Journal of Experimental Psychology:
Human Perception and Performance
2000, Vol. 26, No. 5, 1549–1569

Copyright 2000 by the American Psychological Association, Inc.
0096-1523/00/$5.00 DOI: 10.1037/0096-1523.26.5.1549

An Exemplar-Retrieval Model of Speeded Same–Different Judgments

Andrew L. Cohen and Robert M. Nosofsky
Indiana University Bloomington

R. M. Nosofsky and T. J. Palmeri's (1997) exemplar-based random-walk (EBRW) model of speeded classification is extended to account for speeded *same–different* judgments among integral-dimension stimuli. According to the model, an important component process of *same–different* judgments is that people store individual examples of experienced *same* and *different* pairs of objects in memory. These exemplar pairs are retrieved from memory on the basis of how similar they are to a currently presented pair of objects. The retrieved pairs drive a random-walk process for making *same–different* decisions. The EBRW predicts correctly that *same* responses are faster for objects lying in isolated than in dense regions of similarity space. The model also predicts correctly effects of same-identity versus same-category instructions and is sensitive to observers' past experiences with specific *same* and *different* pairs of objects.

The main tenet of exemplar-based models of cognition is that people store particular instances of events in memory, called exemplars, and that these exemplars are later retrieved to perform a particular task. Exemplar-based models have long been used to model performance in categorization tasks (Medin & Schaffer, 1978). These models assume that people represent categories as a set of exemplars and make categorization decisions by retrieving exemplars from memory. Such models have rendered accurate quantitative accounts of category learning, transfer, and generalization (Hintzman, 1986; Kruschke, 1992; Nosofsky, 1986).

Recently, Nosofsky and Palmeri (1997) proposed an exemplar-based model of classification, which also accounts for the time-course of classification. This exemplar-based random-walk (EBRW) model posits that people store category examples along with their category labels in memory. These exemplars are assumed to reside in a multidimensional psychological space. During a classification judgment, exemplars race to be retrieved. How fast an exemplar races is determined by its similarity to the test item. The winning exemplar adds incremental evidence to a random walk process. A response occurs when the random walk counter reaches a criterion. The EBRW correctly predicts effects of within- and between-category similarity, practice, and familiarity on classification response times and accuracies. This model is discussed in more detail below.

Exemplar-based models have also been successful in other domains. For example, Logan's (1988, 1992) instance theory of automaticity construes skilled action as a race between algorithmic and instance-based processes. He suggests that to perform a task (e.g., deciding whether a four-letter string is a word or a nonword), novices must rely on an algorithmic method. However, on each exposure to the task, a representation of the solution to that task, an instance, is stored in memory. As an alternative to the algorithm, these instances can be retrieved and used to perform the task. As the person becomes practiced, more and more instances are stored. Automaticity of the task has been achieved when the instance-based processes dominate the race. This model can account quantitatively for the power-function decrease in mean and standard deviation reaction times often seen with training (Logan, 1988).

Palmeri (1997) extended such exemplar-based models to account for the effects of exemplar similarity on the development of automaticity. Participants judged the numerosity of between 6 and 11 dots in random patterns. Reactions times (RTs) flattened out after several days of training, signaling that the task had become automatic. Palmeri used the EBRW (Nosofsky & Palmeri, 1997) to model the results. The model correctly accounted for the inverse relationship between RTs and the similarity of new patterns to old exemplars. The EBRW also predicted the facilitation of the development of automaticity for categories with high within-category similarity and low between-category similarity.

Finally, exemplar models have played central roles in accounting for performance in other cognitive domains as well, including individual object identification, old–new recognition memory, and problem solving. Classic models of object identification posit that the probability with which an object is misidentified is based on its similarity to the other exemplars in the set of alternatives (e.g., Lockhead, 1972; Luce, 1963; Shepard, 1958; Townsend & Landon, 1982). Various models of old–new recognition memory posit that the degree to which an item is judged as familiar is based on its summed similarity to all previously experienced exemplars (e.g., Hintzman, 1988; Nosofsky, 1988, 1991). And various theories of problem solving posit that current problems are solved by relying on similarities to past examples of the problems (e.g., Ross, 1987).

In summary, previous research suggests the use of exemplar retrieval in a wide variety of cognitive tasks (for a review of the potential role of exemplar retrieval in still other cognitive tasks,

This work was supported by Grant PHS R01 MH48494-09 from the National Institute of Mental Health and by a Jacob Javits Fellowship.

We thank Nathaniel Blair, Jerome Busemeyer, Daniel Ennis, Richard Shiffrin, and an anonymous reviewer for helpful discussions, criticisms, and suggestions related to this work.

Correspondence concerning this article should be addressed to Andrew L. Cohen or Robert M. Nosofsky, Department of Psychology, 1101 E. 10th St., Indiana University, Bloomington, Indiana 47405-7007. Electronic mail may be sent to alcohen@indiana.edu or to nosofsky@indiana.edu.

see, e.g., Jacoby & Brooks, 1984). In the present research, we continue to explore a potential role of exemplar retrieval by examining performance in tasks of *same–different* perceptual judgment. In these tasks, a pair of objects is presented on a given trial and observers are required to judge whether they are the *same* or *different*. The *same–different* task is a seemingly quintessential example of an on-line perceptual judgment task that involves little or no memory for previous experience. A pair of stimuli is presented concurrently, and if they are judged sufficiently similar, then a *same* response is made, else a *different* response. Memory for specific past exemplars seems to play no role in such a situation. In contrast to this view, a central thesis of this article is that even in the concurrent *same–different* judgment task, memories for past exemplars play an important role. Furthermore, we develop an extended version of Nosofsky and Palmeri's (1997) EBRW to account for performance in *same–different* judgment tasks.

This idea about the role of memory for exemplars in influencing *same–different* judgments was brought out in compelling fashion in a previous study reported by Crist (1981). The theme of Crist's (1981) experiments was to demonstrate that the similarity structure of the total set of stimulus possibilities in a given experimental context plays an important role in *same–different* judgment. In other words, it is not simply the properties of the two objects that are judged on a given trial that determine performance but rather their relationship to the other exemplars in the stimulus set. In particular, Crist (1981) demonstrated that observers are faster to judge two identical objects as *same* if the object in the pair is located in an isolated region of similarity space than if it is located in a dense region of similarity space. This idea is illustrated schematically in Figure 1. In this example, there are four objects in the stimulus set, the lower case letters *a, e, o,* and *f*. (These letters formed a subset of the ones actually used by Crist, 1981, in some of his experiments.) In this illustrative example, note that the letter *f* is located in an isolated region of the similarity space, whereas the letters *e* and especially *o* and *a* are located in a dense region. The type of phenomenon observed by Crist (1981) was that participants were faster to respond *same* to the pair (*f, f*), which is an isolated pair, than to pairs such as (*o, o*) and (*a, a*), which are dense pairs.

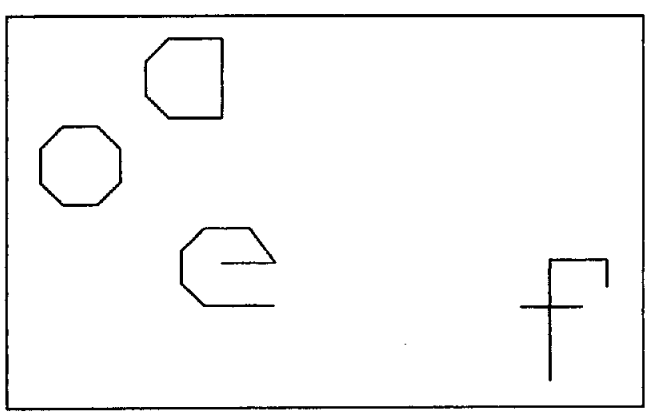

Figure 1. A subset of the stimuli used by Crist (1981). The perceived distance between stimuli is schematically illustrated by physical distance in the figure.

Crist (1981) explained this phenomenon in terms of Lockhead's (1970, 1972) holistic-discriminability model, which had previously been used to explain patterns of results in tasks of stimulus identification and classification. As described by Crist (1981, pp. 271–272), according to the holistic-discriminability model,

> the observer has an internal representation of the stimulus set that is structured by the principle of similarity. . . . In identification of singly presented stimuli, the initial task of the observer is "one of locating the multidimensional stimulus in that [representational] space . . ." (Lockhead, 1970, p. 3). How easily this locus in perceptual space may be located, its discriminability, is determined by the distance to other potential sites in the perceptual space. If there are other expected stimuli that are very similar to the stimulus object and thus many points near to each other in the perceptual space, then identification will be relatively slower and less accurate than if the stimulus items had been isolated in similarity space.

Crist (1981, p. 276) then extended this basic idea to the *same–different* task by suggesting that "performance in the matching task is a joint function of how easily each of the two items is located in the representational space." Thus, it is easy to rapidly locate the objects in the pair (*f, f*) because the object *f* is located in an isolated region, whereas it is difficult to locate the objects in the pair (*o, o*) because *o* is located in a dense region (it is hard to discriminate from *a*). Note that the holistic-discriminability model entails the assumption that memories for exemplars not present on a given trial influence *same–different* judgment. For example, the hard-to-discriminate *a* is not present on trials in which *same* pair (*o, o*) is judged.

The present research follows up on Crist's (1981) earlier work in several ways. Most important, the holistic-discriminability model that Crist (1981) used for explaining the fast responding associated with isolated *same* pairs was presented in broad, conceptual terms, with no attempt at formalization. In addition, as acknowledged by Crist (1981, p. 290), "For the purposes of elucidating and defining specific mechanisms by which objects are identified or classified, the holistic-discriminability approach is unsatisfactory to date. It does not explain how stimulus processing is accomplished." Finally, although the conceptual model described by Crist (1981) addresses the fast responses associated with isolated *same* pairs, it provides no direct account of other fundamental phenomena observed in the *same–different* judgment task. For example, observers are faster to judge two objects as *different* the less similar they are to one another. Yet, because Crist's (1981) account places emphasis solely on the time that it takes to locate each individual object in the stimulus space, it provides no direct explanation for this fundamental, robust phenomenon (for similar comments, see Nickerson, 1981, p. 299).

In the present work, we develop and test a formal model of *same–different* judgment that shares some key ideas with Crist's (1981) holistic-discriminability approach but that goes beyond that approach in the respects outlined above. First, the formalized model is able to make precise quantitative predictions of the speed and accuracy of *same–different* judgments while taking into account both the similarity of the presented pair of objects to each other, as well as their relations to the other exemplars in the

stimulus set. Second, it provides a unified account of the speed and accuracy of making both *same* and *different* judgments while providing a mechanistic account of processing in the *same–different* judgment task.

Our proposed formal model is an extended version of Nosofsky and Palmeri's (1997) EBRW model. In this extended model, two component processes are assumed to drive *same–different* judgments. First, when a pair of objects is presented, observers are assumed to make on-line judgments of perceptual similarity. High values of judged similarity tend to drive a random-walk counter toward a *same* response criterion, whereas low values of judged similarity tend to drive the random walk toward a *different* response criterion. Analogous component processes are incorporated in a variety of previously formalized models for the *same–different* judgment task (e.g., Ennis, Palen, & Mullen, 1988; Link, 1992; Ratcliff, 1981; Takane & Sergent, 1983; Thomas, 1996). Such a process accounts for the well-known result that observers are faster to make a *different* judgment as the similarity between the objects in a pair decreases.

The key novel assumption in the extended EBRW model is that observers also store memories for specific pairs of exemplars that have received *same* and *different* feedback during the course of the experiment. When a pair of objects is presented, various of these exemplar pairs are retrieved from memory according to how similar they are to the current pair of objects. Retrieval of *same* pairs moves the random walk toward the *same* response criterion, whereas retrieval of *different* pairs moves the random walk toward the *different* response criterion.

Consider how such a process might account for the pattern of results observed in Crist's (1981) *same–different* data. Suppose that a *same* pair is presented from an isolated region, for example, pair (*f, f*) in Figure 1. The only exemplar pair in memory that is highly similar to this *same* pair is the *same* pair itself. Thus, the exemplar-retrieval process would consistently move the random walk toward the *same* criterion, so rapid responding and high accuracy would be observed for isolated *same* pairs. By contrast, suppose that a *same* pair is presented from a dense region of the similarity space, for example, pair (*o, o*) in Figure 1. In this case, there are various *different* pairs that are also highly similar to this *same* pair of objects, for example, pairs (*o, a*) and (*o, e*). Retrieval of these *different* exemplar pairs would tend to move the random walk toward the *different* criterion, so the dense *same* pairs would have slower response times and lower accuracy, just as was observed in Crist's (1981) data.

Note that although our extended EBRW model shares some key ideas with Crist's holistic-discriminability model, there are some important conceptual distinctions as well. As described by Crist (1981), a *same–different* judgment takes place by making two successive identifications of the individual objects. The speed with which identification of each object takes place is determined by the similarity structure of the entire set of stimulus possibilities. By contrast, in the EBRW, pairs of objects are retrieved from memory, and performance is determined by how similar the presented pair is to the specific set of pairs on which an observer was trained.

