Reversing the picture superiority effect: A speed–accuracy trade-off study of recognition memory

ANGELA BOLDINI, RICCARDO RUSSO, SAHIBA PUNIA, AND S. E. AVONS
University of Essex, Colchester, England

Speed–accuracy trade-off methods have been used to contrast single- and dual-process accounts of recognition memory. With these procedures, subjects are presented with individual test items and required to make recognition decisions under various time constraints. In three experiments, we presented words and pictures to be intentionally learned; test stimuli were always visually presented words. At test, we manipulated the interval between the presentation of each test stimulus and that of a response signal, thus controlling the amount of time available to retrieve target information. The standard picture superiority effect was significant in long response deadline conditions (i.e., ≥2,000 msec). Conversely, a significant reverse picture superiority effect emerged at short response-signal deadlines (<200 msec). The results are congruent with views suggesting that both fast familiarity and slower recollection processes contribute to recognition memory. Alternative accounts are also discussed.

In standard recognition memory tests, both previously studied and new items are presented and subjects are required to decide whether or not each item was presented at study. Two general classes of theories have been put forward to explain performance in recognition memory tasks. One approach to recognition memory is common to the so-called global matching models, such as TODAM (Murdock, 1982), SAM (Gillund & Shiffrin, 1984), and MINERVA 2 (Hintzman, 1986). Although they offer different representations of memory and retrieval mechanisms, global matching models maintain that subjects make familiarity judgments on the basis of a continuous index of memory strength.[1] When a test item exceeds the criterion of memory strength set by the subject, the stimulus is judged as old. (For a review of global matching models of recognition memory, see Clark & Gronlund, 1996.)

A second class of theories—the so-called dual process models—posits that a single continuous quantitative index of memory strength is not sufficient to account for the results obtained in experiments in which recognition memory tasks are used (see, e.g., Jacoby, 1991; Mandler, 1980; for a review of dual process models, see Yonelinas, 2002). Two processes are proposed: one based on *familiarity* and the other on *recollection*. Despite some differences between models (see Yonelinas, 2002), familiarity is considered to be a rapid, automatic process, often observed to be sensitive to manipulations of the perceptual features of target items (see, e.g., Lamberts, Brockdorff, & Heit, 2002). Recollection, in contrast, is considered to be a slower, intentional process dependent on the retrieval of specific qualitative information about prior occurrence

of the target items and predominantly sensitive to manipulations affecting semantic or conceptual encoding, such as level of processing (see, e.g., Atkinson & Juola, 1973; Boldini, Russo, & Avons, 2004; Jacoby, 1991; Mandler, 1980). Further evidence about the distinction between the time course of familiarity and that of recollection has been obtained in studies in which event-related brain potentials (ERPs) were examined through recognition memory tests. For example, Curran (2000) detected an early ERP component, at about 300–500 msec, associated with familiarity, and a later component, at about 400–800 msec, associated with recollection (for a recent review of the relevance to this issue of ERPs and functional neuroimaging studies, see Rugg & Yonelinas, 2003). It has often been suggested that familiarity and recollection influence recognition memory independently (see, e.g., Jacoby, 1991; Mandler, 1980; but see also Humphreys, Dennis, Chalmers, & Finnigan, 2000).

Various experimental procedures, such as the remember–know procedure (see, e.g., Gardiner & Richardson-Klavehn, 2000) and the process dissociation procedure (see, e.g., Jacoby, 1991), have been used to assess the potential contributions of different processes to recognition memory (see Yonelinas, 2002, for a review of the techniques used to investigate the contributions of different processes to recognition memory). A useful empirical approach to assessing the contribution of different processes to recognition memory is the *response signal procedure* (see, e.g., Boldini et al., 2004; Hintzman & Curran, 1994; Mulligan & Hirshman, 1995). In this procedure, subjects are presented with individual studied and new test items, each

---

R. Russo, rrusso@essex.ac.uk

followed by a response signal indicating that the subjects have to promptly decide whether the item is old or new. By manipulating the interval between the test stimulus and the response signal, it is possible to control the amount of time available to retrieve the target information. Therefore, this technique allows increments in recognition accuracy to be plotted as a function of the time available to process test items. Hence, if familiarity and recollection are indeed two distinct processes that contribute to recognition memory in fast- and slow-acting ways, respectively, it should be possible to trace their contributions by assessing the effects of different variables, which are expected to influence familiarity and recollection selectively, on recognition accuracy at short and at long response deadlines.

In particular, if familiarity is sensitive to the perceptual correspondence between study and test items, then recognition memory should be enhanced by a familiarity process when the perceptual characteristics of study and test stimuli match and reduced when they differ. Furthermore, if familiarity is a fast-acting process, then this effect should be observed at short response-signal lags—that is, when subjects have limited time to make recognition memory decisions at test. Now consider recollection as a slow-acting process based mainly on the semantic or conceptual characteristics of target items and on the retrieval of specific qualitative information about the prior occurrence of targets. If this is the case, then recognition should be more optimal under conditions that promote both elaborative (e.g., semantic) encoding and effective search processes at test than under conditions that do not promote these processes (e.g., those that promote nonsemantic encoding; Boldini et al., 2004). Moreover, if the elaborative advantage reflects the contribution of recollection to recognition memory, this effect should be particularly evident when relatively more time is allowed to make recognition decisions. Because recollection is a slow process, the effect of elaborative encoding should be minimal or absent when subjects have limited time to make recognition memory decisions at test.

The current evidence seems only partially supportive of the hypotheses above. With respect to the effect of perceptual matching, 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 either short or long response-signal lags and in fact reported that asymptotic performance was higher in the mismatch condition.

