

One process is not enough! A speed-accuracy tradeoff study of recognition memory

ANGELA BOLDINI, RICCARDO RUSSO, and S. E. AVONS
University of Essex, Colchester, England

Speed-accuracy tradeoff (SAT) methods have been used to contrast single- and dual-process accounts of recognition memory. In these procedures, subjects are presented with individual test items and are required to make recognition decisions under various time constraints. In this experiment, we presented word lists under incidental learning conditions, varying the modality of presentation and level of processing. At test, we manipulated the interval between each visually presented test item and a response signal, thus controlling the amount of time available to retrieve target information. Study-test modality match had a beneficial effect on recognition accuracy at short response-signal delays (≤ 300 msec). Conversely, recognition accuracy benefited more from deep than from shallow processing at study only at relatively long response-signal delays (≥ 300 msec). The results are congruent with views suggesting that both fast familiarity and slower recollection processes contribute to recognition memory.

In a standard recognition memory task, a list of items is presented at learning, and at test both previously presented and new items are provided so that subjects have to decide whether or not each item was presented at study. Recognition memory accuracy is assessed by measures that take into account the difference between the proportion of previously presented items declared "old" (i.e., hit rate) and the proportion of new items called "old" (i.e., false alarm rate).

Two general classes of theories have been put forward to explain performance in recognition memory tasks. One class suggests that the test item gives rise to a scalar familiarity value, and the stimulus is classified as old if this exceeds some criterion value. This general approach to recognition memory is common to a group of otherwise diverse "global-matching" models (for a review, see Clark & Gronlund, 1996).

A second class of theories, the dual-process models, posit that a single continuous quantitative index of memory strength is not sufficient to explain recognition memory (e.g., Jacoby, 1991; Mandler, 1980). In addition to a familiarity-based component, these models include some form of *recollection* process. Despite disagreements between the models (see Yonelinas, 2002, for a review), familiarity is generally characterized as acting rapidly and automatically on the perceptual features of target items. Recollection is considered to be a slower, intentional, process, usually based on the semantic characteristics of target items and specific qualitative information about their prior occurrence (e.g., Jacoby, 1991; Mandler, 1980).

Some accounts claim that familiarity and recollection influence recognition memory independently (Jacoby, 1991; Mandler, 1980).

Because familiarity and recollection processes are presumed to operate at different rates, one critical manipulation is to regulate the time allowed for retrieval before the subject makes a decision. Various speed-accuracy tradeoff (SAT) techniques exist, all of which force the subject to make speeded or slow recognition decisions and compare recognition as a function of response speed. One method to achieve graded response speeds is to provide a signal after the recognition probe that requires an immediate response, as in the response-signal procedure (e.g., Hintzman & Curran, 1994; Mulligan & Hirshman, 1995). Thus, with this technique it is possible to plot increments in recognition memory accuracy as a function of the time to respond to the recognition probe. Hence, if familiarity and recollection are indeed two distinct processes with different time courses, then variables that selectively affect familiarity or recollection should modify recognition accuracy under different response deadlines. For example, if familiarity is a fast-acting process based on perceptual characteristics, recognition memory should be optimal when the perceptual characteristics of target and test stimuli match and the effect of perceptual matching should be observed at relatively short response-signal lags (e.g., Toth, 1996). Conversely, if recollection is a slow-acting process based on semantic or contextual information, recognition should be optimal under conditions that promote either semantic or elaborative encoding, and these effects should be observed under relatively long response-signal deadlines (e.g., Mulligan & Hirshman, 1995).

With respect to these hypotheses, the available evidence is only partially supportive. With respect to the ef-

Correspondence should be addressed to R. Russo, Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, England (e-mail: rrusso@essex.ac.uk).

fect of perceptual matching, Mulligan and Hirshman (1995) and Toth (1996) reported a recognition advantage at short response-signal lags when the modality of study and test stimuli matched. However, Hintzman and Caulton (1997; see also Jones, Jacoby, & Gellis, 2001) did not find any such advantage at short response-signal lags and reported that asymptotic performance was higher in the mismatch condition. Of the studies reporting a modality match advantage, only Toth's (1996) study reported that the effect was larger at short response-signal lags.

Other studies have used the response-signal deadline procedure while manipulating levels of processing to assess the contribution of recollection. However, the well-known recognition advantage for deep levels of processing interacted with response-signal lag in only one study (Mulligan & Hirshman, 1995), but was present at all lags tested in other studies (the shorter response-signal lag used in these studies was 500 msec) (Gardiner, Ramponi, & Richardson-Klavehn, 1999;¹ Gillund & Shiffrin, 1984; Toth, 1996). Thus, the results obtained using the response-signal procedure do not provide consistent support for a dual-process model of recognition where the two processes operate over different time scales.

