EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2003, *15* (4), 589–605

Remember, know, confidence and the mirror effect: Changes as a function of discriminability conditions

Salvador Algarabel, Arcadio Gotor, and Alfonso Pitarque

University of Valencia, Spain

Recognition memory for Spanish-Catalan cognate and noncognate words was tested at retention intervals of 20 minutes, 1 hour, and 24 hours (Experiment 1) using a remember/know response procedure, and requiring a confidence judgement on the yes/no response. Noncognate words were accompanied by more ''remember'' responses than cognates, and overall A′ was significantly different from remember A′, except in the cognate condition at the longest retention interval. A strong mirror effect for the cognate–noncognate stimulus class was found for overall responding, and for high but not low confidence, indicating a differential use of recollection and familiarity in recognition. In general, the pattern of results was inconsistent with Donaldson's (1996) signal detection model, indicating that, when available, subjects use two different sources of information for discrimination. The examination of individual hits and false alarms as a function of confidence indicated that ''remember'' is uniformly associated with high confidence, but ''know'' shows a bipolar pattern. In Experiment 2, new and old words were repeated at test 2 and 3. Repetition greatly affected the difference between the discrimination indices, indicating that an increase in the familiarity of new words prevented the use of a dual source of information in recognition. Results are discussed in terms of Rajaram's distinctiveness (1996, 1998) and Reder, Nhouy-vanisvong, Schunn, Ayers, Angstadt, and Hiraki's (2000) SAC theories.

The idea that recognition may be mediated by more than one type of information and/or mechanism is hardly new (Jacoby & Dallas, 1981; Mandler, 1980). Of more contemporary novelty is the association between recognition and states of consciousness, brought to mainstream research by Tulving's distinction between remembering and knowing (Tulving, 1985). In this tradition, remembering (r) refers to the recognition of an item plus its encoding context, whereas knowing

Requests for reprints should be addressed to S. Algarabel, Facultad de Psicologia, Universidad de Valencia, Blasco Ibañez, 21, 46010 Valencia, Spain. Email: Salvador.Algarabel@uv.es

Preparation of this paper was supported by Grant PB/97-1379 from The Dirección General de Investigación Científica y Técnica (Spanish Ministry of Education and Technology). Part of the data reported here were presented at the Third International Conference on Memory, Valencia, Spain, July 2001. We acknowledge the comments of Dr Hockley, which were crucial in our analysis of the mirror effect.

http://www.tandf.co.uk/journals/pp/09541446.html DOI:10.1080/09541440340000033

(k) denotes the conscious state when those environmental cues are absent. This two process theory of remember–know is based on a large body of experimental dissociations (for reviews see, Gardiner & Java, 1993; Gardiner & Richardson-Klavehn, 2000), showing that some variables have an effect on remembering and not knowing; others on knowing and not remembering; and others have parallel or opposite effects.

Rajaram's distinctiveness/fluency processing view (1996, 1998; Rajaram & Geraci, 2000) helps to look at the variables that influence one or the other judgement. Her original formulation established that remembering was the result of deep processing of the stimulus, whereas knowing was the product of perceptual processing. However, the observation that perceptual changes had an effect on remember responses (Rajaram, 1996) and that changes in conceptual fluency affected knowing (Rajaram & Geraci, 2000; see also Gardiner, Gregg, Mashru, & Thaman, 2001) led to a reformulation in which remember is supposed to be influenced by any distinctive manipulation, and knowing by fluency of processing. A closely related account, more linked to the explanation of the basic mechanisms, is the source of activation confusion theory (SAC; Reder, Nhouyvanisvong, Schunn, Ayers, Angstadt, & Hiraki, 2000). In this theory, one node represents the actual presented concept, whereas another one represents the encoding episode. ''I remember'' is based on the activation of the episodic node, whereas ''I know'' derives from the activation of the word node. The theory makes strong and interesting predictions about several phenomena, but we are particularly interested in the mirror effect. The mirror effect refers to a pattern of greater hits for low than high frequency words, and a reverse pattern for false alarms. The theory says that as high frequency words have more connections than low-frequency ones, when they are activated, the spreading activation to the context node is less (Reder, Angstadt, Cary, Ericksen, & Ayers, 2002) than when they are less frequent. We can say also, following Rajaram's theory, that high frequency words are less distinctive. In any case, this is a convincing explanation of the mirror effect.

Not all explanations of the recognition data recur to more than one factor. The signal detection theory (Donaldson, 1996; Hirshman & Master, 1997), in the tradition of the single process explanations of recognition memory (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988; Hockley & Murdock, 1987), has been formulated as an alternative explanation of the experimental literature. The theory argues that r/k states can be understood by assuming a single continuum of mnemonic information over which two criteria are set. Items that lie above the first criterion are identified as having previously occurred, and get an ''old'' response. A second criterion is then established that partition these ''old'' responses into those above the new criterion, which are labelled ''remember'', and those below it, which are labelled ''know''.

Although, several predictions can be made from this formulation (Donaldson, 1996), a key one that routinely has been investigated is that the discrimination index calculated from the overall recognition data (A′[y/n]) should be the same as when it is calculated from the remember only data (A′[rem]). An original meta-analysis on a series of published experiments carried out by Donaldson (1996) was not completely clear regarding this point. Gardiner and Gregg (1997) noted that, although the analysis on parametric d′ was in line with single trace expectations, nonparametric A′ differences were in the correct direction, that is, A′(y/n) was significantly better than A′(rem). Since in r/k experiments changes in criterion are expected, A′ is a better index of discrimination than d′. Some authors (e.g., Yonelinas, 2001) have pointed out, though, that this ''standard'' single factor theory can be modified to account for the differences among the discrimination parameters. In particular, if we relax the requirement that both the old and new item distributions have the same variance, then the modified theory could explain the possible differences between discrimination indices calculated from y/n or remember only data. Nonetheless, the traditional literature on recognition shows that variance of the old is usually larger than that of the new distribution (e.g., Glanzer, Kim, Hilford, & Adams, 1999). Therefore, the modified theory would be a credible explanation if the analysis of the published literature shows that the variances of the old and remember distribution follow the same pattern.

