CHAPTER

7 An Exemplar-Based Random-Walk Model of Categorization and Recognition

Robert M. Nosofsky and Thomas J. Palmeri

Abstract

In this chapter, we provide a review of a process-oriented mathematical model of *categorization* known as the *exemplar-based random-walk* (EBRW) model (Nosofsky & Palmeri, 1997a). The EBRW model is a member of the class of *exemplar* models. According to such models, people represent categories by storing individual exemplars of the categories in memory, and classify objects on the basis of their similarity to the stored exemplars. The EBRW model combines ideas ranging from the fields of choice and similarity, to the development of *automaticity*, to response-time models of evidence accumulation and decision-making. This integrated model explains relations between categorization and other fundamental cognitive processes, including individual-object identification, the development of expertise in tasks of skilled performance, and old-new *recognition memory*. Furthermore, it provides an account of how categorization and recognition decision-making unfold through time. We also provide comparisons with some other process models of categorization.

Key Words: categorization, recognition, exemplar model, response times, automaticity, random walk, memory search, expertise, similarity, practice effects

Introduction

A fundamental issue in cognitive psychology and cognitive science concerns the manner in which people represent categories and make classification decisions (Estes, 1994; Smith & Medin, 1981). There is a wide variety of process-oriented mathematical models of *categorization* that have been proposed in the field. For example, according to *prototype* models (e.g., Posner & Keele, 1968; Smith & Minda, 1998), people represent categories by storing a summary representation, usually presumed to be the central tendency of the category distribution. Classification decisions are based on the similarity of a test item to the prototypes of alternative categories. According to *decision-boundary* models (e.g., Ashby & Maddox, 1993; McKinley & Nosofsky, 1995), people construct boundaries, usually assumed to be linear or quadratic in form, to divide a stimulus space into category response regions. If an object is perceived to lie in Region A of the space, then the observer emits a Category-A response. According to *rule-plus-exception* models (e.g., Davis, Love, & Preston, 2012; Erickson & Kruschke, 1998; Nosofsky, Palmeri, & McKinley, 1994), people construct low-dimensional logical rules for summarizing categories, and they remember occasional exceptions that may be needed to patch those rules.

In this chapter, however, our central focus is on *exemplar* models of classification. According to exemplar models, people represent categories by storing individual exemplars in memory, and classify objects on the basis of their similarity to the stored exemplars (Hintzman, 1986; Medin & Schaffer, 1978; Nosofsky, 1986). For instance, such models would assume that people represent the category of "birds" by storing in memory the vast collection of different robins, sparrows, eagles (and so forth) that they have experienced during their lifetimes. If a novel object were similar to some of these bird exemplars, then a person would tend to classify it as a bird.

Although alternative classification strategies are likely to operate across different experimental contexts, there are several reasons why we chose to focus on exemplar models in this chapter. One reason is that models from that class have provided a successful route to explaining relations between categorization and a wide variety of other fundamental cognitive processes, including individual-object identification (Nosofsky, 1986, 1987), the development of *automaticity* in tasks of skilled performance (Logan, 1988; Palmeri, 1997), and old-new *recognition memory* (Estes, 1994; Hintzman, 1988; Nosofsky, 1988, 1991). A second reason is that, in our view, for most "natural" category structures (Rosch, 1978) that cannot be described in terms of simple logical rules, exemplar models seem to provide the best-developed account for explaining how categorization decision-making unfolds over time. Thus, beyond predicting classification choice probabilities, exemplar models provide detailed quantitative accounts of classification response times (Nosofsky & Palmeri, 1997a). We now briefly expand these themes before turning to the main body of our chapter.

One of the central goals of exemplar models has been to explain relations between *categorization* and other fundamental cognitive processes, including old-new *recognition memory* (Estes, 1994; Hintzman, 1988; Nosofsky, 1988, 1991; Nosofsky & Zaki, 1998). Whereas in categorization people organize distinct objects into groups, in recognition the goal is to determine if some individual object is "old" (previously studied) or "new." Presumably, when people make recognition judgments, they evaluate the similarity of test objects to the individual previously studied items (i.e., exemplars). If categorization decisions are also based on similarity comparisons to previously stored exemplars, then there should be close relations between the processes of recognition and categorization.

A well-known model that formalizes these ideas is the *generalized context model* (GCM; Nosofsky, 1984, 1986, 1991). In the GCM, individual exemplars are represented as points in a multidimensional psychological space, and similarity between exemplars is a decreasing function of the distance between objects in the space (Shepard, 1987). The model presumes that both classification and recognition decisions are based on the "summed similarity" of a test object to the exemplars in the space. By conducting similarity-scaling studies, one can derive *multidimensional scaling* (*MDS*) solutions in which the locations of the exemplars in the similarity space are precisely located (Nosofsky, 1992). By using the GCM in combination with these MDS solutions, one can then achieve successful fine-grained predictions of classification and recognition choice probabilities for individual items (Nosofsky, 1986, 1987, 1991; Shin & Nosofsky, 1992).

A significant development in the application of the GCM has involved extensions of the model to explaining how the categorization process unfolds through time. So, for example, the exemplar model not only predicts choice probabilities, but also predicts categorization and recognition response times (RTs). This development is important because RT data often provide insights into cognitive processes that would not be evident based on examination of choice-probability data alone. Nosofsky and Palmeri's (1997a,b) *exemplar-based random-walk* (EBRW) model adopts the same fundamental representational assumptions as does the GCM. However, it extends that earlier model by assuming that retrieved exemplars drive a *random walk* process (e.g., Busemeyer, 1985; Link, 1992; Ratcliff, 1978). This exemplar-based random-walk model allows one to predict the time course of categorization and recognition decision making.

In this chapter, we provide a review of this EBRW model and illustrate its applications to varieties of categorization and recognition choice-probability and RT data. In the section on The Formal EBRW Model of Categorization RTs we provide a statement of the formal model as applied to categorization. As will be seen, the EBRW model combines ideas ranging from the fields of choice and similarity, to the development of automaticity, to RT models of evidence accumulation and decision making. In the section Effects of Similarity and Practice on Speeded Classification, we describe applications of the model to speeded perceptual classification, and illustrate how it captures fundamental phenomena including effects of similarity and practice. In the section Automaticity and Perceptual Expertise we describe how the EBRW accounts for the development of automatic categorization and perceptual expertise. In the section Using Probabilistic Feedback Manipulations to Contrast the Predictions From the EBRW Model and Alternative Models we describe experimental manipulations that have been used to try to

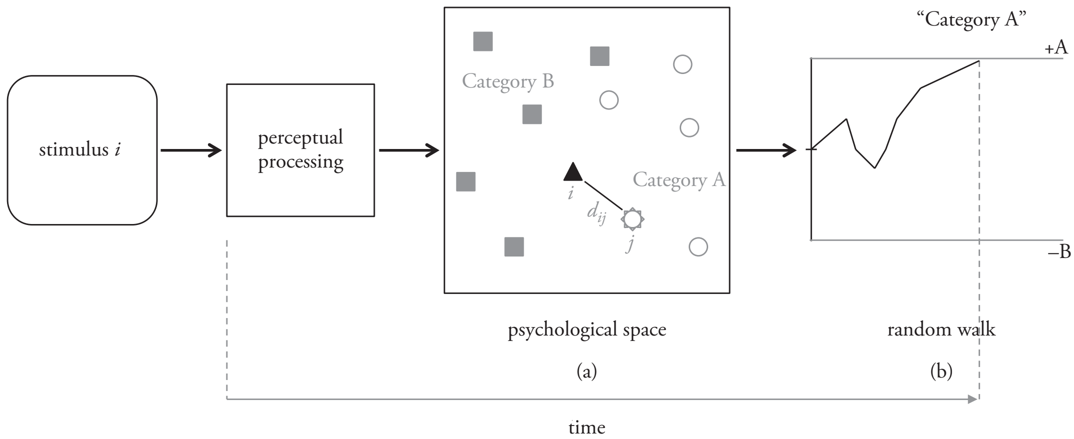

Fig. 7.1 Schematic illustration of the ideas behind the exemplar-based random-walk model. Panel a: Exemplars are represented as points in a multidimensional psychological space. Similarity is a decreasing function of the distance between objects in the space. Presentation of a test probe causes the exemplars to be "activated" in accord with how similar they are to that probe. The activated exemplars "race" to be retrieved. Panel b: The retrieved exemplars drive a random-walk process for making categorization decisions. Each time that an exemplar from Category A is retrieved, the random walk takes unit step towards the Category-A response threshold, and likewise for Category B. The retrieval process continues until one of the response thresholds is reached.

distinguish between the predictions of the EBRW model and some other major models of speeded categorization, including decision-boundary and prototype models. Finally, in the section Extending the EBRW Model to Predict Old-New Recognition RTs, we describe recent developments in which the EBRW model has been used to account for old-new recognition RTs. We then provide conclusions and questions for future research in the section Conclusions and New Research Goals.

The Formal EBRW Model of Categorization RTs

The EBRW model and the GCM build upon classic theories in the areas of choice and similarity (Shepard, 1957, 1987). As described in the introduction, in the model, exemplars are represented as points in a multidimensional psychological space (for an illustration, see Figure 7.1a). The distance between exemplars i and j (d_{ij}) is given by the *Minkowski power model*,

$$d_{ij} = \left[\sum_{k=1}^{K} w_k \left| x_{ik} - x_{jk} \right|^{\rho} \right]^{\frac{1}{\rho}}, \qquad (1)$$

where x_{ik} is the value of exemplar i on psychological dimension k; K is the number of dimensions that define the space; ρ defines the distance metric of the space; and w_k ($0 < w_k$, $\sum w_k = 1$) is the weight given to dimension k in computing distance. In situations involving the classification of holistic or *integral-dimension stimuli* (Garner, 1974), which will be the main focus of the present chapter, ρ is set equal to 2, which yields the familiar Euclidean distance metric. The dimension weights w_k are free parameters that reflect the degree of "attention" that subjects give to each dimension in making their classification judgments (Carroll & Wish, 1974). In situations in which some dimensions are far more relevant than others in allowing subjects to discriminate between members of contrasting categories, these *attention-weight parameters* may play a significant role (e.g., Nosofsky, 1986, 1987). In the experimental situations considered in this chapter, however, all dimensions tend to be relevant and the attention weights will turn out to play a relatively minor role.

The similarity of test item i to exemplar j is an exponentially decreasing function of their psychological distance (Shepard, 1987),

$$s_{ij} = exp(-c \cdot d_{ij}), \qquad (2)$$

where c is an overall *sensitivity parameter* that measures overall discriminability in the psychological space. The sensitivity governs the rate at which similarity declines with distance. When sensitivity is high, the similarity gradient is steep, so even objects that are close together in the space may be highly discriminable. By contrast, when sensitivity is low, the similarity gradient is shallow, and objects are hard to discriminate.

Each exemplar j is stored in memory (along with its associated category feedback) with memory-strength m_j. As will be seen, the precise assumptions involving the *memory strengths* vary with the specific experimental paradigm that is tested. For example, in some paradigms, we attempt to model performance on a trial-by-trial basis, and assume that the memory strengths of individual exemplars decrease systematically with their lag of presentation. In other paradigms, we attempt to model performance in a transfer phase that follows a period of extended training. In that situation, we typically assume that the memory strengths are proportional to the frequency with which each individual exemplar was presented in common with given category feedback during the initial training phase.

