Journal of Memory and Language **45,** 82–102 (2001)
doi:10.1006/jmla.2000.2761, available online at http://www.academicpress.com on IDEAL®

Feature and Conjunction Errors in Recognition Memory: Evidence for Dual-Process Theory

Todd C. Jones

New York University and Victoria University of Wellington, Wellington, New Zealand

and

Larry L. Jacoby

Washington University in St. Louis

Feature and conjunction errors in recognition memory were investigated using a dual-process framework. In Experiment 1, dividing attention at study or test decreased old word recognition but did not influence feature and conjunction recognition errors after correcting for false alarms to new words (baseline). In Experiment 2, a response deadline manipulation influenced old word recognition but not feature and conjunction effects (i.e., feature or conjunction error rate minus baseline). Across Experiments 3 and 4, study repetitions increased the probabilities of feature and conjunction errors for participants under strong pressure to respond quickly. However, no such increases were observed for participants who were given more time to respond, providing evidence that the familiarity underlying feature and conjunction errors can be countered with recollection. Thus study repetition increased both familiarity and recollection. Feature and conjunction errors are based on familiarity in the absence of recollection. An approach that combines an item–associative distinction with a dual-process framework (e.g., Yonelinas, 1997) also can account for these errors. However, an approach that uses a feature-configuration distinction must be modified to account for these results.

Key Words: recognition memory; false recognition; feature errors; conjunction errors; recollection; familiarity; dual-process theory.

False memory phenomena have enjoyed a great surge of interest in the past decade (Roediger, 1996), and one line of this research has focused on false recognition of recombined study stimuli on a later test. For example, participants might study *buckwheat*, *blackmail*, and *jailbird*, and then on a later recognition test, judge whether *buckshot* (partly old) or *blackbird* (both parts old but recombined) occurred on the study list. False recognitions of these types of items have been referred to as feature errors (e.g., *buckshot*) and conjunction errors (e.g., *blackbird*; Reinitz, Lammers, & Cochran, 1992). The results from this type of false recognition experiment typically show that old responses are given to truly old stimuli at a higher rate than that for conjunction stimuli, followed by that for feature stimuli, followed by that for wholly new stimuli (old > conjunction > feature > new). Throughout this paper, we refer to a difference in false alarm rates for feature lures and new words as a feature effect and the difference in false alarm rates for conjunction lures and new words as a conjunction effect. (These effects reflect higher false alarm rates for feature and conjunction lures compared to a new word, which is

We thank Mark Reinitz for sending us the stimuli used in Reinitz, Verfaellie, & Milberg (1996). We thank Naomi Galvez for her help in data collection. Finally, we thank Patrick Dolan, David Harper, Michael Humphreys, Neal Kroll, Johanna Nordlie, Mark Reinitz, and Tom Trainham for their helpful comments on versions of this manuscript. The results of Experiments 1 and 2 were presented at the 38th Annual Meeting of the Psychonomic Society in November, 1997. Part of this research was supported by a grant from the National Science Foundation (SBR 9596209) to Larry L. Jacoby.

Address correspondence and reprint requests to Todd C. Jones, School of Psychology, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand. E-mail: todd.jones@vuw.ac.nz.

0749-596X/01 $35.00

considered to be a baseline.) Finally, feature and conjunction errors have been found to occur with a variety of materials (single words, where syllables are conjoined, Underwood & Zimmerman, 1973; word phrases, Underwood, Kapelak, & Malmi, 1976; compound words, Ghatala, Levin, Bell, Truman, & Lodico, 1978, Underwood et al., 1976, Reinitz et al., 1992; face drawings, Reinitz et al., 1992, Reinitz, Morrisey, & Demb, 1994; photographs, Searcy, Bartlett, & Meman, 1999; abstract figures, Kroll, Knight, Metcalfe, Wolf, & Tulving, 1996).

Recently, a dual-process theory of memory was used to account for feature and conjunction errors (Jones, Jacoby, & Gellis, in press). In dual-process theory, an automatic process, familiarity, and a controlled process, recollection, provide alternative bases for responding (e.g., Atkinson & Juola, 1974; Hay & Jacoby, 1996; Hintzman, Caulton, & Levitin, 1998; Hintzman & Curran, 1994; Jacoby, 1991, 1999; Jacoby, Yonelinas, & Jennings, 1997; Mandler, 1980, 1991; Yonelinas, 1994, 1997). Familiarity is characterized as relatively fast, whereas recollection is characterized as relatively slow (e.g., Atkinson & Juola, 1973, Hintzman & Curran, 1994; McElree, Dolan, & Jacoby, 1999).

Both familiarity and recollection contribute to performance on a recognition test. For old words, familiarity and recollection work in concert—either process can lead to a hit. New words are assessed on the basis of familiarity (e.g., on a strength continuum; Yonelinas, 1994, 1997; Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998). Recognition of feature and conjunction lures is based on both familiarity and recollection, but familiarity and recollection oppose each other. Familiarity pushes one to commit feature and conjunction errors, but recollection can be used to avoid the errors. Conjunction lures possess a higher familiarity strength than feature lures because the familiarity for both elements (instead of one element, as is the case for feature lures) is boosted by the study phase presentations.

For example, if one was to see the study primes *checklist* and *needlepoint* in a first phase but *checkpoint* on a recognition test, familiarity engendered by the components from the study primes (e.g., *check* from *checklist* and *point* from *needlepoint*) would push one to respond "old." If one depended solely on familiarity in making a recognition decision, then one would likely commit a conjunction error. However, one could avoid such an error by recollecting that *checklist* or *needlepoint* was presented in the study phase, not *checkpoint*. Thus feature and conjunction errors are based on familiarity in the absence of recollection.

Feature and conjunction lures are thought to be poor retrieval cues for the corresponding study primes, resulting in very low recollection (Jones et al., 2001). One reason feature and conjunction lures are poor retrieval cues is that there is a lack of transfer of conceptual information from study to test (e.g., Roediger, Weldon, & Challis, 1989). That is, *checkpoint* is conceptually different from *checklist* and *needlepoint*. Without a strong overlap in conceptual processes between study primes and test lures, retrieval of the study prime is a challenge. With familiarity relatively unopposed by recollection, feature and conjunction errors are difficult to avoid (see Kroll et al., 1996, and Reinitz et al., 1992, on the difficulty of avoiding the errors).

Another theory of feature and conjunction errors is that memory judgments are based on a familiarity assessment, and feature and conjunction lures possess greater familiarity than new words (e.g., Rubin, Van Petten, Glisky, & Newberg, 1999). Thus, feature and conjunction lures are identified "old" at higher probabilities than baseline. In cases where recollection of the study primes is nonexistent (i.e., recollection = 0), feature and conjunction error rates provide a good measure for familiarity. However, this approach does not allow the possibility that one could avoid an error by remembering that a different word was presented in the study phase. To avoid an error, a second process is required. Experiments 3 and 4 demonstrate this point.

Dissociations between old item recognition and feature and conjunction errors support the familiarity–recollection explanation. Old word recognition is based on familiarity and recollection working in concert, but feature and conjunction errors are based on familiarity in the absence of recollection. Therefore, a change in the hit rate without an alteration of feature and conjunction effects (feature and conjunction error rates after subtracting out the baseline false alarm rate) can be interpreted as an effect on recollection but not familiarity. (Such an interpretation assumes that, for the feature and conjunction conditions, effects of familiarity and recollection have not offset each other.)

Several manipulations have affected hit rates but not feature and conjunction effects. For example, dividing attention during a study phase reduced recognition for old faces relative to a full attention condition, but feature and conjunction effects were influenced little (Reinitz et al., 1994). Also, imposing a short response deadline during the test has decreased recognition of old words compared to a long deadline condition, but the response deadline manipulation did not affect feature and conjunction effects (Jones et al., 2001). Other variables that have influenced hit rates but not feature and conjunction effects include level of processing during study (e.g., shallow vs. deep; Reinitz, Verfaellie, & Milberg, 1996) and subject group (e.g., amnesics vs. normals, Reinitz et al., 1996, though see Kroll et al., 1996; older adults vs. young adults, Rubin et al., 1999, though see Kroll et al., 1996).

Only a few variables have been found to influence conjunction effects. One example is that individuals with hippocampal damage produced a higher conjunction error rate when the lag between study word primes was relatively short (e.g., 1- vs. 5-word lag; Kroll et al., 1996). In similar findings, normal participants produced a higher conjunction error rate for face drawing stimuli that were presented simultaneously compared to that for faces that were not presented simultaneously (Hannigan & Reinitz, 2000; Reinitz & Hannigan, in press). This latter finding may be based on an effect of attention instead of simultaneity, per se (Reinitz & Hannigan, in press).