On the other hand, although the single process theory reduces r and k to differences in criterion setting, the explicit experimental scrutiny of the relation between confidence and awareness has not thoroughly been investigated beyond the usual signal detection analysis (e.g., Donaldson, 1996; Hirshman & Henzler, 1998; Postma, 1999). Initially, Tulving (1985) asked simultaneously for the evaluation of remember/know and confidence, and found that both were highly correlated. The simultaneous measurement of both responses was later criticised on methodological grounds (Gardiner & Java, 1990). Subjects may base their responses on the recollective experience exclusively, contaminating the rest of the judgements. This possibility precluded further direct, simultaneous, and independent measurement of both responses, with some exceptions (Holmes, Waters, & Rajaram, 1998; Rajaram, 1993). As a result, sure/unsure, as an isolated judgement, was subject to the same variables as r/k judgements, trying to determine whether they behaved identically or differently. If identical, the two process view might be in trouble. For example, Gardiner and Java (1990) found that when ''sure/unsure'' was used as a response category instead of ''remember/know'', the pattern of results in word–nonword recognition was opposite to that obtained with remember–know. Parkin and Walter (1992) found the same result with regard to the age variable. However, in a simultaneous and independent measurement of confidence and r/k, Rajaram (1993) also showed that r/k judgements were not affected in the same way as sure/unsure in a masked primed recognition task. However, a follow-up analysis of both experiments by Inoue and Bellezza (1998) found differences in sensitivity and/or bias among the critical conditions and dismissed them as negative evidence against the trace strength hypothesis. Inoue and Bellezza also noted that ‘‘know’’ responses represented a stricter criterion than ‘‘unsure’’, pointing to the imperfect correlation between the two different types of recollective experience and confidence. The effects of experimental manipulations were not always different on r/k and confidence responses. Perfect, Mayes, Downes, and van Eijk (1996) tried to determine differences in contextual information between r and k judgements and found (Exp. 2) this time that there was a close correlation between both variables. Finally, Holmes et al. (1998) discovered that know and remember judgements showed a linear abstraction effect. That is, using Bransford and Franks’ (1971) semantic integration task, remember and know judgements increased linearly with the number of idea units of the presented sentences, despite a near chance recognition level. More importantly, confidence also increased linearly with sentence complexity and was superior in the case of remember. Furthermore, the comparison of Experiments 1 and 2 by Holmes et al. indicated that the simultaneous request of remember/know and confidence judgements did not disrupt or alter the subject behaviour in the experiment. Overall, we can conclude that the relation of confidence and conscious states, as measured in remember/know judgements, is not completely consistent with neither view, although we find more support in the literature for the dual process account.

On the other hand, the analysis of the absolute or relative discriminability level across experimental conditions is also a variable of important consequences for the production of know and remember judgements. If two conditions are compared, the most discriminable one will produce a greater remember rate. This argument has been used previously (Inoue & Bellezza, 1998) to undermine the support of published experimental data (Gardiner & Java, 1990; Rajaram, 1993) for the two process theory. Hirshman and Lanning (1999), for example, followed this same argument to question the association between self-reference encoding (Conway & Dewhurst, 1995; but see Conway, Dewhurst, Pearson, & Sapute, 2001, for its rebuttal; Hirshman, Lanning, Master, & Henzler, 2002) and a greater number of remember responses. We expect that, as in the case of the relation between confidence and availability of episodic information, there should be a logical relation between discriminability and availability of contextual and item information. If ‘‘given the opportunity’’, and both item and contextual information are available, the ‘‘use’’ of both in recognition is much more likely in high versus low discrimination conditions. We think that an important consideration in r/k experiments is to provide ample opportunities for the encoding of contextual information.

A situation in which contextual information is explicit is in associative recognition. Hockley and Consoli (1999) presented lists of single words (item information) and pairs of words (associative information) for regular recognition and remember/know judgements. Their conclusion, comparing A′(y/n) versus A′(rem), was that item recognition was better explained by a signal detection model. However, in the case of associative information, A′(rem) was greater than A′(y/n), indicating that the postulation of a single source of information for recognition decision was not enough to explain the results. However, Hockley and Consoli's conclusion with regard to item recognition might have been premature. Participants may not have had the possibility of encoding episodic information in their single item condition, despite the fact that they chose items high in imaginability. This possibility is likely given their use of mixed lists of single and paired words. Moreover, their use of the three-response category: ''remember/know/new'' (see Hicks & Marsh, 1999) might have modelled undesirable strategies in the participants.

In the following experiments, we pursue three closely related goals. The first is to test single versus dual explanations of remember/know-based item recognition in situations in which we provide ample opportunities for the use of episodic information in recognition. Second, we intend to analyse the structure of r/k judgements as a function of confidence in line with some previous independent measurements of both variables (see for example, Holmes, Waters, & Rajaram, 1998). Finally, we will look at the response structure for the two stimulus classes (the mirror effect) as an additional way of providing evidence for or against single versus dual explanations of r/k. More than an approach based on dissociations, we base our interpretation in the finding of possible qualitative differences in the task at hand, as a function of the possible action of one or a mixture of two factors (see Joordens & Hockley, 2000).

