

Distinctiveness in Recognition and Free Recall: The Role of Recollection in the Rejection of the Familiar

Ian G. Dobbins, Neal E. A. Kroll, Andrew P. Yonelinas, and Qiang Liu

University of California at Davis

Dual process models of recognition have identified two underlying processes which contribute to recognition performance: recollection, which involves the retrieval of qualitative information regarding an event occurrence, and familiarity, which represents a generalized feeling of prior occurrence. It has been proposed that recognition and free recall may be related because both involve the retrieval of qualitative event information. To examine this possibility, we compared recognition and free recall under different levels of word frequency, presentation frequency, and distinctiveness of semantic encoding. All three variables dissociated across recognition and recall. Most importantly, shifting the semantic orienting task between preexposure and study lists greatly facilitated recognition, yet left free recall unaffected. This benefit occurred primarily because the shift enabled subjects to more efficiently reject distractors that were familiar as a result of preexposure, but not encoded on the appropriate dimension. Since subjects in recall conditions were not prone to intrusions as a function of preexposure, and, in fact, could not intentionally provide sizable numbers of these items, such a selection mechanism was unnecessary. The current findings, in conjunction with those from process dissociation studies, emphasize the role of recollection in terms of selective responding in the presence of highly familiar competitors. Retrieved information which is not distinctive cannot serve as a basis for excluding alternative sources, and therefore will not contribute to performance nor be reflected in estimates of recollection. As a result, recollection estimates may often diverge from free recall performance. © 1998 Academic Press

A considerable body of research has demonstrated that the effects of preexperimental word frequency dissociate across recognition and recall (for review see Gregg, 1976). Such dissociations may yield critical information regarding differences in decision and control processes mediating memory performance on the two types of task. Given that both types of task are often held to rely on a unified declarative memory system (e.g., Haist, Shimamura, & Squire, 1992) and are frequently

used in clinical memory assessment, knowledge regarding dissociations is vital.

In the case of recognition, discrimination improves as word frequency decreases (see Mandler, Goodman, & Wilkes-Gibbs, 1982; and Wixted, 1992, for exceptions), whereas for free recall, performance increases with the frequency of the items. Despite decades of research, it is still unclear why low frequency items should enjoy a recognition advantage over high frequency counterparts. One approach taken has been to apply dual process models of recognition (e.g., Atkinson & Juola, 1974; Hintzman, Curran, & Oppy, 1992; Jacoby, 1991; Mandler, 1980; Tulving, 1985) to the word frequency problem. Although these models have different areas of emphasis, all suggest that recognition judgments may be made on the bases of two cognitive states, processes, or types of information: the retrieval of contextual information (i.e., “recollection” or “remembering”) and/or item en-

This research was supported in part by the National Institutes of Neurological Disorders and Stroke (Grant NS17778 to Neal Kroll). We thank Endel Tulving, Douglas Nelson, Reed Hunt, and Jeffrey Toth for earlier reviews of this manuscript. In addition, we thank all of the subjects, and Susan Landau and Chris Cribari for helping to conduct the experiments.

Address reprint requests to Ian G. Dobbins, Department of Psychology, One Shields Avenue, University of California, Davis, CA. 95616. Fax: 530 752-2087. E-mail: igdobbins@ucdavis.edu.

dorsement in the absence of such qualitative information (i.e., “fluency,” “familiarity,” “knowing,” or “registration without learning”). Recent investigations of the word frequency effect, using the remember/know (Tulving, 1985) and process dissociation (Jacoby, 1991) paradigms, suggest that it is largely the result of recollective processes (Gardiner & Java, 1990; Guttentag & Carroll, 1994; Kinoshita, 1995; Yonelinas, Regehr, & Jacoby, 1995), which presumably rely on the retrieval of qualitative information regarding the study items. However, this viewpoint is in contention with that of Mandler (1982) and MacLeod and Kampe (1996), who assert that the effect is primarily driven by automatic, familiarity processes. These researchers contend that the lower preexperimental frequency items receive a differential increment to their familiarity value, relative to high frequency items during study. In short, these items receive a greater familiarity “boost” during study through increased rehearsal, differential attending, greater integration, etc. In a recognition test using novel distractors, this leads to the prediction of lower false alarm rates for low frequency distractors (because of lower preexperimental familiarity), in conjunction with higher hit rates for low frequency items (due to a greater familiarity increment). Hirshman and Palij (1992) questioned this account on the basis of their failure to find increases in the word frequency effect with increased rehearsal time. Similarly, Kinsbourne and George (1974, Experiment 1) used a multiple alternative forced choice method to address the issue, in which each target was embedded in an array spanning 10 word frequency bands. In this case, differential “boosting” of the target item familiarity during study would not differently affect discrimination performance because both item types would have a distractor of equal familiarity present in the test choices. Despite this manipulation, a word frequency effect was still observed. Finally, both the studies of Gardiner and Java (1990) and Kinoshita (1995) failed to find word frequency effects for “know” responses which presumably directly reflect the operation of familiar-