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

Boldini et al. (2004) pointed out that those studies in which manipulation of perceptual matching or levels of processing information had no effect on the time course of recognition decision face several criticisms. For example, Mulligan and Hirshman (1995, Experiments 2 and 3) and Hintzman and Caulton (1997) both used the standard response-signal procedure, in which response signals vary randomly from trial to trial for each subject. Although this technique equates retrieval strategies at the point at which the recognition probe appears, several sessions of extensive testing are usually required to achieve stable measures. This procedure precludes the possibility of testing under incidental conditions, thus making it difficult to evaluate the effect of the conventional manipulation of levels of processing. In order to overcome some of these difficulties while trying to contribute to a better understanding of recognition dynamics, Boldini et al. implemented the response-signal procedure so that each subject was tested in only one response-signal deadline condition, thus allowing target items to be studied under incidental conditions. Recognition performance was recorded in seven deadline conditions ranging from 100 to 3,000 msec. Learning was incidental. Levels of processing and study–test modality matching were manipulated to assess the contributions of recollection and familiarity processes, respectively, to recognition memory. It appears that the study–test modality match had a beneficial effect on recognition accuracy at short response-signal delays ($\leq$300 msec). Conversely, recognition accuracy benefited more from deep than from shallow processing at study only at relatively long response-signal delays ($\geq$300 msec). This set of results was interpreted as suggesting that at least two different processes, one fast and one relatively slow, contribute to recognition memory.

The present study is a conceptual follow-up of Boldini et al. (2004). Here, we extend our earlier findings by using study items presented either as pictures or as words, whereas test stimuli were always words. Therefore, if familiarity is based on the perceptual features of target stimuli, then recognition accuracy should be impaired for targets studied in pictorial form relative to those studied as words, and this impairment for studied pictures should be observed at short response deadlines. This predicted outcome at short response deadlines is the inverse of the standard picture superiority effect commonly observed in episodic memory tasks. This effect is a well established phenomenon (see, e.g., Madigan, 1983; Paivio, 1971; Rajaram, 1993) whereby stimuli presented as pictures are more likely to be recalled and recognized than stimuli presented as words, even when words are presented during a recognition memory test.

In summary, the aim of the present study was to further investigate the retrieval dynamics and time course of recognition memory processes by manipulating the format of stimuli between study and test. In Experiment 1, we tested six different groups of subjects at different response-signal delays (ranging from 100 to 3,000 msec). At study, the subjects were required to commit to memory two sets of items. Items in one set were displayed as pictures and those in the other set as words. At test, all items were pre-

sented as words. If responses at short response lags are based mainly on familiarity processes that rely on the perceptual characteristics of target items, then a reverse of the picture superiority effect should be observed. On the other hand, the standard picture superiority effect should be apparent at longer response-signal delays.

In addition to running standard analyses, we fitted the empirical data obtained in Experiment 1 using an exponential growth-to-a-limit function to model the increase in recognition with response-signal deadline as a function of study conditions (see, e.g., Hintzman & Caulton, 1997; Mulligan & Hirshman, 1995; Reed, 1976). The curve fitting was obtained by applying the formula

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

where $T > I$, and $d' = 0$ otherwise.

Equation 1 indicates how recognition accuracy (d') is a function of three parameters (A, I, R): A represents the asymptote and indexes how much information is in memory; I represents the intercept (i.e., the time at which performance rises above chance) and indicates the time when mnemonic information begins to accumulate; R indicates the rate of growth to asymptote and the speed at which mnemonic information accumulates. T is the response time. Since these parameters are associated with relatively specific characteristics concerning memory retrieval, hypotheses can be drawn about how the parameters would be expected to vary if specific theoretical conditions are met. In other words, the goal of this analysis is not simply to find the best-fitting curve (and its parameters) per se, but to compare certain patterns of parameters (i.e., certain theoretical assumptions) in order to see which alternatives provide better fits. For instance, if recognition memory depends only on a global familiarity value associated with each test item, then the only parameter that should discriminate between different experimental conditions would be A (i.e., the index of memory strength), and no other improvement in the fitting of the data would be added by allowing either I or R to be free to vary. On the other hand, if more than one process is required to account for recognition memory performance, then a better fit would be obtained when at least one of the other parameters (I or R) is left free to vary. *F* tests were carried out to assess whether or not there was any significant difference in the variability (R^2) accounted for by different models having different numbers of free parameters (Howell, 1997).

EXPERIMENT 1

Method

Subjects. Seventy-two students, mainly undergraduates at the University of Essex, were recruited for this experiment. They were all native English speakers. The subjects were tested individually.

Design. A two-factor mixed design was used. Item format at study (i.e., picture vs. word) was manipulated within subjects, whereas response deadline (100, 150, 200, 700, 2,000, and 3,000 msec) was manipulated between subjects. Twelve subjects were tested at each response deadline.

Materials. A total of 120 items was selected from Snodgrass and Vanderwart (1980) to be used as target and lure stimuli. Another 30 items were selected from the same database as buffer items to be presented at the beginning and at the end of the study lists to reduce

the contribution of primacy and recency effects. Ten of these stimuli were presented at the beginning of each study list and the remaining 20 at the end. Half of the buffer items were presented as pictures and the other half as words. Of the initial 120 items, 40 were presented as words (Set A) and 40 as pictures (Set B) in the study phase. The remaining 40 items (Set C) were presented as lures in the test phase. Sets A, B, and C were counterbalanced across subjects using a Latin square approach, so that each set was presented in each condition (i.e., study picture, study word, and test lure) an equal number of times within each group of subjects. For the training phase, a set of 40 nonwords was used. For the test phase, all 120 stimuli were presented in word form and in a random order that was different for each subject.

All stimuli were presented in the center of a computer screen. Words were displayed in 60-point extended, lowercase Comic Sans MS font. Words and pictures were of comparable sizes when seen on the screen. The response deadline signal was a row of asterisks that appeared two lines below the test word. Finally, 400 msec after each test deadline presentation, a beep alert was played to indicate that the time allowed for responding had elapsed.

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

During the study phase, the subjects were presented with a list with 10 initial buffer items, 80 target items, and 20 buffer items presented at the end. Picture items and word items were presented in alternation. The subjects were simply asked to memorize as many items as they could for a later, unspecified memory test. Each item was displayed for 2 sec with a 1-sec interstimulus interval (ISI).