When a test item is presented, it causes all exemplars to be "activated" (Figure 7.1a). The activation for exemplar j, given presentation of item i, is given by

$$a_{ij} = m_j s_{ij}. \tag{3}$$

Thus, the exemplars that are most highly activated are those that have the greatest memory strengths and are highly similar to the test item. In modeling early-learning behavior, we also presume that "background" elements exist in memory at the start of training that are not associated with any of the categories. These *background elements* are presumed to have fixed activation b, independent of the test item that is presented.

Borrowing from Logan's (1988) highly influential instance theory of automaticity, the EBRW assumes that presentation of a test item causes the activated stored exemplars and background elements to "race" to be retrieved. For mathematical simplicity, the race times are presumed to be independent *exponentially distributed random variables* with rates proportional to the degree to which exemplar j is activated by item i (Bundesen, 1990; Logan, 1997; Marley, 1992). Thus, the probability density that exemplar j completes its race at time t, given presentation of item i, is given by

$$f(t) = a_{ij} \cdot exp(-a_{ij} \cdot t). \tag{4}$$

This assumption formalizes the idea that although the retrieval process is stochastic, the exemplars that tend to race most quickly are those that are most highly activated by the test item. The exemplar (or background element) that "wins" the race is retrieved.

Whereas in Logan's (1988) model the response is based on only the first retrieved exemplar, in the EBRW model exemplars from multiple retrievals drive a random-walk evidence accumulation process (e.g., Busemeyer, 1982; Ratcliff, 1978). In a two-category situation, the process operates as follows (for an illustration, see Figure 7.1b): First, there is a random-walk counter with initial setting zero. The observer establishes *response thresholds* representing the amount of evidence needed to make either a Category-A response (+A) or a Category-B response (–B). Suppose that exemplar x wins the race on a given step. If x received Category-A feedback during training, then the random-walk counter is increased by unit value in the direction of +A, whereas if x received Category-B feedback, the counter is decreased by unit value in the direction of –B. (If a background element wins the race, the counter's direction of change is chosen randomly.) If the counter reaches either threshold +A or –B, the corresponding categorization response is made. Otherwise, the races continue, another exemplar is retrieved (possibly the same one as on the previous step), and the random walk takes its next step.

Given the processing assumptions outlined earlier, Nosofsky and Palmeri (1997a) showed that, on each step of the random walk, the probability (p_i) that the counter is increased in the direction of threshold +A is given by

$$p_i = (S_{iA} + b)/(S_{iA} + S_{iB} + 2b), \tag{5}$$

where S_{iA} denotes the summed activation of all currently stored Category-A exemplars given presentation of item i (and likewise for S_{iB}). (The probability that the counter is decreased in the direction of Category B is given by q_i = 1-p_i.) Thus, as the summed activation of Category-A exemplars increases, the probability of retrieving Category-A exemplars, and thereby the probability of moving the counter in the direction of +A, increases. The categorization decision time is determined jointly by the total number of steps required to complete the random walk and by the speed with which those individual steps are made.

Given these random-walk processing assumptions, it is straightforward to derive analytic predictions of classification choice probabilities and mean RTs for each stimulus at any given stage of the learning process (Busemeyer, 1982). The relevant equations are summarized by Nosofsky and Palmeri (1997a, pp. 269–270, 291–292). Here, we focus on some key conceptual predictions that follow from the model.

A first prediction is that rapid classification decisions should be made for items that are highly similar to exemplars from their own target category and that are dissimilar to exemplars from the contrast category. Under such conditions, all retrieved exemplars will tend to come from the target category, so the random walk will march consistently toward that category's response threshold. For example, if an item is highly similar to the exemplars from Category A, and dissimilar to the exemplars from Category B, then the value p_i in Eq. 5 will be large, so almost all steps in the random walk will move in the direction of +A. By contrast, if an item is similar to the exemplars of both Categories A and B, then exemplars from both categories will be retrieved; in that case, the random walk will meander back and forth, leading to a slow RT.

A second prediction is that practice in the classification task will lead to more accurate responding and faster RTs. Early in training, before any exemplars have been stored in memory, the background-element activations are large relative to the summed-activations of the stored exemplars (see Eq. 5). Thus, the random-walk step probabilities (p_i) hover around .5, so responding is slow and prone to error. As the observer accumulates more category exemplars in memory, the summed activations S_{iA} and S_{iB} grow in magnitude, so responding is governed more by the stored exemplars. A second reason that responding speeds up is that the greater the number of exemplars stored in memory, the faster the "winning" retrieval times tend to be (cf. Logan, 1988). The intuition is that the greater the number of exemplars that participate in the race, the higher is the probability that some particular exemplar will finish quickly. Therefore, as more exemplars are stored, the speed of the individual steps in the random walk increases.

As discussed later in this chapter, many more fine-grained predictions follow from the model, some of which are highly diagnostic for distinguishing between the predictions of the EBRW model and alternative models. We describe these predictions in the context of the specific experimental paradigms in which they are tested.

One other important formal aspect of the model involves its predictions of classification choice probabilities. In particular, in the special case in which the response thresholds $+A$ and $-B$ are set an equal magnitude γ from the starting point of the random walk, the model predicts that the probability that item i is classified in Category A is given by

$$P(A|i) = (S_{iA}+b)^{\gamma}/[(S_{iA}+b)^{\gamma}+[(S_{iB}+b)^{\gamma}]. \quad (6)$$

This equation is the descriptive equation of choice probability that is provided by the GCM, one of the most successful formal models of perceptual classification (for discussion, see, e.g., Nosofsky, 1986; Wills & Pothos, 2012).[1] Thus, besides providing a formal account of classification RTs, the EBRW model provides a processing interpretation for the emergence of the GCM response rule.

Effects of Similarity and Practice on Speeded Classification

In their initial tests of the EBRW model, Nosofsky and Palmeri (1997a) conducted a speeded classification task using the category structure shown in Figure 7.2. The stimuli were a set of 12 Munsell colors of a constant red hue varying in their brightness and saturation[2]. As illustrated in the figure, half the exemplars were assigned by the experimenter to Category A (squares) and half to Category B (circles). On each trial, a single color was presented, the observer classified it into one of the two categories, and corrective feedback was then provided. Testing was organized into 150 blocks of 12 trials (30 blocks per day), with each color presented once in a random order in each block. (See Nosofsky and Palmeri, 1997a, pp. 273–274 for further methodological details.)

The mean RTs observed for one of the participants are shown in Figure 7.3. The top panel illustrates the mean RT for each individual color, averaged across the final 120 blocks. The diameter of the circle enclosing each stimulus is linearly related to the stimulus's mean RT. To help interpret these data, the figure also displays a dotted boundary of equal summed-similarity to the exemplars

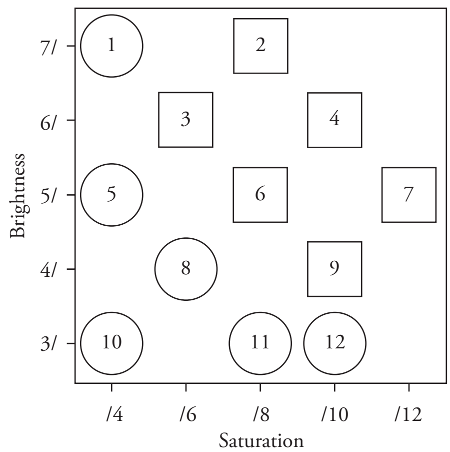

Fig. 7.2 Schematic illustration of the category structure tested by Nosofsky and Palmeri (1997a, Experiment 1).

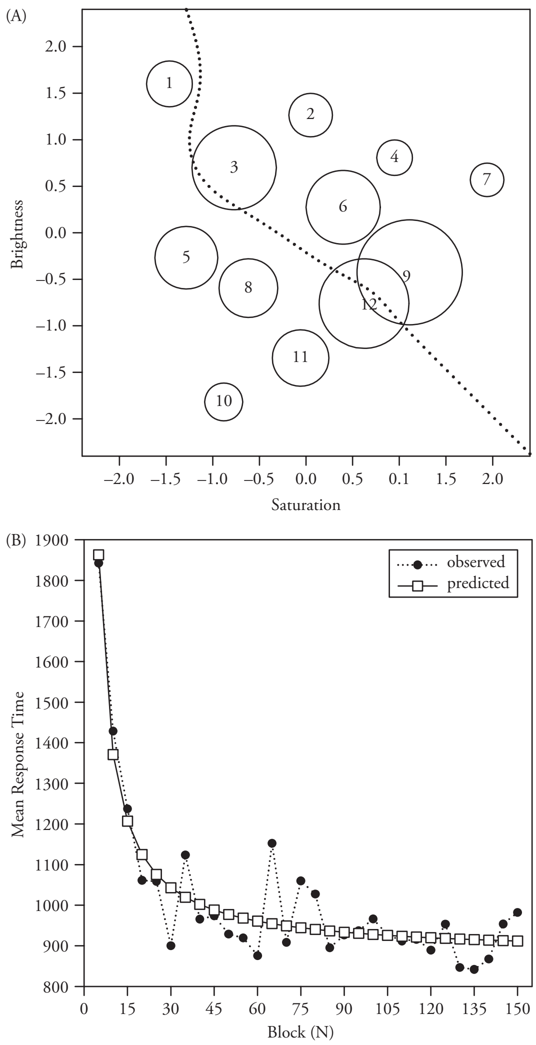

Fig. 7.3 Data from Experiment 1 of Nosofsky and Palmeri (1997a). (Top) Mean RTs for individual colors (RT proportional to the size of the circle). (Bottom) Mean RTs averaged across all colors as a function of grouped blocks of practice.

of Category A and Category B. Points falling to the upper right have greater summed similarity to Category A, and points falling to the lower left have greater summed similarity to Category B. As can be seen, the mean RTs tend to get faster as one moves farther away from the boundary of equal summed similarity. The bottom panel of Figure 7.3 provides another perspective on the data. This panel plots the overall mean RT for all 12 colors for each "grouped-block" of practice in the task, where a grouped-block corresponds to five consecutive individual blocks. It is evident that there is a speed-up with practice, with the lion's share of the speed-up occurring during the early blocks.

To apply the EBRW model to these data, we first derived a multidimensional scaling (MDS) solution for the colors by having the participant provide extensive similarity ratings between all pairs of the colors. A two-dimensional scaling solution yielded a good account of the similarity ratings. The solution is depicted in the top panel of Figure 7.3, where the center of each circle gives the MDS coordinates of the color. We then used the EBRW model in combination with the MDS solution to simultaneously fit the mean RTs associated with the individual colors (Figure 7.3, top panel) as well as the aggregated mean RTs observed as a function of grouped-blocks of practice (Figure 7.3, bottom panel). Specifically, the MDS solution provides the coordinate values x_{ik} that enter into the EBRW model's distance function (Equation 1). For simplicity, in fitting the model, we assumed that on each individual block of practice, an additional token of each individual exemplar was stored in memory with strength equal to one. (To reduce the number of free parameters, we set the background-element activation b equal to zero.) The first step in the analysis was to use the model to predict the mean RT for each individual color in each individual block. Then, for each individual color, we averaged the predictions across Blocks 31–150 to predict the overall individual-color mean RTs. Likewise, we averaged the predicted mean RTs over all colors in each grouped-block of practice to predict the speed-up curves. We fitted the model by searching for the free parameters that minimized the total sum of squared deviations between predicted and observed mean RTs across both data sets. These free parameters included the overall sensitivity parameter c; an attention-weight parameter w_1 (with $w_2 = 1 - w_1$); and a response-threshold parameter $+A$ (with $+A = |-B|$). In addition, we estimated a mean residual-time parameter μ_R; a scaling constant v for transforming the number of steps in the random walk into ms; and an individual step-time constant α (see Nosofsky & Palmeri, 1997a, pp. 268–270 for details).