More critically for the dual-process account, Mulligan and Hirshman (1995) used an exponential growth-to-a-limit function to model the increase in recognition with response-signal lag as a function of level of processing and the match between study and test modality. They found an acceptable fit to their data when the asymptotic level of performance was allowed to vary across conditions, but no further advantage was gained when lag-related parameters such as rate of increase and intercept on the time axis were varied. In their view, the effect of experimental manipulations such as level of processing and the matching of modality between study and test is to change only the total amount of information available in memory, not the time course over which this information accrues. The conclusion from their modeling is that recognition information accumulates over time at a constant growth rate in all conditions, supporting a single-process, familiarity-based model of recognition.

However, studies reporting no difference in the time course of perceptual match or levels-of-processing information at test face several criticisms. Toth (1996) tested only 6 subjects per cell, and thus the failure to detect a levels-of-processing \times response-signal lag interaction, which would provide evidence for a recollection process, may have simply reflected low statistical power. Similarly, in Gardiner et al.'s (1999) study, the effect of levels of processing was numerically smaller in the fast than in the slow response condition used in the experiment, but no statistical analysis on the interaction was provided. Mulligan and Hirshman (1995, Experiment 2) and Hintzman and Caulton (1997) both used the standard response-signal procedure, in which response lag was varied randomly from trial to trial. The advantage of this technique is that it equates retrieval strategies at the point when the recognition probe appears. However, to

achieve stable measures, several sessions of extensive testing are required. This precludes the possibility of testing under incidental conditions and thus also the conventional manipulation of levels of processing (Mulligan & Hirshman, 1995, Experiment 3). Moreover, the use of randomized response lags does not prohibit a shift in response strategy after the recognition probe is presented, nor does it ensure that responding will be optimal at each lag. Subjects may adopt a strategy that is accurate at longer lags and still suffer the failures that occur with fast responses. Thus the use of a within-subjects response-signal procedure may produce encoding and retrieval strategies that encourage uniformity. While acknowledging the merits of the standard response-signal procedure, we propose that alternative SAT procedures such as the one used here are more likely to optimize responding as a function of response delay and are more likely to provide a reflection of the availability of information after probe onset.

The aim of the present study was to reassess the dual-process account, modifying the response-signal procedure and manipulating levels of processing and modality matches. In particular, to overcome some of the objections raised above, target items were studied under incidental conditions and each subject was tested at only one response-signal lag (e.g., Mulligan & Hirshman, 1995, Experiment 1; Toth, 1996). In a further effort to increase the speed and facility of responding, we used a go/no-go paradigm in which subjects responded only if they recognized a probe item as having been presented during the incidental learning phase. Responses of this kind are both faster and more accurate than traditional two-choice responses in a variety of domains (e.g., Perea, Rosa, & Gómez, 2002). Mulligan and Hirshman (1995, Experiment 1) and Toth (1996) used only two response-signal conditions. The fastest responses occurred about 850 msec after the appearance of the test stimuli. In the present study, we used a wider range of response lags, especially short ones, to provide a more thorough assessment of the effect of levels of processing and the study-test modality match at different times after probe onset. To this aim, we used seven response-signal lags (100, 150, 200, 300, 1,500, 2,250 and 3,000 msec). If a familiarity process based on the perceptual characteristics of target items supports recognition, we expect to detect effects of study-test modality match at short response-signal lags, whereas recollection-based effects, influenced by levels of processing, should emerge at longer lags.

METHOD

Subjects

One hundred twelve native English speakers, mainly undergraduate students recruited at Essex University, took part in this experiment. They were either paid or given credits for their participation in the study.

Design

The experiment used a three-factor design. Levels of processing (deep vs. shallow) and study-test modality match (auditory-visual

vs. visual–visual) were manipulated within subjects. The four study conditions made by combining these two variables were given as four blocks of 20 trials, and the order of these was counterbalanced across subjects. Response-signal lags (100, 150, 200, 300, 1,500, 2,250, and 3,000 msec) were manipulated between subjects (16 subjects were tested at each response-signal lag). Levels of processing and study–test modality match were counterbalanced across subjects in an orthogonal way. At study, words were incidentally learned in each of the four possible conditions (i.e., semantic vs. nonsemantic \times visual vs. auditory) in blocks of 20 words each. We used four different orders of the possible ways in which the study conditions could be combined. These were (1) A–S, V–S, A–N, V–N; (2) V–N, A–N, V–S, A–S; (3) V–S, V–N, A–S, A–N; (4) A–N, A–S, V–N, V–S (where A = acoustic; V = visual; S = semantic; N = nonsemantic). Four subjects in each response lag condition were randomly allocated to each of the four orders of conditions.

