	0.487	0.154	0.047	0.451
q_1	0.005	0.009	5.509	9.696	0.011
q_2	0.010	0.004	1.166	7.816	0.005
q_3	0.011	9.999	0.003	0.004	6.108
q_4	0.019	0.092	0.004	0.004	9.816
c^*	11.980	24.167	11.371	20.215	17.713
t_{res}	284.794	555.517	718.539	585.352	480.959
b	0.326	0.776	0.790	0.450	0.667
θ	16.230	19.463	19.395	19.766	19.961

Note: EGCM-RT is the extended generalized context model for reaction times; u_1–u_4 are the utility values; q_1– q_4 are the inclusion rates; c^* the generalization parameter; t_{res} the residual time; b the bias; θ the response-scaling parameter. See Lamberts (2000) for a discussion of these parameters.

Table 13
Fit summaries for the EBRW-PE, EBRW-R, and EGCM-RT as measured by choice probability and reaction time proportion of variance accounted for (PVAF) for each participant when the models were trained on the even trial data and tested on the odd trial data

	Models					
	EBRW-PE		EBRW-R		EGCM-RT	
Statistic	Calibration	Validation	Calibration	Validation	Calibration	Validation
Participant 1						
Choice	0.99	1.00	0.99	0.99	1.00	0.99
RT	0.87	0.78	0.86	0.79	0.92	0.89
Participant 2						
Choice	0.99	0.99	1.00	1.00	0.99	0.99
RT	0.94	0.91	0.90	0.87	0.98	0.94
Participant 3						
Choice	0.95	0.92	0.96	0.93	0.95	0.93
RT	0.88	0.79	0.77	0.74	0.72	0.63
Participant 4						
Choice	0.99	0.99	0.99	1.00	0.99	0.99
RT	0.97	0.96	0.82	0.82	0.99	0.98
Participant 5						
Choice	1.00	0.99	0.98	0.98	1.00	0.99
RT	0.98	0.97	0.86	0.86	0.99	0.97

Note: EBRW-PE is the exemplar-based random walk with perceptual encoding, EBRW-R the restricted exemplar-based random walk, and EGCM-RT the extended generalized context model for response times. The calibration and validation samples were the even- and odd-numbered trials, respectively.

To account for differences in stimulus performance across different trials, mean drift rates are assumed to be normally distributed across trials with standard deviation, η. Thus, not only does v vary within a trial, it can also vary across trials. However, within a trial, the mean of the drift rate for a particular stimulus, v, is always assumed constant. This is not the case in the EBRW-PE, where p_i^A and p_i^B change within a trial based on the set of dimensions that have been encoded at a particular time. Thus, the EBRW-PE gives a process account of how a diffusion process or random-walk mean drift rate may change systematically within a trial.

Acknowledgments

This work was supported by Grant PHS R01 MH48494-09 from the National Institute of Mental Health to Robert M. Nosofsky and by a Jacob Javits Fellowship and an NIMH Training Grant in Mathematical Modeling to Andrew L. Cohen.

The authors thank Jerome Busemeyer and Richard Shiffrin for helpful discussions, criticisms, and suggestions related to this work.

Appendix A

The assignment of each physical dimension to each logical dimension is given in Table 9. The best-fitting parameters for the EBRW-PE, EBRW-R, and EBRW-RT for each participant are given in Tables 10–12, respectively. The results of the cross-validation analysis are given Table 13.

All three models were fitted to the data and evaluated. The results are given in Table 13.

References

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

Ashby, F. G., Boynton, G., & Lee, W. W. (1994). Categorization response time with multidimensional stimuli. *Perception and Psychophysics*, *55*, 11–27.

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology*, *44*, 108–132.

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

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

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

Garner, W. R. (1974). *The processing of information and structure*. New York: Wiley.

Lamberts, K. (1995). Categorization under time pressure. *Journal of Experimental Psychology: General*, *124*, 161–180.

Lamberts, K. (1998). The time course of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 695–711.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*, 227–260.

Lamberts, K., & Freeman, R. P. J. (1999a). Building object representations from parts: Tests of a stochastic sampling model.

Journal of Experimental Psychology: Human Perception and Performance, *25*, 904–926.

Lamberts, K., & Freeman, R. P. J. (1999b). Categorization of briefly presented objects. *Psychological Research/Psychologische Forschung*, *62*, 107–117.

Laming, D. R. J. (1968). *Information theory of choice reaction time*. New York: Wiley.

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

Lockhead, G. R. (1972). Processing dimensional stimuli: A note. *Psychological Review*, *79*, 410–419.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.

Marley, A. A. J., & Colonius, H. (1992). The "horse race" random utility model for choice probabilities and response times, and its competing risks interpretation. *Journal of Mathematical Psychology*, *35*, 1–20.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 87–108.

Nosofsky, R. M., & Alfonso Reese, L. A. (1999). Effects of similarity and practice on speeded classification response times and accuracies: Further tests of an exemplar-retrieval model. *Memory and Cognition*, *27*, 78–93.

Nosofsky, R. M., & Palmeri, T. J. (1997a). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266–300.

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

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

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

Rouder, J. N. (1996). Premature sampling in random walks. *Journal of Mathematical Psychology*, *40*, 289–296.

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

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317–1323.

Shiffrin, R. M., & Thompson, M. (1988). Moments of transition-additive random variables defined on finite, regenerative random processes. *Journal of Mathematical Psychology*, *32*, 313–340.

Van Zandt, T., & Ratcliff, R. (1995). Statistical mimicking of reaction time data: Single-process models, parameter variability, and mixtures. *Psychonomic Bulletin and Review*, *2*, 20–54.