ity processes. Thus, the preponderance of recent evidence suggests that the effects of word frequency on recognition are more consistent with the operation of recollective than non-recollective processes. However, preexperimental word frequency is not the only manipulation shown to dissociate across recognition and recall. There is also evidence that direct manipulation of frequency of items within the experiment may also dissociate recognition and recall performance.

For example, Anderson and Bower (1972), employing a multiple list learning paradigm with a restricted item pool, demonstrated that proactive interference during immediate recall developed relatively slowly in comparison with the benefit afforded by target item repetition in previous lists. Thus, on average, immediate recall improved with the increasing frequency of targets despite the fact that potential intrusion items had also been repeatedly presented. Such is not the case for recognition judgments, where prior presentation of items in the test pool impairs the ability to later discriminate a particular subset of the items (Kinsbourne & George, 1974; Tulving & Kroll, 1995). Furthermore, this preexposure effect may be independent of the effect of word frequency on recognition (Kinsbourne & George, 1974, Experiment 2). These findings suggest that the preexposure of future distractors in a recognition paradigm may have particularly severe consequences relative to similar manipulations in free recall. One might expect such a pattern if recognition procedures encourage item selection based on stimulus driven familiarity when recollection of list source information is absent (cf., Huppert & Piercy, 1978; Koriat & Goldsmith, 1994).

Two recent studies suggest that multiple prior intra-experiment presentations exert a strong influence over non-recollective recognition processes. Hintzman et al. (1992) demonstrated that increasing the presentation frequency of study items (words or line drawings) can be shown to directly increase judgments of frequency for highly similar distractors (plurality changes or mirror reversals). Importantly, Hintzman et al. (1992) dis-

TABLE 1
Results of Tulving and Kroll (1995)

Study condition	Hits	False alarms	Accuracy (H-F/A)
Novel	.77 (1)	.21 (0)	.56
Familiar	.67 (7)	.47 (6)	.20

Note. Proportions of test items judged as belonging to target list with item frequency in parentheses. From *Novelty Assessment in the Brain* by E. Tulving and N. Kroll, 1995, *Psychonomic Bulletin & Review*, 2, pp. 387–390. Copyright 1995 by the Psychonomic Society, Inc. Adapted by permission.

covered that the distribution of responses for similar distractors was bimodal and consistent with a two process model where subjects reject distractors outright (i.e., assign a judgment of frequency 0) if they consciously apprehend that plurality or orientation has been altered, whereas in the absence of this information, the probability of incorrect endorsement is proportional to the number of times the item has been previously seen in the experiment (i.e., a function of familiarity). They termed this latter process “registration without learning” because the increased learning trials failed to improve the ability to detect a qualitative change to the study stimulus while simultaneously increasing the judged frequency of the similar distractor.

A second paradigm displaying a similar effect was fashioned by Tulving and Kroll (1995) in an investigation of the novelty encoding hypothesis. This hypothesis, derived largely from neuroimaging data (e.g., Tulving, Markowitsch, Craik, Habib, et al., 1996; Tulving, Markowitsch, Kapur, Habib, & Houle, 1994) holds that the efficacy of an item’s encoding is directly related to its novelty within the experiment at the time of study. Items which receive large amounts of pre-exposure within the experiment context prior to their inclusion in a study list are predicted to be remembered less well than items whose initial appearance within the experiment is during study. Table 1 is recreated from the original and shows the proportion of items endorsed

as targets, along with the total number of times the items were encountered prior to test.

