Perceptual Processes in Matching and Recognition of Complex Pictures

Koen Lamberts, Noellie Brockdorff, and Evan Heit
University of Warwick

The role of perceptual feature sampling in speeded matching and recognition was explored in 4 experiments. Experiments 1–3 involved a perceptual matching task with pictures of various objects and scenes. In Experiments 2 and 3, same–different judgments were given under time pressure. The main objective of the matching task was to obtain measures of the perceptual processing rates of different object features. Experiment 4 was an old–new recognition experiment, in which the same stimuli as those in the matching task were used. Response signals were used to limit processing time in the recognition task. The results demonstrated that it is possible to predict speeded recognition performance from performance in perceptual matching. A simple stochastic feature-sampling model provides a unified account of the data from the 4 experiments.

The ability to classify perceived objects as familiar or unfamiliar is an essential component of cognition. When people perceive an object, they have a remarkable ability to decide very rapidly whether they have encountered the object before. It is not surprising, therefore, that old–new recognition memory has become a central topic in experimental psychology. Many aspects of recognition memory are well understood, and current theories of recognition explain a wide range of empirical findings.

Even without a detailed conceptual analysis, it is clear that recognition must involve at least three process components. The first is perceptual processing of stimulus information. Stimulus information must be acquired first, to support further processing. This acquisition process is often represented as an information-accumulation process (e.g., Brockdorff & Lamberts, 2000; Bundesen, 1990; Busey & Loftus, 1994, 1998; Lamberts, 1995, 2000; Loftus & McLean, 1999). Second, information must be retrieved from memory. This component is the focus of most recognition theories, which differ in their assumptions about the form of stored information and the mechanisms of retrieval. Finally, an old–new decision must be made, based on the perceived stimulus information and the information retrieved from memory.

In studying the role of these component processes, mapping the time course of processing in recognition is particularly relevant. Recognition response times have been studied extensively (e.g., Ratcliff, 1978), providing valuable information about the time course of retrieval and decision making. In other experiments, response signals or response deadlines have been used to limit the processing time on individual recognition trials (e.g., Brockdorff &

Lamberts, 2000; Dosher & Rosedale, 1991; Gronlund & Ratcliff, 1989; Hintzman & Curran, 1994; Rotello & Heit, 1999). The effects of time pressure on the proportions of "old" responses (or on some derived measure, such as d') provide important information about the availability of different kinds of information after varying amounts of processing time.

In a recent article (Brockdorff & Lamberts, 2000), we have argued that recognition response patterns at different deadlines or signal intervals may reflect the accumulation of perceptual information about the test stimulus rather than the time course of retrieval or decision making. We proposed a model of old–new recognition, called the feature-sampling theory of recognition (FESTHER), which provides a detailed account of the response patterns in a wide range of recognition experiments with response deadlines. FESTHER is based on the extended generalized context model of categorization (EGCM; Lamberts, 1995, 1998, 2000; Lamberts & Brockdorff, 1997; Lamberts & Freeman, 1999), which is derived from Nosofsky's (1986, 1988, 1991) generalized context model.

FESTHER is an exemplar model of recognition memory (see also Estes, 1994; Nosofsky, 1988, 1991). It is assumed that each stimulus in a study set leaves a distinct trace in memory, and that subsequent recognition judgments are based on the total similarity of a stimulus to the traces in memory. In FESTHER, it is further assumed that the earliest stages of recognition involve the construction of a perceptual stimulus representation through a process of stochastic feature sampling. The time course of this perceptual process is assumed to be crucial for understanding results on recognition under time pressure. When a stimulus is presented, its features are processed in a parallel, independent fashion (see Lamberts & Freeman, 1999). The time needed for processing each feature is a random variable, and some features tend to be processed faster than others. The probability that a given stimulus dimension p has been processed at or before a given time t after the start of processing is given by an exponential distribution function:

$$i_p(t) = 1 - \exp(-q_p t), \tag{1}$$

where q_p is the processing rate of dimension p. The probability that any given dimension has been processed increases with processing

time. Each processed feature is integrated into the current perceptual representation of the stimulus. This representation thus becomes more complete as more features are processed. Whenever a feature is processed, the similarity of the current stimulus representation to the traces in memory changes (see Brockdorff & Lamberts, 2000), and these changes determine the probability that the stimulus will be recognized.

In experiments with response signals or response deadlines, responses may have to be initiated before stimulus processing is complete (i.e., before all stimulus features have been processed). In other studies (Lamberts, 1995, 1998, 2000, 2002; Lamberts & Brockdorff, 1997; Lamberts & Freeman, 1999), we have shown that a similar principle explains a wide range of results on timing effects in categorization. In FESTHER, it is assumed that the recognition functions in experiments with response signals also reflect the time course of feature sampling. Brockdorff and Lamberts (2000) have shown that FESTHER can provide accurate accounts of the response patterns for individual stimuli in a number of recognition experiments in which artificial objects and words were used as the stimuli and unpredictable response signals were used to limit processing time. The response patterns were often complex in the sense that the response proportions did not always vary monotonically as a function of available processing time. The model's feature-sampling assumptions proved sufficient to reproduce these complex patterns.

Among process models of recognition memory, FESTHER is quite unique in predicting a strong link between the time course of perception and the time course of recognition (although Loftus & McLean, 1999, have made a related proposal, which we discuss in detail later). Traditional process models of recognition do not assign a central functional role to perceptual processes. Yet, if FESTHER proves to be correct, it has the important implication that we cannot fully understand how people recognize visual stimuli unless we also understand how these stimuli are perceived. The experiments in this article aimed to further explore the role of perceptual processing in recognition. If time-dependent variation in recognition judgments depends on the time course of perceptual processing, it should be possible to use information about the time course of stimulus perception to predict the time course of recognition judgments. That is, if the stimulus dimensions and their processing rates are known, it is possible to make precise, testable predictions about the time course of recognition. In the experiments in Brockdorff and Lamberts (2000), the stimulus dimensions were always known (because artificial stimuli were used), but the processing rates of the dimensions were free parameters in the model. Because perceptual processing rates were not measured independently, the experiments in Brockdorff and Lamberts (2000) did not provide direct evidence that perceptual processing rates of stimulus dimensions affect the time course of recognition judgments. Moreover, the use of relatively simple artificial stimuli, with a small number of well-defined dimensions, restricted the generality of the conclusions that could be drawn from the experiments. In the research we report in this article, we aimed to remedy these shortcomings by attempting to predict the time course of recognition judgments of realistically complex stimuli on the basis of independent measurements of dimensional processing rates.

This task raised two important challenges. First, the dimensions of complex, realistic stimuli (such as photographs or drawings of objects or scenes) are unknown, which makes it difficult to control or manipulate similarity between such stimuli. Second, the determinants of feature processing rates are complex (see Lamberts, 1998; Lamberts & Freeman, 1999), precluding the use of a simple manipulation to produce stimulus sets with reliably different processing rates for specific features. To overcome these difficulties, we used a multiexperiment research strategy. We first carried out three perceptual matching experiments, the aim of which was primarily to provide independent measures of perceptual processing rates of stimulus features. Our fourth experiment was a standard recognition experiment, in which judgments were given under time pressure.

Our experimental stimuli were drawings of various objects and scenes. We started with a set of 240 different pictures. We then added another 48 pictures, which were obtained by modifying 48 pictures from the original set. The modifications always involved the deletion of one minor feature from an original picture. Examples of original and modified pictures are shown in Figure 1 and in Appendix A. Henceforth, we call the modified feature, or the part of the original picture that corresponded to the modified feature, the *critical* feature. These stimuli were used in the four experiments. The first three experiments involved a perceptual matching task, in which the participants were shown pairs of pictures and had to decide whether the pictures in each pair were identical. The fourth experiment was an old–new recognition experiment. In the matching experiments, three types of picture pairs could be presented for matching: identical pairs, different pairs (which consisted of two unrelated pictures), and similar pairs (which consisted of an original picture and its modified version). The primary purpose of the three matching experiments was to measure perceptual processing rates of features. To do this, we applied FESTHER's assumptions about perceptual processing directly to the matching tasks. (Recall that FESTHER states perceptual stimulus processing involves independent, stochastic feature sampling.)
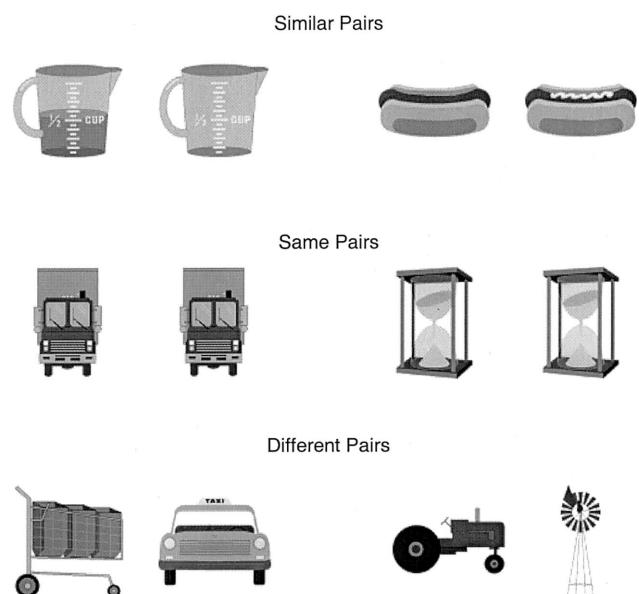

Figure 1. Sample stimuli used in the four experiments.

In the first experiment, two pictures were presented simultaneously on each trial, and the participants were instructed to judge their identity as quickly and accurately as possible. In this experiment, we were primarily interested in the response times (RTs) for "different" judgments of similar pairs of items. We assumed that the participants would respond "different" to these pairs as soon as they detected the critical feature in both pictures. Because the stimuli in the matching task were presented simultaneously on the screen, we interpreted the mean time for correct responses to each similar pair as an ordinal index of the perceptual processing rate of the critical feature. On the basis of the RTs for correct responses, we carried out a median split, dividing the 48 similar pairs into a set of 24 pairs that yielded the fastest responses (the *fast–similar* set) and a set of 24 pairs that produced the slowest responses (the *slow–similar* set). This division was maintained throughout the other three experiments in the series.