Material

The words used were 184 nouns from 4 to 8 letters in length and ranging from 50 to 200 occurrences per million according to the Kučera and Francis (1967) norms. Of these, 24 nouns were used as buffer words in the study phase, 3 at the beginning and 3 at the end of each block. Eighty nouns were used as target stimuli for the study phase and 80 nouns as foils for the test phase. These were randomly selected from among the remaining 160 nouns. This process was repeated 16 times, so that a set of 16 different study and test lists were created to be used at each time lag. The 80 nouns selected for the study phase were then divided in four sets of 20 words for each of the four learning conditions; word frequencies were comparable across sets.

Visual stimuli were presented at the center of a computer screen in 90-point lowercase Trebuchet MS font and preceded by a cross as fixation point. In the training and test phases, the signal for the answers to be given was a row of asterisks appearing two lines below the test word. Four hundred milliseconds after the row of asterisks appeared, a beeping sound alert was played to indicate to the subjects that the time allowed for answering had elapsed. Acoustic stimuli were spoken in a male voice.

Procedure

The experiment was conducted in three phases: the study phase, training at the response-signal procedure, and test. Each section began with appropriate instructions given on the screen.

At study, visual and acoustic target words were presented with a stimulus onset asynchrony (SOA) of 3 sec. Visual targets were displayed for 2 sec with a 1-sec interstimulus interval. The presentation of acoustic words was never longer than 2 sec. The deep orienting task consisted of the rating of the pleasantness of each of the presented words on a 3-point scale (*pleasant*, *neutral*, or *unpleasant*). The shallow orienting task consisted of repeating aloud each presented word. Before each block of 20 words, subjects received detailed task instructions. Responses were spoken aloud by the subject. Subjects were told that their responses would be recorded, but in fact these were not recorded. No mention was made of any subsequent memory test.

The training phase provided the subjects with some practice with the response-signal procedure. The stimuli were 20 nonwords ranging from 6 to 9 letters in length. A set of arrows (>>> <<<<), indicating the future location of the test string, appeared at the center of the screen for 500 msec. Then a nonword appeared and, after the appropriate time lag (the same as for the response-signal lag condition), a row of asterisks appeared two lines under the test string. The task was to decide whether the letter *t* was present in the nonword, and subjects were instructed to respond as soon as possible once they saw the row of asterisks. They were told to press the “z” key if they saw the target letter and to do nothing otherwise. A beeping signal was emitted whenever an answer was given after 400 msec from the appearance of the response signal. Between consecutive

items, there was a pause of 2 sec. This training phase lasted up to 2 min (depending on the response-delay condition), and only when subjects were confident enough with the procedure did they go on to the next phase.

All the words in the test phase were presented visually. This was done to facilitate fast responding, especially at short response-signal delays. Subjects were asked to decide for each test word if it had been presented during the study phase. Test words were presented following the same procedure used during the training phase. Again, subjects had only to respond “yes” (pressing the “z” key) if they thought they had already seen or heard that particular word in the study phase. They were encouraged to be correct, but to respond on time was more important. As in the previous phase, answers were considered as given on time when the latency was not longer than 400 msec.

RESULTS

Recognition Memory Accuracy

Only answers given “on time”—that is, within 400 msec from the response signal, were considered for the analysis. Mean percentages of hits, false alarms, and corrected d' scores (Snodgrass & Corwin, 1988) for each group and condition are presented in Table 1 (see also Figure 1).²

A 2 (levels of processing: deep vs. shallow) \times 2 (study–test modality match: match vs. mismatch) \times 7 (response-signal delay: 100, 150, 200, 300, 1,500, 2,250, and 3,000 msec) mixed analysis of variance (ANOVA) was conducted on the d' scores. This analysis showed a significant main effect of delay [$F(6,105) = 14.90, MS_e = 0.94, p < .01$], indicating that recognition accuracy increased as a function of response-signal lag. Both the levels of processing [$F(1,105) = 113.57, MS_e = 0.16, p < .01$] and the study–test match [$F(1,105) = 90.05, MS_e = 0.10, p < .01$] main effects were significant, indicating that accuracy was higher for deeply encoded words and for words studied and tested in the same modality. More interestingly, both the study–test match \times delay and the levels-of-processing \times delay interactions were significant [$F(6,105) = 3.55, MS_e = 0.1, p < .01$, and $F(6,105) = 9.22, MS_e = 0.16, p < .01$, respectively]. Figures 2 and 3 provide graphs of these interactions. No other interaction approached significance ($F < 1$).

Two sets of paired comparisons were performed to further investigate these significant interactions. A Bonferroni correction was applied to each set of comparisons to keep the family-wise error rate at .05 for each set (i.e., critical $\alpha = .007$). Significantly greater recognition accuracy was detected in the study–test match conditions at 100-, 150-, 200-, and 300-msec response-signal delays but not at longer response-signal delays. On the other hand, significantly greater accuracy was found for deeply encoded items only at response-signal delays ≥ 300 msec. Overall, the form of the two significant interactions appears to be congruent with dual-process accounts suggesting that a fast familiarity process and slower recollection support recognition memory.