To accomplish these goals, we use the same basic experimental design reported by Hockley and Consoli (1999) for item recognition, although changing the basic task and providing greater opportunities for the use of recollection. Our task was quite different and aimed at maximising encoding richness. It consisted of the presentation of a list of Catalan words, half of them Spanish cognates, for which the bilingual participants had to provide the Spanish equivalent. Catalan and Spanish are two very closely related Romanic languages which share many graphemic and phonemic characteristics. We think that this type of task provides for a great deal of distinctive stimuli and processing (Rajaram & Geraci, 2000), particularly when cognates and noncognates are presented to the same participant (Dewhurst & Parry, 2000). The sample of participants is perfectly fluent in Spanish and we expect a less distinctive processing of the Catalan cognate words, given their similarity to Spanish words, than of the noncognates. The literature shows that all variables that increase the distinctiveness of the to be recognised items at any level will increase r responses, and vice versa. This is the case for variables like attention (Gardiner & Parkin, 1990), intentionality in the study phase (Macken & Hampson, 1993), orthographic distinctiveness (Rajaram, 1998), emotionality (Dewhurst & Parry, 2000), and word frequency (Gardiner & Java, 1990), among others. Moreover, depending on the richness of the encoding, an effect may act on remembering or knowing. Such is the case of the size congruency effect under deep (influence on r) or shallow processing (influence on k) as recently shown by Gardiner et al. (2001).

Accordingly, we have the following general expectations: better discriminability of the noncognates than the cognates and, as is the case of the effect of word frequency on r/k judgements (Gardiner & Java, 1990, Exp. 1), a greater level of r responses. But we expect the difference between A′(y/n) and A′(rem) to be significant or not as a function of additional discriminability conditions. With regard to confidence, we expect r to be more confident than k, but we anticipate observing a clear demonstration that, although correlated, r/k is not identical to confidence. We base this expectation on the fact that confidence may arise from a number of sources and not only from the presence of episodic information. Finally, we test recognition at three retention intervals to allow for a possible differential loss of either item or episodic information and shift the use of the dual source of information as the basis of the recognition judgement.

In Experiment 1, the delayed recognition tests were based on different words from the study list, completing a sequence of three tests after 20 minutes, 1 hour, and 24 hours. In Experiment 2, recognition was tested on the same words at the same intervals. Therefore, we expect to observe differences between A′(y/n) and A′(rem), but, as a function of repetition or forgetting, and more likely in the cognate condition, the difference may disappear due to changes in discrimination conditions.

EXPERIMENTS 1 AND 2

Method

Participants. Thirty-six and thirty-five Psychology students participated in each experiment for course credit. All of them had a perfect knowledge of Spanish but were knowledgeable of Catalan as a second language to different degrees.

Materials. We selected 168 Catalan words from Spanish-Catalan norms (Nacher, Gotor, & Algarabel, 1998). Half of them were Catalan-Spanish cognates (as in ‘‘tranvia/tramvia’’—trolley), and the other half were not (as in ‘‘fusta/madera’’—wood). The separation of cognates and noncognates was carried out by computing a similarity ratio. The ratio takes into account the number of common letters in the same positions, and varies between 0 (no similarity) and 1 (perfect similarity). The specific values for this index are listed below.

Procedure. The study phase of the experiment consisted of a list of 94 words with the first and last 5 considered as recency buffers and not tested later. We constructed two separate lists of the same length. Both lists were equated on word frequency (Alameda & Cuetos, 1995) in Spanish and Catalan (Gotor, Borras, & Perea, 1994) and length. Concretely, the specific mean values for the main parameters for both lists were 105 (frequency), 5.6 (number of letters), and 0.1 (similarity ratio), and 121, 5.52, and 0.2 for noncognate lists A and B, respectively. For cognates, the respective values were: 131 (frequency), 5.8 (length in letters), and 1.0 (similarity ratio) for list A, and 115, 5.86, and 1.0 for list B. The test lists were created according to the following procedure. We divided each study list into seven consecutive blocks of 12 words, and from each block we sampled 4 different words for each of the three test lists until the list of 28 words was formed. This process was carried out independently for each participant, who received a different randomised study list printed in a booklet. Participants were told that their task was to provide the Spanish translation for each Catalan word in the blank space provided to the right of each word, or leave the space blank in the case that they did not know the answer. They were also told that they would be tested with an unspecified memory test at a later point, although they did not know that they were going to receive multiple tests. Each word was timed by the experimenter and every 5 s they were told to advance a word, until the end of the list was reached, at which point the booklet was collected. The first test started 20 min after the study phase. The second test started 1 hour later, and the third 24 hours later. The three tests were identical except for the differences in retention interval. In the first experiment, each participant received a different set of old words at test time in a new booklet. In the second, the participants received the same set of old and new words at test time. In both cases, old words were randomly mixed with new words (coming from the second list) in equal proportions. The delays between the first and second test were filled up by the regular activities of a Psychology course. This second test was not announced in advance, and there were no environmental cues suggesting that it was going to take place. The third test was carried out the following day at the start of the regular class and was also unannounced and unexpected. All tests were self-paced but were quickly completed. The instructions given to the participants asked for three responses to each stimulus. The first response was a new/old discrimination, the second (only for yes responses), a remember/know/guess, and the third (only for yes responses), a confidence decision on the new/old according to a 4-point scale. The instructions followed others usually found in the literature (e.g., Rajaram, 1996) with the addition of instructions for ''guess'' (g). If the initial response was ''old'', and they were able to bring to mind specific aspects of the previous occurrence, like details of the physical appearance, place in the list, or any other image or idea occurred at the moment of the presentation, they have now to respond, ''I remember''. If they recognised the word as old but were unable to recollect any other specific event, then they were to say, ''I know''. If undecided, ''I guess''. Additionally, we emphasised the fact that the confidence response was independent from the r/k/g and had to be carried out in relation to the old/new discrimination. We explicitly made clear that a remember response would be possible for strongly remembered or weakly remembered items, and that the same was true of the know response. We gave examples of these possibilities and emphasised again the above difference.