On the basis of RTs alone, it is not possible to obtain a reliable estimate of perceptual processing times, unless the stimulus structure is directly manipulated (see Lamberts, 2000). Experiments 2 and 3 were designed to produce estimates of the average processing rates of the critical features in the fast–similar and slow–similar sets. In these experiments, the participants matched the same pairs of pictures as in Experiment 1, but unpredictable response signals were used to limit the time available for responding on any trial. The patterns of responses at different signal intervals can be used to obtain reliable estimates of processing rates (Brockdorff & Lamberts, 2000; Lamberts, 1998). Experiments 2 and 3 differed only in the sequence of events on each trial. In Experiment 2, the two pictures were presented simultaneously on the screen. In Experiment 3, presentation of the pictures was sequential, with an interstimulus interval (ISI) of 2,000 ms. On the basis of the data from Experiments 2 and 3, we aimed to obtain consistent estimates of the processing rates for the critical features in the fast–similar and slow–similar stimulus sets. The presentation mode of the picture pairs (simultaneous or sequential) was varied between the experiments, because this would allow us to obtain a critical test of the perceptual processing assumptions in the model. The assumption that feature processing times are exponentially distributed led us to predict different functional relations between available processing time and response rates in simultaneous and sequential matching, as explained in detail in the modeling section below.

The recognition task in Experiment 4 used the same stimuli as did the matching tasks in Experiments 1, 2, and 3. After an initial study phase, different types of items were presented for recognition. *Old* items were part of the study set, *new* items were unrelated to the study set, and *similar* items differed on one critical feature from a stimulus in the study set. The similar test items were divided into fast–similar items and slow–similar items, according to the match RTs in Experiment 1. Response signals could occur after various lags, and the proportion of "old" responses after various lags was the variable of interest.

Together, the four experiments provide a critical test of our assumptions about feature sampling and perceptual processing in recognition. If the timing effects that have been found in previous recognition experiments with response deadlines or response signals are due to the time course of feature sampling, then it should be possible to fit the data from the matching and recognition experiments with a single set of processing rate parameters. The results from the experiments may also have more general implications for our understanding of the relation between perceptual matching and recognition of pictures. Although these two tasks have been widely studied, there have been few previous attempts to investigate the theoretical and empirical connections between them (but see Ratcliff, 1981; Van Zandt, Colonius, & Proctor, 2000; also relevant is Cohen & Nosofsky, 2000, which explores the relation between matching and categorization). Yet, even without much analysis, it is clear that same–different matching tasks share many characteristics with standard old–new recognition tasks. In matching, two stimuli are presented simultaneously or in close succession, and the participants judge the identity of the stimuli. In recognition, participants first study a set of stimuli, and then judge whether individual test stimuli occurred in the study set. Both tasks involve a comparison process. In matching, the comparison is between two simultaneously presented stimuli or between a stimulus in short-term memory and a stimulus on screen. In recognition, the comparison is between a test stimulus and a potentially large set of stimuli in long-term memory. Our results should indicate to what extent a common framework for both tasks is feasible.

The remainder of this article is organized as follows. We first present the four experiments, together with a brief discussion of the results. Next, we discuss the modeling principles that provide the link between the four data sets. Finally, we apply the model to the data, and we discuss the implications of the modeling.

Experiment 1

In this experiment, participants were asked to give speeded identity judgments of pairs of drawings. The stimuli were color drawings of various everyday office and household items, places, and people. The pictures were used to construct three classes of stimulus pairs, called *same, different,* and *similar.* The same pairs were made up of two copies of the same picture, the different pairs consisted of two different pictures, and the similar pairs were pairs of pictures produced by making one picture a copy of the other picture but with part of the image deleted (see Figure 1 for examples). Pairs of pictures were presented on a computer screen, and participants were asked to decide as quickly as possible whether the two members of each pair were identical or different.

Method

Participants. A total of 20 undergraduate and graduate students from the University of Warwick were recruited through advertisements placed around campus and paid £2 (approximately $3) each.

Apparatus and stimuli. The experiment was controlled by a Pentium 400 MHz computer with a 17-inch color monitor with a resolution of 1,280 pixels (horizontally) × 1,024 pixels (vertically). Responses were registered by means of two microswitches connected to the computer's parallel port. Responses were timed with an accuracy of 1 ms. The participants viewed the screen from a distance of 60 cm.

A total of 288 drawings was used in the experiment to produce 48 similar pairs, 48 different pairs, and 96 same pairs. All participants received the same pairs of stimuli, but in a different random order. Each picture occurred only in one pair. The drawings were in full color on a white background. Each picture measured 5.7 cm × 5.7 cm, and each stimulus pair consisted of two pictures side by side, centered on a black screen. The distance between the inner edges of the two pictures in a pair was 6 mm.

Design and procedure. Each participant attended one session consisting of 192 trials. Each stimulus pair was shown once. The left–right position of the pictures that made up the similar and different pairs was counterbalanced across participants. On each trial, a cue (a cross) was shown at the center of the screen for 500 ms. The screen was blank (black) for 100 ms, after which a stimulus pair appeared. The stimuli remained on screen until a response was made or until 5 s had elapsed, whichever occurred first.

Participants were instructed to press the *same* button if the two pictures shown were exactly the same and to press the *different* button if the pictures were different in any way. Participants were asked to respond as accurately and as quickly as possible. Accuracy feedback was provided after every trial. Before the experimental session, participants were given 30 practice trials with nonexperimental stimuli. During practice, the need for quick responding was emphasized.

Results and Discussion

The mean RTs on trials with correct responses and the proportions correct on similar, different, and same pairs of pictures are shown in Table 1. A repeated measures analysis of variance (ANOVA) on the mean RTs revealed a reliable main effect of match type, $F(2, 38) = 25.78$, $p < .01$, $MSE = 16,970$. Planned comparisons showed that all pairwise RT differences between the match types were reliable: different versus similar, $t(19) = 11.35$, $p < .01$; different versus same, $t(19) = 5.18$, $p < .01$; same versus similar, $t(19) = 1.83$, $p < .05$. An ANOVA of the proportions correct across the three match types also yielded a reliable main effect, $F(2, 38) = 59.34$, $p < .01$, $MSE = 0.003$. All pairwise accuracy differences between the match types were reliable: different versus similar, $t(19) = 9.38$, $p < .01$; different versus same, $t(19) = 2.34$, $p < .01$; same versus similar, $t(19) = 7.51$, $p < .01$.

Because we were particularly interested in the stimuli that made up the similar pairs, the responses to these pairs were analyzed further. We ranked the 48 similar pairs according to mean RT for correct responses across participants. (See Appendix A for ranking and mean RT for all similar pairs.) The mean RT for the fastest pair was 569 ms, and that for the slowest pair was 1,229 ms. Next, we divided the similar pairs into two sets. The 24 pairs that yielded the fastest mean RTs formed the *fast* set, and the 24 pairs with the slowest mean RTs formed the *slow* set. Within the fast set, mean RTs ranged from 569 ms to 749 ms, with a set mean of 662 ms. Mean RTs in the slow set ranged from 751 ms to 1,229 ms, with a set mean of 936 ms. The total proportion of correct responses was .93 in the fast set and .68 in the slow set.

As we indicated in the introduction, Experiment 1 was primarily carried out to provide data on the relative speed of processing the critical differences in a wide range of similar pairs. The difference in accuracy between the slow and fast sets of similar pairs is potentially important here. The direction of the difference (more

errors on slow pairs) allows us to rule out differential speed–accuracy trading as a primary source of RT differences between the slow and fast sets. However, the accuracy difference introduces the risk that the median split not only separates slow from fast pairs but also separates pairs in which the difference is perceptible from pairs in which the difference simply cannot be seen, regardless of processing time. To rule out this possibility, we asked 2 additional participants to carry out the matching task with exclusive emphasis on accuracy and without the requirement to respond as fast as possible. Both participants obtained perfect scores, which allowed us to assume safely that the median split separated pairs that differed only in the detection speed of the critical difference.

Experiment 2

As in Experiment 1, in this experiment participants were asked to give identity judgments of pairs of simultaneously presented drawings. The stimuli were the same as those in Experiment 1. A response-signal procedure was used to limit the available response time on individual trials. On hearing a signal from the computer, participants were to decide immediately whether the two members of each pair were identical or different. On the basis of the classification into fast–similar and slow–similar pairs from Experiment 1, we expected higher proportions of correct "different" responses to the fast–similar pairs than to the slow–similar pairs, across all response-signal intervals.

Method

Participants. A total of 16 undergraduate students from the University of Warwick participated in partial fulfillment of a course requirement.

Apparatus and stimuli. These were the same as those in Experiment 1.

Design and procedure. Each participant attended one session consisting of 192 trials. Each stimulus pair was shown once. The left–right position of the pictures that made up the similar and different pairs was counterbalanced across participants. A response-signal procedure was used on each trial. On each trial, a cue (a cross) was shown at the center of the screen for 500 ms. The screen was blank (black) for 100 ms, after which a stimulus pair appeared. At variable time lags after stimulus onset (100, 200, 300, 350, 400, or 600 ms), a 1000-Hz tone sounded, and the picture pair disappeared from the screen. Participants were instructed to respond immediately upon hearing the tone. If no response was made within 350 ms from the onset of the tone, or if a response was made before the onset of the tone, an appropriate error message was displayed.

Although all participants received the same match stimuli in the course of the experiment, the presentation order of the match stimuli and the assignment of stimulus pairs to signal intervals were randomized between participants, with the constraint that the number of same (16), different (8), fast–similar (4), and slow–similar (4) pairs was the same in each of the six response-signal conditions.

Participants were instructed to press the *same* button if the two pictures shown were exactly the same and to press the *different* button if the pictures were different in any way. Accuracy feedback was provided after every trial. Before the experimental session, participants were given 48 practice trials with nonexperimental stimuli.