Modeling

The average response times for the four conditions at each response-signal delay are reported in Table 2. Adding

Table 1
Proportions of Hits and False Alarms (FAs), and d' Scores for Each of the Conditions of the Experiment at Each Response-Signal Delay

Response Signals	Deep				Shallow				FAs
	Match		Mismatch		Match		Mismatch		
	Hits	d'	Hits	d'	Hits	d'	Hits	d'	
100 msec	.46	0.684	.33	0.331	.41	0.539	.30	0.236	.23
	.17	0.36	.15	0.28	.18	0.42	.16	0.38	.53
150 msec	.54	0.983	.47	0.763	.51	0.914	.39	0.579	.21
	.17	0.55	.15	0.49	.23	0.49	.16	0.45	.52
200 msec	.60	1.136	.44	0.721	.58	1.072	.35	0.473	.21
	.18	0.57	.13	0.38	.17	0.38	.14	0.40	.48
300 msec	.74	1.572	.58	1.124	.62	1.223	.49	0.909	.19
	.13	0.51	.10	0.39	.14	0.33	.17	0.40	.44
1,500 msec	.85	2.169	.77	1.848	.59	1.289	.49	1.035	.18
	.14	0.86	.20	0.82	.17	0.71	.17	0.49	.56
2,250 msec	.83	2.036	.75	1.845	.63	1.394	.58	1.291	.16
	.11	0.43	.18	0.61	.20	0.43	.22	0.66	.38
3,000 msec	.81	2.004	.77	1.842	.56	1.193	.58	1.245	.18
	.17	0.87	.14	0.91	.18	0.81	.16	0.73	.41

Note—Standard deviations are provided under each relevant score.

each response time to the associated response-signal delay gives the average times to respond “old” to test words. This is the variable t of the exponential to a limit function used to model the data:

$$d' = A\{1 - \exp[-R(t - I)]\},$$

where A is the asymptotic recognition accuracy, R is the rate of growth, and I is the intercept, the time at which recognition performance starts to grow. Because subjects performed only at single response-signal delays, we could apply the function only to fit the mean data shown in Figure 1. Data fitting was done for each of the four experimental conditions of the study. In this process, SPSS was used to estimate the parameters of the exponential function—that is, A , R , and I —to maximize the proportion of variability accounted, that is, r^2 . The values of r^2 for the

overall model, where all the parameters for each of the four experimental conditions were left free to vary, as well as for the reduced models, where some parameters were fixed, were adjusted as a function of the number of free parameters according to the approach suggested by Reed (1976):

$$r^2 = 1 - \frac{\sum_{i=1}^N (d_i - \hat{d}_i)^2}{\sum_{i=1}^N (d_i - \bar{d})^2} \cdot \frac{N - k}{N - 1},$$

where N is the number of data points d_i , each d_i corresponds to an observed d' score, each \hat{d}_i value corre-

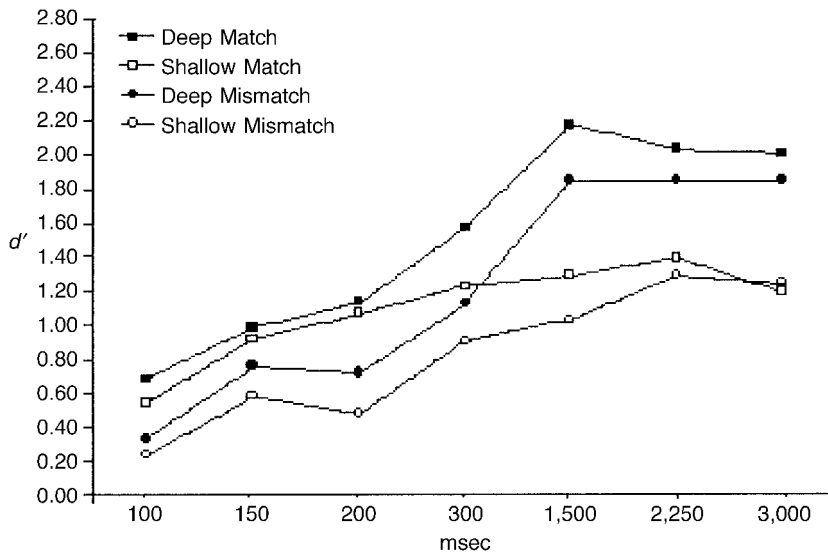


Figure 1. Mean d' scores for each of the conditions of the experiment at each response-signal delay.