Results

In the following analysis, first we look at nonparametric indices of discriminability and bias (Donaldson, 1992, 1996) to be followed by the analysis of the remember versus know responses as a function of confidence. As the responses in the confidence judgements were very unevenly distributed, we decided to group the two highest and two lowest categories in a single high and low confidence response and separately analyse know and remember at high and low confidence. Likewise, guess responses were at a very low level, and we did not carry out any statistical analysis on them. A significance level of $p < .05$ is assumed throughout.

Recollection versus recognition. A′ (Donaldson, 1996) for overall recognition (A′[y/n]) and for remember responses (A′[rem]) and confidence (A′[+/−]) were calculated as a function of retention interval and linguistic status for Experiments 1 and 2.

In Experiment 1 (see upper part of Figure 1), noncognate words were more accurately recognised than cognates, $F(1, 35) = 50.54$, $MSE = 0.021$, performance declined as a function of successive tests, $F(2, 70) = 74.33$, $MSE = 0.006$,

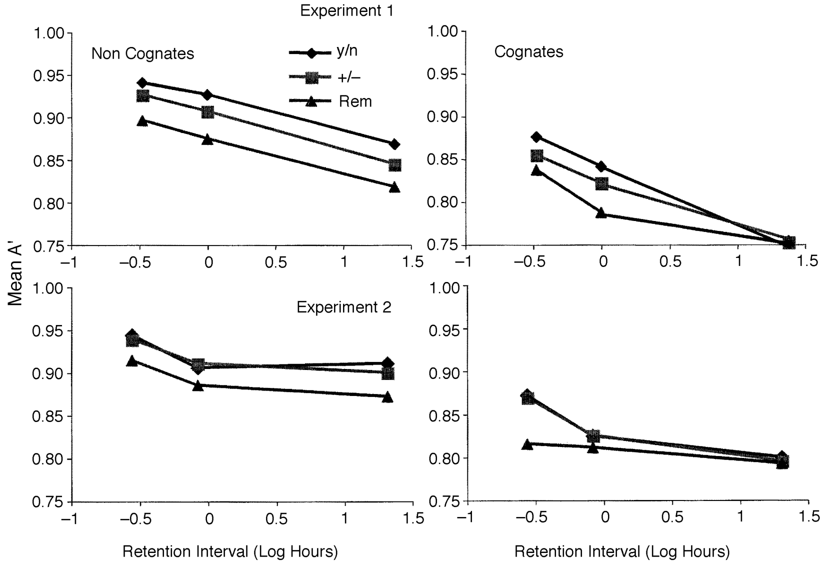

Figure 1. Mean A′ for overall recognition (y/n), confidence (+/−), and remember only responses (Rem) as a function of retention interval for Experiments 1 and 2.

and type of response was also significant, $F(2, 70) = 13.319$, $MSE = 0.006$. Tukey tests revealed that all contrasts between A′(y/n), A′(rem), and A′(+/−) were significant and they were in that decreasing order (.8665, .8509, .8277). Trial by type of response was marginally significant, $F(4, 140) = 2.21$, $p = .071$, as was the higher order interaction among linguistic status, trial, and type of response, $F(4, 140) = 2.36$, $MSE = 0.002$, $p = .056$.

With regard to bias (B''_D), there was a main effect of type of response, $F(2, 70) = 45.63$, $MSE = 0.364$, and linguistic status, $F(1, 35) = 5.72$, $MSE = 0.176$. Tukey tests revealed that the criterion based on overall responses was less strict (.4660) than the criteria based on confidence and on remember only responses (.9460, .9465). More important, the difference in criteria between these last two was not significant. The interaction of response type by linguistic status was also significant, $F(2, 70) = 3.11$, $MSE = 0.113$, as was the interaction between response and trial types, $F(4, 140) = 3.96$, $MSE = 0.079$. The combined analysis on A′ and B''_D indicates that, although remember responses are very high confidence responses, this is not due simply to a difference in criterion placement, as discrimination is not homogeneous.

In Experiment 2 (see lower part of Figure 1), the data looks similar. Noncognate words were more accurately recognised than cognates, $F(1, 34) = 73.43$, $MSE = 0.019$, overall recognition decreased with trials, $F(2, 68) = 21.78$, $MSE = 0.005$, and response type was significant, $F(2, 68) = 5.48$, $MSE = .009$. Tukey tests revealed that A′(y/n) was not different from A′(+/−), but both were different from A′(rem) (.79 and .81 vs. .69). The interaction of trial by response type was also significant, $F(4, 136) = 2.86$, $MSE = .002$, as was the interaction of linguistic status by trial by response type, $F(4, 136) = 3.27$, $MSE = 0.001$. To clarify this three-way interaction, we separately analysed trial by response type at each level of linguistic status. For noncognates, the interaction of trial by response type was not significant, $F(4, 136) < 1$. This interaction, on the other hand, was highly significant for cognate words, $F(4, 136) = 4.67$, $MSE = 0.002$. A simple effect analysis of this interaction showed that response type was different in the first trial, $F(2, 68) = 8.67$, $MSE = .004$, but not in trials 2 and 3, $F(2, 68) < 1$ in both cases.