Results and Discussion

Only responses that were given within a 100- to 350-ms window from the onset of a response signal were analyzed. Overall, a proportion of .15 responses were outside the response window, and a repeated measures ANOVA showed that this did not vary sig-

Table 1

Mean Correct Response Times (RTs; in Milliseconds) and Accuracy as a Function of Match Type in Experiment 1

Match type	Mean RT	Proportion correct
Similar	761	.80
Different	544	.97
Same	839	.93

nificantly with match type, $F(3, 45) < 1$. The proportions of responses outside the window were .15, .15, .13, and .13 for fast–similar, slow–similar, different, and same pairs, respectively. An ANOVA yielded a reliable effect of signal interval on proportions of responses outside the window, $F(5, 75) = 7.29$, $p < .01$, $MSE = 0.030$; and a significant interaction between match type and signal interval, $F(15, 255) = 2.19$, $p < .01$, $MSE = 0.018$. Table 2 summarizes the proportions of responses that were outside the response window, by signal interval and match type. A decrease in the proportion of responses outside the window with increasing signal lags is a common occurrence in response-signal experiments (e.g., Brockdorff & Lamberts, 2000). The interaction between match type and signal interval is potentially important. The data in Table 2 show that the proportions of excluded trials are quite similar for the same, slow–similar, and fast–similar trials, and that the interaction appears to be primarily due to the different trials. For the different trials, the proportion of excluded trials is generally low, except at the longest signal interval (.27), at which it is higher than that for any other match type. Further analyses showed that this proportion consisted primarily of anticipation errors (89%). This is compatible with the notion that the different trials were easiest and that for these trials, the participants sometimes found it difficult to delay a response until the signal at the longest interval. Although there are no reasons to assume that the differences in discarded response proportions could have introduced artifacts in our analyses or modeling, we decided to carry out all analyses on the full data sets as well, without discarding any trials. Because the conclusions from this exercise were identical to those from the standard analyses (with discarded trials), we discuss only the standard analyses in this article.

The proportions of "same" responses for the fast–similar, slow–similar, different, and same pairs are shown in Figure 2. An ANOVA on the proportions of "same" responses yielded a main effect of match type, $F(3, 45) = 247.90$, $p < .01$, $MSE = 0.036$; and an interaction between match type and deadline, $F(15, 225) = 6.04$, $p < .01$, $MSE = 0.039$. As can be seen from the figure, fast–similar picture pairs produced consistently higher proportions of correct responses than did slow–similar pairs. A separate ANOVA on the data from similar pairs showed only a reliable main effect of match type (fast or slow), $F(1, 15) = 102.96$, $p < .01$, $MSE = 0.035$; a main effect of response-signal interval, $F(5, 75) = 4.45$, $p < .01$, $MSE = 0.070$; and an interaction between match type and signal interval, $F(5, 75) = 3.08$, $p < .05$, $MSE = 0.066$.

Table 2
Proportions of Responses Outside the Response Window in Experiment 2, as a Function of Signal Interval (in Milliseconds) and Match Type

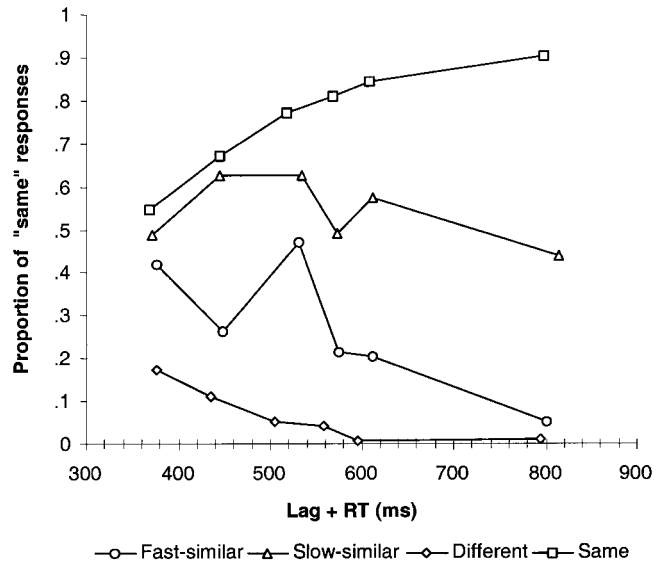|	Match type			
Interval	Fast–similar	Slow–similar	Different	Same
100	.22	.30	.18	.25
200	.17	.13	.08	.12
300	.17	.13	.07	.09
350	.13	.11	.08	.11
400	.08	.05	.08	.10
600	.11	.22	.27	.12

Figure 2. Observed proportions of "same" responses in Experiment 2. Responses to same, fast–similar, slow–similar, and different pairs are plotted against total time elapsed from presentation of stimulus (mean response time [RT] across all participants for all stimuli in each type) in the six response-signal conditions (100, 200, 300, 350, 400, and 600 ms).

The results from Experiment 2 were largely as expected. Accuracy generally increased at longer signal intervals. Of particular interest was the difference between the two types of similar pairs. The fast–similar pairs generally produced higher proportions of correct responses than did the slow–similar pairs, and the difference between the two types became more pronounced at longer signal intervals. This confirms that the division into two groups on the basis of RTs in Experiment 1 was useful, because it did separate similar pairs with a rapidly detectable difference from similar pairs with a slowly detectable difference. The theoretical implications of these data are discussed in the modeling section.

Experiment 3

This experiment was identical to Experiment 2, except for the time course of stimulus presentation on each trial. The pictures that made up each pair were now presented sequentially. The picture on the left-hand side of the display always appeared first, followed by the picture on the right-hand side after a 2.0 s delay. As in Experiment 2, response signals were used to limit processing time on individual trials.

Method

Participants. A total of 12 undergraduate students from the University of Warwick participated in partial fulfillment of a course requirement.

Apparatus and stimuli. These were the same as those in Experiments 1 and 2.

Design and procedure. Each participant attended one session consisting of 192 trials. Each stimulus pair was shown once. The left–right position of the pictures that made up the similar and different pairs was counterbalanced across participants. A response-signal procedure was used. On each trial, the first picture in a pair appeared on the left-hand side of the screen. After 2,000 ms, the second picture of the pair appeared on the

right-hand side of the screen. At variable time lags after the right-hand element of the pair was displayed on screen (50, 100, 200, 300, 400, and 600 ms), a 1000-Hz tone sounded, and the picture pair disappeared from the screen. Although all participants received the same match stimuli in the course of the experiment, the presentation order of the match stimuli and the assignment of stimulus pairs to signal intervals was randomized between participants, with the constraint that equal numbers of same, different, fast–similar, and slow–similar pairs were assigned to each response-signal condition.

Participants were instructed to respond immediately upon hearing the tone. If no response was made within 350 ms from the onset of the tone or if a response was made before the onset of the tone, an appropriate error message was displayed. Instructions, feedback, and practice trials were the same as those in Experiment 2, with obvious modifications related to the sequential nature of the matching task.

Results and Discussion

The data from trials in which a response was given outside the 100- to 350-ms window after signal onset were excluded from the analyses. Overall, a proportion of .14 of responses were excluded in this way. As in Experiment 2, there were no reliable differences between the proportions of excluded trials for different match types, $F(3, 33) < 1$. The proportions of excluded trials were .16, .14, .13, and .13 for fast–similar, slow–similar, different, and same pairs, respectively. Again, the effect of signal interval on proportions of responses outside the window was reliable, $F(5, 55) = 9.89$, $p < .01$, $MSE = 0.203$ (see Table 3), but there was no interaction between signal interval and match type, $F(15, 165) = 1.03$, $p = .43$, $MSE = 0.018$. The proportions of excluded trials were similar to those in Experiment 2, except for a higher proportion of exclusions at the 600-ms signal interval (.21). Most of the exclusions (88%) at this signal interval were anticipatory responses.

The proportions of "same" responses for the fast–similar, slow–similar, different, and same pairs at different signal lags are shown in Figure 3. A repeated measures ANOVA on the proportions of "same" responses yielded a main effect of match type, $F(3, 33) = 123.88$, $p < .01$, $MSE = 0.065$; a main effect of signal interval, $F(5, 55) = 8.44$, $p < .01$, $MSE = 0.035$; and an interaction between match type and signal interval, $F(15, 165) = 3.64$, $p < .01$, $MSE = 0.040$. A separate ANOVA on the similar pairs showed only main effects of match type, $F(1, 11) = 22.49$, $p < .01$, $MSE = 0.089$; and signal interval, $F(5, 55) = 10.45$, $p < .01$, $MSE = 0.059$; and no interaction between type and signal interval, $F(5, 55) = 0.32$, $p = .90$, $MSE = 0.075$. The results from Experiment 3 largely replicated the difference between fast–similar and slow–similar pairs found in Experiment 2. Accuracy

Table 3
Proportions of Responses Outside the Response Window in Experiment 3, as a Function of Signal Interval (in Milliseconds)

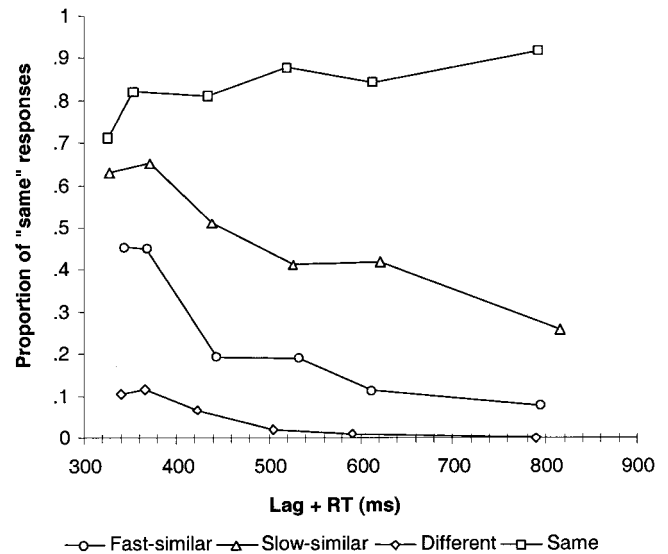| Signal interval	Proportion outside
50	.21
100	.12
200	.09
300	.09
400	.10
600	.21

Figure 3. Observed proportions of "same" responses in Experiment 3. Responses to same, fast–similar, slow–similar, and different pairs are plotted against total time elapsed from presentation of stimulus (mean response time [RT] across all participants for all stimuli in each type) in the six response-signal conditions (50, 100, 200, 300, 400, and 600 ms).

was generally higher in Experiment 3 than in Experiment 2. This is almost certainly due to the sequential nature of the matching task in Experiment 3. The 2,000-ms ISI was probably sufficient to allow complete processing of the first stimulus by the time the second stimulus was presented.