We organize the remainder of our article as follows. In the next section, we report an experiment designed to collect a rich set of *same–different* response time data and to corroborate Crist's (1981) earlier findings that *same* judgments are systematically related to a stimulus' degree of isolation in multidimensional similarity space. We then present our formalized extension of the EBRW model and test its ability to quantitatively fit the complete set of *same* and *different* RTs observed in the task. We follow this experiment with ones that are designed to distinguish our account of the fast isolated *same* responses from those of some alternative models. Finally, holding fixed the set of stimulus possibilities in the experiment, we report additional studies that provide support for the assumption in the EBRW that *same–different* judgments are sensitive to the specific pairs on which observers have been trained.

Experiment 1: Collection of a Rich Set of *Same–Different* RT Data

The purpose of Experiment 1 in this article was to gain further evidence of Crist's (1981) observation of an effect of isolation on the speed of *same* judgments and to obtain a rich set of RT data suitable for quantitative fitting by the EBRW.

In Experiment 1 and the remaining experiments, we used as stimuli sets of colors varying in hue, brightness, and saturation. These stimuli are examples of integral-dimension stimuli, in which the individual component dimensions combine into relatively unanalyzable, integral wholes (Garner, 1974; Lockhead, 1972; Shepard, 1987; Shepard & Chang, 1963). We focused on integral-dimension stimuli in this research because our primary interest was on the role of overall similarity relations among objects in influencing *same–different* judgments. For integral-dimension stimuli, a reasonable simplifying assumption is that the objects are encoded holistically and the mean encoding time for each object is roughly equal. Thus, the main variable that should influence the *same–different* RTs is overall similarity relations among the objects. By contrast, for stimuli composed of highly separable dimensions (Garner, 1974; Shepard, 1964), a key determinant of *same–different* response times undoubtedly involves the process by which the individual component dimensions of the objects are encoded and compared, a topic that goes beyond the scope of the present research.

The stimuli for Experiment 1 were eight colors presented on a computer screen. The colors varied in degree of isolation. Three participants engaged in a 10-day speeded *same–different* judgment task. On each trial, a pair of colors was displayed and participants responded *same* or *different* as quickly and accurately as possible. Feedback was provided on each trial. Although it was hoped that error rates would be nonnegligible for modeling purposes, participants rarely made errors, and thus the important dependent variable was RT. In addition to the *same–different* task, the participants rated the similarity between each pair of colors. These ratings were used to create multidimensional scaling (MDS) solutions for the colors. The MDS solutions were then used in conjunction with the EBRW to predict the data.

Method

Participants. The participants were 3 Indiana University psychology graduate students. They were paid $10 per day. All participants claimed to have corrected-to-normal vision and normal color vision. None of the participants was familiar with the issues under investigation in this study.

Stimuli. The stimuli were eight colors presented on a computer screen. The colors were created using Adobe Photoshop 4.0 on a Macintosh Quadra 840AV and SuperMatch 20-T XL monitor with the red, green, and blue (RGB) values as given in Table 1. They were displayed on a Dell Ultrascan 800HS Series monitor (Model D825HT) using a Dell Dimension P133v CPU. The same hardware was used each day by all 3 participants. The monitor's contrast and brightness settings were 80 and 40, respectively. Each color occupied a 5.08 cm × 5.08 cm square surrounded by a black background. The colors were shown side by side, 1.90 cm apart. Participants entered responses by pressing appropriate keys on the keyboard. Response times were measured using the internal ms-accuracy PC timer.

Procedure. The experiment lasted 10 days and took approximately 1.5 h per day. The experiment involved three tasks: a *same–different* task, a similarity scaling task, and a categorization task.

During the *same–different* task, participants viewed pairs of the eight stimuli and judged whether the two colors were the *same* color or *different* colors. Within each block, each of the 28 different pairs was presented once and each of the 8 same pairs was presented twice. Thus, each block had 44 trials. Both the order of presentation of pairs and the left–right placement of the colors within each pair were randomized. Participants were instructed to rest their index fingers on the *same* and *different* response buttons and were urged to respond as quickly as possible without sacrificing accuracy. Each trial began with a 100 ms black screen. A cue then appeared in the center of the screen for 750 ms. Then, the cue disappeared and a pair of stimuli was shown, one color on each side of the cue location. This display was response terminated. A correct response was followed by a high frequency 100 ms beep, and an incorrect response was followed by a low frequency 800 ms beep. Incorrect feedback also included a message describing why the response was wrong. The stimuli remained on the screen during feedback.

In the similarity-scaling block, each of the 28 different pairs of colors was presented once. Both the order of presentation and left–right placement were randomized. The 5.08 cm × 5.08 cm color squares were presented side by side separated by 1.90 cm. Participants judged the similarity between the colors in the pair using a 9-point scale (1 = *very dissimilar*, 9 = *very similar*). As a reminder, a rating scale was displayed on the bottom of the screen. Participants were urged to use the full range of ratings.

In the categorization tasks, the observers learned to classify each color into each of three levels of a hierarchical category structure. At the subordinate level, each color was assigned its own category label. At an intermediate level, unique labels were assigned to color pairs 1 and 2, 3 and 4, 5 and 6, and 7 and 8. Colors 1, 2, 5, and 6 composed one superordinate level category, and Colors 3, 4, 7, and 8 composed the other superordinate category. Depending on the block, observers classified the colors into either the subordinate, intermediate, or superordinate-level categories. On each trial, a single color was randomly selected and presented on the screen; the participant classified it as rapidly and accurately as possible; and corrective feedback was then provided. The data from the categorization tasks form part of a separate research project and are not discussed further in this article.

The arrangement of blocks over the 10 days was not identical. Day 1 included a *same–different* task block before and after a single categorization block. Days 2 through 4 began with a *same–different* block followed by five alternations of categorization and *same–different* blocks. Days 5 through 10 were identical to Days 2 through 4 except that a similarity-scaling block followed each *same–different* block. Thus, there were 56 *same–different* blocks, involving 2,464 trials of 896 *same* trials and 1,568 *different* trials. Each *different* pair was seen in 56 trials, and each *same* pair was seen in 112 trials. There were 36 similarity-scaling blocks, that is, 36 ratings per pair.

Table 1
Red (R), Green (G), and Blue (B) Values and Approximate Color for Each Stimulus in Experiment 1

Stimulus	R	G	B	Color
1	99	66	57	Light brown-gray
2	82	49	41	Dark brown-gray
3	181	66	49	Pinkish orange
4	165	49	24	Reddish orange
5	140	90	66	Light brown
6	123	74	49	Dark brown
7	189	74	24	Light orange
8	181	57	8	Dark orange

Results

Multidimensional scaling. MDS solutions for the colors were derived by fitting the standard Euclidean model separately to each participant's overall similarity ratings. The data for the scaling were the averaged similarity rating matrices for each participant. The three dimensional results of the scaling after a rotation to be maximally similar to the physical dimensions are shown in Figure 2. These scaling solutions account for 99.9, 100.0, and 99.7% of the variance in the three participant's averaged similarity ratings, respectively. As desired, certain colors occupy dense regions of the space (e.g., Colors 3, 4, 7, and 8) and others occupy isolated regions (e.g., Colors 1 and 2).

Same–different judgments. In the following analyses, the first five trials of each block, the data from Day 1, and RTs of less than 150 ms or greater than 4 s were excluded. The results of the *different* trials, that is, the trials where the two colors viewed were not identical, are shown in Figure 3. The graphs display the mean RT for each color pair plotted against the MDS derived distance between the two colors. As expected, as the distance between the colors in a pair increases, the mean RT for responding *different* tends to decrease.[1]

Recall that one of the main goals of this experiment was to examine the effect of isolation on *same* judgments. Table 2 shows the RT and percent correct data for each color and for each participant averaged over the *same* trials, that is, the trials where the two colors viewed were identical.

First, notice that the participants were near ceiling performance in terms of accuracy. We had hoped to collect both RT and accuracy data for modeling. Unfortunately, although the accuracy data are included in the model fits, most of the data variance involves RTs only.

Second, as predicted, the RTs for the *same* trials are not identical. This observation is confirmed by a statistical test. A one-way analysis of variance (ANOVA) comparing the mean RTs of the colors showed significant differences for participant 1, $F(7, 759) = 8.974$, $p < .001$; participant 2, $F(7, 757) = 17.236$, $p < .001$; and participant 3, $F(7, 759) = 4.668$, $p < .001$.

Finally, the circles in Figure 2 illustrate individual *same* reaction times for the colors. In corroboration of Crist's (1981) earlier results, the figure suggests that stimuli in dense regions of the

[1] Please contact Andrew L. Cohen for a table of the mean RTs and accuracies for each participant and for each color pair.

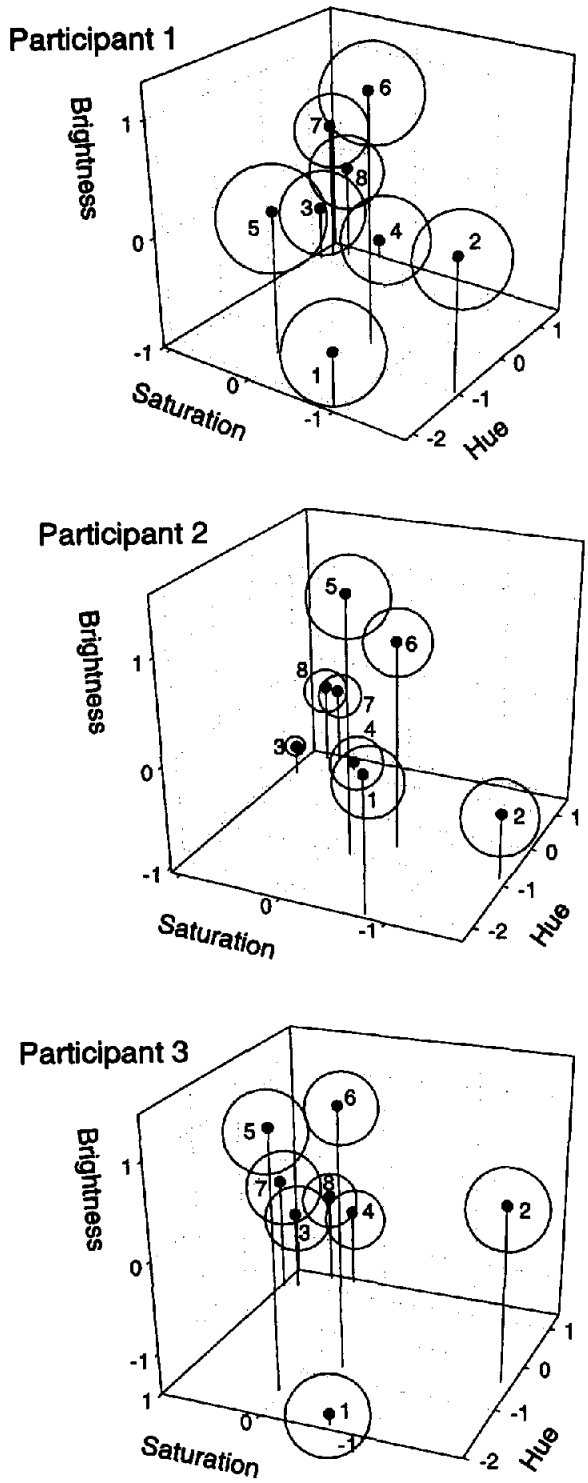

Figure 2. Individual participant multidimensional scaling (MDS) solutions derived from averaged similarity ratings for the color stimuli of Experiment 1. The center of each circle represents the MDS coordinate for each color. The diameter of each circle is a linear decreasing function of the mean reaction time on *same* trials. Thus, smaller circles denote slower responses.

psychological color space have slower *same* reaction times than colors in isolated regions. We tested this observation statistically in two different ways. First, we conducted a one-tailed Pearson correlation on RTs and a nearest-neighbor measure of density. That is, each stimulus was represented by its shortest distance to any other stimulus in the MDS space. The correlations were $r = -.973$, $-.553$, and $-.651$ and $p = .000$, .078, and .040 for the 3 participants, respectively. Second, we conducted a one-tailed Pearson correlation on RT and a summed-similarity measure of density. Summed similarity is determined by summing the similarity of a color to all other colors (where similarity is determined by Equation 2 in the next section).[2] Thus, the lower a color's summed similarity, the more isolated it is. The correlations were $r = .762$, .715, and .883 and $p = .014$, .023, and .002 for the 3 participants, respectively. Thus, as predicted, mean RTs for *same* pairs correlate with their degree of isolation.

The Extended Exemplar-Based Random-Walk Model

The extended version of the EBRW applied to *same–different* judgments builds directly on the ideas from the categorization model. Exemplars are represented as points in a multidimensional psychological space, and similarity between exemplars is a decreasing function of distance in the space. An observer establishes criteria representing the amount of evidence needed to initiate either a *same* or *different* judgment. When a test item is presented, similarity-based processes contribute information that moves the random walk toward either the *same* or *different* criterion.

In the extended model, two separate processes compete to drive the random walk. The first process, the *on-line* process, operates relatively independently of memory and is based only on the perceived similarity of the two items presented. The more similar the two items are judged, the higher the probability that the on-line process will move the random walk toward the *same* boundary. Likewise, the more dissimilar the two items are judged, the more likely the on-line process will move the random walk toward the *different* boundary. It is important to note that because we assume the self-similarities of all stimuli are equal, this process predicts identical response times for all *same* pairs of objects. Later in this article, we consider models that allow for differing degrees of self-similarity.