The analysis of bias showed the same results as in Experiment 1. Criterion was more strict for cognates than noncognates, $F(1, 34) = 4.53$, $MSE = 0.694$, was also marginally more strict for the first trial than for the rest, $F(2, 66) = 2.52$, $MSE = .178$, $p = .09$ and, more importantly, was less strict for overall recognition data than for bias calculated either from confidence or remember data, $F(2, 68) = 40.44$, $MSE = 0.021$. Tukey tests confirmed that the F test was significant because of that contrast. Again, there was no difference between the criterion estimates corresponding to confidence and remember data.

Remember versus know responses as a function of confidence. The mean proportions of correct remember and know responses by linguistic status, retention interval and confidence are given in Table 1 separately for Experiments 1 and 2. Remember and know responses were separately analysed for high and low confidence.

In Experiment 1, for high confidence, the proportion of remember responses was higher for noncognates than for cognates, $F(1, 35) = 49.41$, $MSE = 0.044$, although false alarms were marginally significant in the opposite direction, $F(1, 35) = 3.39$, $MSE = 0.002$. Remember responses decreased with trials, $F(2, 70) = 49.31$, $MSE = 0.019$, but the decrease in false alarms was not significant, $F(2, 70) = 1.93$, $MSE = 0.002$, $p > .15$. No further main effect or interaction was significant.

For low confidence, the only significant effect was the greater proportion of cognate words versus noncognate words, $F(1, 35) = 7.61$, $MSE = 0.001$. In false alarms, the interaction of linguistic status by trial approached significance, $F(2, 70) = 2.54$, $MSE = 0.000$, $p = .09$. The level of responding in these conditions was almost zero, although there was a very slight nonsystematic difference in responding with trial for cognate words.

For high confidence know responses, a main effect of retention interval was found, $F(2, 70) = 4.48$, $MSE = 0.015$, indicating a decrease in responses with time. Cognates produced more false know alarms than noncognates, $F(1, 35) = 12.92$, $MSE = .002$. In low confidence, know responses for noncognates were lower than for cognates, $F(1, 35) = 25.25$, $MSE = 0.015$, and they increased steadily with trials, $F(1, 35) = 12.94$, $MSE = 0.008$. The effect of linguistic status and trial, as well as their interactions, was significant in false alarms, respec-

TABLE 1
Mean proportions of remember and know responses to old and new test words as a function of confidence (high and low), linguistic status (noncognates, NC, and cognates, C), and test (1, 2, and 3)

	Old								*New*							
	High				*Low*				*High*				*Low*			
	Rem		*Know*		*Rem*		*Know*		*Rem*		*Know*		*Rem*		*Know*	
Test	*NC*	*C*	*NC*	*C*	*NC*	*C*	*NC*	*C*	*NC*	*C*	*NC*	*C*	*NC*	*C*	*NC*	*C*
Experiment 1																
1	0.59	0.39	0.13	0.13	0.00	0.02	0.09	0.19	0.01	0.01	0.00	0.03	0.00	0.00	0.03	0.09
2	0.54	0.32	0.13	0.14	0.02	0.02	0.10	0.21	0.01	0.04	0.01	0.03	0.00	0.01	0.04	0.10
3	0.36	0.18	0.08	0.08	0.01	0.02	0.19	0.24	0.01	0.02	0.01	0.03	0.01	0.00	0.05	0.17
Experiment 2																
1	0.69	0.37	0.12	0.18	0.01	0.01	0.06	0.14	0.02	0.02	0.01	0.02	0.00	0.00	0.04	0.08
2	0.67	0.38	0.09	0.12	0.02	0.02	0.06	0.16	0.04	0.04	0.03	0.05	0.01	0.01	0.06	0.11
3	0.64	0.33	0.09	0.10	0.02	0.04	0.10	0.18	0.05	0.05	0.03	0.05	0.01	0.01	0.06	0.14

tively, $F(1, 35) = 28.45$, $MSE = 0.012$, $F(2, 70) = 11.69$, $MSE = 0.004$, $F(2, 70) = 5.62$, $MSE = 0.004$, indicating greater false alarms for cognates, with trials, and for cognates as a function of trials.

In Experiment 2, for high confidence the proportion of correct responses was greater for noncognates than cognates, $F(1, 34) = 90.87$, $MSE = 0.056$, and there was a marginally significant decrease in hits with trials, $F(2, 68) = 3.92$, $MSE = 0.011$, as well as an increase in false alarms, $F(2, 68) = 5.83$, $MSE = 0.003$.

In the case of high confidence know responses, the only significant effect was the decrease in the proportion correct with trials, $F(2, 68) = 7.49$, $MSE = 0.008$. False alarms increased with trials, $F(2, 68) = 4.31$, $MSE = 0.003$, indicating that they were lower in the first versus the second and third trials. For low confidence, again cognates were significantly greater than noncognates, $F(1, 34) = 33.44$, $MSE = 0.012$, and they increased with trials, $F(2, 68) = 3.97$, $MSE = 0.008$. The same pattern of significant effect was found in false alarms, respectively, $F(1, 34) = 10.53$, $MSE = 0.017$, and $F(2, 68) = 6.75$, $MSE = 0.005$.