Experiment 4

As we indicated in the introduction, Experiment 4 was designed as an old–new recognition experiment, in which response signals were used to restrict decision times. The task consisted of standard memorization and testing stages. The pictures from the previous experiments were used to construct four classes of test stimuli: old (presented in the initial memorization stage), fast–similar, slow–similar, and new. One randomly chosen stimulus from each similar pair was used for the study set, and the other stimulus of the pair was presented during test.

Method

Participants. A total of 36 undergraduate and graduate students from the University of Warwick were recruited through advertisements and paid £3.50 ($5) each.

Apparatus and stimuli. The apparatus and stimuli were the same as in the previous experiments. Each study block consisted of 30 pictures, including a primacy buffer and recency buffer, each containing three nonexperimental pictures (leaving 24 experimental pictures). Presentation order of the pictures in the study block was randomized. Each test set consisted of 36 stimuli: 12 old stimuli (taken from the study set), 12 similar stimuli (obtained by alteration of stimuli from the study set, with 6 of the 12 classified as slow–similar and 6 of the 12 classified as fast–similar), and 12 new stimuli (unrelated to the study set). Half of the similar transfer stimuli were pictures with deletions, and half were the original pictures.

Design and procedure. Each participant attended one session. The stimuli in the study block were presented one picture at a time on a

computer screen at a rate of 3 s per picture. The participants were told that their task in the study phase was to try to remember each picture. Following each study block, a recognition test was given on 12 old, 12 similar (6 slow–similar and 6 fast–similar) and 12 new pictures. Participants were instructed to press the *old* button if the test picture was present in the study block and to press the *new* button if the picture was different in any way from any picture seen in the study block.

A response-signal procedure was used on each test trial. Pictures were presented one at a time and appeared in the center of the screen on a black background. Each test picture was shown once in the test phase. On each trial, a cue (a cross) was shown at the center of the screen for 500 ms. The screen was blank (black) for 100 ms, and then a picture appeared. At variable time lags after the stimulus onset on screen (100, 300, 500, 700, 1,000, and 1,250 ms), a 1000-Hz tone sounded, and the stimulus disappeared from the screen. Participants were instructed to respond immediately upon hearing the tone. If no response was made within 350 ms of the onset of the tone, or if a response was made before the onset of the tone, an appropriate error message was displayed. No accuracy feedback was provided after the trials, but the proportion of correct responses and the proportion of responses made within 350 ms of the response signal were shown after each block.

A total of four study–test experimental blocks were presented. Each of the four blocks contained a different set of pictures. Each participant was shown the same old, similar, and new pictures within each block, but the presentation order of the blocks and the order of the stimuli within each block were randomized. An equal number of pictures of each type (old, fast–similar, slow–similar, and new) was assigned (at random for each participant) to each response-signal condition. Before the experimental blocks, participants were given a practice block to familiarize them with the experimental procedure. The practice block was identical to a test block except that the stimuli used were not experimental stimuli.

Results and Discussion

The data were trimmed to remove trials in which participants either took more than 350 ms to respond to the signal or responded faster than 100 ms after the response signal. An ANOVA on the proportions of discarded trials showed the expected main effect of signal interval, $F(5, 175) = 67.29$, $p < .01$, $MSE = 0.073$ (with more trials discarded at the shortest intervals); but also a reliable effect of stimulus type, $F(3, 105) = 5.36$, $p < .01$, $MSE = 0.026$. The effect of stimulus type is potentially problematic, because differences in proportions of discarded trials can introduce artificial shifts in proportions of correct responses. Table 4 shows the proportions of discarded trials for the four stimulus types in the experiment. The proportion was lowest for the old and new stimuli and highest for the similar stimuli. The difference between the fast–similar and slow–similar types was small and nonsignificant. Moreover, this difference was in the opposite direction from what might be expected, with slightly more responses outside the window for fast–similar stimuli than for slow–similar stimuli. It is

Table 4

Proportions of Responses Outside the Response Window in Experiment 4, as a Function of Stimulus Type

Stimulus type	Proportion outside
Old	.24
New	.26
Fast–similar	.30
Slow–similar	.28

unlikely, therefore, that differential rates of discarded stimuli would introduce significant artifacts in the data. As an additional check, we also analyzed the data from Experiment 4 without discarding any responses. This alternative method of analysis did not alter the results or the conclusions from the modeling in any meaningful way.

The proportions of "old" responses for old, fast–similar, slow–similar, and new stimuli are shown in Figure 4.[1] A repeated measures ANOVA on the proportion of "old" responses showed a reliable main effect of stimulus type, $F(3, 105) = 295.52$, $p < .01$, $MSE = 0.430$; and a reliable interaction between stimulus type and response-signal condition, $F(15, 525) = 11.23$, $p < .01$, $MSE = 0.031$.[2] To obtain a better view of the difference between the fast–similar and slow–similar stimuli, we also carried out an additional ANOVA looking at only these two types of stimuli, using response-signal condition and stimulus type (fast or slow) as independent variables. This ANOVA produced a reliable main effect of stimulus type, $F(1, 35) = 109.97$, $p < .01$, $MSE = 0.238$; and a reliable interaction between response-signal condition and stimulus type, $F(5, 175) = 4.94$, $p < .01$, $MSE = 0.362$.

For the old and new stimuli, there was a monotonic increase in the proportion of correct responses with increasing signal lag, as observed in other experiments with response signals (e.g., Brockdorff & Lamberts, 2000). The asymptotic response levels with these stimuli indicate that the recognition task was relatively easy. For our purposes, the responses to the similar items are most relevant. For both the fast–similar and slow–similar items, the response function was nonmonotonic, as is often the case in experiments with test items similar to study items (e.g., Brockdorff & Lamberts, 2000). The initial increase in errors at shorter lags was followed by a decrease at longer lags. Overall, the slow–similar items produced more errors than did the fast–similar items. The response function for the fast–similar items also reached a maximum at a shorter lag than did the response function for the slow–similar items.

Modeling

The primary purpose of the formal modeling was to provide a unifying account of the data from the matching and recognition experiments by applying our assumptions about the time course of feature sampling. We did not attempt to model the data from Experiment 1. This experiment was used only to classify the similar pairs into slow and fast groups. First, we discuss the assumptions that were made in the modeling of the results from Experiments 2–4.

---

[1] Rotello and Heit (1999) have recently argued in favor of a difference score measure (d' or dL) rather than raw scores such as proportion of "old" responses. However, we focus on raw scores because we are primarily interested in showing a difference between false-alarm rates on fast–similar and slow–similar stimuli. Rotello and Heit suggested subtracting a baseline measure of false alarms to completely new items. We note that subtracting the same baseline from both kinds of similar items would not affect the difference between the two kinds of similar stimuli.

[2] In the ANOVAs on the recognition data, we carried out an arcsine transformation before analysis to improve conformity of the data to the standard assumptions of an ANOVA (e.g., Kirk, 1995). Although the F values and mean square errors are based on the transformed data, we report mean proportions without transformation.
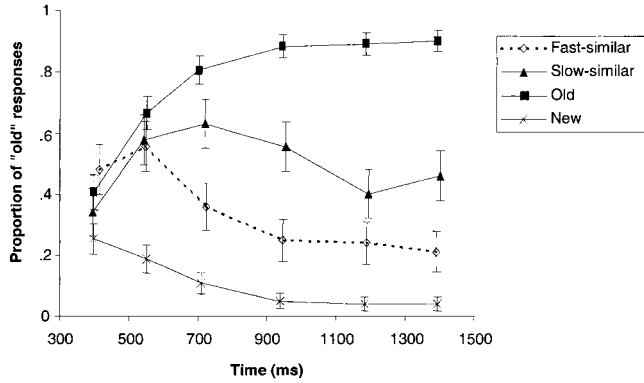
Figure 4. Observed proportions of "old" responses in Experiment 4. Responses to old, fast–similar, slow–similar, and new stimuli are plotted against total time elapsed from presentation of stimulus (mean response time across all participants for all stimuli in each type) in the six response-signal conditions (100, 300, 500, 700, 1,000, and 1,250 ms). Error bars represent 95% confidence intervals.

Experiment 2 involved matching of simultaneously presented stimuli under various response deadlines. We assumed that perceptual processing of the stimuli involves stochastic, parallel, and independent sampling of features and that the processing times are exponentially distributed. Because the stimuli in this experiment were presented simultaneously, we assumed that the feature-sampling process started at the same time for both stimuli (thus assuming complete parallelism within and between stimuli). We further assumed that the response signal interrupted perceptual processing of both stimuli, such that the time available for perceptual processing on a group of trials with a given signal interval i (t_i) was

$$t_i = \mathrm{RT}_i - t_{\mathrm{res}}, \qquad (2)$$

in which RT_i is the mean RT (measured from match stimuli onset) on the group of trials, and t_{res} is the latency period after stimulus onset and the time needed for response production.[3] In the modeling, t_{res} was a single free parameter that applied across all experiments.