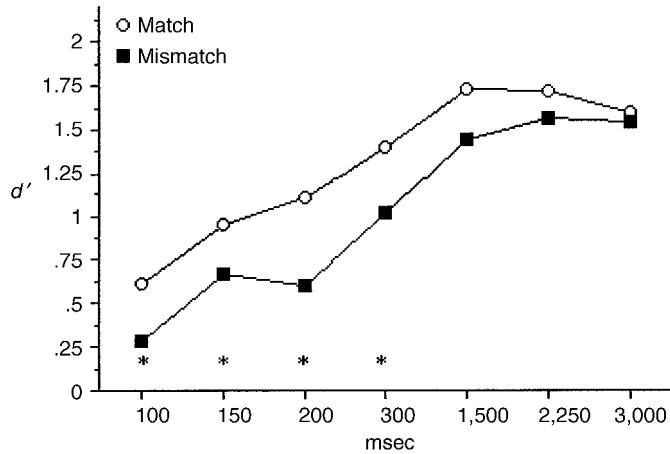


Figure 2. Graph of the study-test match \times response-signal delay interaction. The asterisks indicate that the difference in recognition accuracy between match and mismatch conditions was significant ($\alpha = .007$).

sponds to the d' score predicted by the model, \bar{d} is the overall mean recognition accuracy, and k is the number of free parameters.

Table 3 provides the parameters for the best-fitting full model. This model allowed all 12 parameters to vary and it accounted for about 96.6% of the variability in the data. We then tried to fit the model to the data by allowing only the asymptotic performance parameter to vary. Both the rate of the increment in d' (R) and the intercept (I) were fixed across the four experimental conditions. This asymptote-only model reflects a single process with constant rate and intercept. The asymptote can be taken as a measure of the information available in memory, or equivalently of its strength. Thus, if the asymptote model can adequately fit the accuracy data, it would appear that

the main difference between the experimental conditions resides only in the strength of the memory trace they generated. From Table 3, it appears that the asymptote-only model accounted for about 90.8% of the variability in the data. The difference in the proportion of the variability accounted for by the full and the asymptote-only model was significant [$F(6,15) = 4.30, p = .01$]. Thus the asymptote-only model provided a less accurate fit than the full model.

A second model was then fitted to the data. Two different intercept parameters were fitted to the match and mismatch conditions, reflecting the anticipated earlier onset of recognition under the match condition. Similarly, two different rate parameters were fitted to the two processing conditions, where a slower rate of growth

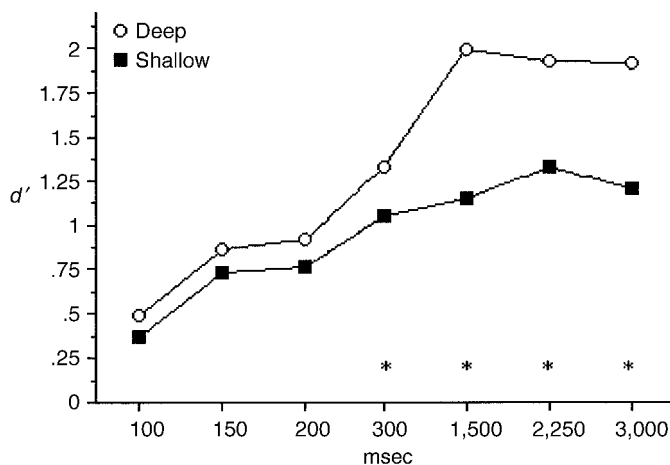


Figure 3. Graph of the levels of processing \times response-signal delay interaction. The asterisks indicate that the difference in recognition accuracy between deep and shallow conditions was significant ($\alpha = .007$).

Table 2
Average Response Times for Each of the Four Experimental Conditions at Each Response-Signal Delay

Response Signals	Deep		Shallow	
	Match	Mismatch	Match	Mismatch
100 msec	318	322	330	306
150 msec	320	324	321	322
200 msec	288	300	291	286
300 msec	235	261	237	254
1,500 msec	244	246	252	249
2,250 msec	270	266	277	272
3,000 msec	276	283	284	285

Note—The total time to respond to test items is obtained by adding the response-signal delay to the response time.

would be expected under deep encoding conditions more likely to support recollection. Finally, we left the asymptote free to vary, as in the asymptote-only model. Thus, this second, two-process model had eight parameters free to vary. From Table 3, it appears that the dual-process model accounted for about 95.6% of the variability in the data. The difference between the proportion of the variability accounted for by the full and the two-process model was not significant ($F < 1.2$), whereas the difference between the proportion of the variability accounted for by the two-process model and the asymptote-only model was significant [$F(2,19) = 10.41, p < .01$] (see Figure 4 for a display of the best-fitting curves obtained with the six- and the eight-parameter models). Thus, the two-process model provided a fit as good as the full model and a significantly better fit than the asymptote-only model.³

DISCUSSION

The purpose of this study was to determine the time course of increases in recognition accuracy arising from the manipulation of the match between study and test modality, and the level of processing at study. Our main result was a clear dissociation between the effect of the study–test modality match and of the levels-of-processing manipulations across response-signal delays. When the modality matched between study and test, recognition accuracy was higher than in the mismatch condition, but only at short response-signal delays. In contrast, words that were deeply processed at study were

remembered significantly better than words processed at a shallow level, but only at long response-signal delays. This dissociation is congruent with a dual-process interpretation of recognition memory, by which recognition can be based on a fast-acting familiarity process and a slow-acting recollective process. Furthermore, when we fitted an exponential growth-to-a-limit function to recognition accuracy, it appeared that a two-process model provided a significantly improved fit over an asymptote-only (familiarity) model.