Analysis of the mirror effect (cognate vs. noncognates) for the unpartitioned and confidence data. With the purpose of analysing the mirror effect we carried out a standard analysis of variance of 2 (linguistic status) $\times$ 3 (trials) separately for hits and false alarms, and for the unpartitioned (overall) and confidence (high–low) data. For the unpartitioned proportions, noncognate hits were significantly higher than cognates for Experiment 1, $F(1, 35) = 16.71$, $MSE = 0.033$, and Experiment 2, $F(1, 34) = 44.83$, $MSE = 0.039$. The false alarm analysis produced the same results but the conditions means were in the opposite direction, $F(1, 35) = 31.98$, $MSE = 0.023$, and $F(1, 34) = 10.68$, $MSE = 0.032$. This is the first demonstration of a mirror effect for the two classes of stimuli defined by the cognate–noncognate distinction. Additionally, in both experiments, the hit rate decreased, $F(2, 70) = 61.15$, $MSE = 0.013$, and $F(2, 68) = 4.04$, $MSE = 0.010$, and the false alarm rates increased, $F(2, 70) = 6.80$, $MSE = 0.010$, and $F(2, 68) = 20.71$, $MSE = 0.010$. This phenomenon, the decrease in hits and the increase in false alarms, is defined as concentering (Glanzer, Adams, & Iverson, 1991; Joordens & Hockley, 2000) and is the result of forgetting.

In the previous section we analysed the judgement categories by confidence. We have carried out the same analysis for remember and know collapsing and eliminating the confidence partition. The results were virtually the same. We are not going to repeat them here, and we base our interpretation on the previous and reported results. As shown in that analysis, the mirror effect is produced by the differences in hit rates reflected in remember responses, and in false alarm rates reflected in the know judgements.

Finally, we repeated the analysis separately for high and low confidence. For high confidence, the typical mirror effect was obtained in Experiment 1 but not in Experiment 2: in Experiment 1, $F(1, 35) = 62.17$, $MSE = 0.033$, and $F(1, 35) = 13.23$, $MSE = 0.004$, and $F(1, 34) = 125.97$, $MSE = 0.031$, and $F(1, 34) = 2.75$, *MSE* = 0.009, $p > .10$, for Experiment 2. The absence of the mirror effect in Experiment 2 was due mainly to an increase in false alarms for the noncognates. However, no mirror effect was observed in the analysis carried out on low confidence: $F(1, 35) = 30.95$, *MSE* = 0.016, and $F(1, 35) = 30.37$, *MSE* = 0.012, for Experiment 1, and $F(1, 34) = 29.06$, *MSE* = 0.016, and $F(1, 34) = 10.99$, *MSE* = 0.017. In this case, the lack of the mirror effect was due to a reverse pattern of hits. That is, cognates showed greater number of hits than noncognates in Experiments 1 and 2.

DISCUSSION

The present study shows that, under the appropriate conditions, subjects use two different sources of information (item and recollective) in recognition, producing a further differentiation between remember and know judgements. In Experiment 1, A′(y/n) was significantly superior to A′(rem) with noncognate and cognate words. Only in the cognate condition and after a 24-hour delay, did the difference between discrimination indices disappear, presumably because there is a balance shift between the use of recollection (context) and familiarity (item). A similar explanation can be applied to the second experiment, although in this case, this is due to repeated testing of the same words. Both experiments taken together, and in conjunction with Hockley and Consoli's data, indicate that both encoding richness (information availability) and discrimination condition at test are important determinants of r/k performance. The analysis on A′ calculated from the confidence judgements indicated that the subjects were perfectly able to emit them independently of the r responses. Nicely, bias estimations were exactly the same for r only and for confidence responses, and much lower for overall recognition (remember plus know). The presence of episodic information drives up confidence, and whereas remember responses are overwhelmingly high confidence responses, know responses are split with regard to it (see also Holmes et al., 1998; Koriat, 1993). The analysis of the individual hits and false alarms as a function of the different conditions is very instructive in the case of Experiment 1. The decrease in hits associated with retention interval is on the order of 23% in the remember category but stayed constant in the know category, when both were not parcelled out by confidence; statistically, $F(2, 70) = 49.14$, *MSE* = 0.020, for overall remember, and $F(2, 70) < 1$, for know. The corresponding false alarms were nonsignificant, $F(2, 70) = 1.72$, *MSE* = 0.002, and almost of negligible value for remember, but increased significantly, $F(2, 70) = 9.02$, *MSE* = 0.006, for know. Interestingly, when the two types of judgement are analysed separately for high and low confidence, the decrease in remember responses is associated exclusively with the high confidence category, and does not transmit to low confidence. In fact, apart from not finding an effect of retention interval on low confidence remember, its level was insignificant. The exactly opposite pattern was observed in know responses. Whereas the previous analysis showed no effect of retention interval on overall know, the analysis of high confidence know indicates that there is a significant decrease with interval, compensated by an increase on low confidence. A high confidence decrease plus a low confidence increase in know leads to an overall noneffect of retention interval on know. Although these differences in confidence and A′ are not explainable by a single factor signal detection account, the theory could still be saved releasing the assumption of equal variances for the old-remember distributions. However, a recent meta-analysis shows (Rotello, Macmillan, & Reeder, 2001) that the ratio of those two variances are very different (about or more than 1.0) to the ratio calculated from the old–new traditional data (about 0.8). We conclude that the introduction of the unequal variance assumption is not by itself sufficient to keep the explanatory power of the theory, and a two factor explanation results more natural for the data.