On trials with similar pictures, computing the probability that the participants would detect a difference between the match stimuli in the processing time available is relatively straightforward. We assumed that similar pairs differed in only a single feature. Detection of the difference requires that the critical feature is processed in both pictures before the response is initiated:

$$p_{\mathrm{det}}(\textit{simultaneous, similar, } t_i) = [1 - \exp(-q_{\mathrm{crit}} t_i)]^2, \qquad (3)$$

in which $p_{\mathrm{det}}(\textit{simultaneous, similar, } t_i)$ is the probability of detecting the difference between two simultaneously presented similar stimuli within perceptual processing time t_i, and q_{crit} is the processing rate of the critical dimension. In the modeling, q_{crit} was a free parameter and allowed to have different values for the slow–similar and fast–similar pairs.

Next, we needed to determine the probability of a "same" or a "different" response. We assumed that the participants would always respond "different" if they had detected a difference be-

tween the stimuli. However, in experiments with response signals, it is not realistic to assume that all "different" responses are actually based on detection of a difference. Especially at shorter signal intervals, in which the time pressure is quite severe, some proportion of "different" responses can be expected even on trials in which the difference between the stimuli had not been detected. The data from trials with identical match stimuli were used to estimate the proportion of "different" responses to similar pairs that were not the result of detection of a difference. For convenience, we call a "different" response to a similar or different pair a *hit* and a "different" response to an identical pair a *false alarm.* We denote the hit and false alarm rates at signal interval i as H_i and F_i, respectively. On trials with a similar or a different pair of pictures, the total hit rate equals the weighted sum of the true hit rate (for which a difference has actually been detected) and the rate of "different" responses when a difference has not been detected (estimated by F_i):

$$H_i = p_{\mathrm{det}}(t_i) + [1 - p_{\mathrm{det}}(t_i)] \times u_i, \qquad (4)$$

in which u_i is equal to F_i.

In this formulation, F_i is effectively used as a bias measure. Formally, this model is equivalent to the single high-threshold theory of discrimination (Macmillan & Creelman, 1990). We also explored other discrimination theories (such as the double high-threshold theory, Macmillan & Kaplan, 1985) in our modeling of the matching and recognition data, but the single high-threshold model provided the best compromise between a good fit and parsimony across the data sets.

Modeling the choice proportions for different pairs in Experiment 2 was somewhat more complex. If we assumed that the stimuli in a different pair differed on all of the stimulus dimensions, we could show that the probability of detecting any difference within a given time interval depends on the number of stimulus dimensions that are involved in the comparison. Because we do not know the dimensional composition of the stimuli, we estimated the number of dimensions for comparison by a free

---

[3] In Experiments 2, 3, and 4, the stimuli disappeared from the screen as soon as the response signal was presented. However, the stimuli were not masked, so it is likely that perceptual processing continued for a short time after signal presentation. This leaves two alternative methods for estimating perceptual processing times. The first method assumes that total perceptual processing time equals presentation time T_i (i.e., the signal interval) plus a constant that represents the duration of iconic memory: $t_i = T_i + t_{\mathrm{visual\ memory}}$. Although this method would be acceptable, it relies on the assumption that stimulus persistence time in memory after signal presentation is a constant. This assumption may not be correct; it is possible, for instance, that persistence is longer at longer presentation times. Therefore, we chose the second method, which is based on observed RTs from which we subtracted a residual time to estimate perceptual processing times. This method does not assume that visual persistence time is constant, but it relies on the assumption that response production always has the same duration. Although this assumption also may not be justified, we have used it consistently in previous work (e.g., Brockdorff & Lamberts, 2000) without difficulties. In this article, as a final check, we also used the first method to model the data. The model fits for the two methods were almost identical (slightly better for the RT-based method), so we do not report the modeling outcomes obtained with the first estimation method.

parameter, δ. The probability of detecting any difference between the stimuli in time interval t_i is given by

$$p_{\text{det}}(simultaneous,\ different,\ t_i) = 1 - \{1 - [1 - \exp(-qt_i)]^2\}^{\delta},$$

(5)

in which q is the processing rate of the dimensions involved (we made the simplifying assumption that all dimensions of both stimuli had the same processing rate), and δ is the number of dimensions (see Appendix B for a derivation of this equation). The detection rates were translated into response rates in the same way as the similar pairs (see Equation 4).

The response proportions from same pairs were not explicitly modeled, because we assumed that these proportions reflected only response bias. However, the results from trials with identical stimuli were taken into account in the modeling of the results from similar and different trials (through Equation 4).

In Experiment 3, the matching task was sequential. We assumed that the interstimulus interval was sufficiently long (at 2,000 ms) to allow complete processing of the first stimulus by the time the second stimulus was also presented. This assumption made it possible to model the results from the similar and different trials. On similar trials, the difference between the stimuli would be detected as soon as the critical dimension of the second match stimulus had been processed:

$$p_{\text{det}}(sequential,\ similar,\ t_i) = 1 - \exp(-q_{\text{crit}}t_i),$$

(6)

in which all symbols have the same meaning as in Equation 5. (Note that t_i now refers to the perceptual processing time for the second stimulus only.) As for the simultaneous matching task, the detection probability is translated into a response rate by Equation 4.

In modeling different trials, we again assumed that the stimuli differed on a number of dimensions, estimated as δ. The probability that any difference between the match stimuli is detected equals one minus the probability that no dimensions of the second match stimulus have been processed in the interval t_i:

$$p_{\text{det}}(sequential,\ different,\ t_i) = 1 - [\exp(-qt_i)]^{\delta},$$

(7)

in which the symbols have the same meaning as in Equation 5.

Finally, we used the model to predict proportions of "old" responses to similar and new test items in the recognition task in Experiment 4. In modeling recognition, we relied only on FESTHER's assumptions about feature sampling, without using the component of the model that translates total similarity to studied items into choice probabilities. Instead, the recognition task was modeled entirely as an analog to the sequential matching task from Experiment 3. For test items that were similar to a studied item, we assumed that the probability of a correct "new" response was contingent on the detection of the difference between the items. The probability of detecting the difference within time interval t_i is given by

$$p_{\text{det}}(recognition,\ similar,\ t_i) = 1 - \exp(-\rho q_{\text{crit}}t_i).$$

(8)

This equation is identical to Equation 6, with the exception of the rate scaling parameter ρ. This parameter reflects that the recognition task is sensitive not only to the perceptual processing rate of the critical stimulus dimension but also to the rate at which

the information about the critical dimension in memory becomes available. This particular model assumes that the total processing rate of the critical dimension (including its perceptual and retrieval components) is proportional to the purely perceptual processing rate of the dimension, as measured by the matching tasks. We justify this assumption later, when we compare this model's performance to that of models with different assumptions about the time course of perception and retrieval. In any case, the best-fitting value of ρ provides information about the contribution of purely perceptual processes to the time course of recognition.

The model for recognition of new items assumed that new items differed from studied items on all dimensions. This leads to the following expression for the probability that any dimension of the test stimulus has been processed in interval t_i:

$$p_{\text{det}}(recognition,\ different,\ t_i) = 1 - [\exp(-\rho q t_i)]^{\delta}. \quad (9)$$

The detection probabilities in the recognition task were translated into response proportions using Equation 4, in the same way as the matching tasks. The only difference with the matching tasks was that a hit in recognition is defined here as a correct "new" response to a new or similar test item, whereas a false alarm was defined as a "new" response to an old item. (Note that this definition differs from the conventional definition of hits and false alarms in recognition tasks; see, e.g., Rotello & Heit, 1999.)

The data from Experiments 2–4 were modeled jointly, using only one set of six parameters. Three dimensional processing rates were estimated: q_{crit}(fast) and q_{crit}(slow) were the processing rates of the critical dimensions for fast–similar and slow–similar stimuli, respectively, and q was the (average) processing rate of the dimensions of the stimuli from different pairs (in matching) or of the new stimuli (in recognition). The fourth parameter was ρ, the rate scaling parameter which applied only to the recognition task. The fifth parameter was the residual time, t_{res}. Finally, δ estimated the number of dimensions on which stimuli from different match pairs or new stimuli from recognition were compared. A maximum-likelihood criterion was used to determine the best-fitting parameter values, assuming that response proportions in all tasks had a joint binomial distribution. (See Lamberts, 1995, for a discussion of the likelihood function that was applied; see also Riefer & Batchelder, 1988.)

Figure 5 shows the observed and predicted response proportions in the three experiments that were modeled. The best-fitting parameter values are shown in Table 5. From Figure 5, it is clear that the model provided a good account of the choice data in the three experiments. Across the experiments, the model explained 93.1% of the variance in the choice proportions, $\ln(L) = -140.594$.