We have used the familiarity–recollection account as a framework for our feature and conjunction experiments (Jones et al., 2001) and our approach accounts for most findings (for exceptions, see Kroll et al., 1996; Hannigan & Reinitz, 2000; Reinitz & Hannigan, in press). We continued the use of this framework in the experiments reported below. For Experiments 1 and 2, variables known to affect recollection but not familiarity were employed (e.g., dividing attention, response deadline). Critically, these two experiments, as well as Experiments 3 and 4, used retrieval manipulations. The expectation was that old–new discrimination would be influenced by these manipulations but that feature and conjunction effects would be affected little or not at all. This prediction was based, in part, on the notion that, under normal conditions (i.e., full attention), recollection for the study primes is low. Without much room to decrease recollection for the study primes, little change in feature and conjunction effects could be expected. In Experiments 3 and 4, we attempted to influence both recollection and familiarity. The idea was that if study items were made more memorable, then recollection could be used to avoid the errors. Thus, the goal in the final two experiments was to demonstrate how the two processes can work in opposition during recognition of feature and conjunction lures.

EXPERIMENT 1

Several studies have shown that dividing attention affects estimates of recollection but not familiarity (Jacoby, 1996, 1999; Jacoby, Toth, & Yonelinas, 1993; Wolters & Prinsen, 1997; see Mulligan & Hartman, 1996 and Schmitter-Edgecombe, 1999, for a further distinction). To date, a divided attention manipulation has been used in only one study on memory feature and conjunction errors. In that study, which used face drawings as stimuli, large divided attention effects were found for old conditions, but small or no effects were obtained on feature and conjunction rates after correcting for baseline differences (Reinitz et al., 1994). For the present experiment, we attempted to demonstrate that an encoding and a retrieval manipulation could produce similar patterns of data by dividing attention during *encoding* or *retrieval* (cf., Jacoby, Woloshyn, & Kelley, 1989). If dividing attention during retrieval produces results similar to those for dividing attention during study, then the problem cannot simply be a failure to encode the study items or to retain them, as suggested by Reinitz and colleagues (1996). That is, if dividing attention during the test decreased old–new discrimination but did not affect feature and conjunction effects, then that finding would reflect a problem of *accessing* study items instead of an encoding failure or a forgetting problem.

Because old word recognition is thought to be based, in part, on recollection, dividing attention at study or test was predicted to hinder recognition of old words. Also, if feature and conjunction lures are poor retrieval cues for their corresponding study primes, then recollection for the study primes should be quite low. Indeed, there may be a floor effect. In this case, dividing attention should have little influence on recollection for the study word primes, leaving the familiarity engendered by the lexical components relatively unopposed. Therefore, dividing attention was predicted to have little or no influence on feature and conjunction effects.

Method

Participants. Ninety-six New York University undergraduates participated in the experiment for credit toward an introductory psychology course. One participant was replaced for a failure to follow instructions.

Materials and equipment. Two hundred fifty-eight compound words were used. Two hundred forty words were used to form 80 critical sets of triplets, and the remaining 18 words were used as buffer study items or practice test items. Each set of triplets contained a target word (e.g., *checkpoint*), which ranged from 6 to 11 letters long, and two word primes that overlapped the target word with regard to its two lexical components (e.g., *checklist*, *needlepoint*). A few of these items were from Reinitz et al. (1996), while the others were constructed. The 80 triplets were assigned to one of four lists with 20 triplets each, and the lists were matched for mean word length of the targets. The experiment was run on a pentium computer with a VGA monitor using Micro Experimental Laboratory software (MEL; Schneider, 1990). An RCA radio stereo dual cassette recorder was used to play a tape of single-digit numbers in a female voice.

Design and procedure. The experiment employed a 3 (Group: full attention, divided attention at study, divided attention at test) × 4 (Item type: old, feature, conjunction, or new word) mixed design. Attention was manipulated between participants, but item type was manipulated within participants. Participants were randomly assigned to one of the three groups, resulting in 32 participants per group. For all three groups, single digit numbers were heard at a 1.5-s rate during the study and test phases. For the attention manipulations, one group tracked the single-digit numbers during the study phase and another group tracked the digits during the test phase, thus dividing attention at study or at test (e.g., Craik, 1982). A third group (full attention) heard the digits but was not required to attend to them. An experimenter monitored each participant's performance on the digit tracking task.

Three of the lists were used in the study phase, while all four lists were used in the test phase. The four lists were rotated through each of the different item types (old, conjunction, feature or new word; see Table 1) such that each item type occurred equally often across participants. A single study order based on item type was constructed such that each studied item type was evenly distributed throughout the study list. Items from three of the four lists (balanced across participants) were then drawn and placed into the appropriate slots of the item type study

TABLE 1

Example of Item Types at Study and Test

Item type	Study	Test
Old	checklist	checklist
Conjunction		
Order 1	blackmail jailbird	blackbird
Order 2	needlepoint checklist	checkpoint
Feature		
1st feature	buckwheat	buckshot
2nd feature	cellmate	playmate
New	—	lumberyard

Note. Study word primes (e.g., *blackmail*, *jailbird*, *buckwheat*) were never presented on the test. Order of the study primes for the conjunction item type was counterbalanced across participants. Presentation of the two study primes (e.g., *buckwheat* or *slapshot*) for each feature target word (e.g., *buckshot*) was counterbalanced across participants.

order. For the sake of completeness, both components of the feature words were presented equally often in the study phase (see Table 1). Also, for half of the conjunction words the first component of the target was displayed first, and for the other half the second component of the target word was displayed first (see Table 1). With counterbalancing, the total number of study lists was 16. Spacing within the study primes for the conjunction condition ranged from one to five intervening words with a mean of three words. A single test list for the 80 target words was constructed such that each item type was evenly distributed throughout the test with the constraint that no more than two trials of a given item type occurred consecutively.

In the study phase, all participants were told to read the words silently in preparation for a later memory test. The participants were not told about the various types of items that they would encounter on the test. The divided attention at study group performed the digit tracking and reading–studying of the words at the same time during the study phase, and the importance of doing well on the digit-tracking test was emphasized. For the digit-tracking task, participants were told to say "hit" after each sequence of three odd numbers. If a sequence was missed the experimenter quietly said "miss" to stress the importance of doing well on the digit-tracking task. The divided attention at test and full attention groups were told to concentrate on the reading–studying of the words and to ignore the digits. For all groups, 22 digits were heard before the word presentations began. This procedure was used to make sure that the divided attention at study participants understood the digit-tracking instructions. If a participant misunderstood the instructions, then the instructions were clarified, and the tape was played again from the beginning. Eighty study words appeared in white letters on a black background in the center of the computer screen for 3 s with an intertrial interval of 1 s, and the critical study trials were bounded by five primacy and five recency compound words (buffers).

Immediately following the study phase, participants were given test instructions, followed by a practice recognition test, and finally, a yes–no recognition test. There were no feature or conjunction lures on the practice test. All groups of participants were informed that some of the test words would be old, while others would be new. However, they were not informed of the feature and conjunction lures. The participants were directed to respond "yes" to old words, to respond "no" to new words, and to guess "yes" if they were unsure whether a word was old by pressing keys labeled *Y* or *N*. (A liberal criterion was encouraged in all four experiments and was used to prevent floor effects for the three false alarm conditions—conjunction, feature, and new.) Items in the test phase were presented for up to 3 s with a 1-s intertrial interval. The program was designed such that when a participant responded before the 3-s time limit the screen went blank for the remainder of the 3 s. Thus, the timing of the study and test trials was equated across all groups. Participants in the divided attention at test group performed the digit tracking task and the recognition test at the same time, and as with the divided attention at study group, the importance of the digit tracking task was emphasized. Both the divided attention at study and the full attention groups were told to ignore the numbers and to concentrate on the recognition test. All participants were told to make their recognition judgments as accurately and as quickly as possible. If a participant failed to respond within 3 s, the computer beeped and the next trial began.

Results and Discussion

First, failures to respond in the allotted time occurred seldom for each group (time-out rates: .00–.03). To calculate the mean proportions of old responses, the number of old responses for each condition was taken as a proportion of the total number of trials within a condition. (This approach was used for all four experiments.) For each of the three groups, the mean rate of responding "old" was highest for the old words, followed (in order) by conjunction lures, feature lures, and new words. Each of these old response rates was significantly different from the others. This same pattern of results was expected and was obtained in Experiments 2–4. Thus, we do not belabor this point in the other experiments. Also, for the sake of space and readability we report only the statistics that were of particular relevance to our predictions.

Statistical comparisons for this experiment, as well as the following three experiments, failed to produce any significant differences between various counterbalancing conditions (i.e., for feature errors, presentation of the first vs. second lexical component in the study primes; for conjunction errors, presentation order of the study word primes). A one-way ANOVA with a follow-up Tukey test showed that a relatively high false alarm rate in the divided attention-study group (.44) was significantly greater than those for the full attention (.25) and divided attention-test groups [.28; $F(2,93) = 12.04$, $MSE = .04$]. (For all analyses, $\alpha = .05$. All post-hoc tests were Tukey tests, unless stated otherwise. All *t*-tests were two-tailed.) The higher base rate for the divided attention-study group relative to the full attention group was similar to findings reported by Reinitz et al. (1994). To accommodate for the different base rates, the means for each of the other item type conditions were corrected for baseline by subtracting out the baseline rate. Also, *d*' scores were calculated, and the *d*' analyses were parallel to the results of the corrected recognition analyses. In the interest of brevity, we report the analyses on the corrected recognition data.