The training phase provided the subjects with some practice with the response-signal procedure. The stimuli were 40 nonwords ranging from five to nine 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 used during testing), a row of asterisks appeared two lines under the test string. The task was to decide whether or not the letter "t" was present in the nonword. The subjects were instructed to respond as quickly 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 a response was given more than 400 msec after the appearance of the response signal. There was a 2-sec pause between consecutive items.

During the test phase, all the stimuli (80 target items and 40 distractors) were presented as words. The subjects were asked to decide whether or not each test item had been encountered at study, irrespective of its format. Test words were presented following the same procedure used during the training phase. The subjects had only to respond "yes" (by pressing the "z" key) if they thought that a stimulus had previously been studied. They were encouraged to be accurate, but they were told it was more important to respond on time. As in the training phase, responses were considered as given on time when the response latency was not longer than 400 msec. The ISIs were 2,500 msec for the 100-, 150-, and 200-msec response-signal conditions; 2,000 msec for the 700-msec condition; and 1,500 msec for the 2,000- and 3,000-msec conditions. This was done to roughly equate the length of the test phase across response lag conditions. At the end of the test phase, the subjects were debriefed and given the opportunity to know their results.

Results and Discussion

Recognition memory accuracy. Only responses given within 400 msec were considered for the analysis.[2] Mean proportions of hits, false alarms, and corrected d' scores for groups and conditions are presented in Table 1. To avoid hit rates of 1 and false alarm rates

Table 1
Experiment 1: Mean Proportions (With Standard Deviations) of Hits, False Alarms
(FAs), and d' Scores for Each Condition and Each Response Deadline

| Response Deadline (msec) | Hits | | | | | | d' | | | |
| | Words | | Pictures | | FAs | | Words | | Pictures | |
	M	SD	M	SD	M	SD	M	SD	M	SD
100	.57	.14	.46	.14	.35	.16	0.60	0.31	0.31	0.31
150	.64	.09	.48	.09	.34	.09	0.77	0.38	0.36	0.28
200	.60	.16	.57	.11	.29	.15	0.87	0.42	0.78	0.39
700	.59	.17	.70	.12	.16	.10	1.28	0.77	1.59	0.72
2,000	.63	.13	.82	.10	.18	.14	1.36	0.71	1.98	0.65
3,000	.61	.17	.76	.12	.21	.13	1.18	0.65	1.62	0.67

of 0, the raw data of each subject were adjusted as suggested by Snodgrass and Corwin (1988): H(%) = (#H + 0.5)/(#old words + 1), and FA(%) = (#FA + 0.5)/(#new words + 1).

A 2 (stimulus type: word vs. picture) × 6 (response deadline: 100 vs. 150 vs. 200 vs. 700 vs. 2,000 vs. 3,000 msec) mixed ANOVA was carried out on d' scores. A significant main effect of stimulus type [$F(1,66) = 4.83, MS_e = 0.07, p < .05$] reflected the fact that pictures were better recognized than words overall. A significant main effect of response deadline [$F(5,66) = 11.55, MS_e = 0.54, p < .01$] indicated that performance increased as a function of lag. More interestingly, the interaction between stimulus type and response deadline was significant [$F(5,66) = 14.92, MS_e = 0.07, p < .01$], indicating that the effect of stimulus type differed in different deadline conditions. A series of t tests was carried out to follow up on this significant interaction; for each t test, the alpha level was set to .01. In both the 100- and 150-msec deadline conditions, recognition accuracy was significantly better for words than for pictures [$t(11) = 4.44$ and $t(11) = 5.23$, respectively]. In the 200-msec deadline condition, there was no significant difference between pictures and words [$t(11) = 1.01$]. In the 700-msec response deadline condition, the picture superiority effect approached significance [$t(11) = 2.36$], and it was clearly significant in the 2,000- and 3,000-msec response deadline conditions [$t(11) = 5.12$ and $t(11) = 3.05$, respectively]. The pattern of the significant interaction suggests that at least two processes may be needed to account for recognition memory performance.

Modeling. As was described in the introduction, the d' scores were fitted using an exponential growth-to-a-limit

Table 2
Experiment 1: Average Response Times (in Milliseconds)
for Words and Pictures at Each Response Deadline
(in Milliseconds)

Response Deadline	Words	Pictures
100	318	311
150	290	300
200	291	303
700	217	209
2,000	262	258
3,000	262	255

Note—The total time to respond is given by adding each response time to the respective response deadline.

function (Equation 1). Accuracy data for each study condition were plotted against the total time elapsed at response, which was calculated by adding the respective response deadline to the average reaction time for each group (see Table 2). If recognition memory is based on a single process (i.e., familiarity), the fitting of recognition accuracy for pictures and words should differ only in terms of the asymptote (i.e., the total amount of information stored in memory). However, if a better fit is obtained by allowing either the rate of growth or the intercept to be free to vary, then this would indicate that a single process alone may not be sufficient to account for the recognition memory performance.

For each model, the proportion of the variability accounted for was calculated by adjusting the value of r^2 as a function of the number of free parameters. Following Reed (1976),

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

where N is the number of data points d_i, each d_i corresponds to an observed d' score, each $\hat{d}_i$ value corresponds to the d' score predicted by the model, $\bar{d}$ is the overall mean recognition accuracy, and k is the number of free parameters.