In accord with the hypothesis, there is a moderate increase in hit rates for items that were novel, as opposed to familiar at study, suggesting that they are remembered better. However, the majority of the difference between conditions is carried in the false alarm rates. Distractors that had six preexposures prior to testing were more than twice as likely to be incorrectly endorsed than completely novel distractors. Given that the novelty encoding hypothesis focuses on efficacy of encoding for studied items, it has little to say with regard to this false alarm behavior. At first glance, the difference in false alarm rates may seem paradoxical. One would think that having seen an item six times outside the target study list, subjects would often recollect its source during testing and not be lured into making a false alarm. Although this might occur when distractor and target lists are mutually exclusive, in this experiment half of the study list items were drawn from the pool of preexposed items. This creates a situation in which recollecting that an item occurred in the preexposure phase does not reduce the uncertainty regarding the item’s potential occurrence in the study list. Furthermore, because the same semantic orienting task is used across lists, recollection of responses, or semantic bases for responses, likewise does not reduce uncertainty regarding list source. Based on these considerations, we suggest that the difference in false alarms rates between preexposed and novel distractors represents a process similar to “registration without learning” proposed by Hintzman et al. (1992); that is, the operation of familiarity based responding in the absence of the contribution of recollection to performance (e.g., Huppert & Piercy, 1976, 1978; Jacoby, 1991).

The goal of the current study is to expand upon the work of Tulving and Kroll (1995) by contrasting performance in the original design, in which recollection was presumed to be minimal, with a condition in which it was predicted to be substantial. This was accomplished using the same basic design, but in

addition, including a condition where the semantic orienting task shifted between preexposure and study trials. Furthermore, we added a level to the preexposure frequency manipulation and included a word frequency manipulation in order to determine the degree to which these variables are related in their effect on recognition. The research reviewed earlier suggests that word frequency primarily exerts its effect on recollection. This being the case, it is possible that our attempt to improve recollection will critically depend on this factor. Finally, a free recall test was administered to other subjects to examine the effects of the same variables on free recall performance. At present, the relationship between recollective processes in recognition and free recall performance are poorly understood. The dual process model of Mandler (1980) asserts that the two are highly related; in fact, recollection is often estimated using recall performance. However, the dual process model of Jacoby and colleagues (e.g., Jacoby, 1991; Jacoby, Toth, & Yonelinas, 1994) makes no such claim. Therefore, the comparison of recognition and recall may prove informative.

It was hoped that by simultaneously analyzing three critical factors (i.e., encoding task, presentation frequency, and preexperimental word frequency) across two dependent measures (recognition and recall), a clearer picture might evolve of the relative contribution of each factor to the two memory tasks. Given that current models of declarative memory (e.g., Squire & Zola, 1996) ostensibly hold that these tasks draw on the same memory system, differences in performance across tasks may yield useful information regarding the contribution of decision-level control processes to performance (cf., Koriat & Goldsmith, 1994).

GENERAL METHOD

The following methods describe the subjects, materials, and study tasks employed in the three separate experiments. Critical differences will be noted prior to each results section.

Participants

Three separate groups of 42 subjects were tested in three separate experiments (total 126). All groups were treated identically prior to testing. The subjects were undergraduate students at the University of California: Davis and participated in return for course credit.

Materials

A total of 126 English nouns were selected from three preexperimental frequency categories (lexical frequency = high, low, and rare). High and low frequency words were selected according to the counts of Kucera and Francis (1967) and had occurrences of greater than forty per million and one to three per million respectively. Rare items were selected from Thorndike and Lorge (1944) with frequencies of four per 18 million. Additional items were added to the rare list with the constraint that they appeared in neither of the frequency references cited above. All words were composed of two to four syllables and were three to ten letters in length.

Each of the 126 words were shown in one of six conditions illustrated in Table 2. The left column of this table shows the condition and whether the item represents a future target or distractor. The next two show the number of times an item occurred in the preexposure phase and whether or not the item appeared in the critical study phase. Summing across columns shows the total number of times the word appeared prior to testing (presentation frequency). These conditions gave us three levels of presentation frequency to contrast. This enabled contrasts across three conditions in which both the target and the distractor items had been seen once (0-1 vs 1-0), twice (1-1 vs 2-0), and five times (4-1 vs 5-0) prior to recognition testing. There were two reasons for employing this method. First, we wanted to be sure that any accuracy differences between presentation conditions were the result of the novelty of the target item at encoding and not confounded with a relative frequency difference between a target and its associated distractor. Second, we wanted to avoid com-