The analysis of the mirror effect provides converging evidence for the dual factor interpretation. First, and most important, the lack of a mirror effect in the low confidence data is due to a reduced use of recollection, as the hit rates for noncognates are lower than for cognates in both experiments. Second, the reason why in the second experiment there is no mirror (confidence data) is because of an increase in familiarity (false alarms) affecting particularly the noncognates. In general, the data shows that in the first experiment there is a steeper decrease in hits (less use of recollection), but in the second, there is a growth in familiarity and the use of recollection stay at a high level because repetition compensates its loss because of forgetting.

Theoretically, these data support the distinctiveness/fluency account (Rajaram, 1996, 1998). As expected, noncognates produced more remember responses than cognates, and this is true even when participants were not able to give a Spanish translation for 19% of the presented words for study. The access to meaning is not the important aspect of the processing of words here. It is the fact that in a mixed list (Dewhurst & Parry, 2000) noncognates are much more distinctive on several grounds previously hypothesised in the introduction. Also worth pointing out is the fact that some past experiments (Gardiner & Java, 1990; Rajaram, 1993) have been criticised (Inoue & Bellezza, 1998) using the argument that there is a difference in absolute discriminabilities between the basic experimental conditions. The present experiment shows that absolute discriminability differences are not sufficient or necessary for rejecting a two process explanation.

Although Rajaram's theory focuses on the factors leading to influences on r and k, it does not indicate which type of relation exists between the information supporting familiarity and recollection. In this respect, Knowlton and Squire (1995; see also Kelley & Jacoby, 2000) described three possible ways in which that relation could take place: redundancy, independence, and exclusivity. In redundancy, item is a prerequisite for episodic information to be acquired. Subsequently, with forgetting, the loss of episodic information leads necessarily to a k judgement. The independence view (for example Jacoby, Yonelinas, & Jenning, 1997; Yonelinas & Jacoby, 1995) differs from the previous explanation in that context and item information behave autonomously, and a remember response may either become a know or a miss. Finally, in exclusivity, r and k are two noninteractive processes, which are subject to the action of different variables. Knowlton and Squire (1995) retested the same subjects with the same stimuli 1 week apart. They found that a significant proportion of stimuli that elicited an r response later led to a k response. In these experiments we can carry out a similar analysis with certain restrictions. First, Experiment 1 gives us an indication of the amount of forgetting from retention interval 1 to 3, and the analysis of the data from Experiment 2 can give us a clue about the rate of conversion of r–k or k–r. As there is a strong repetition effect, given that the retention intervals are short, we are going to look simply at a possible differential change rate from r–k and k–r. We make the assumption that those items showing changes are little or not at all influenced by repetition. A 2 (type of change: from r–k vs. k–r) × 2 (test: first–second vs. second–third) analysis of variance produced a significant main effect of trial, $F(1, 34) = 5.89$, MSE = 0.002, and more importantly, the interaction between both variables, $F(1, 34) =$ 4.58, $MSE = 0.004$. Simple effects showed that this interaction was produced by the fact that the proportion of changes from remember to know doubled from first–second in comparison with second–third (0.04 vs. 0.08, respectively), whereas the proportion of changes of know to remember stayed constant (0.06 vs. 0.05). The equivalent analysis of false alarms showed a small but significant effect of trial, $F(1, 34) = 4.08$, $MSE = 0.001$, produced by a proportion of 0.01 vs. 0.02.

The previous analysis, together with the inspection of the individual hits as false alarms from Experiment 1, gives strong support to the independence model (Jacoby et al., 1997). That is, we expect that, when an item is given an r response and it is correctly recognised, the response be based on the availability of both item and episodic information. Forgetting may act independently on one or the other. If it acts on episodic information, the item may later produce a k response. However, if forgetting strongly affects item information, then it turns out to be not recognised. This is the most plausible explanation when considering the present results globally. First, we find a sizeable 22% decrease in hits from the first to the last test in overall remembering (Experiment 1), but no effect of retention interval on know. On the other hand, the analysis of remember–know as a function of confidence indicates that remember responses always remain highly confident, but they are subject to strong forgetting. However, there is a decrease in high confidence know followed by an increase in low confidence know with trials. Finally, the evidence from remember to know and vice versa indicates that there is a considerable greater flow of responses from r to k. Therefore, the lack of effect of retention interval on know responses is the by-product of the transformation of some of the remember–know responses. This ''transfer phenomenon'' may help to explain the apparent lack of effect of some variables on know responses, as published in the literature (e.g., Gardiner & Java, 1991; Gardiner & Parkin, 1990).

Finally, the current experiments, designed without knowledge of those of Joordens and Hockley (2000) and Reder et al. (2000, 2002), provide very strong convergent evidence for the theoretical positions held by both research groups. In their most recent experiments, Reder et al. (2002) found that with very low level of familiarisation pseudoword recognition did not show the mirror effect. However, with further training the typical pattern of higher hits for low frequent items and more false alarms for the more frequent emerged. We obtained the same pattern of results, but in this case in relation to the low confident cognate and noncognate stimulus class. With further training, Reder et al. obtained the mirror pattern, just as we did with the high confidence stimuli. As in their case, we found more natural to explain the results from dual theories, like the source of activation confusion framework (SAC; Reder et al., 2000). SAC resorts to the balance of recollection and familiarisation as the mechanism explaining the literature on remember–know and the mirror data. To reach the same explanatory power, single factor theories must turn to more indirect and less parsimonious mechanisms (see for example, Joordens & Hockley, 2000).

Manuscript received January 2002
Revised manuscript received July 2002

REFERENCES

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano* [Frequency dictionary of the linguistic units of Spanish]. Universidad de Oviedo, Departamento de Psicología, Oviedo, Spain.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, *2*, 331–350.

Conway, M. A., & Dewhurst, S. A. (1995). The self and recollective experience. *Applied Cognitive Psychology*, *9*, 1–19.