The second process, the *exemplar* process, is the key contribution of this model. It involves the idea that *same* and *different pairs* of objects are stored in memory as exemplars and are later retrieved when people make *same–different* judgments. When an object pair is presented, that *pair* is stored in memory as an exemplar along with its associated *same* or *different* feedback. When a test pair is presented, each exemplar pair in memory races to be retrieved at a rate determined by its similarity to the test pair. If a *same* pair wins the race, this process moves the random walk in the *same* direction. If a *different* pair wins, the process moves the random walk in the *different* direction. Thus, the probability of

[2] We used the best fitting parameters from Table 4 to calculate summed similarity.

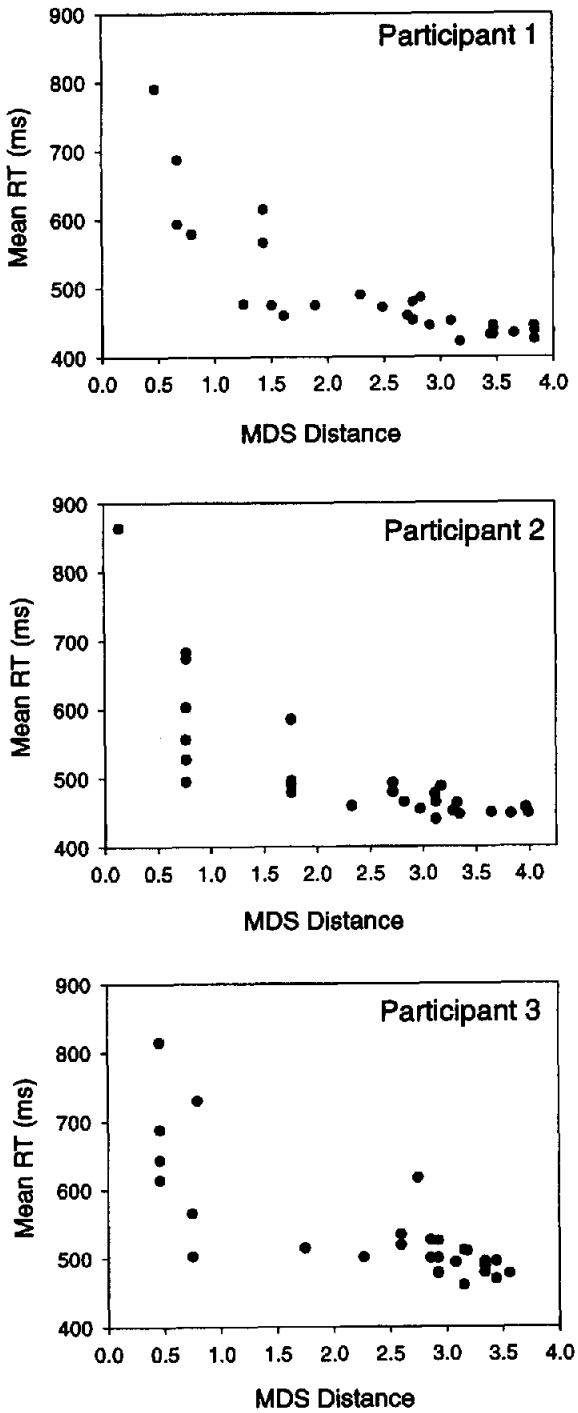

Figure 3. Mean reaction time (RT) for each *different* color pair in Experiment 1 plotted against the multidimensional-scaled distances between the two colors. MDS = multidimensional scaling.

taking a step toward the *same* criterion increases with the similarity of the current pair of stimuli to all of the previously seen *same* pairs. Likewise, the probability of taking a step toward the *different* boundary increases with the similarity of the current pair of stimuli to all previously seen *different* pairs. Note that it is pairs of stimuli that are being compared, not the individual stimuli. As discussed in the introduction, this exemplar-retrieval process can predict nonidentical RTs for *same* pairs.

The on-line and exemplar processes race to contribute information to the random walk counter, which then determines the *same–different* decision.

Formal Properties

As in the standard EBRW, we assume the individual item exemplars are represented as points in an M-dimensional psychological space. Let x_{im} denote the coordinate value of exemplar i on psychological dimension m. Then the distance between exemplars i and j is given by

$$d_{ij} = \sqrt{\sum_m w_m (x_{im} - x_{jm})^2}, \tag{1}$$

where w_m ($0 \leq w_m$, $\Sigma\, w_m = 1$) is the attention weight given to dimension m (Nosofsky, 1986, 1987).

Furthermore, as in the standard EBRW, the similarity of exemplar i to exemplar j (s_{ij}) is an exponential decay function of the distance between these objects,

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

where c is an overall sensitivity parameter used to scale distances in the space. There is a good deal of independent support for these assumptions about the relation between similarity and psychological distance (Shepard, 1987).

In our extension of the EBRW to *same–different* judgments, we assume that on each step of the random walk, there is a cognitive element that races exponentially to be retrieved (with rate λ, a free parameter). This cognitive element represents the on-line process of judged similarity. If the on-line element is retrieved, then the probability that the random walk takes a step toward the *same* criterion is given by

$$\psi_{ij}^{same} = \frac{s_{ij}}{s_{ij} + \beta}, \tag{3}$$

where β is a free parameter that acts as a criterion for translating judged similarity into evidence for a *same* or *different* judgment. Likewise, if the on-line element is retrieved, the probability of taking a step in the *different* direction is given by

$$\psi_{ij}^{diff} = 1 - \psi_{ij}^{same} = \frac{\beta}{s_{ij} + \beta}. \tag{4}$$

The exemplar-retrieval process operates as follows. The similarity between presented pair (i, j) and exemplar-pair (u, v) stored in memory is defined as

$$s_{ij,uv} = \max(s_{iu} \cdot s_{jv}, s_{iv} \cdot s_{ju}), \tag{5}$$

where the individual component similarities (s_{ij}, s_{jv}, s_{iv}, and s_{ju}) are computed as

$$s_{ij} = \exp(-c^* \cdot d_{ij}), \tag{6}$$

Table 2
Observed Proportions Correct, Mean Reaction Times (in Milliseconds), and Reaction Time Standard Deviations for Each Participant and Each Stimulus of Experiment 1

	Participant 1			Participant 2			Participant 3		
Stimulus	*P*(C)	RT	*SD*	*P*(C)	RT	*SD*	*P*(C)	RT	*SD*
1	0.99	534	80	0.97	649	132	0.96	612	184
2	0.98	553	103	0.93	643	142	0.99	610	180
3	0.93	619	177	0.98	860	263	0.99	687	194
4	0.98	603	197	0.99	725	191	0.98	704	261
5	0.98	528	108	0.91	613	121	1.00	600	148
6	0.99	547	94	0.93	671	155	1.00	644	285
7	0.99	643	217	0.96	769	242	1.00	651	191
8	0.99	641	208	0.97	764	256	0.97	726	244

Note. *P*(C) = proportion correct; RT = mean reaction time; *SD* = reaction time standard deviation.

and c^* is a free parameter that acts to scale distances in the same manner as c in Equation 2. Equation 5 formalizes the idea that two pairs of objects are highly similar only if the constituent objects across the pairs are highly similar. The constituent objects within each pair are aligned, however, to allow maximum similarity (cf. Goldstone, 1994). The "activation" of pair (u, v) given pair (i, j) is then given by

$$a_{ij,uv} = M_{uv} \cdot s_{ij,uv}, \tag{7}$$

where M_{uv} is the memory strength of pair (u, v). The memory strength of a pair may vary because of factors such as presentation frequency, recency of presentation, and so forth.

Given these definitions of pairwise similarity and activation, the exemplar-retrieval process in the *same–different* model operates the same way as in the standard EBRW. All exemplar pairs race exponentially to be retrieved with rate given by their activation values. Specifically, the probability density that exemplar-pair (u, v) finishes its race at time t is given by

$$f(t) = a_{ij,uv} \exp(-a_{ij,uv} \cdot t). \tag{8}$$

If exemplar-pair (u, v) wins the race, then it is retrieved and enters the random-walk process. If the retrieved pair (u, v) has been associated with *same* feedback, then the random walk is increased by a unit value in the *same* direction, whereas if it has been associated with *different* feedback, then the random walk is decreased by a unit value in the *different* direction.

Let ϕ_{ij}^{same} denote the summed activation of all *same* exemplar pairs stored in memory given pair (i, j), that is,

$$\phi_{ij}^{same} = \sum_{u} a_{ij,uu}, \tag{9}$$

and likewise, let ϕ_{ij}^{diff} denote the summed activation of all *different* exemplar pairs. Given the processing assumptions outlined above and given the mathematical properties of the exponential distribution, it is straightforward to show that the probability that the on-line element is retrieved on any given step of the random walk is given by

$$P(\text{on-line}|(i, j)) = \frac{\lambda}{\lambda + \phi_{ij}^{same} + \phi_{ij}^{diff}}, \tag{10}$$

whereas the probability that an exemplar-pair associated with *same* feedback is retrieved is given by

$$P(\textit{same}\ \text{exemplar}|(i, j)) = \frac{\phi_{ij}^{same}}{\lambda + \phi_{ij}^{same} + \phi_{ij}^{diff}}. \tag{11}$$

(Recall that λ is a free parameter representing the rate at which the on-line element races to be retrieved.) Thus, the overall probability that the random walk moves in the *same* direction on any step, given pair (i, j), is given by

$$p_{ij}^{same} = P(\text{on-line}|(i, j)) \cdot \psi_{ij}^{same} + \frac{\phi_{ij}^{same}}{\lambda + \phi_{ij}^{same} + \phi_{ij}^{diff}}, \tag{12}$$

and the probability of moving in the *different* direction is

$$p_{ij}^{diff} = 1 - p_{ij}^{same}. \tag{13}$$

Finally, following Nosofsky and Palmeri (1997, pp. 268–269), it can be shown that the expected time to take each step in the random walk is given by

$$E(T|(i, j)) = \alpha + \frac{1}{\lambda + \phi_{ij}^{same} + \phi_{ij}^{diff}}, \tag{14}$$

where α is a constant term associated with each step. Given these values of p_{ij}^{same}, p_{ij}^{diff}, and $E(T)$, the analytic equations from the standard EBRW (Equations 14–17 from Nosofsky & Palmeri, 1997) are then applied directly to predict the probability with which a *same* or *different* judgment is made, as well as the expected time to make each of these judgments.

Submodel

To evaluate the need for combining the two processes, we also tested a model involving only the on-line process. Thus, the probability of taking a step in the *same* direction is given by

$$p_{ij}^{same} = \psi_{ij}^{same}, \tag{15}$$

where ψ_{ij}^{same} is given by Equation 3. The time to take each step is given by

$$E(T) = \alpha. \tag{16}$$

The rest of the model remains unchanged. Note that this submodel arises as a special case of the full model if $\lambda = \infty$.

Experiment 1 Modeling Results

We used the MDS solutions described previously to fit the EBRW to each individual participant's mean RT and accuracy data for each color pair. The parameters in the full model are the individual-item sensitivity parameter, c (Equation 2); a separate sensitivity parameter c^* used for computing similarities between stimulus pairs (Equation 6); the attention weights, w_1 and w_2 (with $w_3 = 1 - w_1 - w_2$; Equation 1); the on-line criterion parameter, β (Equations 3 and 4); the two random walk response criteria, S and D; the on-line rate parameter, λ (Equations 10, 11, and 14); and the step-time constant, α (Equations 14 and 16). In addition, linear regression was used to convert the EBRW predictions, which are in arbitrary units, into milliseconds. Thus, the full version of the model uses 11 free parameters. A single set of parameters was used to simultaneously fit both the RT and accuracy data. A different set was used for each participant.

Because each *different* pair was seen exactly half as often as each *same* pair, we assumed that the memory strength for a *different* pair was exactly half that of a *same* pair. In the model fitting, we set

$$M_{uv} = \begin{cases} 10 & \text{if } u = v \\ 5 & \text{otherwise.} \end{cases} \tag{17}$$

The choice of $M_{uu} = 10$ is arbitrary.

The models were fit to the data by minimizing the (weighted) sum of squared deviations (SSD) between the predicted and observed RTs and accuracies (72 data points). In order to bring the contribution of the accuracy data to this fit measure more in line with the contribution of the RT data, the accuracy SSD was weighted by 10 before being combined with the reaction time SSD.

Figure 4 shows the model-fitting results for each participant's RT data. The model did a reasonably good job of fitting the RT data, accounting for 92, 91, and 83% of the variance in the mean RT data for participants 1–3, respectively.[3] The model did not do a good job of accounting for the accuracy data, however, although the predicted values do come in the ballpark of the observed values. One problem in interpreting the accuracy data fits is that there was relatively little overall variability in these data. Thus, even small perturbations between predicted and observed values can have large effects on the percent variance measure of fit. At this point, it is unclear whether the poor fit to the accuracy data represents a fundamental shortcoming of the model or if it is due to the extremely low variability in the accuracy data. The various summary measures of fit are reported in Table 3, and the best fitting parameters are reported in Table 4.