The model-fitting results for the individual-color RTs are displayed in Figure 7.4, which plots observed against predicted mean RTs for each individual color. The model provides a reasonably good fit ($r = 0.89$), especially considering that it is constrained to simultaneously fit the speed-up curves. The model predicts these results because colors far from the exemplar-based boundary (e.g., 2, 4, 7, and 10—see top panel of Figure 7.3) tend to be similar only to exemplars from their own category.

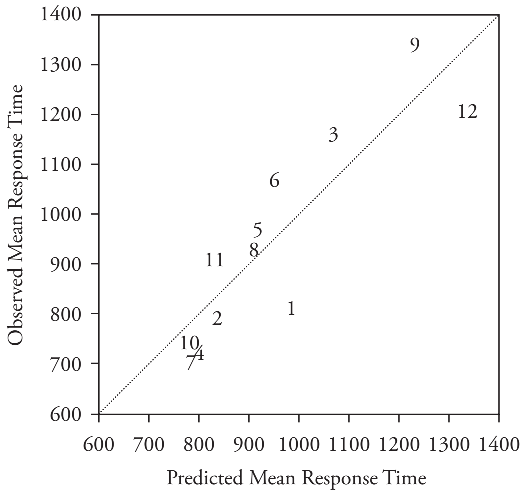

Fig. 7.4 Scatterplot of observed individual-color mean RTs against the predictions from the EBRW model.

Thus, only exemplars from a single category are retrieved, and the random walk marches in rapid fashion to the appropriate response threshold. By contrast, colors close to the boundary (e.g., 3, 9, and 12) are similar both to exemplars from their own category and to the contrast category. Thus, exemplars from both categories are retrieved and the random walk meanders back and forth, leading to slow mean RTs.

The fits to the speed-up curves are shown along with the observed data in the bottom panel of Figure 7.3. Again, the model captures these data reasonably well ($r = 0.931$). It predicts the speed-up for essentially the same reason as in Logan's (1988) model: As practice continues, an increased number of exemplars are stored in memory. The greater the number of exemplars that race to be retrieved, the faster are the winning retrieval times, so the individual steps in the random walk take place more quickly.[3]

Automaticity and Perceptual Expertise

The EBRW provides a general theory of categorization, automaticity, and the development of perceptual expertise (Nosofsky & Palmeri, 1997a; Palmeri, 1997; Palmeri, Wong, & Gauthier, 2004). In some real-world domains, novices learn to categorize objects by first learning to use a set of explicit verbal rules. This novice categorization can be a relatively slow, deliberate, attention-demanding process. With experience, as people develop perceptual expertise, categorization often becomes fast and automatic. Logan (1988) generally attributed the development of automaticity in cognitive skills to a shift from strategic algorithmic processes to exemplar-based memory retrieval. Palmeri (1997) specifically examined the development of automaticity in categorization as a shift from a rule-based process to an exemplar-based process assumed by EBRW (see also Johansen & Palmeri, 2002).

In Palmeri (1997, 1999), subjects were told to use an explicit rule in a dot-counting categorization task.They were asked to categorize random patterns containing between six and eleven dots according to numerosity and did so over multiple sessions. RTs observed in the first session increased linearly with the number of dots, as shown in Figure 7.5A. The dot patternswere repeated across multiple sessions, giving subjects an opportunity to develop automaticity in the task. Indeed, RTs observed in the 13th session were flat with numerosity, a signature of automaticity. In a subsequent transfer test, shown in Figure 7.5B, new unrelated test patterns had categorization RTs that were a linear function of numerosity, just like the first session, and old training patterns continued to be categorized with the same RTs irrespective of numerosity, just like the last session. Dot patterns of low or moderate similarity to the training patterns were categorized with RTs intermediate to the new unrelated and old training patterns.

Categorization RTs over sessions and numerosity were successfully modeled as a horse race between an explicit dot-counting process, whose stochastic finishing time increased linearly with the number of dots in the pattern; and an exemplar-based categorization process determined by the similarity of a dot pattern to stored exemplars of patterns previously categorized—an elaboration of EBRW.[4] When few patterns have been experienced, categorization is based entirely on the finishing times of the explicit counting process, predicting increased RTs with increased numerosity (Figure 7.5C). With experience, the memory strength m_j (Eq. 3) of exemplars in EBRW increases, causing a faster accumulation of perceptual evidence to a response threshold. Faster categorizations based on exemplar retrieval imply increased likelihood that those categorizations finish more quickly than explicit counting. With sufficient experience, EBRW finishes before counting for nearly all categorizations, predicting flat RTs with numerosity. The similarity-based retrieval in EBRW also predicts the transfer results, as shown in Figure 7.5D.

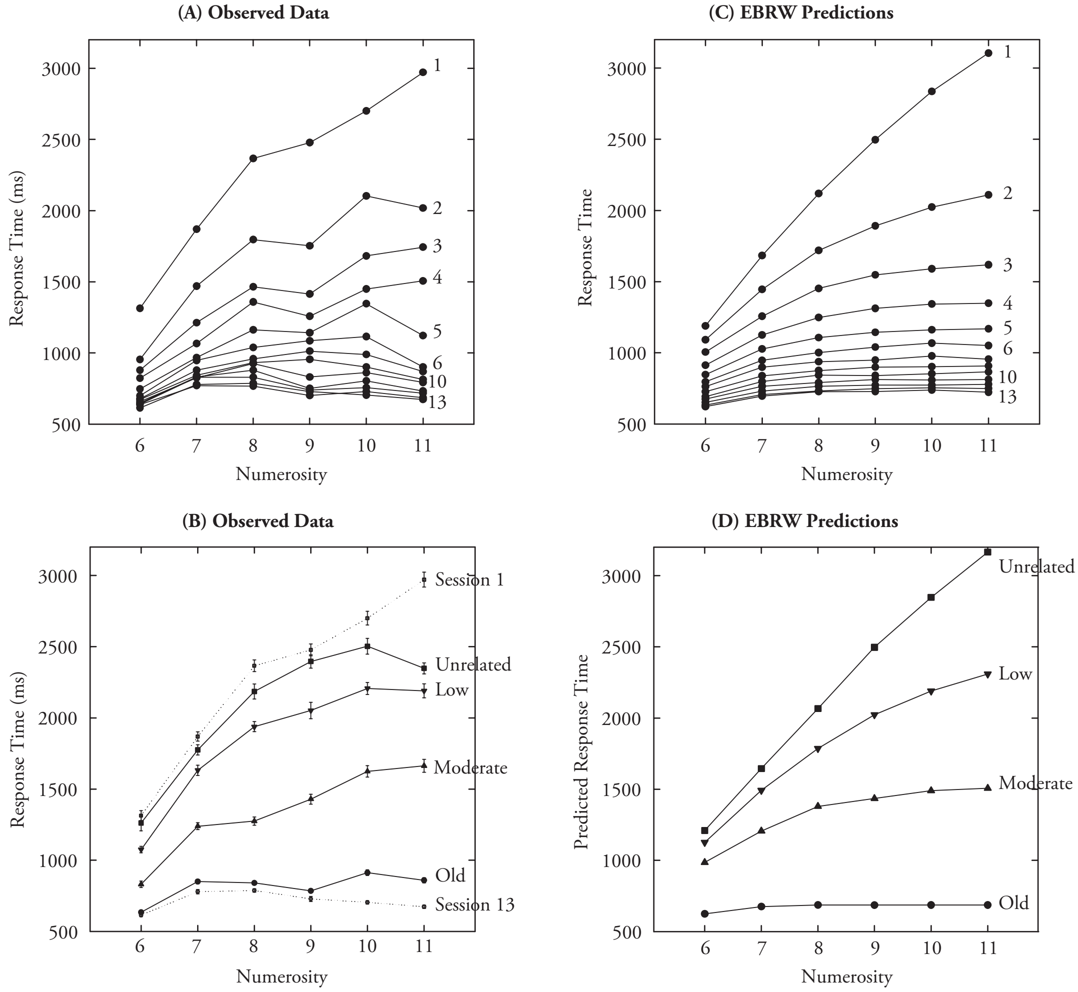

Fig. 7.5 Data from Experiment 1 of Palmeri (1997). Subjects categorized dot patterns varying in numerosity over thirteen training sessions and these were followed by a transfer test. (A) Response time (ms) as a function of numerosity in each training session (1-13). (B) Response time (ms) as a function of numerosity during transfer for old patterns, new unrelated patterns, and patterns of low and moderate similarity to old patterns. (C and D) EBRW predictions.

A shift from rules to exemplars is just one common characteristic of the development of perceptual expertise that EBRW can help explain (see Palmeri et al., 2004; Palmeri & Cottrell, 2009). As one other brief example, consider the well-known *basic-level* advantage. The seminal work of Rosch, Morris, Gray, Johnson, & Boyes-Braem (1976) showed that people are faster to verify the category of an object at an intermediate basic level of abstraction ("bird") than at more superordinate ("animal") or subordinate ("Northern Cardinal") levels. With expertise, subordinate-level categorizations are verified as quickly as basic-level categorizations (Tanaka & Taylor, 1991). One explanation for the basic-level advantage and related findings in novices is that basic-level categorizations reflect an initial stage of visual processing, perhaps as early as object segmentation (Grill-Spector & Kanwisher, 2005; see also Jolicoeur, Gluck, & Kosslyn, 1984). For novices, subordinate categorizations are slow because basic-level categorizations must be made first. For experts, the stage of basic-level categorization is somehow bypassed.

EBRW provides an alternative account (see Mack & Palmeri, 2011; Mack, Wong, Gauthier, Tanoka, & Palmeri, 2007). Faster or slower categorization at different levels of abstraction need not reflect stages of processing but may instead reflect differences in the speed of evidence accumulation in the random walk. Mack et al. (2007) simulated basic-level and subordinate-level categorization by novices and experts. For these simulations, subordinate categories were simply assumed to be individual identities of objects within clusters that were the basic-level categories. For moderate levels of the sensitivity parameter, *c*, which reflects the overall discriminability in the psychological space (Eq. 2), a basic-level advantage was predicted. But for high levels of sensitivity, reflecting the greater discriminability that may come with perceptual expertise, equivalent RTs for the subordinate and basic levels were predicted. Further empirical evidence has supported the hypothesisthat differences in RTs at different levels of abstraction reflect how quickly perceptual evidence accumulates rather than differences in discrete visual processing stages (e.g., Mack, Gauthier, Sadr, & Palmeri, 2008; ~~et al., 2009;~~ Mack & Palmeri, 2010).