Conway, M. A., Dewhurst, S. A., Pearson, N., & Sapute, A. (2001). The self and recollection reconsidered: How a ''failure to replicate'' failed and why trace strength accounts of recollection are untenable. *Applied Cognitive Psychology*, *15*(6), 673–686.

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology*, *12*, 541–551.

Donaldson, S. A. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, *121*, 275–277.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory and Cognition*, *24*, 523–533.

Gardiner, J., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory and Cognition*, *18*, 579–583.

Gardiner, J. M., & Gregg, V. H. (1997). Recognition memory with little or no remembering: Implications for a detection model. *Psychonomic Bulletin and Review*, *4*, 474–479.

Gardiner, J. M., Gregg, V. H., Mashru, R., & Thaman, M. (2001). Impact of encoding depth on awareness of perceptual effects in recognition memory. *Memory and Cognition*, *29*, 433–440.

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory and Cognition*, *18*, 23–30.

Gardiner, J. M., & Java, R. I. (1991). Forgetting in recognition memory with and without recollective experience. *Memory & Cognition*, *19*, 617–623.

Gardiner, J. M., & Java, R. I. (1993). Recognising and remembering. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 163–188). Hove, UK: Lawrence Erlbaum Associates Ltd.

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.

Glanzer, M., Adams, J. K., & Iverson, G. (1991). Forgetting and the mirror effect in recognition memory: Concentering of underlying distributions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 81–93.

Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 500–513.

Gotor, A., Borras, J. R., & Perea, M. (1994). *Diccionari de Frequencies en Catala* [Frequency dictionary of Catalan]. Valencia, Spain: University of Valencia Press.

Hicks, J. L., & Marsh, R. L. (1999). Remember–know judgments can depend on how memory is tested. *Psychonomic Bulletin and Review*, *6*, 117–122.

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

Hirshman, E., & Henzler, A. (1998). The role of decision processes in conscious recollection. *Psychological Science*, *9*, 61–65.

Hirshman, E., & Lanning, K. (1999). Is there a special association between self judgments and conscious recollection? *Applied Cognitive Psychology*, *13*, 29–42.

Hirshman, E., Lanning, K., Master, S., & Henzler, A. (2002). Signal detection models as tools for interpreting judgments of recollections. *Applied Cognitive Psychology*, *16*(2), 151–156.

Hirshman, E., & Master, S. (1997). Modeling the conscious correlates on the remember–know paradigm. *Memory and Cognition*, *25*, 345–351.

Hockley, W. E., & Consoli, A. (1999). Familiarity and recollection in item and associative recognition. *Memory and Cognition*, *27*, 657–664.

Hockley, W. E., & Murdock, B. B., Jr. (1987). A decision model for accuracy and response latency in recognition memory. *Psychological Review*, *89*, 609–626.

Holmes, J. B., Waters, H. S., & Rajaram, S. (1998). The phenomenology of false memories: Episodic content and confidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1026–1040.

Inoue, C., & Bellezza, F. S. (1998). The detection model of recognition using know and remember judgments. *Memory and Cognition*, *26*, 299–308.

Jacoby, L., & Dallas, M. (1981). On the relationship of autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306–340.

Jacoby, L. L., Yonelinas, A. P., & Jenning, J. M. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to the question of consciousness* (pp. 13–47). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Joordens, S., & Hockley, W. E. (2000). Recollection and familiarity through the looking glass: When old does not mirror new. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1534–1555.

Kelley, C. M., & Jacoby, L. L. (2000). Recollection and familiarity process dissociation. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 215–228). New York: Oxford University Press.

Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: Two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 699–710.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*, 609–639.

Macken, W. J. & Hampson, P. (1993). Integration, elaboration, and recollective experience. *Irish Journal of Psychology*, *14*, 279–285.

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

Nacher, M. J., Gotor, A., & Algarabel, S. (1998). Traducciones equivalentes en catalan y castellano de 1533 palabras, y sus valores normativos en concrecion, familiaridad y significatividad [Equivalent translations in Catalan and Spanish of 1533 words together with their concretion, familiarity, and meaningful indices]. *Psicologica*, *19*, 1–26.

Parkin, A. J., & Walter, B. M. (1992). Recollective experience, normal aging and frontal dysfunction. *Psychology and Aging*, *7*, 290–298.

Perfect, T. J., Mayes, A., Downes, J. J., & van Eijk, R. (1996). Does context discriminate recollection from familiarity in recognition memory? *Quarterly Journal of Experimental Psychology*, *49A*, 797–813.

Postma, A. (1999). The influence of decision criteria upon remembering and knowing in recognition memory. *Acta Psychologica*, *103*, 65–76.

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

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

Rajaram, S. (1998). Conceptual and perceptual processes on remembering: The role of salience/distinctiveness. *Psychonomic Bulletin and Review*, *5*, 71–78.

Rajaram, S., & Geraci, L. (2000). Conceptual fluency selectively influences knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1070–1074.

Reder, L., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember–know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 294–320.

Reder, L. M., Angstadt, P., Cary, M., Erickson, M. A., & Ayers, M. S. (2002). A reexamination of stimulus-frequency effects in recognition: Two mirrors for low- and high-frequency pseudowords. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 138–152.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2001, November). *A two-dimensional signal detection model of remember–know judgments*. Paper presented at the 42nd annual meeting of the Psychonomic Society, Tampa, FL.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, *26*, 1–12.

Yonelinas, A. P. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General*, *130*, 361–379.

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language*, *34*, 622–643.