For the simultaneous matching task, the model's predictions of the response proportions for the different pairs were very accurate. The predicted proportions for the fast–similar and slow–similar pairs were not as close to the observed proportions as were the predicted proportions for the different pairs, because the observed proportions for the similar pairs were more irregular. However, the model captured the difference between the fast–similar and slow–similar pairs well. In the model, the only difference between the fast–similar and slow–similar pairs was the processing rate of the critical feature. In all other respects, the model was identical for these two types of stimuli. It is especially interesting that the model captured the difference in overall appearance of the fast and slow response curves (see Figure 5A). The predicted and observed
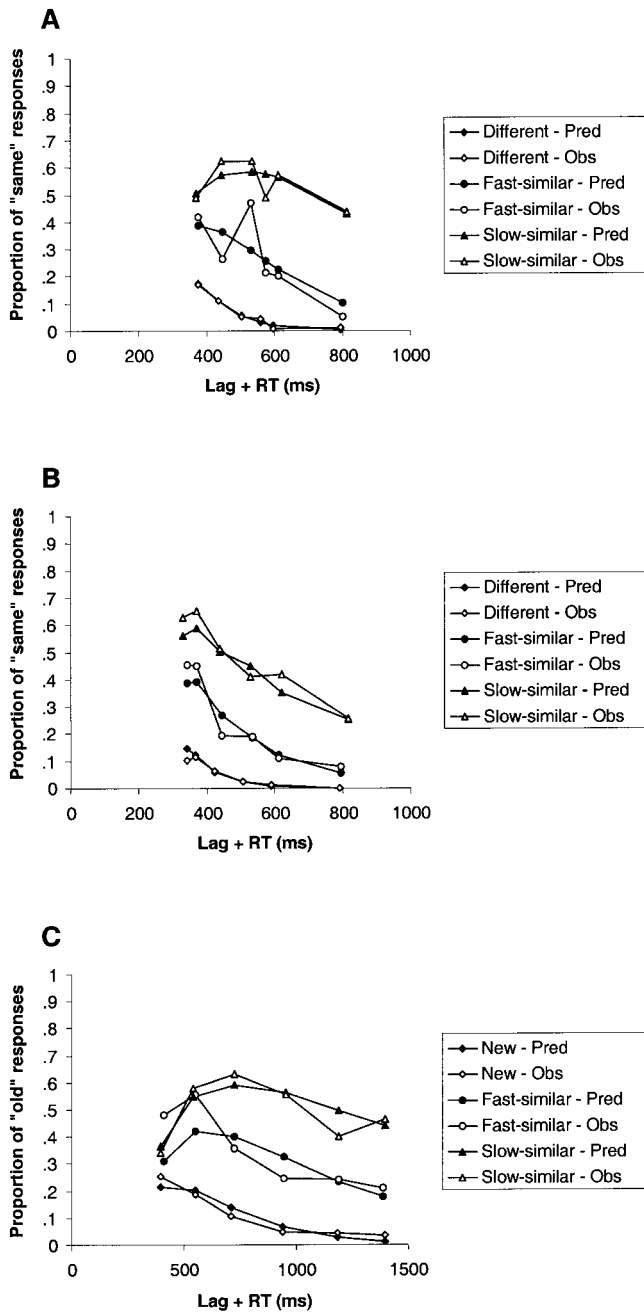
Figure 5. Observed and predicted response proportions in Experiments 2, 3, and 4. A: Simultaneous matching. B: Sequential matching. C: Recognition. Pred = predicted; Obs = observed; RT = response time.

difference between the fast–similar and slow–similar trials was well-predicted, as well as the general shape of the response functions for these trial types. The model predicted a nearly monotonic decline of the proportion of "same" responses with processing time for both of these stimulus types, and this corresponded well to the observed data.

Overall, the model provided a good account of the data from both matching tasks. It is important to note that the fits to the data from both experiments were achieved with a single set of parameters (which also applied to the recognition task). The models for the simultaneous and sequential matching task only differed in their functional form, and this difference appeared to be sufficient to explain the main differences between the response functions for the two tasks. The values of the parameters that applied to the matching tasks (see Table 5) were largely as expected. The estimated processing rate of the critical dimension was higher for the fast–similar pairs than for the slow–similar pairs. The estimated mean perceptual processing time (calculated as $1/q_{crit}$) for the critical dimension was 204 ms for the fast–similar pairs and 455 ms for the slow–similar pairs. The mean processing time for the dimensions of the different pairs was estimated to be 127 ms, and the estimated value of δ shows that the participants took an average of 1.6 dimensions into account in their comparisons. Figure 6 shows the probability of detecting any difference between the different types of stimuli for the simultaneous and sequential matching tasks, as a function of processing time.

For the recognition task (see Figure 5C), the model predictions for the new stimuli are close to the observed values. The same is true for the slow–similar stimuli. The nonmonotonic response pattern is closely mimicked by the model. The model correctly predicts that the response function will peak at the third signal interval. For the fast–similar stimuli, the predictions are not quite so accurate. Again, the model predicts a nonmonotonic response pattern, but it underestimates the proportion of "old" responses at the two shortest signal intervals. Nevertheless, the model correctly predicts a peak in the response function at the second signal interval.

The only model parameter that applied exclusively to the recognition task was ρ, which was used as a multiplier for the processing rates that applied to the matching tasks. The theoretical justification for the inclusion of this parameter was based on the consideration that the matching tasks did not involve the same memory components as did the recognition task. For the purpose of recognition, a dimension is processed only if a comparison with information retrieved from memory has been carried out (see also Lamberts & Freeman, 1999). If recognition performance suggests

curves for the slow–similar pairs are clearly nonmonotonic, with a peak between 400 and 600 ms. The observed curve for the fast–similar pairs was less regular and not obviously nonmonotonic. The predicted curve for these pairs was almost linear, and it closely approximated the best-fitting linear regression line for the data, as well as the best-fitting second- and third-degree polynomials.

The model predictions for the sequential matching task were also good (see Figure 5B). As for the simultaneous matching task, the fit to the data from different trials was nearly perfect. The

Table 5
Best-Fitting Model Parameters for Data From Experiments 2, 3, and 4

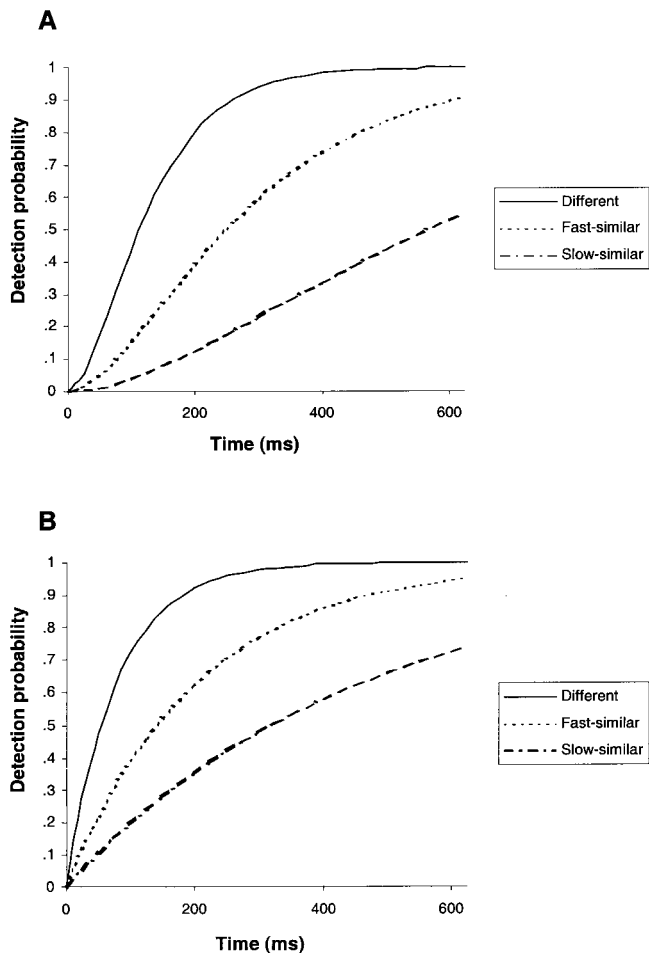| Parameter	Value
q	0.0079
q_{crit}(fast)	0.0049
q_{crit}(slow)	0.0022
t_{res}	218.5
ρ	0.280
δ	1.617

A

B

Figure 6. Detection probabilities of any difference between different, fast–similar, and slow–similar pairs for the simultaneous (A; Experiment 2) and sequential (B; Experiment 3) matching tasks, plotted as a function of processing time.

that a particular dimension has not been processed, this could be due to a failure of perceptual processing (the feature may not have been seen), a failure of retrieval (the feature may not have been stored in the first place, or could somehow not be retrieved), or a failure of both. Therefore, the recognition task measures an aggregate of perceptual and retrieval components. The matching tasks presumably provide a much cleaner measure of purely perceptual processing times, because both match stimuli on any trial were always present when the same–different decision was made. The memory demands in the matching task were thus kept to a minimum.

We tested whether ρ was an essential parameter by carrying out a series of model comparisons. First, we fit a model version to the data from the three experiments, in which ρ was clamped at a value of 1.00. This model version thus assumed that the processing rates in the recognition task were the same as those in the matching tasks. A likelihood-ratio test (see Lamberts, 1995) showed that this version fit the data much worse than the general model version, $\ln(L) = -350.667$, $\chi^2(1) = 420.146$, $p < .01$. This reliable difference indicates that the time course of processing in the

matching and recognition tasks is different and that this difference should be an essential component of the model. Second, we tested whether the difference in time course could be explained in terms of a difference in the residual time between the tasks; perhaps the time course of perceptual processing was identical in matching and recognition, but the time course of decision making (including retrieval) and response production was different. To verify this, we applied a model version in which the same processing rates applied to the matching and recognition tasks, but in which a separate residual time parameter was estimated for the recognition task. This model version also fit the data much worse than the standard version, $\ln(L) = -203.266$, $R^2 = .825$, despite having the same number of free parameters. (Note that a likelihood-ratio test cannot be applied here, because the two model versions that are compared are not hierarchically related.)

The inclusion of ρ as a multiplier of the matching processing rates thus seems essential, and it implies a strong theoretical statement about the relation between purely perceptual processing and combined perception and retrieval. The model states that the combined perception–retrieval time is proportional to the perceptual processing time. The estimated value for ρ of 0.28 indicates that the combined perception–retrieval time is about 3.57 times the perception time. To verify the multiplicative assumption, we applied a model version in which the relation between processing rates in matching and recognition was unconstrained. This was achieved by estimating three separate processing rates for the recognition task and three separate processing rates for the matching tasks. This model, with a total of eight free parameters, did not fit the data reliably better than the standard model with six parameters, $\ln(L) = -140.338$, $\chi^2(2) = 0.512$, *ns*. In fact, the improvement in goodness-of-fit was minimal. Figure 7 presents the processing rates estimated by the standard model and the rates estimated by the more general, eight-parameter model. The estimates are virtually identical, which lends support to the assumption about the multiplicative relation between processing rates in matching and recognition.