Using Probabilistic Feedback Manipulations to Contrast the Predictions from the EBRW Model and Alternative Models

One the major alternative modeling approaches to predicting multidimensional classification RTs arises from a general framework known as "decision-boundary theory" (e.g., Ashby & Townsend, 1986). According to this framework, people construct decision boundaries for dividing a perceptual space into category response regions. Test items are assumed to give rise to noisy representations in the multidimensional perceptual space. If a perceived representation falls in Region A of the space, then the observer classifies the item into Category A.

Most past approaches to generating RT predictions from decision-boundary theory have involved applications of the *RT-distance hypothesis* (Ashby, Boynton, & Lee, 1994). According to this hypothesis, RT is a decreasing function of the distance of a stimulus from the decision boundary. More recent models that formalize the operation of decision boundaries posit perceptual-sampling mechanisms that drive random-walk or diffusion processes, similar to those of the EBRW model (e.g., Ashby, 2000; Fific, Little, & Nosofsky, 2010; Nosofsky & Stanton, 2005). For example, in Nosofsky and Stanton's (2005, p. 625) approach, on each step of a random walk, a percept is sampled randomly from the perceptual distribution associated with a presented stimulus. If the percept falls in Region A (as defined by the decision boundary), then the random walk steps toward threshold A, otherwise it steps toward threshold –B. The perceptual-sampling process continues until either threshold is reached. Such models provide process interpretations for *why* RT should get faster (and accuracy should increase) as distance of a stimulus from a decision boundary increases.

According to the EBRW model and decision-boundary models, the nature of the memory representations that are presumed to underlie categorization are dramatically different (i.e., stored exemplars versus boundary lines). Despite this dramatic difference, it is often difficult to distinguish between the predictions from the models. The reason is that distance-from-boundary and relative summed-similarity tend to be highly correlated in most experimental designs. For example, as we already described with respect to the Figure 7.3 (top panel) structure, items far from the boundary tend to be highly similar to exemplars from their own category and dissimilar to exemplars from the contrast category. Thus, both the distance-from-boundary model and the EBRW model tend to predict faster RTs and more accurate responding as distance from the boundary increases.

In one attempt to discriminate between the RT predictions of the models, Nosofsky and Stanton (2005) conducted a design that aimed to decouple distance-from-boundary and relative summed similarity (for a related approach, see Rouder & Ratcliff, 2004). The key idea in the design was to make use of probabilistic feedback manipulations associated with individual stimuli in the space. The design is illustrated in Figure 7.6. The stimuli were again a set of 12 Munsell colors of a constant hue, varying in their brightness and saturation. As illustrated in the figure, Colors 1–6 belonged to Category A, whereas Colors 7–12 belonged to Category B. To help motivate the predictions, we have drawn a diagonal linear decision boundary for separating the two categories of colors into response regions. Given reasonable assumptions (for details, see Nosofsky & Stanton, 2005), this boundary is the optimal (ideal-observer) boundary according to decision-boundary theory. That is, it is the boundary that would maximize an observer's percentage of correct categorization decisions. In generating predictions, decision-bound theorists often assume that observers will use a boundary with an optimal form. However, we will consider more general possibilities later in this section.

The key experimental manipulation was that, across conditions, either Stimulus Pair 4/8 or Stimulus Pair 5/9 received probabilistic feedback, whereas all other stimuli received deterministic feedback. Specifically, in Condition 4/8, Stimulus 4 received Category-A feedback with probability 0.75

but received Category- B feedback with probability 0.25; whereas Stimulus 8 received Category- B feedback with probability 0.75 and Category- A feedback with probability 0.25. Analogous probabilistic feedback assignments were given to Stimuli 5 and 9 in Condition 5/9. In each condition, we refer to the pair that received probabilistic feedback as the *probabilistic pair* and to the pair that received deterministic feedback as the *deterministic pair*.

The key conceptual point is that, because of the symmetric probabilistic assignment of stimuli to categories, the optimal boundary for partitioning the space into response regions is the same diagonal linear decision boundary that we have already illustrated in Figure 7.6. There is no way to adjust the boundary to achieve more accurate responding in the face of the probabilistic feedback assignments. Furthermore, because the probabilistic and deterministic pairs are an equal distance from the decision boundary, the most natural prediction from that modeling approach is that RTs for the probabilistic and deterministic pairs should be the same.

By contrast, the EBRW model predicts that the probabilistic pair should be classified more slowly (and with lower accuracy) than the deterministic pair. For example, in Condition 4/8, in cases in which stimulus 4 is presented and tokens of exemplar 4 are retrieved from memory, 0.75 of the steps in the random walk will move in the direction of threshold +A, but 0.25 of the steps will move in the direction of threshold –B. By contrast, presentations of the deterministic pair will result in more consistent steps of the random walk, leading to faster RTs and more accurate responding.

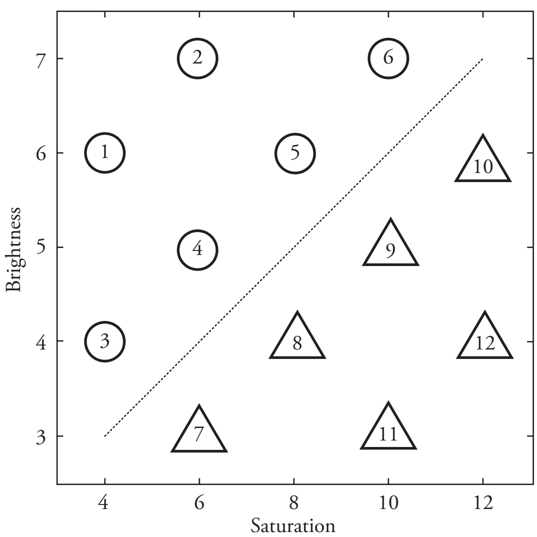

Fig. 7.6 Schematic illustration of the category structure used in Nosofsky and Stanton's (2005) probabilistic-feedback experiment.

Across two experiments (in which the instructions varied the relative emphasis on speed versus accuracy in responding), the qualitative pattern of results strongly favored the predictions from the EBRW model compared to the distance-from-boundary model. In particular, in both experiments, observers responded more slowly and with lower accuracy to the probabilistic pair than to the deterministic pair. These results were observed at both early and late stages of testing and were consistent across the majority of the individual participants. As a converging source of evidence, the EBRW model also provided better overall quantitative fits to the individual-subject choice-probability and RT data than did the distance-from-boundary model, including versions of the latter model in which the slope and y-intercept of the linear boundary were allowed to vary as free parameters.

As suggested by Nosofsky and Stanton (2005, p. 623), the results were particularly intriguing because they pointed toward a stubborn form of suboptimality in human performance: In the Figure 7.6 design, subjects would perform optimally by simply ignoring the probabilistic feedback and classifying objects based on the linear decision boundary. Nevertheless, despite being provided with monetary payoffs for correct responses, subjects' behavior departed from that optimal strategy in a manner that was well predicted by the EBRW model. Similar forms of evidence that favor the predictions from the EBRW model have been reported in related studies that manipulated the overall familiarity of individual study exemplars rather than probabilistic feedback assignments (e.g., Nosofsky & Palmeri, 1997a; Verguts, Storms, and Tuerlinckx, 2003).

Although the focus of Nosofsky and Stanton's (2005) study was to contrast predictions from the EBRW model and decision-bound models, the designs also yielded sharp contrasts between the EBRW model and prototype models.[5] Specifically, Nosofsky and Stanton (2005, p. 610) formulated a *prototype-based random-walk* (PBRW) model, analogous in all respects to the EBRW model, except that the category representation was presumed to correspond to the central tendency of each category distribution rather than to the individual exemplars. It turns out that for the stimulus spacings and probabilistic stimulus-category assignments used in the Figure 7.6 design, the central tendency of each category is equidistant to the probabilistic and deterministic stimulus pairs. Thus, the PBRW model predicted incorrectly that those pairs would show identical choice probabilities and RTs. Not surprisingly, therefore, the PBRW yielded far worse quantitative fits to the full sets of choice-probability and RT data than did the EBRW model.

Extending the EBRW Model to Predict Old-New Recognition RTs

Overview

As noted in the introduction, a central goal of exemplar models is to explain not only categorization, but other fundamental cognitive processes such as old-new recognition. The GCM has provided successful accounts of old-new recognition choice probabilities in wide varieties of experimental situations. When applied to recognition, the GCM assumes that each item from a study list is stored as a distinct exemplar in memory. The observer is presumed to sum the similarity of a test item to these stored exemplars. The greater the summed similarity, the more familiar is the test item, so the greater is the probability with which the observer responds "old." Indeed, the GCM can be considered a member of the general class of "global matching" models of old-new recognition (e.g., Clark & Gronlund, 1996; Gillund & Shiffrin, 1984; Hintzman, 1988; Murdock, 1982). Within this broad class, an important achievement of the model is that it predicts fine-grained differences in recognition probabilities for individual items based on their fine-grained similarities to individual exemplars in the study set (e.g., Nosofsky, 1988, 1991; Nosofsky & Zaki, 2003).

Just as has been the case for categorization, a major development in recent years has involved extensions of the GCM in terms of the EBRW model to allow it to account for recognition RTs (Donkin & Nosofsky, 2012a,b; Nosofsky, Little, Donkin, & Fific, 2011; Nosofsky & Stanton, 2006). In this section we describe these formal developments and illustrate applications to variants and extensions of the classic Sternberg (1966) *short-term probe-recognition paradigm*. In this paradigm, subjects are presented on each trial with a short list of items (the memory set) and are then presented with a test probe. The subjects judge, as rapidly as possible, while minimizing errors, whether the probe occurred in the memory set. Fundamental variables that are manipulated in the paradigm include the size of the memory set; whether the test probe is old (a "positive" probe) or new (a "negative" probe); and, if old, the serial position with which the positive probe occurred in the memory set.

Whereas in the standard version of the Sternberg paradigm the to-be-recognized items are generally highly distinct entities, such as alphanumeric characters, recent extensions have examined recognition performance in cases involving confusable stimuli embedded in a continuous-dimension similarity space (e.g., Kahana & Sekuler, 2002). We focus on this type of extended similarity-based paradigm in the initial part of this section; however, we will illustrate applications of the EBRW model to the more standard paradigm as well.

The Formal Model

The EBRW-recognition model makes the same representational assumptions as does the categorization model: (a) Exemplars are conceptualized as occupying points in a multidimensional similarity space; (b) similarity is a decreasing function of distance in the space (Eqs. 1 and 2); (c) activation of the exemplars is a joint function of their memory strength and their similarity to the test probe (Equation 3); and (d) the stored exemplars race to be retrieved with rates proportional to their activations (Eq. 4).

In our previous applications to categorization, a single global level of sensitivity (c in Eq. 2) was assumed that applied to all exemplar traces stored in long-term memory. In application to short-term recognition paradigms involving high-similarity lures, however, allowance is made for a form of exemplar-specific sensitivity. In particular, an observer's ability to discriminate between test item i and a nonmatching exemplar-trace j will almost certainly depend on the recency with which exemplar j was presented: Discrimination is presumably much easier if an exemplar was just presented than if it was presented in the distant past. In the version of the model applied by Nosofsky et al. (2011), a separate sensitivity parameter c_j was estimated for each individual *lag* j on the study list, where *lag* is counted backward from the presentation of the test probe to the memory-set exemplar. For example, for the case in which memory set-size is 4, the exemplar in the fourth serial position has Lag 1, the exemplar in the third serial position has Lag 2, and so forth.