TABLE 2
Presentation Frequency Conditions Employed in the Study

Presentation frequency	Condition	Preexposures	Study list	Exposures at testing
One	Target	—	1	1
	distractor	1	—	1
Two	Target	1	1	2
	distractor	2	—	2
Five	Target	4	1	5
	distractor	5	—	5

paring a discrimination condition employing a completely novel distractor, to conditions in which the distractors have been seen, based on the possibility that the two conditions may require different cognitive processes. The former judgment can be made solely on the basis of occurrence (i.e., "Have I seen this item in this experiment?"); however, the later requires a source discrimination (i.e., "Where have I seen this item in the experiment?"). Additionally, the recognition test contained a set of completely novel distractors (0-0) to test this hypothesis. Importantly, this means that the frequency of presentation for the targets and distractors at each level is identical, the critical difference is that the target has been seen in the study list. This organization resulted in 18 words per condition, with 6 of each lexical word frequency in each of the conditions. Fifty-four words were in the study list. Each word appeared equally often in each condition resulting in seven lists for counterbalancing.

Procedure

Subjects were tested individually, and responded via a computer keyboard. In the pre-exposure phase they were asked to make semantic living/nonliving judgments regarding individual items presented in the center of the computer monitor in large graphic print. Small reminders were present at the top left and right of the monitor throughout, to prevent subjects from forgetting the appropriate response keys. Subjects were told that many of the items may repeat, however no mention regarding speed or consistency was made. Following this, a

five minute distractor task was given during which subjects entered the names of as many cities they could think of into the keyboard. After the distractor task, they were informed that they were going to be given a 'memory test' for the next set of items appearing on the monitor and should make a strong effort to commit these items to memory. No mention of type of test was made. Two-thirds of these items were drawn from the pre-exposure phase, and subjects were told that they would recognize many of the items from the earlier phase. The items appeared as in the previous phase and subjects were again required to complete a semantic orienting task. However, half of the subjects in each experiment performed the same living/non-living task, and the other half performed a pleasant/non-pleasant judgment task.¹ Therefore, within each ex-

¹ In order to ensure that the obtained results were not dependent upon the particular assignment of encoding tasks, two additional groups of 18 subjects each were tested: A same-question group, which made the pleasant/unpleasant judgments during both list presentation phases, and a different-question group, which made the pleasant/unpleasant judgments during the pre-exposure phase and the living/nonliving judgments during the study phase. These subjects were tested in blocks of six, whereas the original subjects were all tested individually. Both of these new groups were first tested with the recall instructions of Experiment 2, and then tested with the recognition instructions of Experiment 1. The results of the free recall test replicated those of Experiment 2, showing essentially no improvement in recall accuracy with distinctive encoding (mean accuracy was .12 and .15 for the same- and different-question groups). In addition, the results of the recognition test showed the same pattern found in Experiment 1: Greater accuracy for the different-question group

periment, there were two groups, One in which the subjects received the same orienting question for distractors and targets and the other in which the question shifted prior to target study. Study items were presented once, for a total of 54 study list items. Subjects next performed a two minute distractor task in which they typed in names of famous people as rapidly as possible. Finally, subjects were given a yes/no recognition test for all list items (Experiment 1) or a free recall task in which they entered items into the keyboard (Experiments 2 and 3). For recognition, subjects first indicated whether or not each item had appeared in the target/study list, then made a three point confidence judgment in which "1" indicated they were guessing, and "3" that they were very sure that the item was in the study list. Additionally, subjects in Experiments 1 and 2 were paid based on their performance in order to ensure effortful responding. Subjects were paid 15 cents for every hit (or correct recall) and penalized 5 cents for every false alarm (or intrusion). Performance totals were registered in the bottom corner of the monitor during testing so that participants could track their accuracy. Testing time was approximately 45 min.

EXPERIMENT 1: RECOGNITION

The purpose of this experiment was to examine the effects of the three critical variables discussed in the introduction on recognition performance. Forty-two subjects participated. The three variables examined were presentation frequency (the number of times items were presented within the experimental context—1, 2, or 5 times), preexperimental word frequency (i.e., high, low, or rare frequency words), and consistency of semantic encoding. This last variable was a between-subjects manipulation, with half of the subjects ($n = 21$) performing the same semantic encoding task across preexposure and study lists and the

other half switching tasks between lists. During the test phase, subjects were presented with words and required to judge whether each word occurred in the study list and to rate the confidence of their responses.

Results

For ease of interpretation, initial data analyses were conducted on threshold corrected accuracy scores (i.e., hits—false alarms). Following this, hit and false alarm rates were examined to determine if either differentially contributed to relevant accuracy effects between orienting question groups. All *post hoc* comparisons were conducted using Tukey's HSD tests with an α level of .05 unless otherwise noted. Table 3 shows the response proportions across the independent variables. The top row of each level shows hit rates and the lower row false alarms.