Important for our argument was the performance on the various item types across the three groups. The hypothesis was that dividing attention would affect recollection, evidenced by a decrease in the ability to discriminate old from new items, but not familiarity, evidenced by no change in discrimination of feature and conjunction words from new words. In short, the hypothesis was supported. A 3 (Group) × 3 (Item type: old, feature, or conjunction word) mixed ANOVA was conducted on the means corrected for baseline differences, and the analysis yielded significant effects of Group [$F(2,93) = 5.55$, $MSE = .04$] and Item type [$F(2, 186) = 112.39$, $MSE = .02$], as well as a significant interaction [$F(4,186) = 7.81$, $MSE = .02$].

Importantly, the interaction was significant, and this interaction was analyzed further with two ANOVAs. First, a one-way ANOVA was conducted on the corrected hit rates, which demonstrated that there was a difference in old–new discrimination between the three groups [$F(2,93) = 12.11$, $MSE = .04$]. A post-hoc test revealed that old–new discrimination was significantly greater for the full attention group (.48) than for the divided attention at study and divided attention at test groups (.25 and .36, respectively), with an additional difference between the two divided attention groups. Evidence for a difference in feature and conjunction effects across the three groups would be a significant effect of Group or a Group × Item type interaction. A 2 (Item type: feature, conjunction) × 3 (Group) mixed ANOVA was conducted on the corrected feature and conjunction error rates, but neither the Group effect nor the Group × Item type interaction was significant (Fs < 1.77). Thus, as predicted, only old–new discrimination was affected by dividing attention.

Although performance on the old words was better for the divided attention-test group than for the divided attention-study group, this difference is not critical to our argument. Nonetheless, these data are consistent with previous findings (e.g., Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). In part, the difference in corrected recognition might reflect a difference in performance on the digit-tracking task. The identification rate for the numbers sequences was slightly higher for the divided attention at

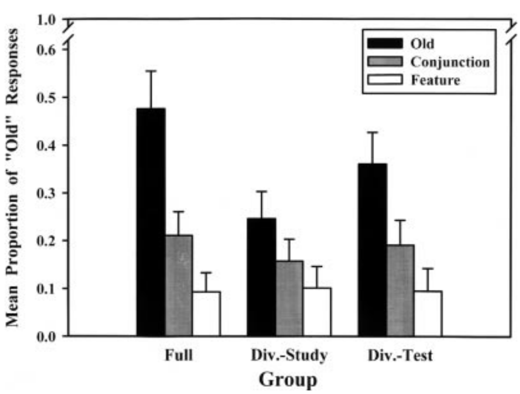

FIG. 1. Experiment 1: Mean proportion of old responses (corrected for baseline) as a function of item type by group. Error bars represent 95% confidence intervals.

study group (.90) than for the divided attention at test group (.84), though this difference was not significant [$t(62) = 1.87$, $SE = .03$, $p = .07$].

For the sake of clarity, we want to emphasize the difference between recollection for old words and the study word primes of the feature and conjunction words. Recollection was relatively easy for old words because the whole word was provided as a cue on the recognition test. Thus, the concept itself was re-presented. In contrast, recollection was relatively difficult for the study word primes because the test cues contained only part of a study word and that part was in a different context as part of a different concept. Thus, the feature and conjunction trials provided a sort of cued recall test. If a participant recalled a study prime word, then the participant would avoid a feature or conjunction error. Note that if recollection of the study primes (e.g., *blackmail*, *jailbird*) were aided substantially by the feature and conjunction lures (*blackbird*) or their lexical components (*black* or *bird*), then the detrimental influence of familiarity would have been overcome, producing lower feature and conjunction error rates in the full attention condition compared to the divided attention conditions. Clearly, this pattern of results did not occur.

Our claim is that dividing attention affected recollection, and the evidence supporting this claim comes from the differences in corrected hit rates. If recollection were a major component of feature and conjunction errors—that is, error reduction—we also would expect dividing attention to influence recollection for the study primes. However, if recollection for the study primes was near zero in the full attention condition, as we have suggested, no direct evidence for the role of recollection on feature and conjunction error production could be obtained. Thus, the evidence for the role of recollection or low level of recollection on conjunction errors is based on indirect evidence from the old word condition.

The results of the experiment are handled well by the familiarity–recollection distinction, and the data for the divided attention at test group particularly, add to the work by Reinitz and colleagues (Reinitz et al., 1992, 1994, 1996). Dividing attention at encoding or retrieval was shown to decrease recollection, evidenced by an attenuation in old–new discrimination, but to leave familiarity invariant, evidenced by equivalent feature and conjunction effects for the three groups. For the full attention group the study words were available and largely accessible. Presumably, for the divided attention at test group, the study words were just as available but not as readily accessible. Thus, feature and conjunction errors may occur due to an inability to access the specific study words through conscious recollection.

EXPERIMENTS 2A AND 2B

In Experiment 1, a retrieval manipulation influenced old–new discrimination but not feature or conjunction effects. This finding for a retrieval manipulation adds to a finding where a response deadline was manipulated (Jones et al., 2001). In that response deadline experiment, decreasing the amount of time to respond to recognition test probes decreased old–new discrimination but not feature and conjunction effects. Experiments 2A and 2B reproduced the prior results with a response deadline manipulation by using a slightly different procedure.

As already noted, information accrues more slowly from recollection than from familiarity (Atkinson & Juola, 1973; Dosher, 1984; Jacoby, 1999; Jacoby, Jones, & Dolan, 1998; McElree et al., 1999). The present experiments used a response deadline manipulation that was similar to one used by Jacoby and colleagues (Jacoby, 1999; Jacoby et al., 1998; see also Jones et al., 2001). Participants were pushed to respond quickly (within 850 ms; Experiment 2A) or wait briefly (1400 ms) before responding within a response window (750 ms; Experiment 2B).[1] Because decreasing the amount of time to make the recognition decision decreases the ability to use recollection, the probability of a hit was expected to drop. In contrast, decreasing the amount of time to respond should not influence the ability

[1] A 750-ms short deadline group ($N = 16$) demonstrated a boundary condition. This shorter deadline resulted in a failure to discriminate feature lures from new items (produce feature errors above baseline).

to use familiarity; therefore, a response deadline manipulation was expected to have little or no influence on feature and conjunction effects.

The procedure differed from that for a similar experiment of ours (Jones et al., 2001) on feature and conjunction errors. In our previous experiment, response deadline was a within-participant variable. In the present experiments, response deadline was a between-participants variable. Also, in the prior experiment, participants entered a response at the appearance of a response signal for both short and long deadline conditions. For the present experiments, no response signal was given for the short deadline conditions. The test probe appeared, and participants attempted to enter a response before the deadline.

Method

Participants. Two groups of 16 New York University undergraduates participated in Experiments 2A and 2B (32 total students) for credit toward an introductory psychology course.

Design and procedure. The experiments used a two-variable, mixed design. Item type (old, feature, conjunction, or new word) was manipulated within participants and response deadline (short or long) was manipulated between participants. The materials and the lists were the same as those in Experiment 1. Also, the general study phase and general test instructions were the same as those used in the first experiment (without the portions regarding divided attention). The intertrial interval at study was 500 ms. On each test trial, a row of plus symbols (++++++) was displayed for 500 ms as a focal point; then the test word was presented. In the short deadline group (Experiment 2A), participants were told to respond to each word before the 850-ms deadline. For the long deadline group (Experiment 2B), participants were instructed to wait for a response signal (******) to appear below the test word and then respond within 750 ms after the response signal appeared. The delay (wait period) between the onset of the test word and the response signal was 1400 ms, and responses were not allowed during this delay. Thus the deadline for the long deadline group was 2150 ms. The wait component of the deadline procedure was used to encourage the use of recollection and to avoid nonoptimal responding (e.g., responding based relatively more on familiarity). When participants responded before the time limit, the program continued on to the next trial. If a response was not made before the deadline (short deadline) or during the response window (long deadline), the computer beeped, and the program continued to the next test item. For the test, the intertrial interval was 500 ms.