Likewise, the memory strengths of the individual exemplars (m_j) were also assumed to depend on lag j: Presumably, the more recently an exemplar was presented on the study list, the greater its memory strength (e.g., McElree & Dosher, 1989; Monsell, 1978). (Although the effects are smaller, in modeling short-term recognition with the EBRW, allowance is also typically made for a primacy effect on memory strength. The memory strength of the item in the first serial position of the memory set is given by $P_M \cdot m_j$, where m_j is the memory strength for an item with lag j, and P_M is a primacy-multiplier parameter.)

To adapt the EBRW model to the domain of old-new recognition, it is assumed that the observer establishes what are termed "*criterion elements*" in the memory system. These elements are similar to the "background elements" used for modeling the early stages of category learning. Just as is the case for the stored exemplars, upon presentation of a test probe, the criterion elements race to be retrieved. However, whereas the retrieval rates of the stored exemplars vary with their lag-dependent memory strengths and their similarity to the test probe, the retrieval rates of the criterion elements are independent of these factors. Instead, the criterion elements race with some fixed rate β, independent of the test probe that is presented. The setting of β is presumed to be, at least in part, under the control of the observer.

Finally, the retrieved exemplars and criterion elements drive a random-walk process that governs old-new recognition decisions. The observer sets response thresholds $+OLD$ and $-NEW$ that establish the amount of evidence needed for making an "old" or a "new" response. On each step of the random walk, if an old exemplar wins the retrieval race, then the random-walk counter takes a step in the direction of the $+OLD$ response threshold; whereas if a criterion element wins the race, then the counter takes a step in the direction of the $-NEW$ threshold. The retrieval process continues until one of the thresholds is reached.

Given the processing assumptions outlined earlier, then on each step of the random walk, the probability that the counter steps in the direction of the $+OLD$ threshold is given by

$$p_i = F_i/(F_i + \beta), \tag{7}$$

where F_i gives the summed activation ("familiarity") of the test probe to all old exemplars on the study list (and β is the fixed setting of criterion-element activation). Note that test probes that match recently presented exemplars (with high memory strengths) will cause high summed familiarity (F_i), leading the random walk to march quickly to the $+OLD$ threshold and resulting in fast *old* RTs. By contrast, test probes that are highly dissimilar to the memory-set items will not activate the stored exemplars, so only criterion elements will be retrieved. In this case, the random walk will march quickly to the $-NEW$ threshold, resulting in fast *new* RTs. Through experience in the task, the observer is presumed to learn an appropriate setting of the criterion-element activation β, such that summed activation (F_i) tends to exceed β when the test probe is old, but tends to be less than β when the test probe is new. In this way, the random walk will tend to step toward the appropriate response threshold on trials in which *old* versus *new* probes are presented.

Experimental Tests

In Nosofsky et al.'s (2011) initial experiment for testing the model, the stimuli were a set of 27 Munsell colors that varied along the dimensions of hue, brightness, and saturation. Similarity-scaling procedures were used to derive a precise MDS solution for the colors.

The design of the probe-recognition experiment involved a broad sampling of different list structures to provide a comprehensive test of the model. There were 360 lists in total. The size of the memory set on each trial was either 1, 2, 3 or 4 items, with an equal number of lists at each set size. For each set size, half the test probes were old and half were new. In the case of old probes, the matching item from the memory set occupied each serial position equally often. To create the lists, items were randomly sampled from the full set of stimuli, subject to the constraints described earlier. Thus, a highly diverse set of lists was constructed, varying not only in set size, old/new status of the probe, and serial position of old probes, but also in the similarity structure of the lists.

Because the goal was to predict performance at the individual-subject level, three subjects were each tested for approximately 20 one-hour sessions, with each of the 360 lists presented once per session. As it turned out, each subject showed extremely similar patterns of performance, and the fits of the EBRW model yielded similar parameter estimates for the three subjects. Therefore, for simplicity, and to reduce noise in the data, we report the results from the analysis of the averaged subject data.

In the top panels of Figure 7.7 we report summary results from the experiment. The top-right panel reports the observed mean RTs plotted as a function of: (a) set size, (b) whether the probe

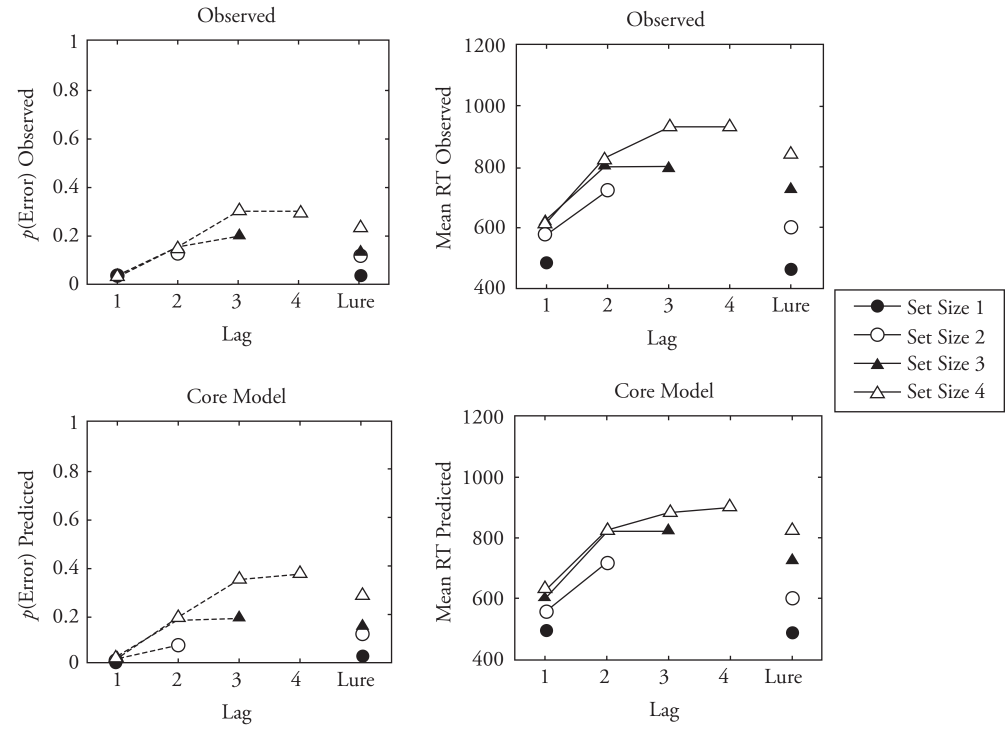

Fig. 7.7 Summary data from the short-term memory experiment of Nosofsky, Little, Donkin, and Fific (2011). (Top) Observed error rates and mean RTs. (Bottom) Predictions from the EBRW model.

was old or new (i.e., a lure), and (c) the lag with which old probes appeared in the memory set. For old probes, there was a big effect of lag: In general, the more recently a probe appeared on the study list, the shorter was the mean RT. Indeed, once one takes lag into account, there is little remaining effect of set size on the RTs for the old probes. That is, as can be seen, the different set size functions are nearly overlapping. The main exception is a persistent primacy effect, in which the mean RT for the item at the longest lag for each set size is "pulled down." (The item at the longest lag occupies the first serial position of the list.) By contrast, for the lures, there is a big effect of set size, with longer mean RTs as set size increases. The mean proportions of errors for the different types of lists, shown in the top-left panel of Figure 7.7, mirror the mean RT data just described.

The goal of the EBRW modeling, however, was not simply to account for these summary trends. Instead, the goal was to predict the choice probabilities and mean RTs observed for each of the individual lists. Because there were 360 unique lists in the experiment, this goal entailed simultaneously predicting 360 choice probabilities and 360 mean RTs. The results of that model-fitting goal are shown in the top and bottom panels of Figure 7.8. The top panel plots, for each individual list, the observed probability that the subjects judged the probe to be "old" against the predicted probability from the model. The bottom panel does the same for the mean RTs. Although there are a few outliers in the plots, overall the model achieves a good fit to both data sets, accounting for 96.5% of the variance in the choice probabilities and for 83.4% of the variance in the mean RTs.

The summary-trend predictions that result from these global fits are shown in the bottom panels of Figure 7.7. It is evident from inspection that the EBRW does a good job of capturing these summary results. For the old probes, it predicts the big effect of lag on the mean RTs, the nearly overlapping set-size functions, and the facilitation in RT with primacy. Likewise, it predicts with good quantitative accuracy the big effect of set size on the lure RTs. The error-proportion data (left panels of Figure 7.7) are also well predicted, with the main exception that a primacy effect was predicted but not observed for the size-2 lists.

The explanation of these results in terms of the EBRW model is straightforward. According to the best-fitting parameters from the model (see Nosofsky et al., 2011, Table 2), more recently

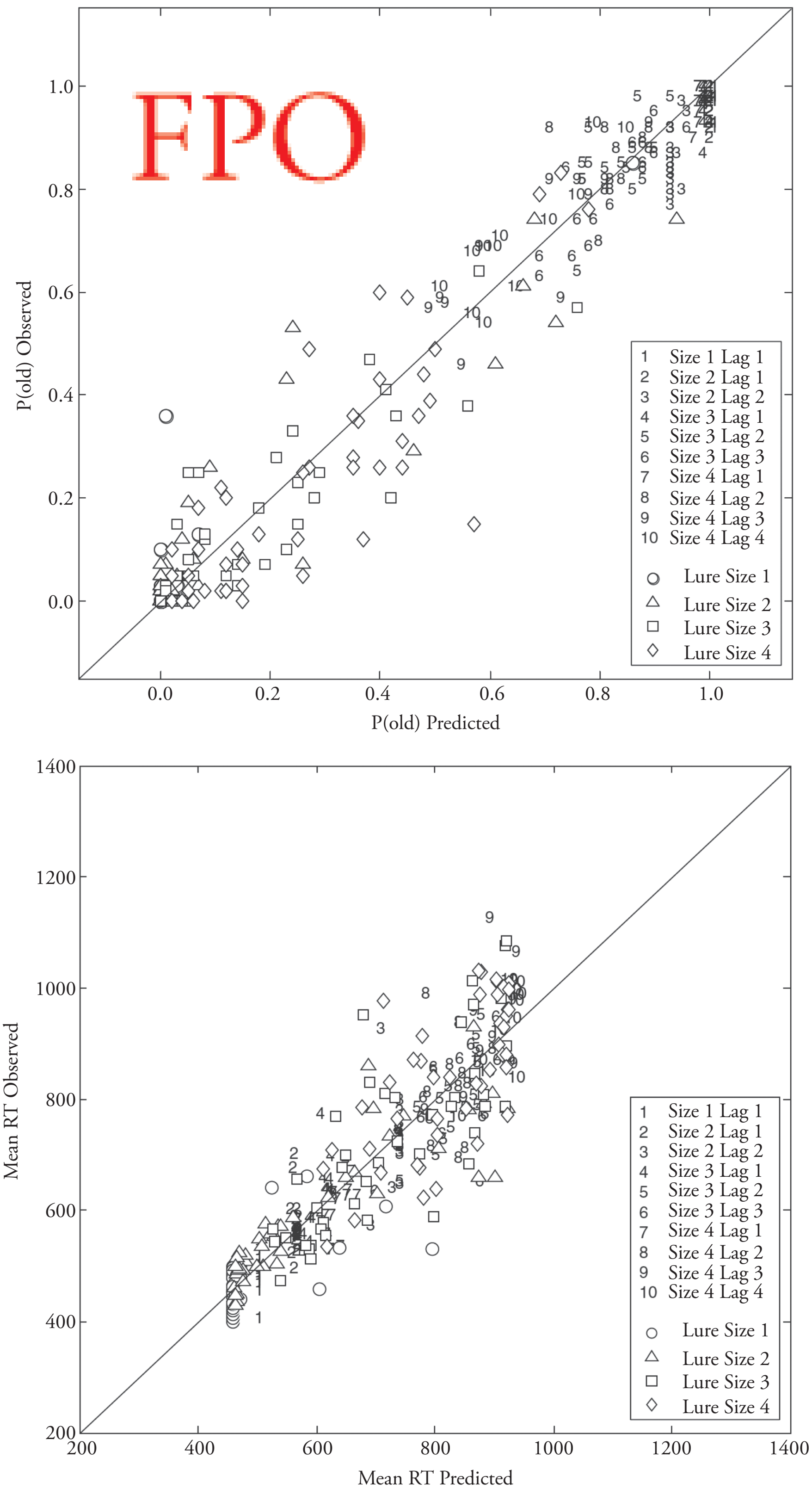

Fig. 7.8 Scatterplots of observed and EBRW-predicted ~~error rates~~ and mean RTs associated with individual lists from the short-term memory experiment of Nosofsky, Little, Donkin, and Fific (2011).