In processing terms, it is not entirely clear why this multiplicative relation emerges. A first possibility is that the retrieval time of
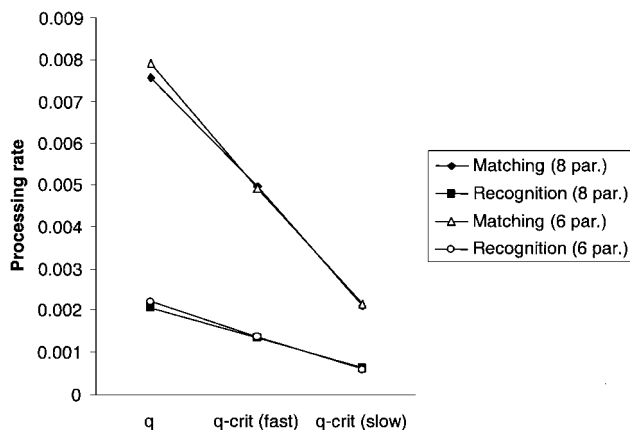

Figure 7. Processing rates for critical stimulus dimensions of different, fast–similar, and slow–similar stimuli estimated by the standard six-parameter model and by the more general, eight-parameter model. par. = parameters.

features is proportional to their perceptual processing time, such that features that are quickly perceived (e.g., very salient features) are also quickly retrieved. A second possibility is that retrieval times are constant across features but that perceptual processing times themselves differ between matching and recognition. In the sequential matching task, processing of the second stimulus may have been primed by the first stimulus. In the simultaneous matching task, some form of copriming may also have induced faster processing rates. Note that a similar principle could explain the ubiquitous fast–same effect in matching (see Van Zandt et al., 2000), which was incidentally not observed in Experiment 1. In recognition, any priming effects are bound to be much weaker, because the stimuli that are compared are separated by a longer time interval and by other stimuli.

The difference in absolute processing rates between the matching and recognition tasks should certainly not be taken as evidence against our assumption about the role of perceptual processing in recognition. On the contrary, the proportionality of the rates in the different tasks suggests that the tasks are intimately linked, in the sense that the variables that determine processing rates in matching also determine processing rates in recognition. The classification into fast–similar and slow–similar items on the basis of matching RTs in Experiment 1 proved to have great predictive value for response patterns in all the other experiments.

Alternative Accounts

An explanation of the data in terms of perceptual processing rates is certainly not the only possibility. Perhaps the link between perceptual matching and recognition memory is not as direct as we assumed but exists only by virtue of a hitherto unexplored process or mechanism. One obvious possibility is that the response patterns at different signal intervals in the recognition task do not reflect the time course of perceptual processing but reflect the time course of a different process.

This assumption is central to dual-process accounts of recognition memory (e.g., Atkinson & Juola, 1973, 1974; Hintzman & Curran, 1994; Mandler, 1980). Dual-process models assume that recognition judgments reflect the outcome of two distinct processes. The first process is a fast familiarity-based process that yields information very quickly but that may not be very accurate. The second process is a slower recall process, which is assumed to be more accurate than the familiarity-based process. Dual-process models are compatible with a wide range of findings (e.g., Hintzman & Curran, 1994; Hintzman, Curran, & Oppy, 1992; Jacoby, 1991; Ratcliff, Clark, & Shiffrin, 1990; Ratcliff, Sheu, & Gronlund, 1992; Yonelinas, 1994; Yonelinas & Jacoby, 1994). Although dual-process models have not been implemented formally (but see Brockdorff & Lamberts, 2000, for an initial attempt), a dual-process account could probably explain some aspects of the data from our recognition experiment. The nonmonotonic response functions for similar items that we observed are similar to the nonmonotonic functions in the experiments of Hintzman and Curran (1994). The dual-process interpretation of these functions is that the initial rise in "old" responses is due to the fast familiarity process, whereas the subsequent decline of "old" response rates reflects the slower but more accurate recall process. In addition, to explain the difference in recognition response functions between fast–similar and slow–similar items, a dual-process model would

have to assume that the recall process takes longer for the slow–similar items than for the fast–similar items. However, within a dual-process account, this assumption would be entirely ad hoc and without theoretical foundation. It seems fair to conclude, therefore, that a dual-process account does not explain the relation between perception and recognition, and that dual-process models provide at best an incomplete account of the time course of recognition.

The main advantages of our perceptual-processing account compared with the dual-process explanation are its parsimony and its productivity. The perceptual-processing model is parsimonious, because it relies on only a single process that is not substantially different between tasks. The account is also productive, because it predicts very specific and systematic links between tasks (in this case, matching and recognition, but one could also include categorization) that would not be seen as related in other accounts. Therefore, we would propose that a single-process account of our data is quite sufficient.

Another alternative explanation of our results would be in terms of storage and/or retrieval probabilities of critical features. Perhaps features that are processed slowly in the matching tasks are less likely to be stored in memory in the recognition task or less likely to be retrieved. Whereas this explanation has some intuitive appeal, it cannot explain the complex response patterns for similar stimuli in the recognition task. Without an additional process, this account could not predict nonmonotonic response functions for the similar stimuli in the recognition task. And if an additional process were included (such as the fast familiarity-based process from the dual-process account), the model would still be bound to predict that the recognition response functions for the fast–similar and slow–similar stimuli are linearly related. This is clearly not the case; the response function for the fast–similar items peaks earlier than that for the slow–similar items, for instance. We have carried out formal model comparisons, which confirm that an explanation in terms of differential storage and/or retrieval probabilities is not feasible.

Our proposal for a perceptual front-end to a theory of recognition is related (but not identical) to a theory of visual information acquisition developed by Busey, Loftus, and McLean (e.g., Busey & Loftus, 1994, 1998; Loftus & McLean, 1999). Of particular interest is the recent article by Loftus and McLean (1999) in which they apply the theory to picture recognition. According to the theory, the initial stages of the visual system act as a linear temporal filter that generates a sensory response from the physical stimulus. This is followed by an information-sampling process, the sampling rate of which is based in part on the magnitude of the sensory response. Loftus and McLean (1999) tested the theory in four recognition experiments in which they manipulated stimulus duration, stimulus contrast, the duration of a gap between successive stimulus presentations, and the presence of a mask following stimulus presentation. Unlike our own experiments, these manipulations were carried out during only the memorization stage of the experiments. Subsequent old–new recognition testing involved no time pressure. Loftus and McLean's (1999) model provided an excellent account of their experimental data.

We see Loftus and McLean's (1999) model as largely complementary to our own account. The perceptual processing component in the Loftus model is more sophisticated than our own and can almost certainly better account for fine-grained effects in the time

course of perception. However, FESTHER contains a more elaborate decision and performance component, which is necessary to explain many of the results that we have reported previously (see Brockdorff & Lamberts, 2000). Therefore, a promising avenue for future research would be to integrate the perceptual assumptions of Loftus et al.'s theory into FESTHER.

The fact remains, however, that FESTHER in its current form does not specify the nature of perceptual processing in recognition in great detail. This relative lack of specificity is inevitable, given our limited understanding of the time course of information accumulation in the perception of complex visual stimuli. The model's assumption that perception involves stochastic accumulation of information about stimulus features is a very general one, which is compatible with a wide range of possible specific processes. For instance, the model can accommodate psychophysical evidence for temporal asynchrony in visual perception, which shows that different attributes of visual stimuli are consciously perceived at different times after stimulus onset (e.g., Zeki & Bartels, 1998). When participants are asked to pair two rapidly alternating states of two attributes, such as color and orientation of a line, they make consistent binding errors. Using this technique, it has been shown that color is perceived about 30 or 40 ms before orientation, which in turn is perceived faster than motion (Moutoussis & Zeki, 1997). These asynchronies have been attributed to the functional organization of the visual system, with different visual attributes being processed in different brain regions (Moutoussis & Zeki, 1997). However, little is known about asynchronies in the processing of more complex visual properties in complex visual stimuli. The data we reported in this article, together with the findings from previous experiments (e.g., Lamberts & Freeman, 1999), suggest that these asynchronies typically involve larger timescales than those for simple visual properties.

Finally, we would like to discuss some of FESTHER's specific processing assumptions and consider alternatives to these assumptions. First is the assumption concerning the exponential distribution of feature sampling times. The model-fitting results from Experiments 2–4 suggest that this distribution is appropriate; the exponential model predicts response functions that have the correct shape for the data. Although it is possible that other distributions could also provide good fits, the exponential distribution has many advantages. It has been used successfully in other models of information accumulation in perception and decision making (see Townsend & Ashby, 1983). Lamberts (2000) has shown that exponential processing functions can provide the basis not only for the correct prediction of choice proportions and mean RTs in categorization but also for the prediction of RT distributions and latency–accuracy functions. Lamberts and Freeman (1999) explicitly compared an exponential-distribution model of the time course of feature sampling in categorization with an alternative model in which processing probabilities of features were required only to increase monotonically over time. Despite the latter model's greater flexibility and larger number of free parameters, it produced only a negligible increase in goodness-of-fit compared with the standard exponential model. Therefore, we believe that the assumption about exponential distribution of processing times is reasonable and that it is unlikely that an alternative distribution would yield significantly better model fits.

The second model assumption concerns the stochastic independence of the sampling processes for different features. In previous research on feature sampling in object perception, the assumption of sampling independence has often been supported (e.g., Lamberts & Freeman, 1999; Townsend, Hu, & Ashby, 1981; Wandmacher, 1976; Wandmacher, Kammerer, & Glowalla, 1980), although there have also been reports indicating that sampling independence does not hold universally (e.g., Townsend & Ashby, 1982; Townsend, Hu, & Evans, 1984). The data from the current experiments did not allow us to test the independence assumption (see Lamberts & Freeman, 1999, for a description of the methodology of such a test). The independence assumption can be tested by comparing the goodness-of-fit of models that make the assumption with models that do not make the assumption. Such a comparison requires a data set in which stimuli contain different combinations of features (as in a factorial design), which cannot be achieved with the realistically complex stimulus materials that we used in the current experiments. We do not see this as an important shortcoming, however. The critical result of our experiments, which concerns the systematic relation between the time course of perception and the time course of recognition, could also have been predicted from a model in which feature sampling was not independent. In any case, an independence model is simpler than a model that allows for dependencies in feature sampling (see Lamberts & Freeman, 1999), and the fact that the independence model that we applied provided a good account of the data suggests that little could be gained from exploring models that abandon the independence assumption.