When all parameters (i.e., A, I, and R) were left free to vary for each study condition (i.e., full model, in which $k = 6$), 93.7% of the variability was accounted for by the model. (Table 3 provides the values of the parameters for the different models.) A second model, in which only A was left free to vary whereas I and R were fixed across the two study conditions, was fitted to the data (asymptote model, in which $k = 4$; see Figure 1A). This model accounted for 82.2% of the variability of the data. The difference between the proportion of the variability accounted for by the two models was marginally significant [$F(2,5) = 4.56, p = .074$]. We also fitted our data to two other models, in which, in addition to A, either I or R was left free to vary. The model in which both A and I were left free to vary ($k = 5$; see Figure 1B) accounted for 94.1% of the variability in the data, which, again, was significantly

Table 3
Experiment 1: Least Squares Fits of the Exponential Growth-to-a-Limit Function
to Recognition Memory Accuracy Data

| | Model | | | | | | | | | | | |
| | Full (6) | | | Asymptote (4) | | | Asymptote and Intercept (5) | | | Asymptote and Rate (5) | | |
Condition	A	R	I	A	R	I	A	R	I	A	R	I
Words	1.28	0.0073	327	1.39	0.0045	344	1.30	0.0043	250	1.27	0.0109	364
Pictures	1.80	0.0040	374	1.70	0.0045	344	1.79	0.0043	378	1.81	0.0036	364
r^2	0.937			0.822			0.941			0.941		

Note—The names of the models reflect the parameters left free to vary. Numbers in parentheses represent the numbers of parameters free to vary in the respective models. A, asymptote; R, rate of growth to asymptote; I, intercept.

greater than the variability accounted for by the asymptote model [$F(1,6) = 12.09, p < .05$]. The model in which both A and R were left free to vary ($k = 5$; see Figure 1C) accounted for 94.1% of the data's variability. This was significantly greater than the variability accounted for by the asymptote model [$F(1,6) = 12.14, p < .05$]. The latter two models were not significantly different from the full model in fitting the data (i.e., both Fs < 1). From the results of the present analysis, it could be argued that at least two different processes are required to account for the recognition

memory data. Further discussion of this interpretation will be provided in the General Discussion section.

In summary, in consistency with our previous results (Boldini et al., 2004), we obtained a clear interaction between changes in the format of items between study and test and response deadlines. On one hand, matching format between study and test resulted in significantly better recognition performance at short response deadlines (i.e., <200 msec). On the assumption that familiarity is a fast-acting process, this result supports the generally
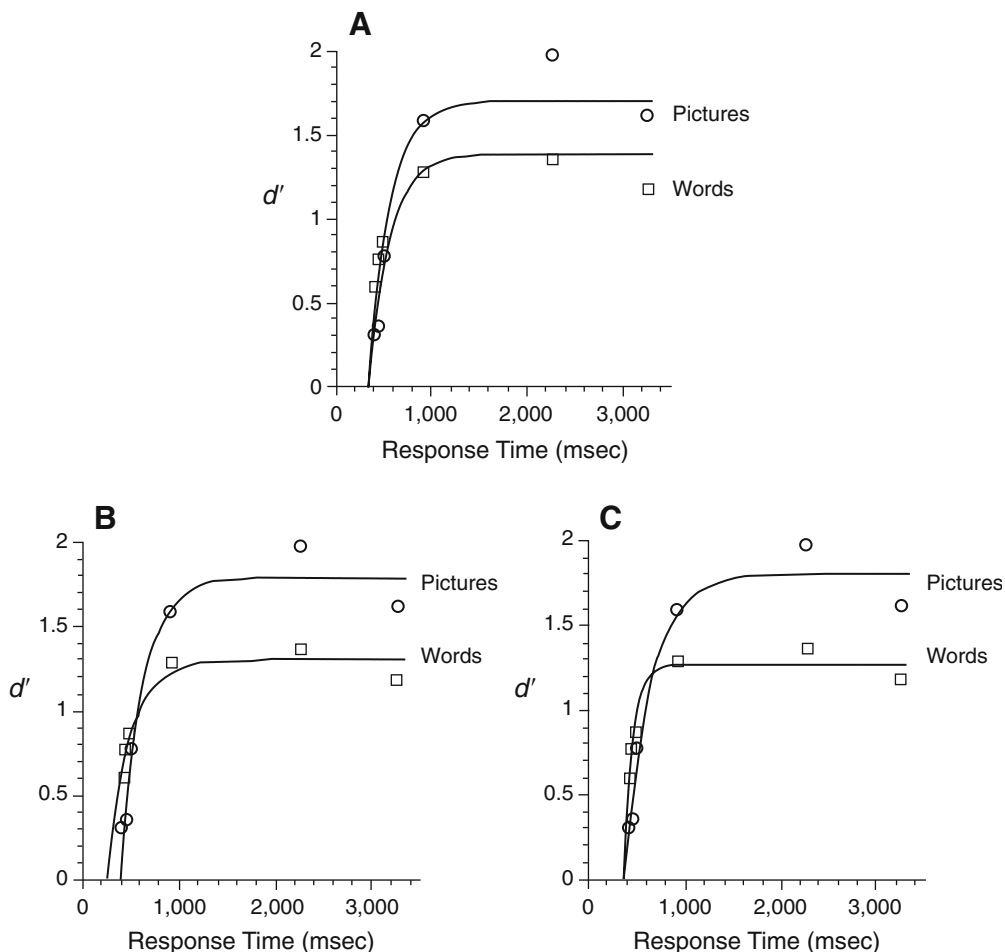

Figure 1. Best-fitting curves obtained with the four-parameter model (A) and the two five-parameter models, (B) different intercepts, and (C) different rates. See Table 3 for details.

accepted view that the perceptual match between study and test items affects familiarity processes in recognition memory. On the other hand, the richness of encoding associated with the pictorial format (see, e.g., Mintzer & Snodgrass, 1999) led to significantly better performance at late response deadlines (i.e., >700 msec).

Our implementation of the response-signal deadline method diverged from that used in other studies. Instead of asking subjects to enter either a "yes" or a "no" recognition response for each test item, we asked them simply to respond if they recognized the item (see also Boldini et al., 2004). The reason for using a go/no-go decision instead of the conventional binary old–new decision is that we wanted to minimize response times by encouraging responses within the deadline while keeping error rates low. Go/no-go responses serve this purpose well and preserve the properties observed in binary choice tasks when used in other tasks as well (e.g., lexical decision tasks; see, e.g., Perea, Rosa, & Gómez, 2002). One potential disadvantage is that with this approach it is difficult to interpret nonresponses to target items at test. These nonresponses may occur for two reasons: Subjects may make the wrong decision and withhold a response, or they may fail to meet the deadline. Strictly speaking, 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 be minimized by our procedures. However, because only "yes" response trials are required to derive estimates of performance such as d', we think that the use of a go/no-go response procedure did not bias the outcome of the experiment.