presented exemplars had greater memory strengths and sensitivities than did less recently presented exemplars. From a psychological perspective, this pattern seems highly plausible. For example, presumably, the more recently an exemplar was presented, the greater should be its strength in memory. Thus, if an old test probe matches the recently presented exemplar, it will give rise to greater overall activation, leading to faster mean old RTs. In the case of a lure, as set size increases, the overall summed activation yielded by the lure will also tend to increase. This pattern arises both because a greater number of exemplars will contribute to the sum, and because the greater the set size, the higher is the probability that a least one exemplar from the memory set will be highly similar to the lure. As summed activation yielded by the lures increases, the probability that the random walk takes correct steps toward the –*NEW* threshold *decreases*, so mean RTs for the lures get longer.

Beyond accounting well for these summary trends, inspection of the detailed scatterplots in Figure 7.8 reveals that the model accounts for fine-grained changes in choice probabilities and mean RTs depending on the fine-grained similarity structure of the lists. For example, consider the choice-probability plot (Figure 7.8, top panel) and the Lure-Size-4 items (open diamonds). Whereas performance for those items is summarized by a single point on the summary-trend figure (Figure 7.7), the full scatterplot reveals extreme variability in results across different tokens of the Lure-Size-4 lists. In some cases the false-alarm rates associated with these lists are very low, in other cases moderate, and in still other cases the false-alarm rates exceed the hit rates associated with *old* lists. The EBRW captures well this variability in false-alarm rates. In some cases, the lure might not be similar to any of the memory-set items, resulting in a low false-alarm rate; whereas in other cases the lure might be highly similar to some of the memory-set items, resulting in a high false-alarm rate.

The application reviewed earlier involved a version of a short-term probe-recognition paradigm that used confusable stimuli embedded in a continuous-dimension space. However, the EBRW model has also been applied successfully to more standard versions of such paradigms that involve easy-to-discriminate stimuli such as alphanumeric characters. In those applications, instead of adopting MDS approaches, a highly simplified model of similarity is used: The similarity of a probe to itself is set equal to one, whereas the similarity between a probe and any nonmatching item is set equal to a free parameter s. Not only has this simple version of the EBRW model captured many of the classic patterns of results involving mean RTs in such paradigms, it also accounts successfully for the detailed RT-distribution data that have been observed (Nosofsky et al., 2011; Donkin & Nosofsky, 2012a,b).

Furthermore, applications of the EBRW model to both continuous-similarity and discrete versions of the probe-recognition paradigm have led to the discovery of an interesting regularity involving memory strength. As noted earlier, in the initial tests of the model, separate-memory strength parameters were estimated corresponding to each individual lag on the study list. It turns out, however, that the estimated memory strengths follow almost a perfect power function of this lag. For example, in an experiment reported by Donkin and Nosofsky (2012a), participants studied 12-item lists consisting of either letters or words, followed by a test probe. Separate RT-distribution data for hits and misses for positive probes were collected at each study lag. (RT-distribution data for false alarms and correct rejections for negative probes were collected as well.) The EBRW model provided an excellent quantitative account of this complete set of detailed RT-distribution and choice-probability data.[6] The discovery that resulted from the application of the model is illustrated graphically in Figure 7.9. The figure plots, for each of four individual participants who were tested, the estimated memory-strength parameters against lag. As shown in the figure, the magnitudes of the memory strengths are extremely well captured by a simple power function. Interestingly, other researchers have previously reported that a variety of *empirical* forgetting curves are well described as power functions (e.g., Anderson & Schooler, 1991; Wickelgren, 1974; Wixted & Ebbesen, 1991). For example, Wixted and Ebbesen (1991) reported that diverse measures of forgetting, including proportion correct of free recall of word lists, recognition judgments of faces, and savings in relearning lists of nonsense syllables, were well described as power functions of the retention interval. Wixted (2004) considered a variety of possible reasons for the emergence of these empirical power-function relations and concluded that the best explanation was that the strength of the memory traces themselves may exhibit power-function decay. The model-based results from Donkin and Nosofsky (2012a,b) lend support to Wixted's

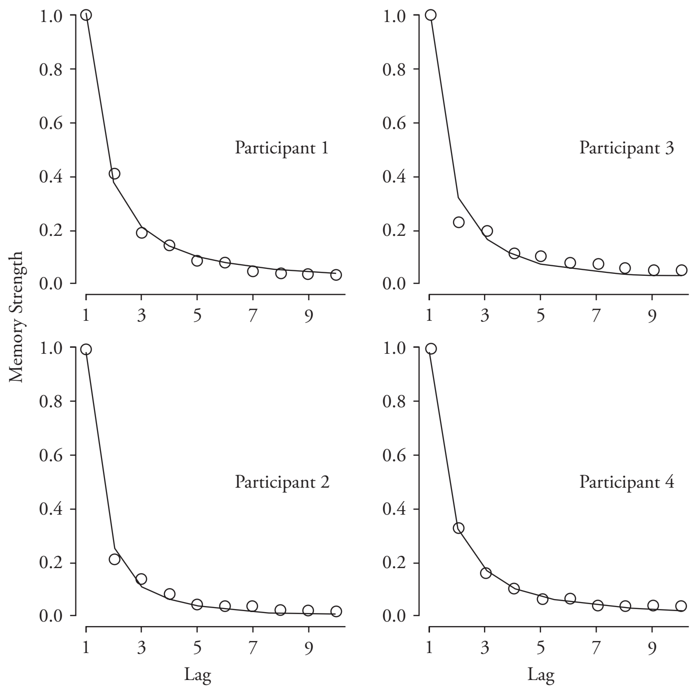

Fig. 7.9 Model-based results from the probe-recognition experiment of Donkin and Nosofsky (2012a). Estimated memory strengths (open circles) are plotted as a function of lag, along with the best-fitting power functions.
Donkin, C., & Nosofsky, R.M. (2012). A power-law model of psychological memory strength in short-term and long-term recognition. *Psychological Science*, 23, 625–634. Adapted with permission of SAGE Publications.

suggestion and now motivate the new research goal of unpacking the detailed psychological and neurological mechanisms that give rise to this discovered power law of memory strength.

Conclusions and New Research Goals

In sum, the EBRW is an important candidate model for explaining how the process of categorization unfolds over time. The model combines assumptions involving exemplar-based category representation and processes of evidence accumulation within a unified framework to account for categorization and recognition choice-probability and RT data. As reviewed in this chapter, it accounts successfully for a wide variety of fundamental effects in these domains including effects of similarity, distance-from-boundary, familiarity, probabilistic feedback, practice, expertise, set size, and lag. Although we were able to sample only a limited number of example applications in this single chapter, we should clarify that the exemplar model has been applied in a wide variety of stimulus domains and to varied category and study-list structures. The stimulus domains include colors, dot patterns, multidimensional cartoon drawings, geometric forms, schematic faces, photographs of faces, alphanumeric characters, words, and pictures of real-world objects. The category structures include small collections of continuous-dimension stimuli separated by boundaries of varying degrees of complexity, normally distributed category structures, high-dimensional stimuli composed of discrete dimensions, and categories generated from statistical distortions of prototypes (e.g., see Richler & Palmeri, 2014). Furthermore, growing neural evidence ranging from single-unit records to functional brain imaging supports a number of the processing assumptions embodied in models like EBRW (see Box 1). Finally, as illustrated in our chapter, a key theme of the theoretical approach is that, despite the dramatically different task goals, the processes of categorization and old-new recognition may be closely related (but see Box 2 for discussion of a major theoretical debate regarding this issue in the cognitive-neuroscience literature). A likely reason for the model's success is that it builds on the strengths of classic previous approaches for understanding processes of choice and similarity, the development of automaticity, and evidence accumulation in decision-making and memory.

Beyond accounting for categorization and recognition, we believe that the EBRW model can serve

Box 1 Neural Evidence Supports Mechanistic Assumptions in EBRW.

EBRW proposes many mechanistic assumptions, such as exemplar representations, attention weights along relevant dimensions, and accumulation of perceptual evidence. In most modeling, these are supported by evaluating predictions of behavioral data like accuracy and RTs. But we can now evaluate particular mechanistic assumptions using relevant neural data from brain regions hypothesized to instantiate those mechanisms.

For example, Mack, Preston and Love, (2013; see also Palmeri, 2014) turned to patterns of brain activity measured with fMRI to evaluate whether the brain represents categories using exemplars or prototypes. Subjects learned to classify objects into one of two categories and then in the scanner were tested on training and transfer objects without feedback (Medin & Schaffer, 1978). Typical fMRI analyses would correlate brain activity with stimuli or responses, for example highlighting regions that modulate with categorization difficulty. Instead, Mack and colleagues first fitted exemplar and prototype models to individual subjects' categorization responses; despite the fact that these models made fairly similar behavioral predictions, they differed in patterns of summed similarity to their respective exemplar or prototype representations. Mack and colleagues showed that patterns of individual subjects' brain activity were more consistent with patterns of summed similarity predicted by an exemplar model than those predicted by a prototype model. According to exemplar models, learning to categorize objects can cause selective attention to relevant psychological dimensions, stretching psychological space to better allow subjects to discriminate between members of contrasting categories. Neurophysiology and fMRI have suggested that category-relevant dimensions can be emphasized in visual cortex. After monkeys learned to categorize multidimensional objects, neurons in inferotemporal cortex were more sensitive to variations along a relevant dimension than an irrelevant dimension (De Baene, Ons, Wagemans, & Vogels, 2008; Sigala and Logothetis 2002; see also Gauthier & Palmeri, 2002). Similarly, after people learned object categories, psychological stretching of relevant dimensions was accompanied by neural stretching of relevant dimensions measured by fMRI (Folstein, Palmeri, & Gauthier, 2013; see also Folstein, Gauthier, & Palmeri, 2012).