Response-signal methods of manipulating response speed are usually implemented so that response-lags are randomly intermixed within each test session. Subjects are thereby prevented from adopting different retrieval strategies in advance of the recognition test (e.g., Hintzman & Caulton, 1997). In the present experiment, in which incidental learning was studied in two different orienting tasks, each subject was tested under only one response-signal deadline. Given that in our study the target information was incidentally learned, it is unlikely that subjects devised any encoding strategy that could have affected recognition memory performance differentially across the various deadline conditions.

However, the possibility that different response strategies were used at different lags needs to be addressed. One possibility is that at short lags, some subjects adopted the “irresponsible” strategy of responding on a large proportion of trials to beat the deadline. This seems unlikely, since false alarm rates were fairly consistent and showed similar variability across all response lags (Table 1).

As a further check on potential differences in the use of different retrieval strategies at short and long response lags, we performed a post hoc analysis on recognition accuracy as a function of the proportion of false alarms committed. In particular, we wanted to assess whether subjects who tended to commit a large number of false alarms showed a different recognition profile from that of more conservative subjects. To do this, we first collapsed the subjects in the 100-, 150-, and 200-msec response-signal conditions into one larger sample. These were the “fast response” conditions in which there was a clear effect of modality match, but in which the effect of levels of processing was not significant (hence the exclusion of the 300-msec condition). Then the subjects in the 1,500-, 2,250-, and 3,000-msec conditions were collapsed into a second sample. These two groups were fur-

Table 3
Least-Square Fits of the Exponential Growth-to-a-Limit Function to Recognition Memory Accuracy Data

Condition	Full Model 12 Parameters			Asymptote Model 6 Parameters			Two-Process Model 8 Parameters		
	A	R	I	A	R	I	A	R	I
Deep match	2.07	0.0072	367	2.09	0.0073	375	2.09	0.0059	349
Shallow match	1.30	0.0191	402	1.43	0.0073	375	1.34	0.0088	349
Deep mismatch	1.85	0.0050	381	1.71	0.0073	375	1.83	0.0059	393
Shallow mismatch	1.19	0.0065	376	1.17	0.0073	375	1.16	0.0088	393
r^2	0.966			0.908			0.956		