Results and Discussion

The data are presented in Table 2 (corrected for baseline). The time-out rates for the two groups were similar (short deadline, .03; long deadline, .02) Although the likelihood of responding “old” to conjunction, feature, and new items was slightly higher for the short deadline group, the main difference between the two groups was that the hit rate was lower for the short deadline than for the long deadline group. The participants were not randomly assigned to the 850-ms deadline and long deadline groups, but they were drawn from the same subject pool and at the same time of the semester. We report statistics comparing the two groups, though conclusions from these analyses should be considered cautiously and with the lack of random assignment in mind. (However, these results were replicated in both Experiments 3 and 4 with random assignment of participants.) As in Experiment 1, analyses on *d*’ scores produced

TABLE 2

Mean Corrected Recognition Rates for Each Group by Item Type

	Deadline group	
Item type	Long	Short
Old	.55 (.20)	.34 (.21)
Conjunction	.23 (.18)	.23 (.17)
Feature	.12 (.15)	.14 (.11)

Note. The mean proportions of old responses appear with the baselines subtracted. The corresponding baseline false alarm rates were .23 and .28 for the long and short deadline conditions, respectively. Standard deviations appear in parentheses.

results similar to the analyses carried out on the corrected recognition rates and, hence, are not reported.

The false alarm rate for new words did not differ between the two groups [$F(1,30) = .83$, $MSE = .02$]. However, to be consistent with the statistical approach in Experiment 1, we analyzed the data corrected for baseline performance. A 2 (Group) × 3 (Item type) mixed ANOVA produced a significant effect of Item type and a significant Group × Item type interaction [$F(2,60) = 54.56$, $MSE = .02$, and $F(2,60) = 8.95$, $MSE = .02$, respectively]. The interaction was followed up with a one-way ANOVA on the corrected hit rates and a 2 (Group) × 2 (Item type) mixed ANOVA on the corrected feature and conjunction error rates. The advantage in old–new discrimination for the long deadline group was significant [$F(1,30) = 8.72$, $MSE = .04$]. For the mixed ANOVA, neither the Group nor the Group × Item type interaction was significant. Thus, the corrected feature and conjunction error rates did not differ between the two groups.

These results are a reproduction of the findings by Jones and colleagues (2001) and provide another example of the importance of retrieval factors for theory on recognition memory. Without sufficient time to use recollection, old–new discrimination suffered but feature and conjunction effects did not. In other words, the response deadline decreased recollection but left familiarity invariant. With regard to the feature–configuration viewpoint, presumably configurative information was available but not accessible under the time pressure.

EXPERIMENT 3

The first two experiments demonstrated that two factors are needed to account for the effects of retrieval manipulations. Evidence for the role of recollection, however, was indirect in both experiments. Anecdotal evidence suggested that, although test instructions did not encourage recollection through cued recall, many participants were aware of the discrepancy between the study word primes and the feature and conjunction lures, and many participants stated that they tried to avoid falling for the feature and conjunction lures. Experiment 3 was designed to obtain direct evidence that recollection could be used to avoid feature and conjunction errors. Such a finding would demonstrate the importance of using two processes to account for conjunction errors. Two manipulations were employed in Experiment 3: number of study presentations and response deadline.

Research conducted under the familiarity–recollection framework has shown study repetition to increase both familiarity and recollection (Jacoby, 1999; Jacoby et al., 1998; McElree et al., 1999), and there is some evidence that repetition can boost conjunction errors (Bartlett & Searcy, 1998; Underwood et al., 1976; Underwood & Zimmerman, 1973). We describe an experiment from Jacoby and colleagues (Jacoby et al., 1998; also see Jacoby, 1999) because its results motivated the following two experiments.

In Experiment 3 of Jacoby and colleagues' (1998) paper, there were three phases. First, a list of study words was presented visually. Some of the words were presented once, and some of the words were repeated. In the second phase, a list of study words was presented aurally, and each word was presented once. (No words occurred in both the visual and the auditory phases.) In the third phase, words that were read or heard in the first or second phase were presented in a test list that included new words. The task for participants was to identify words as old only if the word was heard in the second phase of the experiment. Participants were instructed to exclude the visually presented words (first phase) from old responses. Thus, on the exclusion test, recollection and familiarity were placed in opposition for the words that had been presented on the visual study list. Familiarity for those words would push participants to respond "old," but recollection of those words as visually studied words would allow participants to correctly exclude them as having not been heard.

Two groups of undergraduate participants in Jacoby and colleagues' (1998) experiment received different response deadline procedures on the test. One group of participants was pushed to respond under a short deadline, while the other group waited briefly before responding in a response window (long deadline). The long deadline group showed a *decrease* in exclusion errors for repeated relative to nonrepeated words. In contrast, the short deadline group showed an *increase* in exclusion errors for repeated compared to nonrepeated words. The interpretation of those results was that study repetition increased recollection, evidenced by the ability of the long deadline group to miscategorize fewer repeated than nonrepeated words as heard. In the absence of recollection, study repetition was found to increase familiarity, evidenced by an increase in exclusion errors for repeated compared to nonrepeated words for the short deadline group.

We applied an approach similar to one used by Jacoby and colleagues (Jacoby et al., 1998) in our final two experiments. In the present false recognition experiments, recollection and familiarity oppose one another for judgments on feature and conjunction words. Familiarity pushes one to commit feature or conjunction errors, whereas recollection protects one from those errors; thus, for feature and conjunction lures, familiarity and recollection are placed in opposition. We manipulated the number of study repetitions to increase recollection and familiarity. In addition, we used a variable, response deadline, that has been shown to influence recollection but not familiarity. There were two differences between the present approach and the prior approach by Jacoby and colleagues (1998). First, in the present study, there was only one study phase. Second, because there was only one study list, the test did not involve list discrimination.

There were two groups of participants: a group with a short deadline procedure and a group with a wait period–response window test procedure (long deadline). The expectation was that the short deadline group would exhibit lower recollection, characterized by lower old–new discrimination, than the long deadline group. Two effects of study repetition were expected: an increase in familiarity and an increase in recollection. Thus, a repetition effect was expected on the hit rates for both groups. For the feature and conjunction conditions, an increase in familiarity of stimulus components should push one to commit more feature and conjunction errors. However, a parallel increase in recollection would successfully oppose the increase in familiarity. By limiting the role of recollection with a short deadline, we created the opportunity to observe an increase in familiarity. A relatively low level of recollection in the short deadline group would decrease the likelihood that recollection could be used to counter an increase in familiarity. Therefore, for the short deadline group, feature and conjunction errors were predicted to increase with study repetitions. On the other hand, the long deadline group was expected to use an increase in recollection to counter an increase in familiarity, resulting in no increase, or perhaps a decrease, in feature and conjunction error rates. To encourage the use of recollection, all participants were informed of the feature and conjunction test lures in the test instructions.

Method

Participants. Sixty-four New York University undergraduates participated for credit toward a course in introductory psychology.

Design and procedure. The design was a 2 (Group: deadline, wait) × 7 (Item type: old word-1 presentation, old word-4 presentations, feature condition-1 presentation, feature condition-4 presentations, conjunction condition-1 presentation, conjunction condition-4 presentations, new) mixed design. The students were assigned randomly to one of the two groups (short or long deadline).

The 80 triplets that were used in Experiments 1 and 2 were divided into eight lists of ten triplets. Six lists corresponded to study items: old word-1 presentation, old word-4 presentations, feature condition-1 presentation, feature condition-4 presentations, conjunction condition-1 presentation, and conjunction condition-4 presentations. The other two lists were used as new items on the recognition test. The study list contained 200 critical trials with six primacy and six recency buffer trials. The order of the 200 critical study trials was the same for each participant with respect to item type. Items from the six study lists were drawn and placed into the predetermined item type slots, and each of the eight lists served as each item type equally often across participants. The study order was composed of five blocks of 40 words, such that the presentation of two items from each item type occurred in a block. Thus the distribution of item types throughout the study list was equivalent. For repeated items, the spacing ranged from 7 to 12 intervening items, with mean spacings of 9.3, 9.7, and 9.4 words for the old words, feature primes, and conjunction primes, respectively. The spacing within each pair of study primes for the single presentation conjunction condition was two words. There were two types of spacings for the conjunction primes in the repeated condition (see Table 3). First there was the number of intervening words within a given presentation of the pair (e.g., first presentation of each study prime). This spacing ranged from one to three words with a mean of two words. Of course, because there were multiple presentations of these study primes, there were intervening words between presentations of each pair (e.g., the first presentation of the second word of a pair and the second presentation of the first word of the pair). The mean spacing between presentations of the pairs was 6.3 words (range: 4–10 words). As in the previous two experiments, the design was counterbalanced for the presentation of the first or second component of the feature words and the order of conjunction word components across participants. The study words were presented for 2 s with a 250-ms interstimulus interval, and each study word was presented one or four times. Participants read each study word aloud as it was presented.

The deadline procedures were similar to those for Experiments 2A and 2B. The response deadline for the short deadline group was set at 1000 ms. The 150-ms increase from 850 ms was used to limit the number of time-outs in the short deadline group. For the long deadline group, the wait period was 1500 ms, and the response window was set to 1000 ms. Thus, the deadline for the long deadline group was 2500 ms. For the long deadline group, the increases in the wait period (100 ms) and the response window (250 ms) were expected to give slightly more time to utilize recollection. Unlike Experiments 1 and 2, participants were informed of the feature and conjunction lures in the test instructions. The detailed test instructions described how, if an element of a test word could be remembered as an element of a different word in the study phase, then the test word was certain to be new. An example with two study primes and a conjunction lure was provided.