Finally, mathematical psychology and systems neuroscience have converged on accumulation of perceptual evidence as a general theoretical framework to explain the time course of decision making (see Palmeri, Schall, & Logan, this volume). Some neurons show dynamics predicted by accumulator models, other neurons show activity consistent with encoded perceptual evidence to be accumulated over time, and an ensemble of neurons predicts the time course of decisions made by awake behaving monkeys (e.g., Purcell, Schall, Logan, & Palmeri, 2012; Zandbelt, Purcell, Palmeri, Logan, Schall, 2014).

Box 2 The Exemplar Model Accounts for Dissociations Between Categorization and Recognition Demonstrated in the Cognitive-Neuroscience Literature.

Interestingly, in contrast to the theme emphasized in this chapter, the prevailing view in the cognitive neuroscience literature is that separate cognitive/neural systems mediate categorization and recognition (Smith, 2008). The main source of evidence involves the demonstration of intriguing dissociations between categorization and recognition. For example, studies have demonstrated that amnesics with poor recognition memory perform at normal levels in categorization tasks involving the same types of stimuli (e.g., Knowlton & Squire, 1993). Nevertheless, formal modeling analyses have indicated that even these dissociations are consistent with the predictions from the exemplar model (e.g., Nosofsky, Denton, Zaki, Murphy-Knudson, & Unverzagt 2012; Nosofsky & Zaki, 1998; Palmeri & Flanery, 1999, 2002; Zaki, Nosofsky, Jessup, & Unverzagt, 2003). The general approach in the modeling was to assume that amnesics have reduced ability to discriminate among distinct exemplars in memory. This reduced discriminability is particularly detrimental to old-new recognition, which may

Box 2 Continued

require the observer to make fine-grained distinctions between old versus new items. However, the reduced discriminability is not very detrimental in typical tasks of categorization, which may require only gross-level assessments of similarity to be made. A more direct challenge to the exemplar-model hypothesis comes from brain-imaging studies that show that distinct brain regions are activated when observers engage in recognition vs. categorization tasks (Reber et al., 1998a,b). Exemplar theorists have responded, however, by providing evidence that these brain-imaging dissociations may not reflect the operation of separate neural systems devoted to categorization versus recognition per se. Instead, the brain-imaging dissociations may reflect changes in stimulus-encoding strategies across task situations (Gureckis, James, & Nosofsky, 2011), differences in the precise stimuli that are tested (Nosofsky, Little, & James, 2012; Reber et al., 2003), as well as adaptive changes in parameter settings that allow observers to meet the competing task goals of categorization versus recognition (Nosofsky et al., 2012).

as a useful analytic device for assessing human performance. For example, note that Ratcliff's (1978) diffusion model has been applied to analyze choice behavior in various special populations, including elderly adults, sleep-deprived subjects, and so forth (see Chapter 2 of this volume). The model-based analyses provide a deeper understanding of the locus of the cognitive/perceptual deficits in such populations by tracing them to changes in diffusion-model drift rates, response-threshold settings, or residual times. The EBRW model has potential to reveal even more fine-grained information along these lines. For example, in that model, the random-walk step probabilities (i.e., drift rates) emerge from cognitive/perceptual factors such as overall sensitivity, attention-weight settings, and memory strengths of stored exemplars, each of which can be measured by fitting the model to data obtained in suitable categorization and recognition paradigms. Although exemplar models have been applied to help interpret the behavior of amnesic subjects and patients with mild memory-based cognitive impairment (see Box 2), we have only scratched the surface of such potential applications to many more clinical groups.

Finally, an important theme in the categorization literature is that there may be multiple systems of categorization (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Erickson & Kruschke, 1998; Johansen & Palmeri, 2002; Nosofsky, Palmeri, & McKinley, 1994). A classic idea, for example, is that many categories may be represented and processed by forming and evaluating logical rules. In some modern work that pursues this avenue, researchers have considered the RT predictions from logical-rule models of classification. Furthermore, such approaches have been used to develop sharp contrasts between the predictions of the ERBW model and rule-based forms of category representation and processing (Fific, Little, & Nosofsky, 2010; Lafond, Lacouture, & Cohen, 2009; Little, Nosofsky, & Denton, 2011). In domains involving highly separable-dimension stimuli in which the category structures can be described in terms of exceedingly simple logical rules, evidence has been mounting that the logical-rule models provide better accounts of the detailed patterns of RT data than does the EBRW model. An important target for future research is to develop a deeper understanding of these multiple forms of categorization, to learn about the experimental factors that promote the use of each strategy, and to explain the manner in which exemplar-based and rule-based systems may interact.

Acknowledgments

This work was partially supported by NSF grants SMA-1041755 and SBE-1257098 and by AFOSR grant FA9550-14-1-0307.

Notes

1. In the context of the GCM, the parameter γ is referred to as a *response-scaling parameter*. When $\gamma=1$, observers respond by "probability-matching" to the relative summed similarities. As γ grows greater than 1, observers respond more deterministically with the category that yields the greater summed similarity (Ashby & Maddox, 1993; McKinley & Nosofsky, 1995).

2. The version of the EBRW model described in this chapter is applicable to "integral-dimension" stimuli, which are encoded and perceived holistically. A common example of such integral-dimension stimuli are colors varying in hue, brightness, and saturation. Because there has been extensive previous scaling work indicating that similarity relations among these stimuli are extremely well described in terms of these dimensions, we often use these stimuli in our tests of the EBRW model. An extended version of the EBRW model has also been developed that is applicable to separable-dimension stimuli (Cohen & Nosofsky, 2003). In this version, rather than encoding stimuli in holistic fashion, the encoding of individual stimulus dimensions is a stochastic process, and similarity

relations between a test item and the stored exemplars change dynamically during the time course of processing (see also Lamberts, 2000).

3. Because accuracies were near ceiling in the present experiment, we focused our analysis primarily on the patterns of RT data. However, in a follow-up study, Nosofsky and Alfonso-Reese (1999) tested conditions that allowed examination of how both speed *and* accuracy changed during the early stages of learning. By including the background-activation parameter b in its arsenal, the EBRW model provided good quantitative fits to not only the speed-up in mean RTs, but to the improvements in choice-accuracy data as well. (As noted by Nosofsky and Palmeri, 1997a, p. 291, with b=0, the EBRW model does not predict changes in response accuracy.)

4. The original EBRW model (Nosofsky & Palmeri, 1997a) applies to two-alternative forced- choice responses. There is a single accumulator whose value increases or decreases as evidence accumulates in the random walk until an upper or lower response threshold is reached. Numerosity judgments in Palmeri (1997) permitted six possible responses, so EBRW was extended to allow multiple alternatives. Each response alternative was associated with its own counter, so with six numerosity responses there were six counters. Whenever an exemplar was retrieved with the label associated with a particular counter, the value of that counter was incremented. A response was made whenever the value of one of the counters exceeded all the rest by some relative amount. With only two alternatives, this multiple counter model with a relative threshold response rule generally mimics a standard random-walk model with one accumulator with a positive and negative threshold.

5. There is a long history of debate, too extensive to be reviewed in this chapter, between the proponents of exemplar and prototype models. For examples of recent research that has argued in favor of the prototype view, see Minda and Smith (2001), Smith and Minda (1998, 2000) and Smith (2002). For examples from the exemplar perspective, see Nosofsky (2000), Nosofsky and Zaki (2002), Palmeri and Flanery (2002), and Zaki and Nosofsky (2007).

6. The version of the exemplar-recognition model reported by Donkin and Nosofsky (2012a) assumed a linear-ballistic accumulation process (Brown & Heathcote, 2008) instead of a random-walk accumulation process. However, the same evidence for a power-law relation between memory strength and lag was obtained regardless of the specific accumulation process that was assumed. We should note as well that, in fitting complete RT distributions for correct and error responses, such as occurred in the Donkin and Nosofsky (2012a) experiment, the exemplar model makes provision for drift-rate variability and response-threshold variability across trials (e.g, Donkin & Nosofsky, 2012a; Nosofsky & Stanton, 2006), in a manner analogous to the approach used in Ratcliff's diffusion model (e.g, Ratcliff, Van Zandt, & McKoon, 1999).

Glossary

Attention-weight parameters: a set of parameters in the *GCM* and *EBRW* models that describe the extent to which each dimension is weighted when computing distances among objects.

Automaticity: ability to perform some task at a satisfactory level without requiring conscious attention or effort and without limits in capacity.

Background element: a hypothetical construct in the *GCM* and *EBRW* models that describes initial background noise in people's memories for members of alternative categories.

Basic-level of categorization: an intermediate level of a category hierarchy that is hypothesized to lead to privileged forms of cognitive processing.

Categorization: process in which observers classify distinct objects into groups.

Criterion element: a hypothetical entity in the *EBRW* recognition model. Retrieval of criterion elements leads the *random walk* to step in the direction of the NEW *response threshold*. Biases in the random-walk step probabilities are determined by the strength of the criterion elements.

Decision-boundary model: model of *categorization* that assumes that people form boundaries to divide a stimulus space into category-response regions.

Exemplar model: model of *categorization* that assumes that observers store individual examples of categories in memory.

Exemplar-based random-walk (EBRW) model: an extension of the *generalized context model* that explains how the processes of *categorization* and *recognition* unfold over time. Exemplars stored in memory "race" to be retrieved, and the retrieved exemplars drive a *random-walk* decision-making process.

Exponential distribution: a probability distribution that describes the time between events, in which the events occur continuously and independently at a constant average rate.

Generalized context model (GCM): a member of the class of *exemplar models*. In the GCM, exemplars are represented as points in a multidimensional psychological space, and similarity is a decreasing function of distance in the space.

Integral-dimension stimuli: stimuli composed of individual dimensions that combine into unitary, integral wholes.

Memory strength parameters: parameters in the *GCM* and *EBRW* models that describe the strength with which the exemplars are stored in memory.

Minkowski power model: a model for computing distances between points in a space.

Multidimensional scaling: a modeling technique for representing similarity relations among objects. The objects are represented as points in a multidimensional psychological space and similarity is a decreasing function of the distance between the points in the space.

Prototype model: model of categorization that assumes that observers represent categories by forming a summary representation, usually assumed to be the central-tendency of the category distribution.

Glossary

Prototype-based random-walk (PBRW) model: a model that is analogous in all respects to the *EBRW* model, except that the category representation corresponds to the central tendency of each category distribution rather than to the individual exemplars.

Random walk: a mathematical model that describes a path of outcomes consisting of a sequence of random steps.

Response thresholds: parameters in evidence-accumulation models that determine how much evidence is required before a response is initiated.

Recognition memory: process in which observers decide whether objects are "old" (previously experienced) or "new."

Response-scaling parameter: a parameter in the *GCM* that describes the extent to which observers respond using probabilistic versus deterministic response rules.

RT-distance hypothesis: hypothesis that categorization RT is a decreasing function of the distance of a stimulus from the decision boundary.

Rule-plus-exception model: model of *categorization* that assumes that people classify objects by forming simple logical rules and remembering occasional exceptions to those rules.

Sensitivity parameter: a parameter in the *GCM* and *EBRW* models that describes overall discriminability among distinct items in the multidimensional psychological space.