To check for a possible source of bias in this sense, we took into consideration the total percentages of "yes" responses given within both 400 and 500 msec of the response signal for each deadline condition, allowing for some "late" responses to be considered. For the 400-msec timeout, the percentage of all "yes" responses varied from 46.2% (for the 100-msec response deadline condition) to 52.7% (for the 3,000-msec deadline condition). A linear contrast analysis indicated that the increment in the percentage of "yes" responses was significant [$F(1,66) = 4.30$, $p < .05$]. A similar analysis was performed on the responses given within 500 msec from the response signal. In this case, there was no significant increment in the proportion of "yes" responses from short to long response deadline conditions; these varied from 59.7% for the 100-msec response deadline condition to 54.3% for the 3,000-msec condition [$F(1,66) = 2.84$]. Therefore, under the

assumption that in long response-signal deadline conditions subjects always have time to reach a decision before the deadline, the similar proportions of "yes" responses at short and at long delays suggest that most responses have been made within 500 msec from the response signal. As was previously indicated, statistical analyses on recognition accuracy were carried out not only on responses given on time (within 400 msec of the response signal), but also on those given with an extended timeout period of 500 msec. The two analyses yielded 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 nonresponses.

The goals of Experiment 2 were to provide converging evidence about the viability of a go/no-go procedure and to provide a replication of the present experiment. In Experiment 2, only two response deadlines were used (150 and 2,000 msec) in conjunction with a yes–no test response procedure. In light of the results obtained in the present experiment, we expected to detect a standard picture superiority effect in the 2,000-msec response deadline condition, whereas a reverse picture superiority effect was predicted to occur in the 150-msec response deadline condition.

EXPERIMENT 2

Method

Subjects. Forty-eight students, mainly undergraduates and postgraduates at the University of Essex, were recruited for this experiment. They were all native English speakers. The subjects were tested individually.

Design. A two-factor mixed design was used. Item format at study (pictures vs. words) was manipulated within subjects, whereas response deadline (150 vs. 2,000 msec) was manipulated between subjects. Twenty-four subjects were tested at each response deadline.

Materials and Procedure. We used the same materials and procedure as in Experiment 1. The only difference was that, at both training and test, the subjects were required to make either an "old" or a "new" response to each presented item by pressing the "z" or the "/" key, respectively, on the computer keyboard.

Results and Discussion

Only responses given within 400 msec were considered for the analysis. The percentages of invalid responses (i.e., those given more than 400 msec after the response signal) ranged from 6.1% to 7.7%. Mean proportions of hits, false alarms, and corrected d' scores for groups and conditions are presented in Table 4. A 2 (stimulus type: word vs. picture) × 2 (response deadline: 150 vs. 2,000 msec) mixed ANOVA was carried out on the d' scores. There was a

Table 4
Experiment 2: Mean Proportions (With Standard Deviations) of Hits, False Alarms
(FAs), and d' Scores for Each Condition and Each Response Deadline

Response Deadline (msec)	Hits				FAs		d'			
	Words		Pictures				Words		Pictures	
	M	SD	M	SD	M	SD	M	SD	M	SD
150	.72	.10	.62	.10	.44	.13	.76	.41	0.47	0.46
2,000	.69	.12	.79	.07	.35	.16	.95	.61	1.23	0.58

significant main effect of response deadline [$F(1,46) =$ 11.54, $MS_e = 0.47$, $p < .01$], indicating that performance was better at the longer response deadline, whereas the main effect of stimulus type was not significant ($F < 1$). More interestingly, the interaction between stimulus type and response deadline was significant [$F(1,46) = 25.45$, $MS_e = 0.08$, $p < .01$], indicating that the effect of stimulus type differed in different response deadline conditions. In the 150-msec deadline condition, recognition accuracy was significantly better for words than for pictures [$t(23) = 3.66$, $p < .01$]. In the 2,000-msec deadline condition, the picture superiority effect was significant [$t(23) = 3.47$, $p < .01$].

Overall, the results obtained in this experiment replicated the patterns of results obtained in Experiment 1 using old–new response decisions instead of go/no-go responses. From a comparison of the performance obtained in the 150-msec response deadline conditions in Experiments 1 and 2, it appears that recognition accuracy was significantly greater for words than for pictures [$F(1,34) = 30.9$, $MS_e = 0.06$, $p < .01$], whereas there was no significant difference in overall accuracy between the two experiments and no significant interaction between study conditions and experiments (both $Fs < 1$). It therefore appears that the old–new and go/no-go response procedures provided comparable results in the short response deadline condition examined. The same analysis for the 2,000-msec response condition showed that accuracy was greater for pictures than for words [$F(1,34) = 39.5$, $MS_e = 0.08$, $p < .01$]. However, although a significant picture superiority effect was found in both experiments, performance was more accurate in Experiment 1 [$F(1,34) = 7.66$, $MS_e = 0.70$, $p < .01$] and the picture superiority effect was significantly larger in Experiment 1 [$F(1,34) = 5.47$, $MS_e = 0.08$, $p < .05$]. We do not have a clear explanation for these results. In summary, these analyses further indicate that the go/no-go response procedure is a viable alternative to old–new response decisions in recognition memory experiments in which response-signal procedures are adopted.

It could be argued that the between-subjects implementation of the response-signal procedure may lead to interpretative ambiguities. As an anonymous reviewer suggested, subjects in the different response lag conditions could have adopted different strategies. For instance, the subjects in the short lag conditions might have chosen to focus on perceptual features of the words at test because they were given little time to process anything else. On the other hand, the subjects in the long lag conditions may have focused on either semantic or distinctive features. Hence, time-course differences could simply be a product of different strategies rather than reflecting separate and naturally unfolding fast and slow processes of recognition memory.

In order to assess the above possibility, we ran another experiment, in which the two response lags used in Experiment 2 were manipulated within subjects. In particular, subjects were tested using both response lags used in Experiment 2, but these were randomly intermixed within the same test session. If the outcome of the previous experiment is due to the implementation of differential strategies at different response lags, we should observe a difference between the pattern of results obtained in Experiment 3 and the pattern of those obtained in the previous two experiments.