Results and Discussion

The mean proportions of old responses (corrected for baseline) are shown as a function of item type in Table 4. As in the prior experiments, the studied item types were identified as old at higher rates than that for new words. The two groups differed slightly in their false alarm rate to new words (long deadline group, .13; short deadline group, .16), but this difference was not significant [$F(1,62) = 1.29$, $MSE = .02$]. There were very few time-outs (mean rate = .01 for both groups).

The results reproduced those of Experiments 2A and 2B: a short response deadline reduced the probability of correct responses to old words. Overall, repetition increased the likeli-

TABLE 3

Example Spacing of Repeated Conjunction Study Primes

	1st Word, 1st Pres. blackmail		2nd Word, 1st Pres. jailbird		1st Word, 2nd Pres. blackmail		2nd Word, 2nd Pres. jailbird
List position	1		5		13		15
Number of intervening words		3		7		1	

Note. The 3 and 1 on the lower line represent spacings within a presentation of a study prime pair, while the 7 is a spacing between presentations of that pair. The mean spacing between repetitions (e.g., *blackmail*, *blackmail*) was 9.4 words.

TABLE 4

Mean Corrected Recognition Rates for Each Group by Item Type and Number of Study Presentations in Experiment 3

	Deadline Group			
	Long		Short	
Item type	1P	4P	1P	4P
Old	.53 (.19)	.67 (.16)	.33 (.24)	.51 (.21)
Conjunction	.18 (.13)	.22 (.19)	.16 (.19)	.27 (.15)
Feature	.08 (.14)	.10 (.16)	.07 (.13)	.17 (.16)

Note. P, study presentation(s). The mean proportions of old responses appear with the baselines subtracted. The baseline error rates were similar for the two groups (long deadline, .13; short deadline, .16). Standard deviations appear in parentheses.

hood of responding "old." The only difference between the two groups was that the repetition effect on feature and conjunction errors was smaller for the long than for the short deadline group. A 2 (Group) × 2 (Repetition) × 3 (Item type) mixed ANOVA on the corrected recognition scores yielded main effects of Item type [$F(2,124) = 221.89$, $MSE = .03$] and Repetition [$F(1,62) = 38.57$, $MSE = .02$], and the interaction of those factors was significant [$F(2,124) = 6.55$, $MSE = .01$]. The Group × Item type interaction also was significant [$F(2,124) = 16.27$, $MSE = .03$]. Both groups were expected to show a repetition effect for old words, though overall performance on old words was predicted to be higher for the long deadline group than for the deadline group. A 2 (Group) × 2 (Repetition) mixed ANOVA on the corrected old word scores confirmed this expectation [Group, $F(1,62) = 15.65$, $MSE = .07$; Repetition, $F(1,62) = 49.06$, $MSE = .02$; the interaction was not significant].

Specific predictions concerning repetition effects in the feature and conjunction conditions were made for the two groups. Study repetition was expected to increase feature and conjunction effects for the short deadline but not the long deadline group. Although the data appear to support this prediction, a 2 (Group) × 2 (Item type) × 2 (Repetition) mixed ANOVA did not produce a significant Group × Repetition interaction [$F(1,62) = 3.24$, $MSE = .02$, $p < .08$; the repetition main effect was significant, $F(1,62) = 12.65$, $MSE = .02$]. On the other hand, the outcome of planned comparisons between the repetition conditions for the separate groups did support our predictions. A 2 (Item type) × 2 (Repetition) repeated measures ANOVA was conducted for each group, and the repetition effect was significant for the short deadline group [$F(1,31) = 15.38$, $MSE = .02$] but not the long deadline group [$F(1,31) = 1.45$, $MSE = .02$]. Thus, while the long deadline group appeared to have gained some control over false recognition of the feature and conjunction lures, this control was not particularly strong. If the ability to control these errors had been strong, then feature and conjunction errors would have decreased as the number of study presentations increased, statistically producing a significant Group × Repetition interaction.

Recollection and familiarity were increased by the repetition of study words. The increase in familiarity resulted in more feature and conjunction errors for the short deadline group. In fact, the finding that feature errors increased with study repetition provides additional evidence that these errors need to be accounted for in theory. The increase in recollection was demonstrated in the long deadline group by showing only a very slight increase in feature and conjunction errors with study repetition. The lack of a repetition effect for feature and conjunction conditions for the long deadline group does not appear to be due to a distrust of familiarity. If participants in the long deadline group were unwilling to respond on the basis of familiarity, despite the strong encouragement in the instructions, then a general distrust of familiarity would have produced lower feature, conjunction, and new word error rates for the long deadline group, and significant repetition effects would have been present for the feature and conjunction conditions.

EXPERIMENT 4

The use of recollection to avoid feature and conjunction errors in Experiment 3 was not particularly strong. Therefore, we attempted to design an experiment with study repetitions where recollection might be used more successfully.

The logic was the same as for Experiment 3, but there were a number of changes in the procedure. The first change was a decrease in the length of the study list. The goal was to increase recollection by using shorter lists. The overall number of critical items was maintained by using four study–test sequences for each participant. Also, to decrease the length of the study lists, words in the repetition condition were presented three times rather than four. There were 52 trials in each study set and 24 trials on each test. On each test, there were five item types: old word-1 presentation, old word-3 presentations, conjunction lure with components presented once, conjunction lures with components presented three times, and new word lures. Feature lures were not used.

A major concern was that with the shorter list length, less time might be needed for recollection. We also wanted to keep time-out rates low. Therefore, we adopted a response-signal procedure for both the short deadline and the long deadline groups. For the deadline group, a response signal appeared 400 ms after the onset of the test probe, and the signal was present for 500 ms. For the long deadline group, the signal was present for the same amount of time but did not appear until after the test probe was present for 2 s. The increase in the wait period was employed to increase the likelihood that recollection would be used to avoid the conjunction errors. If a participant did not respond while the response signal was present, a warning signal appeared for 400 ms, and participants were instructed to respond immediately if the warning signal appeared. The study and test instructions and total time per trial were identical for the two groups throughout the experiment. The only difference between the two groups was the time before the appearance of the response signal (400 vs. 2000 ms). Finally, before the first study block participants were told that conjunction lures would appear on the recognition tests, and an example was given.

Method

Participants. The participants were 40 undergraduates at Victoria University of Wellington. The students were tested in a large room with several individual computer stations. Each student was assigned to a computer station, and the number of students tested in a session ranged from 7 to 14. Each participant received \$5 as a gift.

Materials. As in the previous three experiments, 80 sets of triplets were used. Some of the triplets were modified or replaced with words that were known to New Zealanders.

Design and procedure. The experiment used a 2 (Response deadline: short or long) × 5 (Item type: old word-1 presentation, old word-3 presentations, conjunction condition-1 presentation, conjunction condition-3 presentations, new word) design. Response deadline was manipulated as a between-participants variable, but item type was manipulated as a within-participant variable. Twenty participants were assigned randomly to each group.

The 80 triplets were divided into five lists of 16 triplets. Four different study and test orders were made for each study–test block. The orders were based on item type. There were 56 study trials in each study block, but the first 4 and final 4 trials were buffer trials. Study words were presented for 2 s with a 250-ms interstimulus interval. The mean number of words intervening repeated study words was 6.1 and 6.2 for the old words and conjunction primes, respectively (range: 3–12 words). The order of the study word primes for the conjunction conditions was not counterbalanced. Prior experiments in this paper and others (e.g., Jones et al., 2001) have not found any difference in conjunction error rates based on the order of the study word primes; thus, this decision was viewed as trivial. The spacing within a pair of conjunction word primes was 2.75 words (range: 1–6 words), and the spacing between presentations of the pairs was 2.53 words (range: 1–6 words). Both the short and the long deadline groups went through a practice session, which included a practice recognition test on words (but not compound words), in order to get used to responding at the appearance of the response signal.

For the experiment proper, the participants were instructed to read each word silently and to try to remember it for the memory test. Before the first study–test block the conjunction condition was explained, and an example conjunction lure was given. Prior to each test the participants were reminded that if they remembered part of a test word as having appeared in a different word in the study set, then they could be certain the test word was new. However, if they were unsure whether a word was old or new, they were to guess "old." There were 24 trials on each test, 4 for each of the critical item types and 4 fillers (all old, from the buffer trials in the study block). Each of the five lists of 16 triplets was assigned to each item type equally often across participants, and each of the items occurred in each study block (as each item type) equally often across participants. Twenty participants were used in each group to counterbalance the experiment.