Of greatest interest here, notice from Figure 4 that the model accounts for much of the variation in *same* RTs. Those colors that are relatively isolated (e.g., Colors 1, 2, 5, and 6) are predicted to have consistently faster RTs than the dense colors (i.e., Colors 3, 4, 7, and 8). As explained previously, the model predicts this result because isolated *same* pairs are similar only to themselves so tend to retrieve only their own memory traces. By contrast, dense *same* pairs are similar to various *different* pairs as well, so the random walk marches less consistently toward the *same* boundary.

As a source of comparison, we also fitted the on-line-only version of the model to the data. This model uses the same parameters as the full model without λ and c^*, so that it has 9 free parameters. The best fitting parameters are shown in Table 4, and the summary fits are reported in Table 3. The fits of the model to the RT data are shown in Figure 5. The percentage of variance accounted for was 84, 83, and 68 for participants 1–3, respectively. The key shortcoming of this submodel is that it does not predict the variability in the *same* reaction times.

Discussion

In summary, the results from Experiment 1 corroborate past findings regarding *same–different* judgments. As expected, mean RTs for *different* judgments got faster as the similarity between members of a pair decreased. More interesting to the theme of this investigation, identical pairs of objects located in isolated regions of multidimensional similarity space tended to have faster *same* RTs than objects located in dense regions. Most important, our proposed EBRW model did a reasonably good job of quantitatively predicting the complete matrix of mean *same–different* RTs. The results therefore provide support for the idea that retrieval of exemplar memories may be an important component of perceptual *same–different* judgment.

Nevertheless, an important limitation of Experiment 1 is that it was inherently correlational in nature. Thus, strong conclusions cannot be reached that degree of isolation per se was a causal factor in influencing the *same* RTs. For example, it is possible to argue that stimulus-specific factors such as the salience of an individual color may have influenced both its degree of isolation in the MDS solution as well as the speed with which observers were able to make their *same* judgments for that color. Likewise, it is possible to argue that participants realized that certain pairs of colors led to longer RTs, and this knowledge influenced their similarity judgments. One purpose of Experiment 2, therefore, was to explicitly manipulate the degree of isolation of individual objects as an experimental variable instead of relying on the correlational method from Experiment 1. A second purpose was to rule out some alternative accounts of why degree of isolation might influence *same–different* RTs.

Experiment 2a: Explicit Manipulation of Isolation

The design of Experiment 2a is illustrated schematically by the insets in Figure 6. As illustrated in the figure, four colors were selected such that they created two dense regions in the color space: Colors 1 and 2 were highly similar to one another, as were Colors 3 and 4; however, the two groups were highly distinct. The

[3] The model fit referred to here for Participant 3 is labeled *Solution 1* in the tables. The results from Solution 1 are graphed in Figure 4. Unfortunately, this solution does a poor job of accounting for Participant 3's accuracy data. However, with different parameters the model can provide a better fit to these data. The parameters that provide a good fit to the accuracy data, however, cause the model to predict no variation in the RTs of Participant 3's *same* judgments. This model fit is referred to as *Solution 2* in the tables.

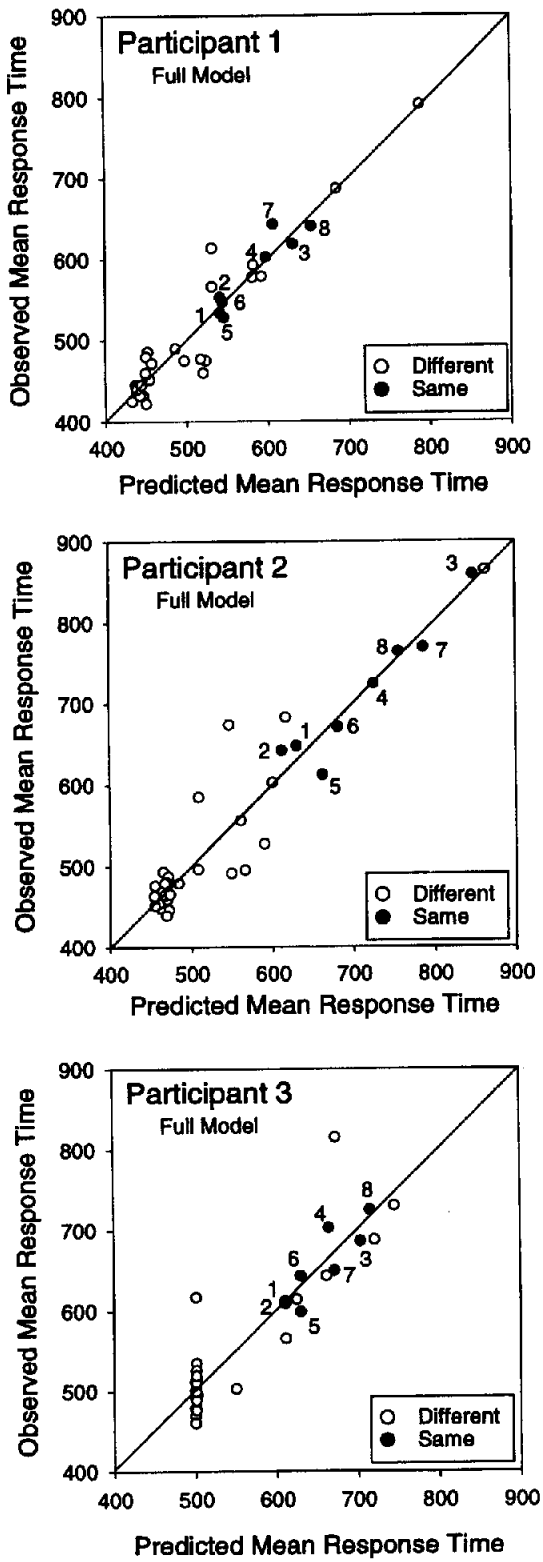

Figure 4. Observed mean response times (ms) for each color pair of Experiment 1 plotted against the predicted mean response times (ms) from the full version of the exemplar-based random-walk model. The numbers that appear along the diagonal are used to label the individual *same* colored pairs.

stimuli were color pairs shown on a computer screen. The task was a speeded *same–different* judgment.

In any condition, participants saw only three of the four colors. Half of the participants did not see Color 4, and the other half did not see Color 1. Thus, for half of the participants, Colors 1 and 2 were in a dense region and Color 3 was isolated. For the other half, Colors 3 and 4 were in a dense region and Color 2 was isolated. It is important to note that two of the stimuli, Colors 2 and 3, were common to all participants. Thus, if the *same* RTs for the isolated color are systematically faster than for the dense color across conditions, it cannot be due to stimulus-specific properties and must be due to the relative isolation of the colors.

A second manipulation involved the *same–different* instructions. Half of the participants received *same-stimulus* instructions. These participants responded *same* only when the two colors were identical, otherwise they responded *different*. The other half of the participants received *same-category* instructions. These observers were instructed to respond *same* for any identical pair or for any pair from the dense region. For example, in Condition 4, *same* responses would be given for pairs (2, 2), (3, 3), (4, 4), and (3, 4).

The purpose of testing the same-category instructions conditions was to evaluate some alternative accounts of effects of isolation on *same–different* judgments. In considering the predictions of a pure on-line perceptual similarity process, we have assumed that all stimuli are judged equally similar to themselves so that *same* RTs for *same* pairs should not differ. Various models have been proposed in the literature, however, that allow for differences in self-similarity among objects (e.g., Ennis et al., 1988; Krumhansl, 1978; Tversky, 1977). For example, according to Krumhansl's (1978) distance-density model of similarity, objects in isolated regions of a psychological space have greater self-similarity than do objects in dense regions of the space. Likewise, in Ennis et al.'s (1988) stochastic theory of similarity, objects give rise to distributions of points in perceptual space. In applying this model to account for a set of *same–different* data reported by Rothkopf (1957), Ennis et al. posited that stimuli in isolated regions of space give rise to less variable perceptual distributions than do stimuli in dense regions. This reduced variability leads to greater self-similarity, which could also account for faster *same* responses associated with isolated stimuli.

The same-category instructions conditions were used to distinguish between the explanations provided by such models and the exemplar-retrieval process in the EBRW. Note that stimulus conditions are identical across the same-stimulus and same-category instructions. Therefore, if self-similarity is greater for isolated stimuli in the same-stimulus instructions conditions, it should also be greater for these stimuli in the same-category instructions conditions. Thus, models that explain the fast *same* responses solely in terms of differences in self-similarity predict faster *same* responses for isolated stimuli in both the same-stimulus and same-category conditions.

By contrast, the EBRW no longer predicts faster *same* responses for isolated stimuli in the same-category instructions condition. Note that in the same-category condition, when a *same* pair from a dense region is presented, there are no longer any *different* exemplar pairs to which it is similar. For example, in Condition 4, pair (3, 4) is now a *same* pair. Thus, suppose that the *same* pair (3, 3) from the dense region is presented. Note that pair (3, 3) is highly

Table 3
Fit Summaries for the Full Extended EBRW and On-Line Process Only as Measured by Percentage of Variance Accounted for (PVAF) and Sum of Squared Deviation (SSD) for Each Participant in Experiment 1

					Participant 3			
	Participant 1		Participant 2		Solution 1		Solution 2	
Statistic	PVAF	SSD	PVAF	SSD	PVAF	SSD	PVAF	SSD
				Full model				
Reaction time	92	22,122	91	47,762	83	51,375	76	72,143
Accuracy	63	170	35	236	−12	3,160	25	184
Weighted SSD		23,821		50,126		82,977		73,980
				On-line model				
Reaction time	84	43,394	83	89,619	68	93,829		
Accuracy	62	174	66	123	0	325		
Weighted SSD		45,136		90,847		97,078		

Note. The weighted SSD weights the accuracy data by a factor of 10. EBRW = exemplar-based random walk.

similar only to pairs (3, 3), (4, 4), and (3, 4), all of which receive *same* feedback during training. The random walk will therefore march consistently toward the *same* criterion when dense *same* pairs are presented. In fact, because *same* pairs are more numerous in dense regions than in isolated regions in the same-category condition, it turns out that the EBRW predicts faster *same* responses for dense pairs than for isolated pairs in these conditions.

Method

Participants. Thirty-nine Indiana University undergraduate students participated in this experiment for course credit. They also received a bonus based on both the speed and accuracy of their responses. All participants claimed to have normal color vision. There were 10 participants in Conditions 1, 2, and 3, and 9 in Condition 4.

Stimuli. The stimuli were four colors. The RGB values of these colors are listed in the first four rows of Table 5. Each color occupied a 5.08 cm × 5.08 cm square on a white background. Colors 1 and 2 were brownish-grey, with Color 2 slightly darker than Color 1. Colors 3 and 4 were reddish-orange. Color 3 was slightly darker than Color 4. The colors were chosen such that Color 1 was very similar to Color 2 and Colors 3 and 4 were highly similar, but colors from one group were not similar to colors from the other group. Two colors were displayed side by side 1.90 cm apart. The hardware used was the same as in

Table 4
Best-Fitting Parameters for the Full Extended EBRW and On-Line Process Only for Each Participant When Simultaneously Fit to the Response Time and Accuracy Data of Experiment 1

	On-line model			Full model			
						P3	
Parameter	P1	P2	P3	P1	P2	Sol'n 1	Sol'n 2
c	1.07	2.17	1.67	0.12	1.11	5.00	12.50
c^*				3.45	7.61	5.33	156.55
w_1	0.32	0.50	0.07	0.26	0.22	0.28	0.00
w_2	0.41	0.10	0.06	0.37	0.04	0.04	0.16
w_3	0.27	0.04	0.87	0.36	0.74	0.68	0.84
β	0.91	0.91	0.95	1.05	1.20	0.36	0.00
k	1.00	1.00	5.18	2.70	32.34	1.51	13.77
μ	393.51	429.00	412.84	351.25	382.01	373.41	119.92
S	10.46	12.01	8.46	25.62	8.09	5.00	1.42
D	47.40	26.00	38.99	40.43	210.91	43.00	1.00
λ				94.54	98.17	16.84	1.40
α	0.89	1.39	0.26	0.18	0.00	20.99	27.19

Note. P1–P3 = participants 1–3, respectively; EBRW = exemplar-based random walk; c = individual item sensitivity parameter; c^* = item pair sensitivity parameter; w_1–w_3 = attention weights on dimensions 1–3, respectively; β = on-line scaling parameter; k = reaction time scaling constant; μ = reaction time intercept parameter; S = same response criterion; D = different response criterion; λ = on-line rate parameter; α = step-time constant; sol'n = solution.