EXPERIMENT 3

Method

Subjects. Twenty-three new subjects were tested in this experiment. These were mainly undergraduates at the University of Essex. All the subjects were native English speakers and were tested individually. Sixteen additional subjects were also tested, but their data had to be discarded due to a very large proportion of test responses given after the time out (see the Results and Discussion section).

Design. A two-factor repeated design was used. Item format at study (picture vs. word) and response deadline (150 vs. 2,000 msec) were manipulated within subjects.

Materials. A total of 120 items was selected from Snodgrass and Vanderwart (1980) to be used as target and lure stimuli. Another 30 items were selected from the same database to be presented as buffer items at the beginning and end of study lists to reduce the contribution of primacy and recency effects. Ten buffer stimuli were presented at the beginning of each study list, and the remaining 20 at the end. Half of the buffer items were presented as pictures and the other half as words. At test, all 120 stimuli were presented in word format. Response delay was counterbalanced within item type in that half of the items studied as pictures, half of the items studied as words, and half of the lures were presented at the 150-msec response delay. The remaining items were presented at the 2,000-msec response delay.

The original list of 120 items was randomized 23 times. In each case, 40 items were assigned to the study-as-words category, 40 to the study-as-pictures category, and 40 to the distractor category. Each of the 23 lists was then used once for each subject. For the training phase, a set of 100 names of cities (half of which were state capitals and half of which were not state capitals) was used. Response delay (150 vs. 2,000 msec) was counterbalanced within city type.

All stimuli were presented in the center of a computer screen. Words were displayed in 36-point lowercase Gill Sans regular font. Words and pictures were of comparable sizes when seen on the screen. The response deadline signal was a row of asterisks that appeared two lines below the test word. A beep alert was played 500 msec after each test deadline presentation to indicate that the time allowed for responding had elapsed.

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

The study phase was identical to that in Experiments 1 and 2.

The training phase provided the subjects with some practice with the response-signal procedure. The stimuli were 100 city names, and the subjects had to decide whether each city was a state capital or not. They were instructed to press the "a" key for "yes" and the "l" key for "no." Each city name appeared at the center of the screen, and after either 150 or 2,000 msec a row of asterisks appeared two lines under the city name. The subjects were told to respond as soon as they saw the asterisks and that they would hear a beep if they were not quick enough in responding. The beeping signal was sounded 500 msec after the asterisks appeared. The ISI was 2,500 msec. Presentation order of city names with response deadlines at 150 and 2,000 msec was random.

During the test phase, all the stimuli (80 target items plus 40 distractors) were presented as words. The subjects were asked to decide whether or not each test word had been studied during the study phase, regardless of its format. Test words were presented following the same procedure used during the training phase; the subjects had to press the "a" key to respond "yes" and the "l" key to respond "no." They were encouraged to be accurate but were told that it was more

Table 5
Experiment 3: Mean Proportions (With Standard Deviations) of Hits, False Alarms
(FAs), and d' Scores for Each Condition and Each Response Deadline

Response Deadline (msec)	Hits				FAs		d'			
	Words		Pictures				Words		Pictures	
	M	SD	M	SD	M	SD	M	SD	M	SD
150	.76	.17	.69	.13	.49	.24	0.78	0.60	0.53	0.61
2,000	.63	.20	.82	.12	.20	.12	1.23	0.76	1.81	0.73

important to respond on time. As in the training phase, responses were considered as given "on time" when the response latency was not longer than 500 msec. Presentation of test items with response delay at 150 or 2,000 msec was random, with an ISI of 2,500 msec. At the end of the test phase, the subjects were debriefed and given the opportunity to know their results.

Results and Discussion

In order for a subject's data to be included in the analyses, at least 50% of his or her test responses had to be given on time for each type of test item (i.e., targets studied as words, targets studied as pictures, and foils). Responses given within 400 msec after the response signal failed to meet this criterion, so we had to consider performance obtained within 500 msec.

Mean proportions of hits, false alarms, and corrected d' scores for groups and conditions are presented in Table 5. A 2 (stimulus type: word vs. picture) × 2 (response deadline: 150 vs. 2,000 msec) repeated measures ANOVA was carried out on the d' scores. There was a significant main effect of response deadline [$F(1,22) = 40.88$, $MS_e = 0.42$, $p < .01$], indicating that performance was better at the longer response deadline. The main effect of stimulus type was also significant [$F(1,22) = 5.13$, $MS_e = 0.12$, $p < .05$], indicating that performance was better for studied pictures than for studied words. More interestingly, the interaction between stimulus type and response deadline was significant [$F(1,22) = 27.04$, $MS_e = 0.15$, $p < .01$], indicating that the effect of stimulus type differed in different response deadline conditions. In the 150-msec deadline condition, recognition accuracy was significantly better for words than for pictures [$t(22) = 2.21$, $p < .04$]. In the 2,000-msec deadline condition, the picture superiority effect was significant [$t(22) = 5.68$, $p < .01$].[3]

From the results obtained, it appears that the performance of the subjects in Experiment 3 is comparable to that obtained in the previous two experiments. This result rules out the hypothesis that the different performance profiles obtained in the previous experiments at different response lags simply reflect the use of different strategies at specific response lags rather than inherent properties of memory processes. Therefore, from a methodological point of view, the results obtained in Experiments 1, 2, and 3 support the view that implementing the response-signal procedure with either a between-subjects or a within-subjects manipulation of the response signal provides comparable results, at least when intentional learning conditions are used.

Finally, we would like to add the following consideration regarding the within-subjects implementation of the response-signal procedure. Although our subjects were given an extensive training phase of 100 trials in order to

learn to respond on time following the signal to respond, 16 subjects had to be discarded given the large number of trials in which they could not respond on time—that is, within 500 msec. Thus response times were longer and more variable when response deadlines varied unexpectedly across trials. Apart from this, the within-subjects and between-subjects implementations of the response-signal procedure provided comparable outcomes. However, the between-subjects implementation appeared less problematic than the within-subjects implementation, at least in the present study. Hence, between-subjects manipulations of response lags appear to be a viable alternative to within-subjects implementations of the response-signal procedure.