On the tests, a focal point (+++++) was presented for 500 ms followed by a blank screen for 250 ms; then a test word appeared. For the deadline group, 400 ms after the onset of a test word a response signal (a series of asterisks, *****) appeared below the word. The response signal was present for 500 ms. If a participant did not respond while the response signal was present, a late signal (!!!!!) appeared for 400 ms where the response signal had been located. Participants were told that if they saw the late signal, then they had waited too long to respond and that they should enter a response immediately. If a response was not entered while the late signal was present, the trial timed out. The same procedure was used for the long deadline group except that the onset of the response signal did not occur until a test word was present for 2000 ms. The rate of the test trials was held constant across the two groups (3.5-s rate). For both groups, after a participant responded the computer screen was blank for the remainder of the 3.5-s period. If a response was not entered the trial timed out, and the screen was blank for the remainder of the 3.5-s period.

Results and Discussion

Each of the studied conditions (old and conjunction) was identified "old" at a higher probability than baseline, and the means for each study condition (corrected for baseline) appear in Table 4 as a function of group. (The time-out rates were low: short deadline group, .01; long deadline group, .003). Despite the changes in materials, participants, and procedure, the patterns of the data resembled the patterns from Experiment 3. The long deadline group had an overall higher hit rate than the short deadline group, and both groups showed a repetition effect for old words. The two groups showed about the same conjunction effect for a single presentation. However, the short deadline group demonstrated a repetition effect for the conjunction condition, whereas the long deadline group did not. The false alarm rate was different for the two groups [$F = 6.77$, $MSE = .01$].

A 2 (Response Group) × 2 (Item type) × 2 (Repetition) mixed ANOVA on the corrected hit and conjunction error rates gave significant main effects of Group, Item type, and Repetition [$F(1,38) = 4.26$, $MSE = .07$; $F(1,38) = 226.06$, $MSE = .03$; and $F(1,38) = 40.76$, $MSE = .02$, respectively]. The three-way interaction of Group, Item type, and Repetition was not significant [$F(1,38) = 1.37$, $MSE = .01$] but was followed up in a manner consistent with Experiment 3. Separate 2 (Response group) × 2 (Repetition) ANOVAs were run on the old word and conjunction error data. The ANOVA for old word recognition produced significant effects of Group and Repetition [$F(1,38) = 24.94$, $MSE = .06$ and $F(1,38) = 49.60$, $MSE = .01$, respectively], but the interaction was not significant ($F = 2.18$). Thus, the response deadline decreased the probability of correctly recognizing

TABLE 5

Mean Corrected Recognition Rates for Each Group by Item Type and Number of Study Presentations in Experiment 4

	Deadline Group			
	Long		Short	
Item type	1P	3P	1P	3P
Old	.73 (.17)	.88 (.11)	.44 (.20)	.66 (.23)
Conjunction	.19 (.15)	.22 (.12)	.22 (.19)	.38 (.24)

Note. P, Study presentation(s). The mean proportions of old responses appear with baseline error rates subtracted. The baseline error rates were slightly different for the two groups (long deadline, .04; short deadline, .12). Standard deviations appear in parentheses.

old words. On the other hand, study repetition increased the likelihood of correctly recognizing old words. For the corrected conjunction error data, the effect of Repetition was significant [$F(1,38) = 12.82$, $MSE = .02$], but this effect was qualified by a significant interaction [$F(1,38) = 7.07$, $MSE = .02$]. The repetition effect was significant for the short deadline group [$F(1,19) = 12.42$, $MSE = .03$] but not for the long deadline group [$F(1,19) = .98$, $MSE = .01$].

Similar to Experiment 3, study repetition increased both familiarity and recollection. In fact, the increase in familiarity with study repetition was a bit more dramatic in the present experiment. This increase was demonstrated by the repetition effect on conjunction errors for the short deadline group. However, recollection also increased, allowing the long deadline group to counter the increase in familiarity with the increase in recollection. The result was that repetition effects on familiarity and recollection offset each other for the long deadline group.

GENERAL DISCUSSION

In Experiment 1, dividing attention at study or test decreased old–new discrimination but not feature and conjunction effects. In Experiments 2–4, decreasing the amount of time to respond to the recognition probes decreased old–new discrimination but not feature and conjunction effects for single study word presentations. Study repetition increased the hit rate and feature and conjunction effects in a short deadline condition (Experiments 3 and 4). However, a long deadline group provided evidence that feature and conjunction errors could be avoided through recollection (Experiments 3 and 4). For those groups, no repetition effect was observed on feature and conjunction error rates. From a standpoint of controlling feature and conjunction errors, a decrease in conjunction errors with study repetitions would be desirable. However, the lack of a repetition effect on conjunction error rates for the long deadline groups in Experiments 3 and 4 is encouraging and demonstrates that these errors can be avoided to some degree. The results for the long deadline groups also show how offsetting effects of familiarity and recollection can occur, resulting in what appears to be a null effect. Without the data from the short deadline groups, one might incorrectly conclude that study repetition does not increase familiarity or recollection (see also Jacoby et al., 1998).

A two-factor theory (e.g., dual-process theory) is needed to account for these data, and the theory needs to take retrieval dynamics into consideration. In the dual-process framework, familiarity and recollection are alternative bases for responding and are characterized by different retrieval rates (familiarity is relatively fast whereas recollection is relatively slow). Feature and conjunction errors are based on familiarity, but recollection can be used to avoid the errors. Therefore, feature and conjunction errors are based on the influence of familiarity in the absence of recollection. In contrast, a single process model based on a continuum of familiarity strength (e.g., Rubin et al., 1999) cannot account for our data. For example, the simple model cannot account for the dissociations between the corrected recognition scores for old words and feature–conjunction lures (a more complicated model with multiple decision criteria would accommodate these results). Experiments 3 and 4 show the importance of a second process (e.g., recollection). The familiarity account cannot explain the lack of a significant repetition effect for the long deadline groups without an additional process. (This finding would present a challenge for a two-threshold model, as well.)

Reinitz and colleagues (e.g., Reinitz & Hannigan, in press; Reinitz et al., 1992) have argued that intentional retrieval (what we refer to as recollection) can produce conjunction errors. A prediction in line with this reasoning is that there should be a higher conjunction error rate for long deadline conditions compared to short deadline conditions because more time should provide a better opportunity to use recollection. A second prediction is that study repetition should increase the likelihood of producing a conjunction error, even for a long deadline condition. The results from our experiments show patterns that are directly opposite to these predictions. Our findings indicate that recollection is not used to produce conjunction errors. Instead, recollection is used to prevent conjunction errors.

In the present case, the ease of recollection differs for an old word condition and feature and conjunction conditions because the retrieval cues are markedly different. A cue in the old word condition can be characterized as an exact conceptual match to the study episode, whereas a cue in the feature and conjunction conditions can be characterized as a conceptual mismatch. Recognition performance typically relies heavily on conceptually driven processes (e.g., Roediger et al., 1989). When a retrieval cue is not a good conceptual match for an earlier study episode, then it should not be too surprising that recognition for the targeted study episode is poor. We suggest that the mismatch in conceptual processing between the study and test phases for the feature and conjunction conditions makes recollection of the concepts from the study phase particularly difficult (though not impossible). In Experiments 3 and 4, study repetition increased the likelihood that, given enough time during the test, a study prime word (concept) could be recollected despite the poor retrieval cues. This boost in recollection allowed participants to avoid committing a feature or conjunction error, despite a parallel increase in familiarity that also was produced by study repetition. However, as already noted, this benefit in recollection was not very strong. A stronger pattern would have been a negative repetition effect for the long deadline groups.

The possibility exists that a higher number of study repetitions would not decrease the conjunction error rate relative to a single study presentation condition. That is, study repetition might not produce a negative repetition effect on conjunction error rates. Our hunch that some number of study presentations will be sufficient to produce a negative repetition effect is similar to a hunch made by Hintzman, Curran, and Oppy (1992) for the rejection of highly similar items (singular vs. plural forms of a word) in a frequency estimation procedure. In Hintzman and colleagues' experiments, after 25 study presentations of a word similar to, but different from, a test word, the number of frequency estimations of zero, which reflected the realization that the test stimulus was not presented during the study phase, was no different from that when the number of study presentations was one. (This same result was obtained under exclusion instructions.) Overall, frequency estimation increased for both old and similar item conditions. Thus, their hunch regarding a negative repetition effect was not borne out, at least with 25 study presentations. The maximum number of presentations in our experiments was four for Experiment 3 and three for Experiment 4. Further research will be necessary to determine if the predicted outcome from our hunch is equally elusive.