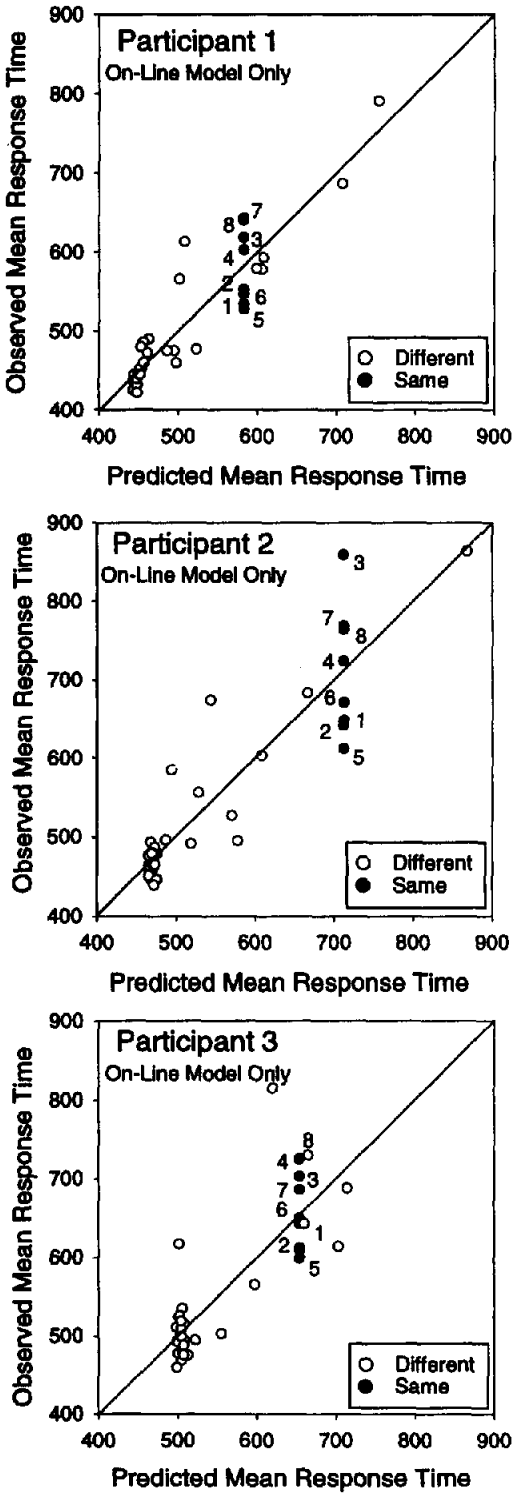

Figure 5. Observed mean response times (ms) for each color pair of Experiment 1 plotted against the predicted mean response times (ms) from the on-line process only version of the exemplar-based random-walk model. The numbers that appear along the diagnoal are used to label the individual *same* colored pairs.

Experiment 1. However, both the brightness and contrast of the monitor were set to 80.

Procedure. A graphical summary of both the color space and classification structure is illustrated by the insets in Figure 6. As explained previously, the conditions differed in which region was dense and which was isolated and on whether same-stimulus or same-category instructions and feedback were provided.

The experiment consisted of 600 trials. Each trial began with a 100-ms white screen. This was followed by a cue in the center of the screen. After 750 ms, the cue disappeared and one color appeared on either side of the cue location. The participant then responded *same* or *different* via a button press. If the response was correct, the next trial was displayed. If the response was not correct, the participant received 2 s of feedback explaining why the response was incorrect. Participants were urged to respond as quickly and accurately as possible. They received a short break after 300 trials.

In each condition, there were six possible color pairs. On each trial, the probability of viewing a particular color pair was ⅙. Thus, on each trial, the probability that a participant viewed an identical (i.e., exactly the same) pair of colors was ½. Likewise there was a ½ chance of viewing a nonidentical pair. Note that stimulus conditions were identical across Conditions 1 and 3 and across Conditions 2 and 4 with only the instructions and feedback being varied. Because of the changed nature of the feedback, however, *same* trials occurred with probability .50 in the same-stimulus conditions but with probability .67 in the same-category conditions. The order of trials, left–right placement, and choice of color pair were randomized for each participant.

Results

In all analyses, RTs below 150 ms and above 4 s were excluded, as were the first 20 trials for each participant and the first five trials after the midway break. Accuracy was near ceiling, so we concentrate on the RT data.[4]

In a preliminary analysis, we considered the mean *different* RTs. In Condition 1, the mean RT for dense pair (1, 2) was 914 ms, whereas the mean RTs for isolated pairs (1, 3) and (2, 3) were 707 ms and 701 ms, respectively. In Condition 2, the mean RT for dense pair (3, 4) was 701 ms, whereas the mean RTs for isolated pairs (2, 3) and (2, 4) were 574 ms and 583 ms, respectively. The much larger *different* RTs for the dense pairs confirms that the structure of the similarity space meets the goals of the experimental design. (For completeness, we report as well that in Conditions 3 and 4 the mean *different* RTs were 613 ms and 659 ms, respectively.)

Recall that the main goal of this experiment was to explicitly test the effect of isolation on *same* trials. The mean RT data for *same* trials are illustrated in Figure 6. It is clear from the figure that in the same-stimulus conditions, Conditions 1 and 2, *same* responses were fastest for the isolated color. This observation is confirmed by a statistical test: Color 3 was significantly faster than

[4] In Condition 1, the mean accuracies for pairs (1, 1), (2, 2), (3, 3), (1, 2), (1, 3), and (2, 3) were 0.99, 0.99, 0.98, 0.98, 0.99, and 1.00, respectively. In Condition 3, the mean accuracies for these pairs were 1.00, 1.00, 0.97, 1.00, 0.98, and 0.98. In Condition 2, the mean accuracies for pairs (2, 2), (3, 3), (4, 4), (2, 3), (2, 4), and (3, 4) were 0.96, 0.97, 0.98, 0.99, 0.99, and 0.95, respectively. The mean accuracies for these pairs in Condition 4 were 0.97, 1.00, 1.00, 0.98, 0.98, and 0.99.

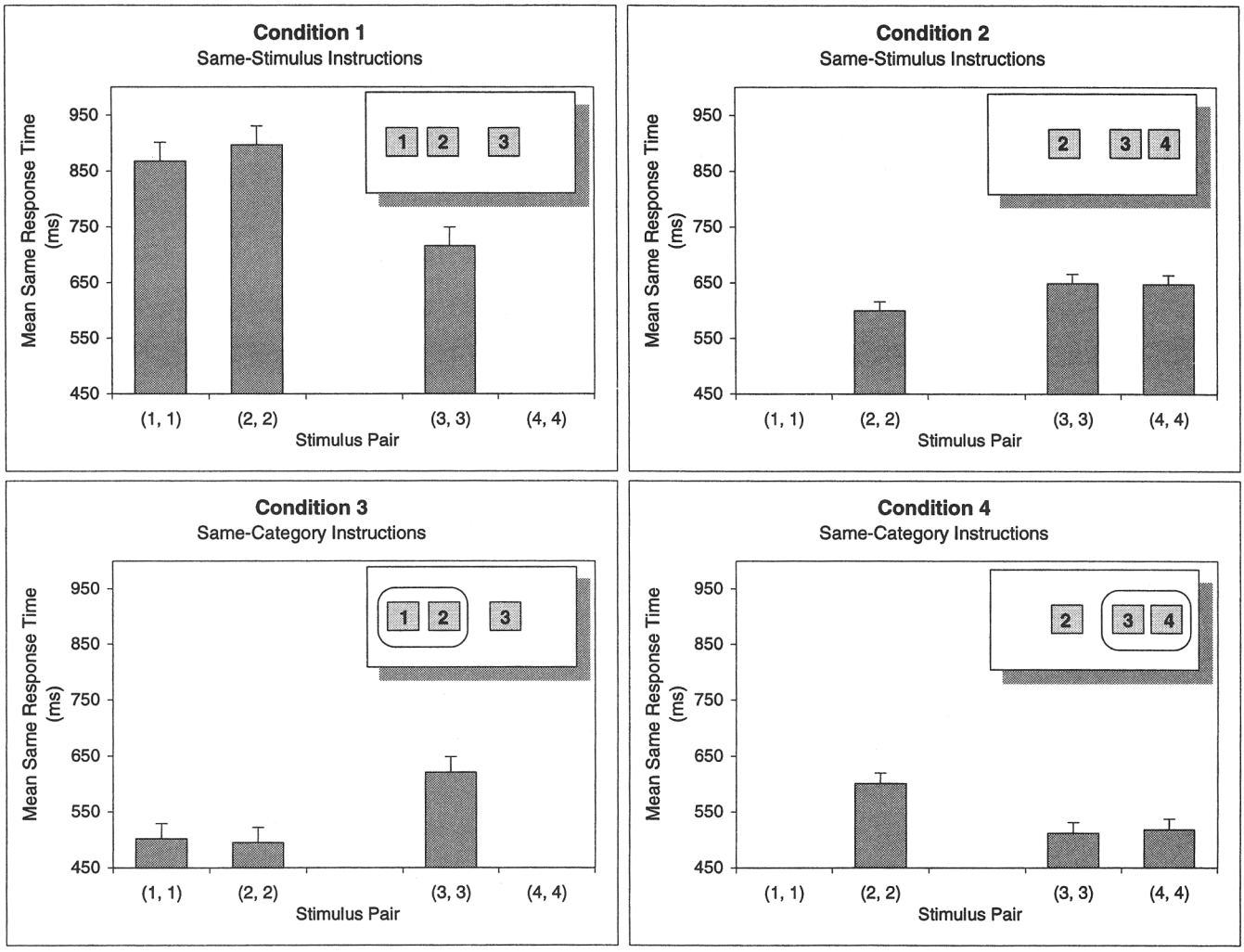

Figure 6. Mean *same* response times for identical pairs of stimuli, Experiment 2a. The error bars reflect the standard error of the mean of the difference scores for the tests reported in the text. The insets give a schematic illustration of the color space and classification structure of Experiment 2a. The more similar two colors are, the closer they appear in the space. In the same-stimulus instruction conditions participants responded *same* only to identical pairs of colors. However, in the same-category conditions, participants responded *same* to any identical pair or any pair from the circled regions.

Color 2 in Condition 1, $t(9) = 5.373$, $p < .001$, whereas Color 2 was faster in Condition 2, $t(9) = 2.994$, $p < .05$. Most important, however, the opposite pattern of RTs was observed in the same-category conditions. Color 3 was significantly slower than Color 2 in Condition 3, $t(9) = 4.597$, $p < .01$, whereas Color 2 was slower in Condition 4, $t(8) = 4.691$, $p < .01$.

Discussion

Same-stimulus conditions. The results from the same-stimulus conditions provide further evidence that the isolation of a stimulus affects its *same* RT. In these conditions, the isolated stimuli had consistently faster *same* RTs than did the dense stimuli. Because Colors 2 and 3 were held constant across these two conditions and only their relative density was manipulated, we remove the pos-

Table 5
Red (R), Green (G), and Blue (B) Values and Approximate Color for Each Stimulus in Experiments 2 and 3

Stimulus	R	G	B	Color
1	102	68	61	Light brown-gray
2	87	53	47	Dark brown-gray
3	176	60	14	Dark orange
4	191	75	28	Light orange
5	0	102	0	Green
6	0	0	200	Blue
7	255	255	0	Yellow
8	252	0	255	Violet

Note. Experiment 2a uses colors 1–4; Experiments 2b, 3a, and 3b use all eight colors.

sibility that the fast *same* responses may be reflecting stimulus-specific factors.

Crist (1981) performed a similar manipulation to obtain experimental evidence for the role of isolation on *same* RTs, although it relied on a cross-experimental comparison and on conditions in which name-matches on letters were being performed rather than physical matches. Specifically, in Crist's (1981) Experiment 4 he found that *same* letter pair (*e, e*) was responded to more slowly than pairs (*E, E*), (*F, F*), and (*f, f*) in a situation in which the letters *o, O, a,* and *A* were also in the stimulus set. By comparison, in Experiment 5, Crist (1981) deleted the letters *o, O, a,* and *A* from the set, and found that pair (*e, e*) was now responded to more rapidly than pairs (*E, E*), (*F, F*), and (*f, f*). Because the letters *o* and *a* were quite similar to the letter *e* and highly dissimilar from the letters *E, F,* and *f,* this crossover in mean RTs was consistent with the idea that the similarity structure of the stimulus set had exerted an important impact on performance. The results from the present same-stimulus conditions are analogous to those obtained by Crist (1981), although in our study the size of the stimulus set was equated and alternative objects were explicitly located in dense versus isolated regions across conditions.

Same-category conditions. The more important contribution of Experiment 2a comes from the comparison of results from the same-category conditions to those in the same-stimulus conditions. Whereas in the same-stimulus conditions, the isolated color had faster *same* RTs than did the dense colors; in the same-category conditions, the reverse result was obtained. As explained previously, models such as those of Krumhansl (1978) and Ennis et al. (1988), which explain the fast *same* RTs in terms of the greater self-similarity of the isolated color, fail to predict this observed crossover. This switch in RT order, however, is consistent with the predictions from the EBRW. We emphasize that our results do not rule out the valuable ideas that the self-similarity of objects may differ or that objects give rise to variable internal representations across trials. Rather, the important point is that these ideas, in and of themselves, are insufficient to account completely for the full range of our *same–different* data.

Alternative accounts of the isolation effects. Regarding Ennis et al.'s (1988) model, one possibility to consider is that the degree of perceptual variance associated with an object depends not only on the prevailing stimulus conditions but on the task goals as well. For example, an observer might try to place greater attention in that region of the perceptual space that requires more fine-grained discriminations. In the same-stimulus conditions, greater attention would need to be placed in the dense region, whereas in the same-category conditions, this need for greater attention in the dense region does not arise (because fine-grained discriminations are no longer required in this region). Note, however, that this attentional hypothesis predicts the wrong pattern of *same* RTs in the same-stimulus conditions. According to most of the theories of attentional focus that address the issue of perceptual variability, placing attention in a given region results in reduced perceptual variability in that region (e.g., Ashby, Prinzmetal, Ivry, & Maddox, 1996; Luce, Green, & Weber, 1976; Tsal & Shalev, 1996). Such an attentional process would be adaptive for better discriminating among the highly similar objects that occur in the dense regions, for example, pair (1, 2) in Condition 1. If attention were placed in the dense region, however, then isolated Color 3 would have greater perceptual variability than would dense Colors 1 and 2. By contrast, Ennis et al.'s (1988) stochastic theory explains the fast and accurate responding for isolated colors in terms of their reduced perceptual variability (which results in greater self-similarity). Thus, it is unclear to us how to explain the complete set of results in terms of adaptive shifts of attention across the same-stimulus and same-category conditions. Nevertheless, we acknowledge that alternative models that allow for other forms of stimulus-specific perceptual adjustments that vary with task goals could conceivably also account for these results.