GENERAL DISCUSSION

To summarize the results obtained in this study, with a go/no-go response procedure the picture superiority effect was reversed in response deadline conditions up to 150 msec, whereas a conventional picture superiority effect emerged at long response deadlines (2,000 and 3,000 msec). Comparable results were then obtained with conventional binary (old–new) decisions. Moreover, when in the last experiment response signals were manipulated within subjects in a random manner, the results obtained were consistent with those obtained in the two previous experiments. This result rules out the possibility that time-course differences in the first two experiments were simply the product of the use of different strategies at different response lags rather than a reflection of inherent properties of memory processes. Overall, it thus appears that when the perceptual characteristics of target stimuli were changed between study and test, recognition memory was negatively affected, but only when subjects were required to quickly respond to test stimuli. This result (see also Boldini et al., 2004) indicates that, as Mandler (1980) first suggested, familiarity is based on the perceptual characteristics of target items. A plausible account of the reverse picture superiority effect is that sensory/perceptual characteristics of targets are encoded during learning. When the same features recur at test, the match with the encoded target is increased, thus leading to increased recognition (see Mulligan & Hirshman, 1995).

The modeling of the results obtained in the first experiment, in which a growth-to-a-limit function was used, are compatible with accounts suggesting that more than one process is required to account for the retrieval dynamics in the recognition task across study conditions and response deadlines (e.g., familiarity and recollection). Whereas the

main aim of this study was to provide a further characterization of the nature of familiarity processes, we could speculate on the nature of other process(es) involved in recognition memory decisions.

Several hypotheses have been put forward to account for the picture superiority effect in recognition memory. One of the first explanations was provided by Paivio (1971), who suggested a *dual coding account*, claiming that elaborating pictures entails the activation of both an image code and a verbal code, whereas words can activate only a verbal code. Since pictures lead to the formation of two memory traces rather than one, pictures are more likely to be retrieved than words. Paivio's theory, however, has been challenged by other types of approaches that focus more on the general idea of an intrinsic "distinctiveness" that pictures should have in comparison with words. Along these lines, Nelson (1979) claimed that the picture superiority effect is caused by a superior sensory code: a higher visual distinctiveness that allows pictures to be uniquely encoded in memory. Weldon, Roediger, and Challis (1989) instead argued that the picture superiority effect is due to a more extensive semantic processing of pictures than of words during study.

Although the debate is still open on the precise mechanism underlying the picture superiority effect (see Kinjo & Snodgrass, 2000; Mintzer & Snodgrass, 1999), there is empirical evidence at variance with the dual coding account and with the view that a more extensive semantic analysis occurs with pictures than with words. Mintzer & Snodgrass (1999; see also Stenberg, Radeborg, & Hedman, 1995) showed that pictures do not spontaneously evoke verbal recoding during learning. Hence, the picture superiority effect cannot be due to pictures' enjoying both verbal and pictorial coding. There is also evidence that enhanced semantic processing of pictures over words is unlikely (see, e.g., Amrhein, McDaniel, & Waddill, 2002; Weldon & Coyote, 1996). For example, using two different conceptual implicit memory tasks, Weldon and Coyote did not find a picture superiority effect in priming for items learned as pictures over those learned as words. Hence, stimuli presented as pictures during study do not lead to greater semantic encoding than do those presented as words.

From the considerations discussed above, plus the finding that high visual similarity among pictures reduces or even reverses the picture superiority effect (Nelson, Reed, & Walling, 1976), it could be argued that the picture superiority effect is primarily due to a distinctive sensory/perceptual coding of pictures. This, along with the observation, with the use of the remember–know paradigm, that the picture superiority effect is associated with an increment in "remember" responses (Dewhurst & Conway, 1994; Rajaram, 1993, 1996), could tentatively suggest that the picture superiority effect in long response deadline conditions is primarily supported by the recollection of, for example, distinctive visual images corresponding to the studied pictures (but see Dunn, 2004, for a critique of the remember–know paradigm as a tool for investigating recollective processes in recognition memory). However, since in the present study we did not directly assess the extent to which recollection underlies the picture superiority effect, the conclusion stated above is only tentative. Further empirical evidence is required to support this suggestion.

Alternatively, the present results could, in principle, be accounted for without the need for a recollective process to complement familiarity. For example, Brockdorff and Lamberts (2000; see also Lamberts et al., 2002) proposed a formal feature-sampling theory of recognition (FESTHER) to account for the time course of recognition memory judgments. According to these authors, "recognition response patterns at different deadlines may reflect the accumulation of perceptual information about the test stimulus rather than the time course of retrieval" (Lamberts et al., 2002, p. 1176). On the assumption that some perceptual features of test targets are processed faster than others, FESTHER was shown to account for a series of recognition memory experiments using the response-signal procedure, in which perceptual characteristics of target items were manipulated. For example, using test distractors that elicited faster responses and test distractors that elicited slower responses in a perceptual matching task, Lamberts et al. found, in a recognition memory test, that the former distractors were also more easily classified as new at short response deadlines.

In principle, FESTHER could be extended to account for the present data. Consider first recognition judgments at short response lags. Let us assume that the speed of processing of the perceptual attributes is similar for all words presented at test. Let us also assume that the processing of perceptual features of words is faster than the processing of other attributes of words (e.g., semantic features; features coding the form of the object to which each word refers). It then follows that the relative fast accumulation of perceptual information about test words should lead to more accurate and faster recognition judgments of words studied as words than of words studied as pictures. On the other hand, performance at long response deadlines is more likely to reflect the processing of semantic features and of those features that code the form of the object to which each word refers. Since the encoding of these distinctive features of pictures appears to be at the root of the picture superiority effect (see, e.g., Nelson, 1979), recognition memory discrimination at long response deadlines, being based on the availability of this type of information, is greater for studied pictures than for studied words. A revision of the FESTHER model (which bases recognition judgments on familiarity only) that takes into account perceptual as well as other features of studied stimuli may then be able to account for the results of the present experiment. It is important to note, however, that in order to account for the present data familiarity must be considered to be a signal characterized by qualitatively different features that become available at different times during the retrieval process. Hence, whereas familiarity may be characterized as a continuous index, this index is nevertheless characterized by the contribution of the differential processing of multiple and qualitatively different features of test stimuli.