Finally, a recall process in recognition also has been considered for word pair stimuli (e.g., Clark, 1992; Clark & Gronlund, 1996; Gronlund & Ratcliff, 1989; Hockley, 1992; Humphreys & Bain, 1983; Mandler, 1980) or very similar study items and test lures (Clark, 1992; see also Hintzman et al., 1992, and Yonelinas, 1997). For example, pairs of study words (e.g., A–B, C–D) often are rearranged on a later recognition test (A–D), and the rejection of a rearranged pair by recalling an earlier target pair is referred to as recalling-to-reject (Clark & Gronlund, 1996). In the feature and conjunction error paradigm, whether recollection (e.g., recall) of the study primes precedes, follows, or runs parallel to an accumulation of information for a familiarity assessment is an open question (see also Hintzman et al., 1992). Others have implied that that when a recall-to-reject approach is employed, recall necessarily follows a familiarity assessment (e.g., late correction; Hintzman & Curran, 1994). Our position is that, while late correction occurs, early selection should not be ruled out (see also McElree et al., 1999). That is, information that is easily accessed with a recollection process (e.g., fast recollection) may obviate the need for a familiarity assessment.

Other Two-Factor Theories

Underwood and colleagues (1976) used a two-factor frequency distinction to account for feature and conjunction errors. That proposal included two independent frequency counters: one for semantic information and one for nonsemantic information (e.g., visual, phonemic, articulatory). A greater situational frequency was hypothesized to increase the likelihood of endorsing a test word as old. Underwood et al. reasoned that feature and conjunction errors represent mostly nonsemantic errors, and evidence for this conjecture was later reported by Ghatala and colleagues (1978). The increase for feature and conjunction errors with study repetitions in our Experiments 3 and 4 provides further support for the frequency theory. An increase in the study presentations increased the situational frequency for those components, which increased the false alarm rate for feature and conjunction lures. Underwood et al. suggested that the situational frequency of nonsemantic information may accrue independently of semantic information but did not argue that the influences from the two types of information could oppose one another (see also Ghatala et al., 1978). Therefore, the two-factor frequency theory was not developed to predict the results from Experiments 3 and 4 where participants in a long deadline condition showed some control over feature and conjunction errors.

A second two-factor theory of memory that has been used to account for conjunction errors, but not feature errors, has proposed two types of representations: features, which are low level representations, and configurations, which are based upon the binding of feature representations. For example, for the study words *checklist* and *needlepoint*, *check*, *list*, *needle*, and *point* would be encoded as features. Those features could then be bound to form the configurations *checklist* and *needlepoint*. Conjunction errors are thought to occur because “subjects selectively forget or fail to encode, the global structures [configurations] of the stimuli that were originally studied” (Reinitz et al., 1996, p. 287). During the test, features (e.g., *check* and *point*) are retrieved and conjoined in the process of that retrieval (e.g., forming *checkpoint*). This mental conjunction process creates the illusion that the conjunction word was presented in an earlier study phase.

One critical flaw with this approach is that it does not account for feature errors (Rubin et al., 1999). In fact, feature errors have not been considered to be different from errors for wholly new items (e.g., Reinitz et al., 1992). Although feature errors do not always occur at a likelihood that is significantly above chance (e.g., Reinitz et al., 1992, 1994), feature errors nearly always occur at a rate above baseline (errors for new items). We surveyed existing results from studies conducted by Reinitz and colleagues (Reinitz & Demb, 1994; Reinitz et al., 1992, 1994, 1996) and found that 14 out of 14 groups falsely recognized feature word or face lures at a higher rate than baseline. Nine out of 12 comparisons from Kroll and colleagues’ (1996) paper showed feature errors higher than baseline. (This tally excluded a 1-participant “group” with bilateral damage to the hippocampus.) Three out of three conditions yielded more feature errors than baseline in a paper by Rubin and colleagues (1999). Finally, 17 out of 17 conditions showed feature error rates above baseline in our own research (Jones et al., 2001, and the present experiments). In Rubin and colleagues’ and our own research, the differences were all significant. The total comes to 43 out of 46 comparisons showing a higher rate for feature errors than baseline false alarms. For experiments using verbal stimuli, which we used in the experiments here, 26 of 28 comparisons show this pattern. Thus, feature errors occur reliably and need to be explained.

A second problem with the feature–configuration approach is a claim about the availability of configurative representations. When conjunction errors are made, configurations for the corresponding study primes are thought to have not been encoded or to have been selectively forgotten, leaving only feature representations. The crux of the problem lies in the distinction between availability and accessibility (e.g., Tulving & Pearlstone, 1966). Configurative representations could be available but not accessible, and the results from the retrieval manipulations in this paper demonstrate this point. Certainly, the feature–configuration approach could be modified to account for feature errors and to account for the results in our experiments. However, in our thinking, most of these modifications would produce a theory that resembles dual-process theory. In fact, one approach would be to combine the feature–configuration distinction with dual-process theory. The result would be an approach quite similar to, or perhaps the same as, one taken for the combination of an item–

associative information and a dual-process viewpoint (Yonelinas, 1997; see below). In this case, the resulting theory might be redundant.

We turn now to a two-factor theory that has not been used to account for feature and conjunction errors. In this final approach, a distinction is made between item and associative information (or relational information; e.g., Humphreys, 1976, 1978; Hunt & Einstein, 1981), and this framework has been used in formal models of memory (SAM, Clark & Shiffrin, 1987; Gillund & Shiffrin, 1984, Raaijmakers & Shiffrin, 1980; MINERVA2, Hintzman, 1984, 1988; TODAM, Murdock, 1982, 1983; matrix models, Humphreys, Bain, & Pike, 1989; Pike, 1984; see Clark & Gronlund, 1996, for a cogent discussion on the different models).[2] In much of this work word pairs are used as stimuli. Item information comes from individual words of a word pair, and associative information concerns the relationship or link between the two words. Unlike the feature–configurative distinction, the item–associative distinction has focused on retrieval dynamics of the two types of information. For example, associative information is thought to accrue more slowly than item information (Gronlund & Ratcliff, 1989) but decay at a slower rate (e.g., Hockley, 1992).

Two studies that have involved study repetition of word pairs are relevant to our present work. In a study by Hockley and Cristi (1996) frequency judgments for word pairs (A–B) were obtained after a study phase where A–B was presented. However, the two words (A and B) also were repeated in separate pairs (A–C, D–B) during the study phase. Frequency of occurrence of A and B in different pairs boosted the estimated frequency for the pair A–B, and this effect was greater than that for when just one of the words (A or B) was repeated in a different pair. This boost in frequency estimation is similar to our boost in false recognition in the short deadline groups.

In the second study, which was conducted by Kelley and Wixted (1997), study repetition of word pairs (A–C, D–B) did not influence false recognition of rearranged pairs (A–B).[3] Recognition errors for repetition conditions neither increased nor decreased relative to a single study presentation. This result is quite like our results in the long deadline conditions. In fact, the interpretation of that result was the same as that in our work: Two opposing processes were equally affected by study repetition but balanced each other, resulting in a null effect. In addition, Yonelinas (1997) has argued that for word pairs recognition of item information is based on familiarity and recollection but that recognition of associative information is based on recollection (see Yonelinas, Kroll, Dobbins, & Soltani, 1999, for a different conclusion for face drawing stimuli).

Our results bolster the proposal that familiarity can underlie item information (Yonelinas, 1997) or associative information (Yonelinas et al., 1999). However, it is not clear whether feature and conjunction lures should be characterized as containing item information or item and associative information. Arguably, compound words are processed as conceptual units instead of two separate concepts with an additional association that is typical of word pairs. In this sense, the compound words appear to contain only item information. On the other hand, a broader definition has been used for associative information. In a broad definition, associative information is the information that allows one to discriminate between old stimuli and rearranged stimuli (cf., Clark & Gronlund, 1996). By this definition, associative information would allow for discrimination of old words and feature or conjunction lures. In either case, familiarity underlies the errors.

[2] The item–associative distinction and the feature–configuration distinction are similar, but the former is more specific than the latter (also see Chalfonte & Johnson, 1996). For example, the feature–configuration distinction was used to account for conjunction errors for both verbal and nonverbal stimuli (e.g., Reinitz et al., 1992). On the other hand, the item–associative distinction has undergone more development with respect to verbal stimuli (though see Yonelinas et al., 1999).

[3] We recently learned through personal communication with J. Wixted that a manuscript describing this work has been accepted for publication in *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

The avoidance of the errors in the long deadline conditions of Experiments 3 and 4 reflects the use of recollection. Again, whether the recollection of a compound word is based on item information or associative information (or both) depends on the definitions of item and associative information. Formal memory models may be able to produce conjunction errors for compound words, but the definitions of what constitutes item or associative information may need to be modified or a dual-process approach may need to be incorporated. Nevertheless, a recall-to-reject operation (e.g., Hintzman & Curran, 1994; Hintzman et al., 1992; see also, Clark & Gronlund, 1996; Yonelinas, 1997) certainly could be used to account for avoidance of feature and conjunction errors through conscious control.