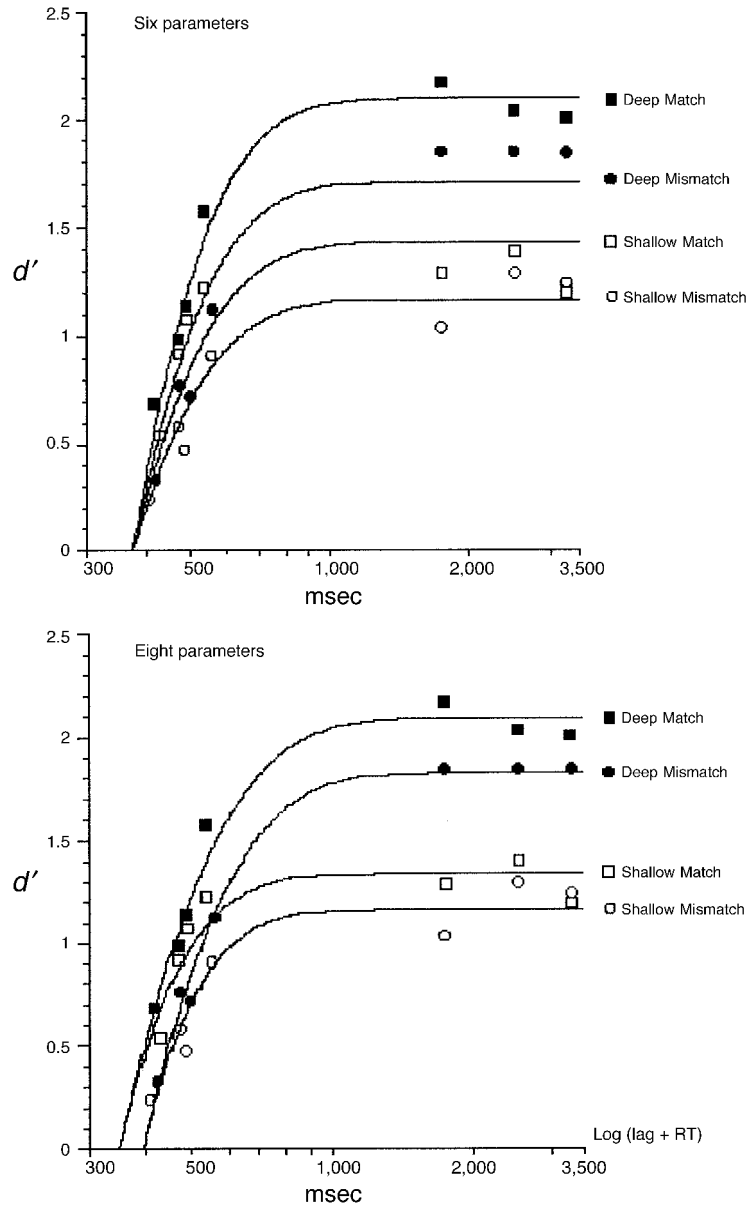


Figure 4. Best fitting curves obtained with the six- and the eight-parameter models. The x-axis refers to the logarithm of the response-signal lag + the response time.

ther divided by a median split on the number of false alarms, forming subgroups of conservative and liberal responders within the short and long lag groups. This approach allowed us to create four groups of relatively large size, and hence we had more power to assess the effect of response strategies on recognition memory profile. We then analyzed whether the conservative and liberal groups differed in their sensitivity to the levels-of-processing or modality match manipulations. This analysis indicated that being either a liberal or a conservative responder made no difference on the effect of the levels-of-

processing variable or on the effect of the study-test match variable across different response lags (i.e., no interactions involving the response criterion approached significance, $p > .10$). Therefore, it appears that the between-subjects manipulation of the response-signal variable provides a viable approach to study the influence of levels-of-processing and study-test modality match on recognition memory accuracy.

Our implementation of the response-signal deadline method diverged from other studies in requiring a go/no-go approach instead of the conventional binary old/new de-

cision. The reason for this is that we wanted to minimize response times, to encourage responses within the deadline, while keeping error rates low. Go/no-go responses have these properties, and in other tasks such as lexical decision, they preserve the properties observed in binary-choice responses (Perea et al., 2002). One disadvantage is that with this approach, it is difficult to interpret non-responses to test items, which may occur for two reasons: Subjects may make the wrong decision and withhold a response, or they may fail to meet the deadline. Strictly, the latter cases should be excluded from the analysis. We acknowledge that this misclassification of late nonresponses might have occurred in some cases, although the number of late responses should have been minimized by our procedures. However, because only "yes" response trials are required in order to derive estimates of performance such as d' , we do not think that the use of a go/no-go response procedure biased the outcome of the experiment.

To investigate further, for each response deadline condition we calculated the percentage of "yes" responses, made to all 160 old and new test items, within 400 msec of the response signal (i.e., strictly on time). We also calculated the percentage of "yes" responses given within 500 msec of the response signal. In the 400-msec case, the proportion of responses increased significantly across response lags from 30.2% in the 100-msec response-signal condition to 43.1% in the 3,000-msec condition ($p < .01$). This could be taken as evidence that at short delays, a relatively large proportion of "yes" responses lay outside the deadline. In contrast, with a 500-msec deadline, there was no significant increase in the percentage of "yes" responses from the 100-msec response-signal condition (40.4%) to the 3,000-msec response condition (45.5%, $p > .1$). Under the assumption that at long response-signal delay conditions subjects always have time to reach a decision before the deadline, the similar proportion of "yes" responses at short delays suggests that most responses have been made within 500 msec from the response signal. As has been indicated above, statistical analyses on recognition accuracy were carried out not only on answers given on time, within 400 msec of the response signal, but also on answers given with an extended deadline of 500 msec. The two analyses gave comparable results. Therefore, there is no reason to believe that the analysis was biased at short delays because a greater proportion of late responses were considered as no responses. Moreover, this type of explanation cannot account for the variations in sensitivity to study-test modality or levels of processing observed at short and long response-signal delays.

Overall, the present findings are congruent with approaches suggesting that recollection contributes to single-item recognition memory discrimination. Similar claims on the basis of the response-signal delay approach were made by Hintzman and Curran (1994; see also Joordens & Hockley, 2000). They showed that when discriminating between targets and highly similar foils (e.g., *cartoon* vs. *car-*

toons), subjects made more false alarms at short response-signal delays. However, at long delays, similarity had less effect on the false alarm rate. Their results suggest that familiarity is available earlier at test than item-specific information; responses at short lags are based on familiarity, whereas responses to highly similar new items at longer lags are influenced by recalled information.

With respect to the contribution of familiarity to recognition memory, the present results indicate that familiarity is based on the perceptual characteristics of target items, as Mandler (1980) first suggested. Recognition accuracy was enhanced when the modality of targets at test matched the modality used to display targets during learning. A plausible account of this effect is that sensory-perceptual characteristics of targets are encoded during learning. At test, if the same features recur, the match with the encoded target is increased, leading to increased recognition (cf. Mulligan & Hirshman, 1995).