Conclusions

The experiments in this article and the modeling results lead to the following conclusions. First, there is a systematic empirical relation between the time course of perceptual matching and the time course of recognition judgments. Stimuli that are compared rapidly are also recognized more accurately when processing time is limited. Second, this relation between the two tasks can be understood in terms of the processing rates of the stimulus dimensions. The model applications confirm that our assumptions about feature sampling are sufficient to explain the results from all the experiments. An important assumption in the modeling was that processing rates in recognition are proportional to those in matching.

Our results have general implications for process models of recognition. They indicate that the time course of perceptual processing could be an important element of an account of the time course of recognition. Stimuli that have the same logical status (such as the similar stimuli in our recognition experiment) can still produce very different response patterns, depending on how quickly their features are processed. This implies that process models of recognition can be complete only if they include some form of analysis of the perceptual structure of the stimuli.
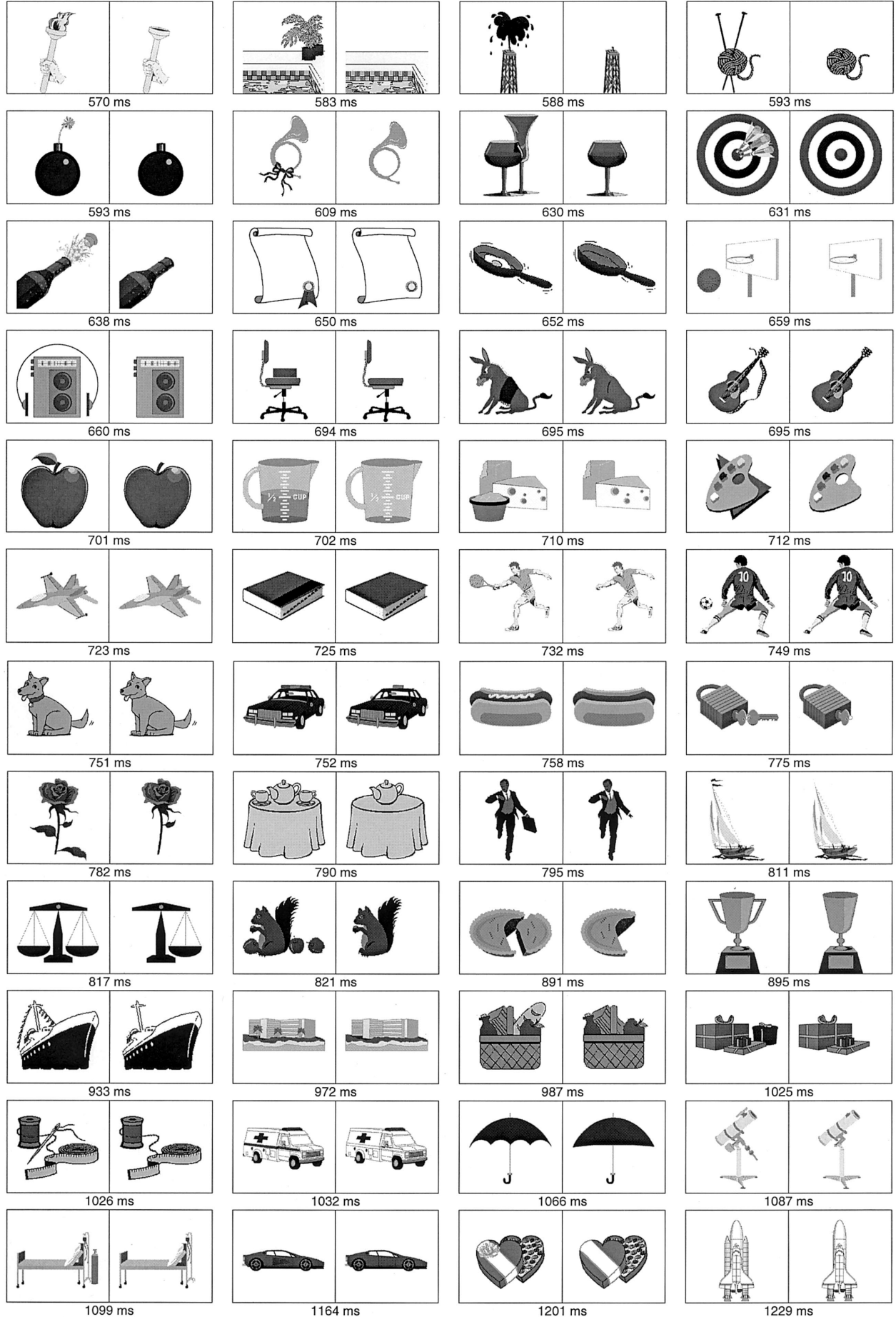
References

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 583–612). New York: Academic Press.

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychol-*

ogy: Vol. 1. Learning, memory and thinking (pp. 243–293). San Francisco, CA: Freeman.

Brockdorff, N., & Lamberts, K. (2000). A feature-sampling account of the time course of old–new recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 77–102.

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

Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review, 101,* 446–469.

Busey, T. A., & Loftus, G. R. (1998). Binocular information acquisition and visual memory. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1188–1214.

Cohen, A. L., & Nosofsky, R. M. (2000). An exemplar-retrieval model of speeded same–different judgments. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1549–1569.

Dosher, B. A., & Rosedale, G. (1991). Judgments of semantic and episodic relatedness: Common time-course and failure of segregation. *Journal of Memory and Language, 30,* 125–160.

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

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

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language, 33,* 1–18.

Hintzman, D. L., Curran, T., & Oppy, B. (1992). Effects of similarity and repetition on memory: Registration without learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 667–680.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences.* Pacific Grove, CA: Brooks/Cole.

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

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

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

Lamberts, K. (2002). Feature sampling in categorization and recognition of objects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55*(A), 141–154.

Lamberts, K., & Brockdorff, N. (1997). Fast categorization of stimuli with multi-valued dimensions. *Memory & Cognition, 25,* 296–304.

Lamberts, K., & Freeman, R. P. J. (1999). Building object representations from parts: Tests of a stochastic sampling model. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 904–926.

Loftus, G. R., & McLean, J. E. (1999). A front end to a theory of picture recognition. *Psychonomic Bulletin & Review, 6,* 394–411.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin, 107,* 401–413.

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98,* 185–199.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

Moutoussis, K., & Zeki, S. (1997). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society of London, Series B, 264,* 1407–1414.

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

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.

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

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

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 163–178.

Ratcliff, R., Sheu, C.-F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99,* 518–535.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95,* 318–339.

Rotello, C. M., & Heit, E. (1999). Two-process models of recognition memory: Evidence for recall-to-reject? *Journal of Memory and Language, 40,* 432–453.

Townsend, J. T., & Ashby, F. G. (1982). Experimental test of contemporary mathematical models of visual letter recognition. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 834–864.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes.* Cambridge, England: Cambridge University Press.

Townsend, J. T., Hu, G. G., & Ashby, F. G. (1981). Perceptual sampling of orthogonal straight line features. *Psychological Research, 43,* 259–275.

Townsend, J. T., Hu, G. G., & Evans, R. J. (1984). Modeling feature perception in brief displays with evidence for positive interdependencies. *Perception & Psychophysics, 36,* 35–49.

Van Zandt, T, Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review, 7,* 208–256.

Wandmacher, J. (1976). Multicomponent theory of perception: Feature extraction and response decision in visual identification. *Psychological Research, 39,* 17–37.

Wandmacher, J., Kammerer, E.-M., & Glowalla, U. (1980). Context dependence in visual feature processing. *Psychological Research, 42,* 335–351.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1341–1354.

Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology, 48,* 516–534.

Zeki, S., & Bartels, A. (1998). The asynchrony of consciousness. *Proceedings of the Royal Society of London, Series B, 265,* 1583–1585.

(Appendixes follow)

Appendix A

Mean RTs on Correct Responses Across Participants for Similar Pairs in Experiment 1

570 ms	583 ms	588 ms	593 ms
593 ms	609 ms	630 ms	631 ms
638 ms	650 ms	652 ms	659 ms
660 ms	694 ms	695 ms	695 ms
701 ms	702 ms	710 ms	712 ms
723 ms	725 ms	732 ms	749 ms
751 ms	752 ms	758 ms	775 ms
782 ms	790 ms	795 ms	811 ms
817 ms	821 ms	891 ms	895 ms
933 ms	972 ms	987 ms	1025 ms
1026 ms	1032 ms	1066 ms	1087 ms
1099 ms	1164 ms	1201 ms	1229 ms

Appendix B

Derivation of Probability of Detection of Any Difference Between Two Simultaneously Presented Stimuli

The probability that dimension x has been processed in Stimulus A within time interval t_i is given by an exponential distribution function:

$$p(x\ processed,\ A) = 1 - \exp(-q_x t_i).$$

Given that the probability of dimension x having been processed in Stimulus A is independent of dimension x having been processed in Stimulus B, the probability that x has been processed in both A and B is given by

$$p(x\ processed,\ A\ and\ B) = [1 - \exp(-q_x t_i)]^2.$$

It is assumed that if any difference exists between two stimuli on a dimension x, then the difference will be detected if dimension x has been processed in both stimuli. It follows that the probability that no difference is detected on dimension x on both stimuli is

$$1 - p(x\ processed,\ A\ and\ B) = 1 - [1 - \exp(-q_x t_i)]^2.$$

If it is assumed that the processing rates for all δ dimensions are equal and independent, it follows that the probability that no difference is detected between two simultaneously presented different stimuli in time interval t_i is

$$p_{not\ det}(simultaneous,\ different,\ t_i) = \{1 - [1 - \exp(-q t_i)]^2\}^\delta.$$

From the above, it follows that the probability that any difference between two simultaneously presented stimuli has been detected in time t_i is given by

$$p_{det}(simultaneous,\ different,\ t_i) = 1 - p_{not\ det}(simultaneous,\ different,\ t_i)$$

$$= 1 - \{1 - [1 - \exp(-q t_i)]^2\}^\delta.$$