CONCLUSION

In summary, the results from the present retrieval manipulations can be explained by dual-process theory. The need for a second process to account for feature and conjunction errors was highlighted in Experiments 3 and 4. Study repetition increased both familiarity and recollection, evidenced by repetition effects on feature or conjunction errors for the short response deadline groups but not the long response deadline groups. Thus, feature and conjunction errors are based on familiarity in the absence of recollection. Although other approaches (e.g., a feature–configuration distinction or an item–associative information distinction) may not be able to account for our results, they probably could be modified to do so. In fact, by taking the approach of Yonelinas (1997), an item–associative distinction that is combined with a dual-process viewpoint accounts for these results with ease. In the sort of mental chemistry of J. S. Mill (1872), the conjunction of elements in feature and conjunction research generates an altogether new concept instead of a simple association between unrelated concepts. Because of the conceptual changes from the study stimuli to the test stimuli, gaining control over feature and conjunction errors (through recollection) is difficult, though possible. While some research has been conducted to identify the underlying source of familiarity (Jones et al., 2001), the current experiments were not designed to shed light on this issue. Thus, the source of familiarity remains unspecified. An important avenue for future research will be to determine the source of familiarity that underlies feature and conjunction errors.

REFERENCES

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

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Supps (Eds.), *Contemporary developments in mathematical psychology: Vol. 1. Learning, memory, and thinking* (pp. 243–293), San Francisco: Freeman.

Bartlett, J. C., & Searcy, J. H. (1998, November). *Configural processing and binding in recognition of faces*. Paper presented at the 39th Annual Meeting of the Psychonomic Society, Dallas, TX.

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, **24,** 403–416.

Clark, S. E. (1992). Word frequency effects in associative and item recognition. *Memory & Cognition*, **20,** 231–243.

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.

Clark, S. E., & Shiffrin, R. M. (1987). Recognition of multiple-item probes. *Memory & Cognition*, **15,** 580–598.

Craik, F. I. M. (1982). Selective changes in encoding as a function of reduced processing capacity. In F. Klix, J. Hoffman, & E. van der Meer (Eds.), *Cognitive research in psychology* (pp. 152–161), Berlin: Springer-Verlag.

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, **125,** 159–180.

Dosher, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology*, **16,** 519–555.

Ghatala, E. S., Levin, J. R., Bell, J. A., Truman, D. L., & Lodico, M. G. (1978). The effect of semantic and nonsemantic factors on the integration of verbal units in recognition memory. *Journal of Experimental Psychology, Human Learning and Memory*, **4,** 647–655.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91,** 1–65.

Gronlund, S. D., & Ratcliff, R. (1989). A time course of item and associative information: Implications for global matching models. *Journal of Experimental Psy-*

chology: Learning, Memory, and Cognition, **15,** 846–858.

Hannigan, S. L., & Rienitz, M. T. 2000. Influences of temporal factors on memory conjunction errors. *Applied Cognitive Psychology*, **14,** 309–321.

Hay, J. F., & Jacoby, L. L. (1996). Separating habit and recollection: Memory slips, process dissociations and probability matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **22,** 1323–1335.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, **16,** 96–101.

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

Hintzman, D. L., Caulton, D. A., & Levitin, D. J. (1998). Retrieval dynamics in recognition and list discrimination: Further evidence for separate processes of familiarity and recall. *Memory & Cognition*, **26,** 448–462.

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

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

Hockley, W. E. (1992). Item versus associative information: Further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18,** 1321–1330.

Hockley, W. E., & Cristi, C. (1996). Tests of the separate retrieval of item and associative information using a frequency-judgement task. *Memory & Cognition*, **24,** 796–811.

Humphreys, M. S. (1976). Relational information and the context effect in recognition memory. *Memory & Cognition*, **4,** 221–232.

Humphreys, M. S. (1978). Item and relational information: A case for context independent retrieval. *Journal of Verbal Learning and Verbal Behavior*, **17,** 175–187.

Humphreys, M. S., & Bain, J. D. (1983). Recognition memory: A cue and information analysis. *Memory & Cognition*, **11,** 583–600.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review*, **96,** 208–233.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, **20,** 497–514.

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

Jacoby, L. L. (1996). Dissociating automatic and consciously controlled effects of study/test compatibility. *Journal of Memory and Language*, **35,** 32–52.

Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25,** 3–22.

Jacoby, L. L., Jones, T. C., & Dolan, P. O. (1998). Two effects of repetition: Support for a dual-process model of know judgments and exclusion errors. *Psychonomic Bulletin & Review*, **5,** 705–709.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Attention, awareness, and control. *Journal of Experimental Psychology: General*, **122,** 139–154.

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, **118,** 115–125.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. Cohen, & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 13–47), Mahwah, NJ: Erlbaum.

Jones, T. C., Jacoby, L. L., & Gellis, L. 2001. Cross-modal feature and conjunction errors in recognition memory. *Journal of Memory and Language*, **44,** 131–152.

Kelley, R., & Wixted, J. T. (1997, November). *Associative information following mixed-strength lists*. Paper presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia, PA.

Kelley, R., & Wixted, J. T. (in press). On the nature of item and associative information. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Kroll, N. E. A., Knight, R. T., Metcalfe, J., Wolf, E. S., & Tulving, E. (1996). Cohesion failure as a source of memory illusions. *Journal of Memory and Language*, **35,** 176–196.

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

Mandler, G. (1991). Your face looks familiar but I can't remember your name: A review of dual-process theory. In W. E. Hockley, & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 207–225), Hillsdale, NJ: Erlbaum.

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating familiarity and recollective processes: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25,** 563–582.

Mill, J. S. (1872). *A system of logic, ratiocinative and inductive, being a connected view of the principles of evidence and the methods of scientific investigation (8th ed.)*, London: Longmans, Green, Reader, and Dyer.

Mulligan, N. W., & Hartman, M. (1996). Divided attention and indirect memory tests. *Memory & Cognition*, **24,** 453–465.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Journal of Experimental Psychology*, **64,** 482–488.

Murdock, B. B. (1983). A distributed memory model for serial-order information, *Psychological Review*, **90,** 316–338.

Pike, R. (1984). A comparison of convolution and matrix distributed memory systems. *Psychological Review*, **91,** 281–294.

Raaijmaker, J. G., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 207–262), New York: Academic Press.

Reinitz, M. T., & Demb, J. B. (1994). Implicit and explicit memory for compound words. *Memory & Cognition*, **22,** 687–694.

Reinitz, M. T., & Hannigan, S. L. (in press). Effects of simultaneous stimulus presentationa and attention switching on memory conjunction errors. *Journal of Memory and Language*.

Reinitz, M. T., Lammers, W. J., & Cochran, B. P. (1992). Memory conjunction errors: Miscombination of stored features can produce illusions of memory. *Memory & Cognition*, **20,** 1–11.

Reinitz, M. T., Morrisey, J., & Demb, J. B. (1994). The role of attention in face encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20,** 161–168.

Reinitz, M. T., Verfaellie, M., & Milberg, W. P. (1996). Memory conjunction errors in normal and amnesic subjects. *Journal of Memory and Language*, **35,** 286–299.

Roediger, H. L., III (1996). Memory illusions. *Journal of Memory and Language*, **35,** 76–100.

Roediger, H. L., III, Weldon, M. S., & Challis, B. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger, III, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3–41), Hillsdale, NJ: Erlbaum.

Rubin, S. R., Van Petten, C., Glisky, E. L., & Newberg, W. M. (1999). Memory conjunction errors in younger and older adults: Event-related potential and neuropsychological data. *Cognitive Neuropsychology*, **16,** 459–488.

Schmitter-Edgecombe, M. (1999). Effects of divided attention on perceptual and conceptual memory tests: An analysis using a process-dissociation approach. *Memory & Cognition*, **27,** 512–525.

Schneider, W. (1990). *Micro Experimental Laboratory* [Computer program, Version 1.0]. Pittsburgh, PA: Psychology Software Tools, Inc.

Searcy, J. H., Bartlett, J. C., & Memon, A. (1999). Age differences in accuracy and choosing in eyewitness identification and face recognition. *Memory & Cognition,* **27,** 538–552.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, **5,** 381–391.

Underwood, B. J., Kapelak, S. M., & Malmi, R. A. (1976). Integration of discrete verbal units in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, **2,** 293–300.

Underwood, B. J., & Zimmerman, J. (1973). The syllable as a source of error in multisyllable word recognition. *Journal of Verbal Learning and Verbal Behavior*, **12,** 701–706.

Wolters, G., & Prinsen, A. (1997). Full versus divided attention and implicit memory performance. *Memory & Cognition*, **25,** 764–771.

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

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, **25,** 747–763.

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I. G., Lazzara, L., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember/know, process dissociation, and receiver operating characteristics. *Neuropsychology*, **12,** 1–17.

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I. G., & Soltani, M. (1999). Recognition memory for faces: When familiarity supports associative recognition judgments. *Psychonomic Bulletin & Review*, **6,** 654–661.

(Received May 10, 2000)
(Revision received August 21, 2000)
(Published online April 12, 2001)