Although the levels-of-processing manipulation was not significant at short delays, the present results do not necessarily preclude the possibility that some other form of familiarity may be based on conceptual factors. For example, in a recognition test where subjects had to respond only to items that were acoustically presented, visually presented words that were deeply processed at study were more likely to be wrongly reported relative to words processed at a shallow level, but only at a short response delay (Toth, 1996). One interpretation is that semantically processed words appeared to be more familiar when rapid responses were required at test. However, at longer delays, when recollection could be operative, the number of false alarms to semantically encoded words dropped to the level of targets that had received a shallower level of analysis during learning. Evidence that semantic processing may facilitate fast responses is also found in the present data. When the 100- and 150-msec response-signal delays were collapsed in order to increase statistical power, there was an advantage for words receiving deep processing ($d' = 0.123$; the 95% CI ranged from 0.002 to 0.245). Thus, fast recognition processes based on familiarity may be based on deep semantic or conceptual processing, although the main effect of such processing occurs later, when recollective processes should be more likely to operate.

Finally, and as a cautionary note, we wish to point out that the results obtained in the present study are congruent with dual-process approaches to recognition memory emphasizing a contribution of both familiarity and recollection. However, the present data do not exclude the possibility that recognition memory could be supported by two different familiarity signals (e.g., perceptual and semantic) with different time courses.

In sum, the main contribution of the present study consisted in a detailed analysis of the retrieval dynamics of recognition memory over several response-signal delays at test. The results strongly suggest that at least two different processes are required in order to account for recognition memory.

REFERENCES

- CLARK, S. E., & GRONLUND, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, **3**, 37-60.
- GARDINER, J. M., RAMPONI, C., & RICHARDSON-KLAVEHN, A. (1999). Response deadline and subjective awareness in recognition memory. *Consciousness & Cognition*, **8**, 484-496.
- GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model of both recognition and recall. *Psychological Review*, **91**, 1-67.
- HINTZMAN, D. L., & CAULTON, D. A. (1997). Recognition memory and modality judgments: A comparison of retrieval dynamics. *Journal of Memory & Language*, **37**, 1-23.
- HINTZMAN, D. L., & CURRAN, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, **33**, 1-18.
- JACOBY, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.
- JONES, T. C., JACOBY, L. L., & GELLIS, L. A. (2001). Cross-modal feature and conjunction errors in recognition memory. *Journal of Memory & Language*, **44**, 131-152.
- JOORDENS, S., & HOCKLEY, W. E. (2000). Recollection and familiarity through the looking glasses: When old does not mirror new. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 1534-1555.
- KUČERA, H., & FRANCIS, V. W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- MANDLER, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, **87**, 252-271.
- MULLIGAN, N., & HIRSHMAN, E. (1995). Speed-accuracy trade-offs and the dual process model of recognition memory. *Journal of Memory & Language*, **34**, 1-18.
- PEREA, M., ROSA, E., & GÓMEZ, C. (2002). Is the go/no-go lexical decision task an alternative to the yes/no lexical decision task? *Memory & Cognition*, **30**, 34-45.
- REED, A. V. (1976). List length and the time course of recognition in immediate memory. *Memory & Cognition*, **4**, 16-30.
- SNODGRASS, J. C., & CORWIN, J. (1988). Pragmatics of measuring recognition memory: Application to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.
- TOTH, J. P. (1996). Conceptual automaticity in recognition memory: Levels-of-processing effects on familiarity. *Canadian Journal of Experimental Psychology*, **50**, 123-138.
- YONELINAS, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, **46**, 441-517.

NOTES

1. The Gardiner et al. (1999) study does not provide a statistical test on the levels of processing by response-signal lag interaction. However, from the cell means it appeared that the levels-of-processing effect was numerically smaller in the fast than in the slow response condition used in the experiment.

2. The same analyses carried out on answers given within 500 msec yielded results analogous to the one based on answers given strictly "on time." Notice that all analyses on recognition memory were also carried out on A' scores. Analyses on d' and A' scores yielded equivalent results. Percentages of hits for the various experimental conditions were calculated using a denominator of 20. Percentages of false alarms were calculated using a denominator of 80.

3. An intermediate model between the asymptote only (i.e., six parameters) and the two-process model (i.e., eight parameters) was also fitted to the data. This intermediate model had seven parameters free to vary (i.e., the four asymptotes; two rate parameters as in the two-process model; and the intercept was fixed as in the asymptote-only model). This model appeared to account for 92.7% of the variability in the data. The difference in the proportion of the variability accounted for by the two-process model and this intermediate model was significant [$F(1,19) = 12.76, p < .01$]. Thus, the intermediate model provided a less accurate fit than the two-process model. Hence, to accurately describe the obtained data, it appears that a model should require at least two different intercepts and two different rate parameters.

(Manuscript received August 6, 2002;
revision accepted for publication February 21, 2003.)