Short-term probe-recognition task: task in which observers are presented with a short list of to-be-remembered items followed by a test probe. The observers judge as rapidly as possible, while trying to minimize errors, whether the probe is old or new.

References

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2,* 396–408.

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

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review,* 105, 442–481.

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

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology, 37,* 372–400.

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

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice reaction time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153–178.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97,* 523–547.

Busemeyer, J. R. (1982). Choice behavior in a sequential decision-making task. *Organizational Behavior and Human Performance, 29,* 175–207.

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.

Carroll, J. D., & Wish, M. (1974). Models and methods for three-way multidimensional scaling. In D. H. Krantz, R. C. Atkinson, R. D. Luce, and P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2). San Francisco: W. H. Freeman.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3,* 37–60.

Cohen, A. L., & Nosofsky, R. M. (2003). An extension of the exemplar-based random-walk model to separable-dimension stimuli. *Journal of Mathematical Psychology, 47,* 150–165.

Davis, T., Love, B. C., & Preston, A. R. (2012). Learning the exception to the rule: Model-based fMRI reveals specialized representations for surprising category members. *Cerebral Cortex, 22,* 260–273.

De Baene, W., Ons, B., Wagemans, J., & Vogels, R. (2008). Effects of category learning on the stimulus selectivity of macaque inferior temporal neurons. *Learning & Memory, 15*(9), 717–727.

Donkin, C., & Nosofsky, R. M. (2012a). A power-law model of psychological memory strength in short-term and long-term recognition. *Psychological Science, 23,* 625–634.

Donkin, C., & Nosofsky, R. M. (2012b). The structure of short-term memory scanning: An investigation using response-time distribution models. *Psychonomic Bulletin & Review, 19,* 363–394.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127,* 107–140.

Estes, W. K. (1994). *Classification and cognition*. New York: Oxford University Press.

Fific, M., Little, D. R., & Nosofsky, R. M. (2010). Logical-rule models of classification response times: A synthesis of mental-architecture, random-walk, and decision-bound approaches. *Psychological Review, 117,* 309–348.

Folstein, J., Gauthier, I., & Palmeri, T. J. (2012). Not all morph spaces stretch alike: How category learning affects object perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(4), 807–820.

Folstein, J., Palmeri, T. J., & Gauthier, I. (2013). Category learning increases discriminability of relevant object dimensions in visual cortex. *Cerebral Cortex, 23*(4), 814–823.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, Md.: Earlbaum.

Gauthier, I., & Palmeri, T. J. (2002). Visual neurons: Categorization-based selectivity. *Current Biology, 12,* R282–284.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–65.

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science, 16,* 152–160.

Gureckis, T. M., James, T. W., & Nosofsky, R. M. (2011). Re-evaluating dissociations between implicit and explicit category learning: An event-related fMRI study. *Journal of Cognitive Neuroscience, 23,* 1697–1709.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93,* 411–428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528–551.

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology, 16,* 243–275.

Johansen, M. K., & Palmeri, T. J. (2002). Are there representational shifts during category learning? *Cognitive Psychology, 45,* 482–553.

Kahana, M. J., & Sekuler, R. (2002). Recognizing spatial patterns: A noisy exemplar approach. *Vision Research, 42,* 2177–2192.

Knowlton, B., & Squire, L. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science, 262* (5140), 1747–1749.

Lafond, D., Lacouture, Y., & Cohen, A. L. (2009). Decision tree models of categorization response times, choice proportions, and typicality judgments. *Psychological Review, 116,* 833–855.

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

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

Little, D. R., Nosofsky, R. M., & Denton, S. E. (2011). Response -time tests of logical-rule models of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1–27.

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

Logan, G. D. (1997). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review, 103,* 603–649.

Mack, M., Gauthier, I., Sadr, J., & Palmeri, T. J. (2008). Object detection and basic-level categorization: Sometimes you know it is there before you know what it is. *Psychonomic Bulletin & Review,* 15(1), 28–35.

Mack, M. L., Preston, A. R., & Love, B. C. (2013). Decoding the brain's algorithm for categorization from its neural implementation. *Current Biology, 23*(20), 2023–2027.

Mack, M. L., & Palmeri, T. J. (2010). Decoupling object detection and categorization. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1067-1-79.

Mack, M. L., & Palmeri, T. J. (2011). The timing of visual object categorization. *Frontier in Psychology,* 2:165.

Mack, M. L., Wong, A. C.-N., Gauthier, I., Tanaka, J. W., & Palmeri, T. J. (2009). Time-course of visual object categorization: Fastest does not necessarily mean first. *Vision Research*, *49,* 1961–1968.

Mack, M. L., Wong, A. C.-N., Gauthier, I., Tanaka, J. W., & Palmeri, T. J. (2007). Unraveling the time-course of perceptual categorization: Does fastest mean first? In the *Proceedings of the Twenty-Ninth Annual Meeting of the Cognitive Science Society.*

Marley, A. A. J. (1992). Developing and characterizing multidimensional Thurstone and Luce models for identification and preference. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 299–333). Hillsdale, NJ: Earlbaum.

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General, 118,* 346–373.

McKinley, S. C., & Nosofsky, R. M. (1995). Investigations of exemplar and decision bound models in large, ill-defined category structures. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 128.

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

Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 775–799.

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology, 10,* 465–501.

Murdock, B. B., Jr. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, *89*, 609-626.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 104–114.

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, & Cognition, 13,* 87–109.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 700–708.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 3–27.

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology, 43,* 25–53.

Nosofsky, R. M. (2000). Exemplar representation without generalization? Comment on Smith and Minda's (2000) "Thirty categorization results in search of a model". *Journal of Experimental Psychology: Learning, Memory and Cognition, 26,* 1735–1743.

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 & Cognition*, 27(1), 78–93.

Nosofsky, R. M., Denton, S. E., Zaki, S. R., Murphy-Knudsen, A. F., & Unverzagt, F. W. (2012). Studies of implicit prototype extraction in patients with mild cognitive impairment and early Alzheimer's disease. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 860–880.

Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review, 118,* 280–315.

Nosofsky, R. M., Little, D. R., & James, T. W. (2012). Activation in the neural network responsible for categorization and recognition reflects parameter changes. *Proceedings of the National Academy of* Sciences, *109*, 333–338.

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 & Psychophysics, 59,* 1027–1048.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101,* 53–79.

Nosofsky, R. M., & Stanton, R. D. (2005). Speeded classification in a probabilistic category structure: Contrasting exemplar-retrieval, decision-boundary, and prototype models. *Journal of Experimental Psychology: Human Perception and Performance*, *31,* 608–629.

Nosofsky, R. M., & Stanton, R. D. (2006). Speeded old-new recognition of multidimensional perceptual stimuli: Modeling performance at the individual-participant and individual-item levels. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 314–334.

Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science*, *9*(4), 247–255.

Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 924–940.

Nosofsky, R. M., & Zaki, S. R. (2003). A hybrid-similarity exemplar model for predicting distinctiveness effects in perceptual old-new recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1194–1209.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 324–354.

Palmeri, T. J. (1999). Theories of automaticity and the power law of practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 543–551.

Palmeri, T. J. (2014). An exemplar of model-based cognitive neuroscience. *Trends in Cognitive Sciences, 18*(2), 67–69.

Palmeri, T. J., & Cottrell, G. (2009). Modeling perceptual expertise. In D. Bub, M. Tarr, & I Gauthier (Eds.), *Perceptual expertise: bridging brain and behavior*. Oxford University Press.

Palmeri, T. J., & Flanery, M. A. (1999). Learning about categories in the absence of training. *Psychological Science, 10*(6), 526–530.

Palmeri, T. J., & Flanery, M. A. (2002). Memory systems and perceptual categorization. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 41, pp. 141–189). San Diego: Academic.

Palmeri, T. J., Schall, J. D. & Logan, G. D. (2014). Neurocognitive modeling of perceptual decision making. In J. R. Busemeyer, J. Townsend, Z. J. Wang, & A. Eidels (Eds.), *Oxford handbook of computational and mathematical Psychology*. Oxford University Press.

Palmeri, T. J., Wong, A. C.-N., & Gauthier, I. (2004). Computational approaches to the development of perceptual expertise. *Trends in Cognitive Science, 8,* 378–386.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.

Purcell, B. A., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2012). From salience to saccades: multiple-alternative gated stochastic accumulator model of visual search. *Journal of Neuroscience, 32*(10), 3433–3446.

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

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

Reber, P. J., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2003). Dissociating explicit and implicit category knowledge with fMRI. *Journal of Cognitive Neuroscience, 15*(4), 574–583.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998a). Cortical areas supporting category learning identified using fMRI. *Proceedings of the National Academy of Sciences*, *95*, 747–750.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998b). Contrasting cortical activity associated with category memory and recognition memory. *Learning and Memory*, *5*, 420–428.

Richler, J. J., & Palmeri, T. J. (2014). Visual category learning. *Wiley Interdisciplinary Reviews in Cognitive Science, 5*, 75–94.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Rouder, J. N., & Ratcliff, R. (2004). Comparing categorization models. *Journal of Experimental Psychology: General, 133,* 63–82.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, *22,* 325–345.

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

Shin, H. J., & Nosofsky, R. M. (1992). Similarity scaling studies of dot pattern classification and recognition. *Journal of Experimental Psychology: General, 121,* 278–304.

Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature, 415*(6869), 318–320.

Smith, E. E. (2008). The case for implicit category learning. *Cognitive, Affective, and Behavioral Neuroscience*, *8*, 3–16.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, J. D. (2002). Exemplar theory's predicted typicality gradient can be tested and disconfirmed. *Psychological Science, 13,* 437–442.

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of*

Experimental Psychology: Learning, Memory and Cognition, 24, 1411–1436.

Smith, J. D., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26,* 3–27.

Sternberg, S. (1966). High speed scanning in human memory. *Science, 153,* 652–654.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: is the basic level in the eye of the beholder? *Cognitive Psychology, 23,* 457–482

Verguts, T., Storms, G., & Tuerlinckx, F. (2003). Decision-bound theory and the influence of familiarity. *Psychonomic Bulletin & Review, 10,* 141–148.

Wickelgren, W. A. (1974). Single-trace fragility theory of memory dynamics. *Memory & Cognition, 2,* 775–780.

Wills, A. J., & Pothos, E. M. (2012). On the adequacy of current empirical evaluations of formal models of categorization. *Psychological Bulletin*, *138*, 102–125.

Wixted, J. T. (2004). On common ground: Jost's (1897) law of forgetting and Ribot's (1881) law of retrograde amnesia. *Psychological Review, 111,* 864–879.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science, 2,* 409–415.

Zaki, S. R., & Nosofsky, R. M. (2007). A high-distortion enhancement effect in the prototype-learning paradigm: Dramatic effects of category learning during test. *Memory & Cognition, 35,* 2088–2096.

Zaki, S. R., Nosofsky, R. M., Jessup, N. M., & Unverzagt, F. W. (2003). Categorization and recognition performance of a memory-impaired group: Evidence for single-system models. *Journal of the International Neuropsychological Society*, *9*(3), 394–406.

Zandbelt, B. B., Purcell, B. A., Palmeri, T. J., Logan, G. D., & Schall, J. D. (2014). Response times from ensembles of accumulators. *Proceedings of the National Academy of Sciences, 111*(7), 2848–2853.