Finally, recall that the overall proportion of pairs defined as *same* was greater in the same-category condition compared with the same-stimulus condition in this experiment. This factor would be expected to lead to a reduced criterion for responding *same* and therefore to a general speeding of *same* RTs in the same-category condition (Ratcliff & Hacker, 1981). However, such a process does not explain the crossover in mean RTs for isolated and dense *same* pairs across the same-stimulus and same-category conditions (because all of these pairs were assigned to the *same* response category). Rather, some type of exemplar-specific or region-specific mechanism appears to be necessary for explaining the observed crossover.

Experiment 2b: Expanding the Stimulus Space

In Experiment 2a participants were exposed to only six color pairs. Perhaps memory for exemplar pairs plays a role in *same–different* judgments only when the number of exemplar pairs is small (cf. Macmillan & Creelman, 1991, p. 148). Experiment 2b was conducted as a follow-up to Experiment 2a by increasing the number of colors in the set.

The design of Experiment 2b was the same as in the same-stimulus conditions of Experiment 2a, except that four extra colors were added to the original four. These colors were chosen to be very distinctive from each other and the original four colors. As before, either Color 1 or Color 4 was left out of a participant's color set. Thus, there was still one dense region per set, and there were two colors that alternately occupied dense and isolated regions, depending on the experimental condition. Note that with seven colors there are 28 color pairs, thereby providing a more challenging test of the exemplar retrieval hypothesis. Participants responded *same* only for identical pairs.

Method

Participants. Twenty-four Indiana University undergraduate and graduate students participated in this experiment, 12 in each condition. They received $5 and a bonus based on both the accuracy and speed of their responses. All participants claimed to have normal color vision and were unaware of the issues under examination in this study.

Stimuli. The RGB values of the stimuli are shown in Table 5. The first four colors are identical to the colors used in Experiment 2a. Four additional colors were included that were highly distinctive both from the four original colors and from each other. Colors 5–8 were green, blue, yellow, and violet, respectively. Each color occupied a 5.08 cm × 5.08 cm square on a black background. Color pairs were shown next to each other separated by 1.90 cm. The hardware and monitor settings were identical to those in Experiment 2a.

Procedure. This experiment had two conditions. Participants in Condition 1 saw pairs of colors involving all colors except Color 4. In

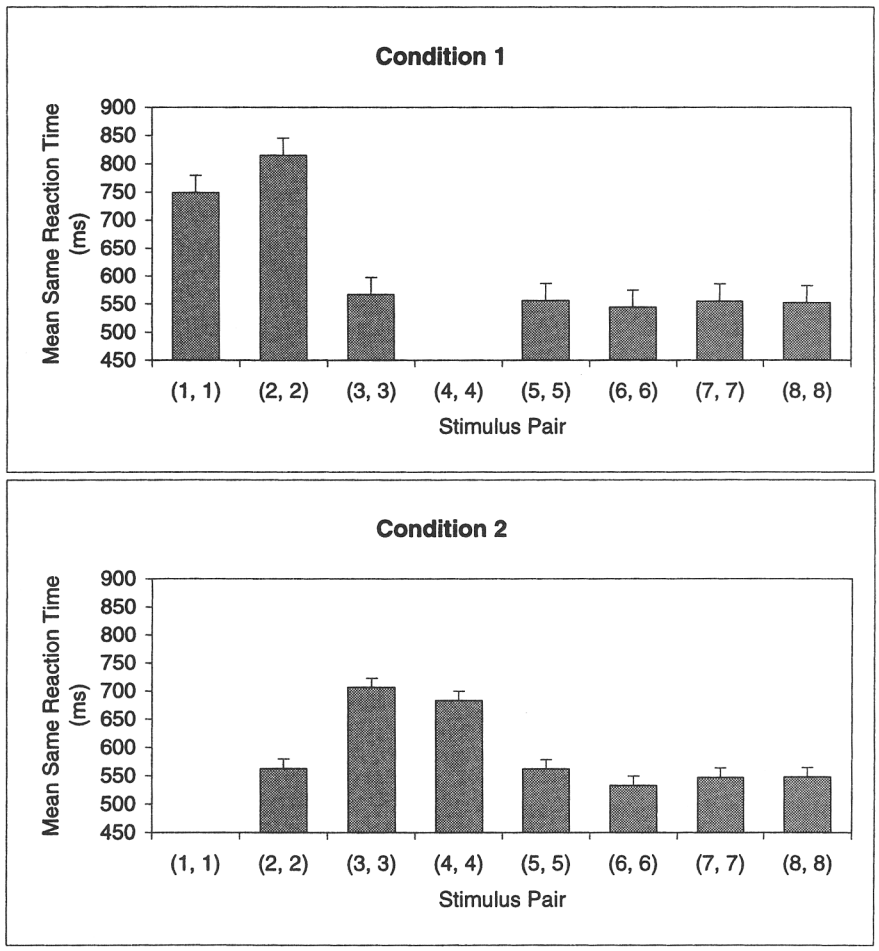

Figure 7. Mean *same* response times for identical pairs of stimuli, Experiment 2b. The error bars reflect the standard error of the mean of the difference scores for the tests reported in the text.

Condition 2, participants viewed pairs involving all colors except Color 1. In both conditions, participants were asked to respond *same* only to identical colors.

The experiment consisted of 600 trials. Each trial proceeded as in Experiment 2a. Participants were asked to respond as quickly as possible without sacrificing accuracy. They received a short rest after 300 trials. On ¼ of the trials, *same* color pairs were chosen from the appropriate set of Colors 1–4, for example, pair (2, 2). On ¼ of the trials, *same* color pairs were chosen from Colors 5–8, for example, pair (5, 5). On ¼ of the trials, one color was chosen from Colors 5–8, and the second was chosen from Colors 1–8, for example, pairs (5, 1), (5, 5), or (5, 8). 1/12 of the trials were the dense pair, that is, pair (1, 2) in Condition 1. On 1/12 of the trials, any pair that included the isolated color as one of the members was presented, for instance, this includes pairs (3, 1), (3, 3), and (3, 5) in Condition 1. The remaining 1/12 of the trials included any pair involving only one of the dense colors. For instance, in Condition 1, this case covered color pairs (1, 1), (1, 3), and (1, 5).[5] Note that there is some overlap between these choices.

Results

The first 20 trials, the first 5 trials after the midway break, and RTs less than 150 ms and greater than 4 s were excluded from analysis. Only *same* trials are discussed. Because accuracies were near ceiling, we discuss only RTs.[6] The results, illustrated in Figure 7, mirror those of Experiment 2a. The relevant comparison is between the manipulated isolated and dense stimuli. As expected, on *same* trials, the isolated stimuli were faster than the dense stimuli. In Condition 1, the *same* RT for Color 3 was faster than for Color 2, $t(11) = 8.149, p < .001$, whereas in Condition 2, Color 2 was faster than Color 3, $t(11) = 8.502, p < .001$.

Discussion

The results of this experiment replicated and extended those of the same-stimulus conditions of Experiment 2a. The dense colors had slower *same* RTs than any of the other colors, including the

[5] Unfortunately, this arrangement sometimes allows the excluded stimulus to be displayed in a *different* pair. Note that this should only attenuate the expected effect.

[6] Averaged across conditions, the *same* mean proportions correct for the isolated stimuli, the dense stimuli, and the filler stimuli were 0.97, 0.94, and 0.98, respectively.

critical isolated color. Furthermore, this result was obtained in a situation involving 28 distinct exemplar pairs, thereby demonstrating some generality for the role of an exemplar-retrieval process in *same–different* judgments.

Experiment 3a: Manipulation of Training in Dense Regions

In Experiment 3a we provided further tests of the role of exemplar retrieval on *same–different* judgments by manipulating whether observers had experience with *same* or *different* pairs in specific regions of the similarity space.

The task was a speeded *same–different* task similar to the previous experiment. Participants viewed pairs of colors created from the same eight colors of Experiment 2b. Recall that there were two dense regions (Colors 1 and 2 and Colors 3 and 4) and four isolated regions (Colors 5, 6, 7, and 8). During a training phase, participants were exposed to all possible color pairs, except they saw only *same* pairs from one of the dense regions and only *different* pairs from the other dense region. Specifically, in Condition 1 participants did not see color pairs (1, 2), (3, 3), or (4, 4). So, these participants saw only *same* pairs from the dense region consisting of Colors 1 and 2 and only *different* pairs from the dense region consisting of Colors 3 and 4. In Condition 2, participants were not exposed to pairs (1, 1), (2, 2), or (3, 4). So, these participants saw only *different* pairs from the Colors 1 and 2 dense region and only *same* pairs from the Colors 3 and 4 dense region. Following the training phase, participants were tested in a transfer phase that included the color pairs that were not seen during training. For ease of discussion, we refer to the pairs that had been experienced during training as *familiar* pairs and pairs that had not been experienced as *unfamiliar* pairs.

The key comparison is the relative RTs and accuracies of the familiar and unfamiliar pairs from the dense regions. The EBRW predicts that during transfer, the unfamiliar pairs should have either slower RTs or lower accuracies than the familiar pairs. For example, in Condition 1, pair (1, 2) is an unfamiliar *different* pair, whereas pair (3, 4) is familiar. When pair (1, 2) is presented at time of transfer, the exemplar process will tend to retrieve *same* pairs (1, 1) and (2, 2), but there will be no exemplars of pair (1, 2) stored in memory. Thus, the exemplar-retrieval process will tend to push the random walk incorrectly toward the *same* boundary, leading to slow RTs or low accuracies. By contrast, when familiar *different* pair (3, 4) is presented, there will be numerous exemplars of this pair stored in memory, and no exemplars of the competing *same* pairs (3, 3) and (4, 4). Thus, the random walk will move efficiently toward the correct *different* boundary. The reverse pattern should be observed for *different* pairs (1, 2) and (3, 4) in Condition 2. Analogous predictions arise for the unfamiliar and familiar *same* pairs across Conditions 1 and 2.

Method

Participants. Ninety-one Indiana University undergraduates participated in this study for course credit, 46 in Condition 1 and 45 in Condition 2. They also received a bonus based on the speed and accuracy of their responses. All claimed to have normal color vision.

Stimuli. The stimuli and hardware used were identical to that of Experiment 2b. However, participants saw pairs involving all eight colors.

Procedure. Participants saw color pairs involving all eight colors. They were shown each of the individual colors before beginning the trials. The experiment was broken into two phases, training and transfer, and had two conditions. In Condition 1, participants experienced all possible color pairs, except for the unfamiliar pairs (1, 2), (3, 3), and (4, 4), which they saw only during transfer. In other words, from the dense region of Colors 1 and 2, participants saw only *same* pairs during training and saw *same* and *different* pairs during transfer. From the dense region of Colors 3 and 4, participants viewed only *different* pairs during training and both *same* and *different* pairs during transfer. Analogously, in Condition 2, the unfamiliar pairs were (1, 1), (2, 2), and (3, 4).

The training phase had 336 trials. Participants were asked to respond as accurately and quickly as possible. Each trial began with a 100-ms blank screen. A cue was then displayed for 750 ms. The cue disappeared and a color pair was shown, one color on each side of the cue location. This display was response terminated. The participant responded *same* or *different* with a button press. If the response was correct, the next trial was displayed. Otherwise, 2 s of feedback was given explaining why the response was incorrect. The color pair remained on the screen during feedback.

During the training phase, *same* pairs from the set of Colors 5–8 were presented on 3/8 of the trials. On another 3/8 of the trials, participants were shown any familiar *different* pair.[7] On 1/8 of the trials, the familiar *different* pair from the dense region was shown. Finally, on 1/8 of the trials, the familiar *same* pairs from the dense regions were shown.

Participants were not informed of the transition to the transfer phase. The transfer phase consisted of four blocks of 16 trials each. Each trial proceeded in the same manner as in the training phase. During a transfer block each *same* pair was shown once, pairs (1, 2) and (3, 4) were shown twice each, and the remaining four trials were any familiar *different* pairs. Within each block the trial order and left–right placement of the colors within pairs were randomized.

Results

The analysis focuses on four data points per participant. Each data point involves stimulus pairs presented from the dense regions of the stimulus space. Each data point is classified as familiar or unfamiliar and as *same* or *different*. Thus, familiar *same* pairs are those *same* pairs from a dense region that were seen during training and familiar *different* pairs are those dense *different* pairs that were seen during training. For instance, in Condition 1, pairs (1, 1) and (2, 2) are the familiar *same* pairs, pair (3, 4) is the familiar *different* pair, (1, 2) is the unfamiliar *different* pair, and (3, 3) and (4, 4) are the unfamiliar *same* pairs.