An addendum to the argument above is relevant at this point. If the encoding of distinctive features of pictures is at the root of the picture superiority effect, the suggested

modification of FESTHER may be too parsimonious to account for the picture superiority effect at long response deadlines. The reason for this is that the presentation of a target word at test does not necessarily elicit the corresponding studied picture with its distinctive features. And of course, if it does, this may be regarded as recollection.

In conclusion, the present study has shown that the picture superiority effect can be reversed in single probe recognition memory when people have limited time to respond to target stimuli at test. On one hand, the outcome of the present study can be accounted for by assuming the contributions of a fast-acting familiarity process based on the perceptual characteristics of target stimuli and those of a slow, recollective process based on distinctive features of studied pictures. Nevertheless, we have also indicated that the results obtained in the present study could be accounted for, in principle, by models that do not require the contribution of recollective processes. Finally, although it is still a matter of debate which of the above classes of models may provide a better account of the picture superiority/inferiority effect in recognition memory, the present study further highlights the need to take into account the different processes associated with different attributes of test stimuli, in order to provide a better understanding of the processes supporting the performance in recognition memory tasks.

AUTHOR NOTE

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

REFERENCES

Amrhein, P. C., McDaniel, M. A., & Waddill, P. (2002). Revisiting the picture-superiority effect in symbolic comparisons: Do pictures provide privileged access? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 843-857.

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance* (Vol. 6, pp. 583-612). New York: Academic Press.

Boldini, A., Russo, R., & Avons, S. E. (2004). One process is not enough! A speed–accuracy tradeoff study of recognition memory. *Psychonomic Bulletin & Review*, **11**, 353-361.

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

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.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, **28**, 923-938.

Dewhurst, S. A., & Conway, M. A. (1994). Pictures, images and recollective experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1088-1098.

Dunn, J. C. (2004). Remember–know: A matter of confidence. *Psychological Review*, **111**, 524-542.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1999). Response deadline and subjective awareness in recognition memory. *Consciousness & Cognition*, **8**, 484-496.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 229-244). New York: Oxford University Press.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-67.

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

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.

Howell, D. (1997). *Statistical methods for psychology*. Belmont, CA: Duxbury.

Humphreys, M. S., Dennis, S., Chalmers, K. A., & Finnigan, S. (2000). Dual processes in recognition: Does a focus on measurement operations provide a sufficient foundation? *Psychonomic Bulletin & Review*, **7**, 593-603.

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.

Kinjo, H., & Snodgrass, J. G. (2000). Is there a picture superiority effect in perceptual implicit tasks? *European Journal of Cognitive Psychology*, **12**, 145-164.

Lamberts, K., Brockdorff, N., & Heit, E. (2002). Perceptual processes in matching and recognition of complex pictures. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 1176-1191.

Madigan, S. (1983). Picture memory. In J. C. Yuille (Ed.), *Imagery, memory, and cognition: Essays in honor of Allan Paivio* (pp. 43-69). Hillsdale, NJ: Erlbaum.

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

Mintzer, M. Z., & Snodgrass, J. G. (1999). The picture superiority effect: Support for the distinctiveness model. *American Journal of Psychology*, **112**, 113-146.

Mulligan, N., & Hirshman, E. (1995). Speed–accuracy trade-offs and the dual process model of recognition memory. *Journal of Memory & Language*, **34**, 1-18.

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

Nelson, D. L. (1979). Remembering pictures and words: Appearance, significance, and name. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 45-76). Hillsdale, NJ: Erlbaum.

Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning & Memory*, **3**, 485-497.

Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 523-528.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.

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.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, **21**, 89-102.

Rajaram, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 365-377.

Reed, A. V. (1976). List length and the time course of recognition in immediate memory. *Memory & Cognition*, **4**, 16-30.

Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Sciences*, **7**, 313-319.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 174-215.

STENBERG, G., RADEBORG, K., & HEDMAN, L. R. (1995). The picture superiority effect in a cross-modality recognition task. *Memory & Cognition*, **23**, 425-441.

TOTH, J. P. (1996). Conceptual automaticity in recognition memory: Levels-of-processing effects on familiarity. *Canadian Journal of Experimental Psychology*, **50**, 123-138.

WELDON, M. S., & COYOTE, K. C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 670-686.

WELDON, M. S., ROEDIGER, H. L., III, & CHALLIS, B. H. (1989). The properties of retrieval cues constrain the picture superiority effect. *Memory & Cognition*, **17**, 95-105.

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. Note that global memory models can be complemented with a recall mechanism. This is often not required to account for recognition memory data.

2. Statistical analyses on d' scores were carried out not only on responses given on time, but also on those given with an extended timeout of 500 msec. The two analyses yielded comparable results.

3. An anonymous reviewer was concerned that a large number of subjects was excluded from the Experiment 3 analyses. The main concern was that the included subjects could somehow differ from the entire sample. To address this concern, we conducted a further ANOVA that also included data from 13 of the 16 discarded subjects (unfortunately, data of 3 subjects were lost). The analysis of this substantially larger subsample showed the same pattern of results reported in the main analysis. The interaction between stimulus type and response deadline was significant [$F(1,35) = 34.05$, $MS_e = 0.16$, $p < .01$], indicating that the effect of stimulus type differed in different response deadline conditions. In the 150-msec deadline condition, recognition accuracy was significantly better for words ($d' = 0.581$) than for pictures ($d' = 0.365$) [$t(35) = 2.79$, $p < .02$]. In the 2,000-msec deadline condition, the picture superiority effect was significant (d' for pictures $= 1.782$; d' for words $= 1.221$) [$t(35) = 6.85$, $p < .01$].