For each participant, we selected the first transfer presentation of each color pair type: familiar *same*, familiar *different*, unfamiliar *same*, and unfamiliar *different*.[8] We then averaged the data from each of the four categories, familiar *same* and *different* and unfamiliar *same* and *different*. The results are shown in Table 6.

[7] Because of a programming error, pair (3, 4) might rarely show up in Condition 2. Note that this event should only lessen the hypothesized effect.

[8] Once an unfamiliar pair is presented during transfer, attention would likely be focused on it because of the high probability that it would result in an error response. Thus, a salient representation of the unfamiliar pair would be formed in memory, so the predictions of the EBRW for that unfamiliar pair would not likely hold after the first block.

Table 6
Mean Response Times (in Milliseconds) and Mean Proportions Correct for All Familiar Same, Familiar Different, Unfamiliar Same, and Unfamiliar Different Stimuli From the First Block of Transfer, Experiment 3a

	Condition 1		Condition 2	
Trial type	*P*(C)	RT	*P*(C)	RT
Familiar same	0.70	954	0.96	976
Unfamiliar same	0.13	1124	0.27	1108
Familiar different	0.72	1008	0.82	840
Unfamiliar different	0.07	1072	0.09	914

Note. Unfamiliar pairs are those pairs not shown during training. The data from the familiar pairs are restricted to familiar pairs from the dense regions. *P*(C) = proportion correct; RT = mean reaction time.

Consistent with the predictions from the EBRW, the RTs for the unfamiliar pairs were slower than for their familiar counterparts. The major effect of the manipulations, however, appears to have been on the accuracy data: Accuracies for the unfamiliar pairs were always dramatically lower than for their familiar counterparts.

Discussion

The results of this experiment provide further evidence that exemplar-retrieval processes may play a role in the concurrent perceptual *same–different* judgment task. During training, participants saw only *same* pairs from one dense region of the color space and only *different* pairs from the other dense region. During transfer, participants performed relatively poorly on those dense color pairs not seen during training. Thus, exposure to specific color pairs during training affected transfer performance in a manner consistent with the predictions of the EBRW.

In this experiment, training was manipulated in only the dense regions of the color space. The results suggest that exemplar-retrieval played a role in these dense regions. However, although there were 33 training pairs shown to participants, only three of these pairs were from the difficult dense regions, and only two of these were similar to each other. It could be argued that people used an exemplar strategy to aid performance on the few difficult pairs but that they used a pure on-line strategy for the remaining 30 color pairs. Experiment 3b was designed to explore this possibility. As explained in more detail below, another purpose of Experiment 3b was to test the prediction from the EBRW that *same* responding can be facilitated by familiarity with the specific *same* pair being judged, while holding fixed the degree of interference from competing *different* pairs.

Experiment 3b: Manipulation of Training in Isolated Regions

Experiment 3b was designed to extend the results of Experiment 3a by discounting the possibility that exemplar retrieval occurs only in the difficult, dense regions of the color-pair space. The task was the *same–different* task of Experiment 3a, and participants viewed pairs constructed from the same eight colors. There were four conditions. During the training phase of Condition 1, participants were shown all color pairs except pair (5, 5). That is, in this condition, participants were not exposed to one of the isolated *same* pairs. Likewise, participants in Conditions 2, 3, and 4 did not see pairs (6, 6), (7, 7), and (8, 8), respectively. In a transfer phase, participants saw all color pairs, including the previously unseen pairs.

The EBRW again predicts that, at transfer, the unfamiliar *same* pairs will have slower RTs and/or lower accuracies than the familiar pairs. There is an important difference, however, regarding the basis for these predictions in Experiments 3a and 3b. In Experiment 3a (as well as in Experiments 1 and 2), our interpretation was that a major basis for slowed *same* RTs was that there were highly similar *different* exemplar pairs that were stored in memory and that interfered with an observer's ability to respond *same*. For example, in Experiments 1 and 2, we noted that *same* pairs from dense regions were highly similar to numerous *different* pairs, whereas isolated *same* pairs were highly similar only to themselves. Likewise, our interpretation was that a major contributing factor for the slowed RTs associated with unfamiliar *same* pairs in Experiment 3a was that participants had been trained on highly similar *different* pairs from this region, which caused major interference.

In Experiment 3b, however, interference from competing *different* pairs cannot differentially affect performance on the familiar versus unfamiliar *same* pairs. The reason is that isolated stimuli 5, 6, 7, and 8 all occur the same proportion of times in the various *different* pairs that are presented. Thus, any interfering effect from *different* pairs that contain these stimuli as members is held constant. It is also worth pointing out that because the unfamiliar *same* pairs in this experiment are all isolated pairs, any interfering effect from *different* pairs should be relatively weak (i.e., none of the *different* pairs stored in memory is very similar to the isolated *same* pairs).

The key reason that the EBRW predicts somewhat slowed RTs for the unfamiliar *same* pairs in this experiment is that there is a lack of *facilitation* associated with such pairs. Participants have experienced the familiar *same* pairs during training, so when a familiar *same* pair is presented at time of transfer, the exemplar process in the EBRW may retrieve the pair, thereby increasing the probability of the random walk taking steps toward the *same* response criterion. By contrast, when an unfamiliar pair is presented at time of transfer, there are no exemplar memories of this *same* pair to be retrieved, so the exemplar-based facilitation is unavailable. Note that the *individual* isolated colors have all been presented with regularity during the training phase, so participants are not experiencing a "surprise effect" involving a completely novel color. Rather, it is the *pairing* of the unfamiliar color in a *same* trial that is novel.

Finally, we emphasize that although we predict slowed RTs and/or lower accuracies of the unfamiliar pairs compared with the familiar ones in this experiment (because of a lack of facilitation), the effects should be much weaker than we observed in Experiment 3a. The reason is that there are no (highly similar) interfering

different pairs in Experiment 3b, and any such interference that exists is equated across stimuli.

Method

Participants. Seventy-two Indiana University undergraduates participated in this study for course credit, 18 in each condition. They received a bonus based on the speed and accuracy of their responses.

Stimuli. The stimuli were identical to those used in Experiment 3a. Three of the computers used in Experiment 3b were identical to those of Experiment 3a. Two additional hardware configurations were also used, a CompuAdd 325s CPU with a CompuAdd 51109 monitor and a Dell Dimension 433v CPU with a CompuAdd 51109 monitor.

Procedure. The procedure was similar to that of Experiment 3a. Participants saw color pairs involving all eight colors. The experiment was broken into training and transfer phases. There were four conditions. In Condition 1, participants saw all possible color pairs during training except pair (5, 5). Similarly, participants in Conditions 2, 3, and 4 saw all possible pairs during training except pairs (6, 6), (7, 7), and (8, 8), respectively. That is, each participant saw all possible pairs except one of the *same* pairs from an isolated region. During transfer, participants were exposed to all possible color pairs, including the unfamiliar pair from the training phase.

The training phase had 315 trials. Half of the trials were *same* trials and half were *different* trials. Each familiar *same* pair had an equal chance of being displayed during a *same* trial. During a *different* trial, each *different* pair had an equal probability of being shown. The trials were randomized, as were the left–right placements of the colors.

Participants were not explicitly informed of the transition to the transfer phase. The transfer phase consisted of three blocks of 43 trials each (129 total trials). During each block, each *different* pair was shown once, each familiar *same* pair was shown twice, and the unfamiliar *same* pair was shown once. Within each block, trial order and left–right placement were randomized. All other aspects of the procedure were the same as in Experiment 3a.

Results

Our analyses focused on the data from the first transfer block. For each participant we compared performance on the single presentation of the unfamiliar pair with the average of the two transfer presentations of each of the familiar, isolated *same* pairs. The results are shown in Table 7. By inspection, it is clear that, during the first block of transfer, a color pair had either a slower RT or a lower accuracy when it was not seen during training. Participants were slowest to respond to pairs (5, 5), (7, 7), and (8, 8) in precisely those conditions in which these pairs were not shown during training (Conditions 1, 3, and 4, respectively). Likewise, accuracies for pairs (5, 5), (6, 6), (7, 7), and (8, 8) were lowest in precisely those conditions in which each pair was not seen during training. Although, in most cases, the results were not statistically significant when each condition was considered in isolation, the results are compelling when the entire pattern is considered together.

As a simple test for the overall effect of the manipulation, we first standardized the RT and accuracy scores for each stimulus. The scores were standardized using the means and standard deviations for each isolated *same* pair pooled across participants and conditions. The main reason for using this standardization was that, regardless of condition, RTs associated with pair (5, 5) were larger than for the other color pairs. For each participant, we then computed a difference score between the mean RT for the unfamiliar *same* pair and the mean RT for the familiar *same* pairs. Analogous difference scores were computed for the accuracy data. A two-tailed t test indicated that these scores differed significantly from zero: for RT, $t(71) = 2.739$, $p < .01$, and for accuracy, $t(71) = 3.002$, $p < .01$. Thus, there is supporting statistical evidence that the unfamiliar *same* pairs were responded to more slowly and less accurately than the familiar *same* pairs.

Discussion

Through manipulation of training, Experiment 3b provided further evidence for the role of memory in perceptual *same–different* judgments. RTs were slower and accuracies were lower for those *same* color pairs not seen during training. Because the unfamiliar pairs resided in an isolated region of the color-pair space, the results of this experiment discount the possibility that participants in Experiment 3a were using an exemplar-retrieval strategy only in the more difficult, dense regions. In addition, the results of the experiment confirm the predictions from the EBRW that *same* judgments can be facilitated by familiarity with the particular *same* pairs being judged. Slowed *same* RTs do not arise solely from differential interference with competing *different* pairs.

Krueger (1973) reported a similar experiment (using letters as stimuli) in which the frequency with which individual *same* and *different* pairs were experienced during training was manipulated. Although he observed a frequency effect on the *different* judgments, with the more familiar *different* pairs being judged more quickly than the unfamiliar ones, the results for the *same* judg-

Table 7
Mean Reaction Time (in Milliseconds) and Mean Percentage Correct for Each Isolated Same Stimulus Pair in Each Condition, Experiment 3b

	Condition 1: No (5, 5)		Condition 2: No (6, 6)		Condition 3: No (7, 7)		Condition 4: No (8, 8)	
Stimulus pair	*P*(C)	RT	*P*(C)	RT	*P*(C)	RT	*P*(C)	RT
(5, 5)	**0.72**	**1,047**	0.97	728	0.94	981	0.97	784
(6, 6)	0.92	596	**0.83**	**631**	1.00	629	0.92	629
(7, 7)	1.00	593	0.97	533	**0.89**	**657**	1.00	575
(8, 8)	0.97	633	0.97	604	0.89	616	**0.83**	**710**

Note. Cells in boldface mark the unfamiliar *same* pairs in each condition. *P*(C) = proportion correct; RT = mean reaction time; No = pair was not presented in condition.

ments were not statistically significant. There were some important differences, however, between our experimental procedure and the one used by Krueger (1973). First, in our study, the unfamiliar *same* pairs were completely novel at time of test, whereas in Krueger's study they had been presented with one fourth the probability of the familiar *same* pairs. Thus, our study involved a more extreme experimental manipulation of stimulus frequency. Second, in Krueger's (1973) design, when the frequency of particular *same* pairs was increased, the frequency of highly similar *different* pairs was simultaneously increased. According to the EBRW, the increased frequency of the competing *different* pairs should tend to offset the benefit of experiencing a particular *same* pair more frequently, which may explain the lack of a significant frequency effect on *same* judgments in Krueger's (1973) design.

General Discussion

Summary

In this article, we presented evidence that memory for stored exemplars may play a role in the perceptual *same–different* task. Also, we proposed a hybrid exemplar-based model, an extended version of the EBRW (Nosofsky & Palmeri, 1997), to formalize these ideas.

In the extended version of the EBRW, two processes race to add incremental evidence into a random walk. The on-line process is based solely on the perceptual similarity between the objects in the presented pair. The more similar the two objects, the more likely the on-line process will increment the random walk counter in the *same* direction. The second process involves the retrieval of exemplar pairs stored in memory. The probability of moving in either the *same* or *different* direction is based on the similarity of the presented pair of objects to all *same* and *different* pairs stored in memory.

As noted in our introduction, our extended EBRW model builds on some earlier ideas advanced by Crist (1981), who provided evidence that the similarity structure of the set of stimulus possibilities, and not just the pair of objects presented on a given trial, plays an important role in *same–different* judgment. Our contribution goes beyond Crist's work, however, by providing an explicit formalized model that makes precise quantitative predictions of *same–different* RTs and accuracies, that takes simultaneous account of the similarity of a presented pair of objects to each other and of their relation to other exemplars in the set, and that provides a mechanistic account of processing in the *same–different* judgment task.

One key prediction of the extended EBRW is that it should be more difficult to respond *same* to an identical object pair that resides in a dense region of the multidimensional similarity space than to a pair from an isolated region. Because there are no *different* pairs near an isolated *same* pair, the exemplar process will tend to retrieve only the isolated *same* pair itself and thus move the random walk consistently toward the *same* boundary. This consistent movement results in faster *same* RTs and higher *same* accuracies for isolated pairs. By contrast, in a dense region, the exemplar process will retrieve both *same* and *different* pairs and so the random walk will wander before it reaches a boundary, thereby leading to slower *same* RTs and lower *same* accuracies. Importantly, however, if feedback conditions are rearranged so that observers are instructed to respond *same* to similar (same-category) objects, rather than to only identical objects, then the model makes the reverse prediction. In particular, it then predicts that dense *same* pairs will be responded to more efficiently than isolated *same* pairs.

A second class of predictions of the EBRW is that the experience an observer has with *same* and *different* pairs in a specific region of the similarity space should affect later performance. For example, if, in a particular region of the space, only *same* pairs are stored, then, when presented with a *different* pair from that region, the exemplar process will tend to retrieve the stored *same* pairs and so push the random walk toward the *same* boundary, resulting in slow RTs or low accuracies. Likewise, increased familiarity with specific *same* pairs in a given region of the multidimensional perceptual space should facilitate *same* responding on these specific pairs. The results of Experiments 1–3 supported all of these predictions.

A further contribution of this work is that we were able to distinguish the predictions of the extended EBRW from some alternative extant models of *same–different* judgment. Well known models that explain *same–different* responding solely in terms of the degree of match between the members of a presented pair (e.g., Link, 1992; Ratcliff, 1981; Takane & Sergent, 1983) are unable to account for the numerous context effects observed in our studies. Context-sensitive models, such as those of Krumhansl (1978) and Ennis et al. (1988), explain fast isolated *same* responding in terms of the greater self-similarity of isolated colors. However, these models fail to account for the finding that the reverse pattern of results was observed in our conditions that used same-category instructions instead of same-stimulus instructions. Models that posit that attention may be drawn to regions of the similarity space that require fine-grained discriminations also do not seem to provide a natural account of the fast *same* responding observed for isolated objects.

Even Crist's (1981) holistic-discriminability framework, which does account for the role of the similarity structure of the set of stimulus possibilities on *same–different* performance, fails to make precise predictions regarding the effect of various manipulations in our study. For example, from the perspective of the holistic-discriminability framework, it is unclear why in our Experiment 3b, novel *same* pairs were responded to more slowly than familiar ones. According to the holistic-discriminability framework, *same* responding is determined by the time that it takes to locate each individual object in the multidimensional similarity space. However, in our Experiment 3b, all of the individual objects in the manipulated pairs were isolated and presumably equally easy to locate—it was only the pairing of the objects that was familiar or unfamiliar. To account for such results, Crist's framework would need to be extended so as to be sensitive to the specific set of pairs on which observers are trained.

Nickerson (1981) pointed to another limitation of Crist's (1981) holistic-discriminability framework that we believe is overcome by the EBRW. In one condition of Crist's (1981) Experiment 1, the stimulus set consisted of the letters *o*, *a*, and *f*, whereas in a comparison condition, the stimulus set consisted of the letters *o*, *x*, and *f*. In the type font used in this experiment, the letters *o* and *a* were highly similar, whereas *o* and *x* were highly dissimilar. As

expected from the predictions of the holistic-discriminability framework, *same* RTs were much longer for the pair (*o, o*) when the set contained the highly similar *a* than the highly dissimilar *x*. However, as noted by Nickerson (1981, p. 301), it was also the case that *same* RTs were longer for the pair (*f, f*) when the set contained *a* than when it contained *x*. As summarized by Nickerson (1981, p. 301),

> To decide that two fs are the same takes longer when the set contains two similar letters (a and o) than when it contains only one of them . . . . The question is why [this effect] should be obtained at all. Why should it take longer to decide that f-f is *same* just because the set of alternatives includes two letters that are similar to each other, neither of which bears any resemblance to f?

Nickerson (1981, p. 301) went on to argue that Crist's holistic-discriminability framework failed to account for the phenomenon by asking, "Why should the fact that the representational space contains two points that are close together affect the time required to deal with a point that is far away from both of them?"

Nickerson (1981, p. 302) suggested that the results were interpretable in terms of an accumulative process with an adjustable decision criterion in which the criterion that was used applied to all pairs in the set. In particular, in the *hard* condition involving the letters *a, o,* and *f,* participants would need to set a strict criterion in order to collect sufficient evidence to judge correctly difficult-to-discriminate pairs such as (*o, a*). But in the *easy* condition involving *o, x,* and *f,* participants could set a lax criterion because there are no difficult-to-discriminate pairs. As pointed out by Nickerson (1981, p. 302),

> Increasing the strictness of the decision criterion would have a general effect; that is, it would lengthen the time for *any* comparison. In contrast, the rate of evidence accumulation should depend on the particular items being processed and their relationship to other items in the set; evidence of difference would accumulate more slowly, for example, for two similar items than for two that are quite different.

The extended EBRW model that we have proposed in this article indeed incorporates into its formal machinery both factors suggested by Nickerson (1981). The rate of accumulation (probability of taking steps toward either the *same* or *different* boundary) in the random walk is determined jointly by the similarity of the objects in a pair to each other and by their similarity to previously experienced exemplar pairs. In addition, adjustable decision criteria are included in the model to account for general effects of task difficulty that extend to all stimuli in the set.

The alternative modeling approach that seems most likely to be able to account for our results is one that allows separate criteria to be maintained for individual exemplars depending on their location in the multidimensional similarity space, the history of presentation of specific pairs of objects from the space, and the general difficulty of the task. For example, a participant might establish a strict criterion for responding *same* to a dense pair and establish a lax criterion for responding *same* to an isolated pair. At present, it is unclear to us how all of these multiple influences listed above should be coordinated within a single exemplar-specific criterion-shift model. Future research will be needed to more thoroughly develop and investigate such models and to distinguish their predictions from the EBRW. In any case, it is important to recognize that because separate criteria are maintained for individual exemplars, such models share the fundamental principle with the EBRW that some form of exemplar-based memory plays an important role in determining *same–different* performance.

Finally, it is worth pointing out explicitly that in our research we have conceived of the *same–different* judgment task, at least in part, as a type of categorization task. According to this view, there are two categories of objects, the *same* category and the *different* category. The exemplars of these categories now correspond to pairs of objects rather than to single objects, and a modified rule is used for computing similarities among these exemplar pairs. Conceptually, however, the extension of the EBRW to *same–different* judgments operates the same way as the original version of the model operates in standard tasks of categorization.

Abstraction and Same–Different Judgment

It is interesting to compare our research with related research from the animal-learning literature. Wasserman, Hugart, and Kirkpatrick-Steger (1995) trained pigeons to peck one button when presented with 16 distinct 4 × 4 arrays of nonidentical pictures, the *different* arrays, and to peck a second button when presented with 16 distinct 4 × 4 arrays of identical pictures, the *same* arrays. The pigeons were later tested on 16 novel *same* and 16 novel *different* arrays. Performance on the untrained stimuli exceeded chance levels, which suggests that the pigeons learned the abstract concepts of *same* and *different*; however, performance on the training stimuli consistently exceeded that of the unfamiliar stimuli.

Implicit in the design of this study was the assumption that the pigeons would learn to correctly label the training stimuli, that is, that the pigeons could learn to associate specific responses to particular stimuli. The data of most interest to Wasserman et al. (1995) however, were whether the pigeons could also respond appropriately to the novel stimuli and thereby display abstract conceptualization of *sameness* and *difference.* Note that this focus on abstraction is opposite from the focus of the present approach. It is a standard assumption in modeling of human *same–different* judgments that people understand the abstract concepts of *same* and *different.* The current article focused instead on the hypothesis that seems so natural to animal *same–different* studies: that memory for specific exemplars also plays a role in *same–different* judgments. In this light, our conclusions mirror those of Cook, Cavoto, and Cavoto (1995) regarding their *same–different* pigeon studies:

> The less than perfect transfer to novel stimuli found in the Wasserman et al. (1995) study and in our study indicates that these birds were sensitive to the specific identity, as well as the relational properties, of the displays. If only a perceptual abstraction were involved, then the identity of the display should have made no difference. (p. 259)

It is interesting to speculate, however, on a more extreme possibility. Our interpretation of the on-line process in the EBRW is that it involves an abstract similarity-judgment comparison between two currently presented pairs of objects without reference to the other objects presented in the experiment. The on-line process was included because it is clear that people can respond

same and *different* in a sensible and consistent manner in the complete absence of feedback in a given experiment. Perhaps, however, even the on-line process may be based on comparisons to previously stored exemplars. Prior to entering the experiment, the observer has experienced multitudes of pairs of objects during his or her lifetime that have been labeled as *same* or *different*. Conceivably, this prior lifetime of experience, for which we have no record, may be called into play each and every time an observer is asked to make a *same–different* judgment. In a similar vein, Heit (1994) has suggested a role of previously stored exemplars to account for prior knowledge effects in other tasks of categorization. The on-line process that we have formalized in the EBRW may simply be a "stand-in" for the mass action of these prior exemplars. Under this interpretation, rather than using dual processes, the EBRW is using only a single, exemplar-based process to model *same–different* judgments.

References

Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, W. T. (1996). A formal theory of feature binding in object perception. *Psychological Review, 103,* 165–192.

Cook, R. G., Cavoto, K. K., & Cavoto, B. R. (1995). Same-different texture discrimination and concept learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 253–260.

Crist, W. B. (1981). Matching performance and the similarity structure of the stimulus set. *Journal of Experimental Psychology: General, 110,* 269–296.

Ennis, D. M., Palen, J. J., & Mullen, K. (1988). A multidimensional stochastic theory of similarity. *Psychological Review, 32,* 449–465.

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

Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 3–28.

Heit, E. (1994). Models of the effects of prior knowledge on category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1264–1282.

Hintzman, D. L. (1986). "Schema abstraction" in multiple-trace memory. *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.

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. *The Psychology of Learning and Motivation, 18,* 1–47.

Krueger, L. E. (1973). Effect of stimulus frequency on speed of "same"-"different" judgements. In S. Kornblum (Ed.), *Attention & performance* (Vol. IV, pp. 497–506). New York: Academic Press.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and density. *Psychological Review, 85,* 445–463.

Kruschke, J. K. (1992). Alcove: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44.

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

Lockhead, G. R. (1970). Identification and the form of multidimensional discrimination space. *Journal of Experimental Psychology, 85,* 1–10.

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

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

Logan, G. D. (1992). Shapes of reaction time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 883–914.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 103–189). New York: Wiley.

Luce, R. D., Green, D. M., & Weber, D. L. (1976). Attention bands in absolute identification. *Perception and Psychophysics, 20,* 49–54.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

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

Nickerson, R. S. (1981). Context is important but it does not explain everything: A comment on "matching performance and the similarity structure of the stimulus set" by Crist. *Journal of Experimental Psychology: General, 110,* 297–302.

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

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 87–109.

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

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

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review, 88,* 552–572.

Ratcliff, R., & Hacker, M. J. (1981). Speed and accuracy of same and different responses in perceptual matching. *Perception & Psychophysics, 30,* 303–307.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 629–639.

Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology, 53,* 94–101.

Shepard, R. N. (1958). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review, 65,* 242–256.

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

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

Shepard, R. N., & Chang, J. J. (1963). Stimulus generalization in the learning of classifications. *Journal of Experimental Psychology, 65,* 94–102.

Takane, Y., & Sergent, J. (1983). Multidimensional scaling models for reaction times and same-different judgments. *Psychometrika, 48,* 393–423.

Thomas, R. D. (1996). Separability and independence of dimensions within the same-different judgment task. *Journal of Mathematical Psychology, 40,* 318–341.

Townsend, J. T., & Landon, D. E. (1982). An experimental and theoretical investigation of the constant-ratio rule and other models of visual letter confusion. *Journal of Mathematical Psychology, 25,* 119–162.

Tsal, Y., & Shalev, L. (1996). Inattention magnifies perceived length: The attention receptive field hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 233–243.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.

Wasserman, E. A., Hugart, J. A., & Kirkpatrick-Steger, K. (1995). Pigeons show same-different conceptualization after training with complex visual stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 248–252.

Received April 2, 1999
Revision received November 1, 1999
Accepted March 13, 2000 ■

Call for Nominations

The Publications and Communications Board has opened nominations for the editorships of ***Journal of Applied Psychology, Journal of Consulting and Clinical Psychology, Journal of Educational Psychology, Psychological Bulletin,*** and ***Journal of Personality and Social Psychology: Interpersonal Relations and Group Processes*** for the years 2003–2008. Kevin R. Murphy, PhD, Philip C. Kendall, PhD, Michael Pressley, PhD, Nancy Eisenberg, PhD, and Chester A. Insko, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2002 to prepare for issues published in 2003. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations are also encouraged.

To nominate candidates, prepare a statement of one page or less in support of each candidate and send to

- Margaret B. Spencer, PhD, for the ***Journal of Applied Psychology***
- Donna M. Gelfand, PhD, and Lucia Albino Gilbert, PhD, for the ***Journal of Consulting and Clinical Psychology***
- Lauren B. Resnick, PhD, for the ***Journal of Educational Psychology***
- Janet Shibley Hyde, PhD, and Randi C. Martin, PhD, for ***Psychological Bulletin***
- Sara B. Kiesler, PhD, for ***JPSP: Interpersonal Relations and Group Processes***

Address all nominations to the appropriate search committee at the following address:

[Name of journal] Search Committee
c/o Karen Sellman, P&C Board Search Liaison
Room 2004
American Psychological Association
750 First Street, NE
Washington, DC 20002-4242

The first review of nominations will begin December 11, 2000.