Accuracy. A $2 \times 3 \times 3$ mixed-design ANOVA with independent variables of orienting question (same or different), word frequency (high, low, and rare), and presentation frequency (one, two, or five exposures prior to test) was conducted on accuracy scores. Results yielded main effects of orienting question ($F(1,40) = 18.93$, $MS_e = .33$, $p < .001$), word frequency ($F(2,80) = 12.32$, $MS_e = .05$, $p < .001$) and a modest effect of presentation frequency ($F(2,80) = 2.9$, $MS_e = .04$, $p < .06$). Additionally, there was an orienting question by word frequency interaction ($F(2,80) = 8.05$, $MS_e = .05$, $p < .001$). The main effect of orienting question revealed that subjects who received a different question during target list learning trials ($M = .53$) were more accurate than subjects receiving the same question ($M = .27$). The main effect of presentation frequency showed a trend toward decreasing accuracy with increased presentation frequency. Least significant difference (LSD) tests suggested that the accuracy for the first two levels of experimental frequency did not differ ($M = .42$ and $.41$, respectively), but both were greater than the accuracy for the highest frequency (i.e., least novel) items ($M = .36$). The main effect of word frequency was not interpreted because of the significant

(.40) than for the same-question group (.21). Thus, it would appear that the particular encoding tasks used did not matter; what did matter was whether or not the tasks differed between preexposure and study phases.

TABLE 3
Hit and False Alarm Rates for the Variables in Experiment 1

Word frequency	Same-question condition				Different-question condition			
	Presentation frequency							
	Zero	One	Two	Five	Zero	One	Two	Five
High		.71	.83	.80	P(Hit)	.87	.89	.87
	.04	.48	.51	.60	P(FA)	.04	.17	.27
Low		.75	.78	.79		.80	.83	.83
	.02	.37	.48	.53		.00	.22	.27
Rare		.51	.63	.79		.56	.69	.63
	.04	.26	.40	.56		.06	.19	.22

orienting question by word frequency interaction. Simple effects analysis of the orienting question by word frequency interaction showed that accuracy did not change as a function of word frequency when orienting questions remained the same between lists, but did when orienting questions changed, $F(2,40) = 18.69, MS_e = .05, p < .001$. For the different-question group, both high ($M = .63$) and low ($M = .55$) frequency words were more accurately discriminated than rare words ($M = .39$), but did not significantly differ from each other.

These effects are illustrated in the top panels of Fig. 1. The superior accuracy of the different-question group is evident when examined across either presentation (panel A) or word (panel B) frequency. Panel A demonstrates that the improvement resulting from shifting the orienting question is independent of the modest impairment that occurs with increased presentation frequency. This is not the case for the word frequency. In panel B, it can be seen that the improvement resulting from shifting the orienting question is dependent on preexperimental word frequency, with the greatest improvement occurring for the highest frequency items.

Hits and false alarms. Given that accuracy estimates can obscure the relative contributions of hit and false alarm data, further analysis was conducted to see if either differentially contributed to the main effect of orienting

question and the orienting question by word frequency interaction noted above. Hit and false alarm rates were examined simultaneously as a factor (item-type) to avoid the independent interpretation of either which can lead to spurious conclusions due to between group shifts of response criterion (Gillund & Shiffrin, 1984; Hirshman & Palij, 1992).

Item-type (hit and false alarm rates) interacted significantly with the main effect of orienting question, $F(1,40) = 18.93, MS_e = .17, p < .001$. Simple effects analyses demonstrated that although the groups did not differ in hit rates, the same-question group ($M = .46$) made significantly more false alarms than the different-question group ($M = .25$), $F(1,40) = 15.05, MS_e = .30, p < .001$. Therefore, the majority of the difference in accuracy between the groups results from the same-question group being almost twice as likely to commit false alarms to preexposed distractors.

Figure 2 shows the hit and false alarm data that underlie the orienting group by word frequency accuracy interaction depicted in panel B of Fig. 1. This interaction occurred because accuracy increased with word frequency for the different-question group, but did not change as function of word frequency for the same-question group. Closer inspection of the hit and false alarm data for each question group shows why this occurred (see Fig. 2).

For the same-question group (left panel, Fig. 2), word frequency did not interact with

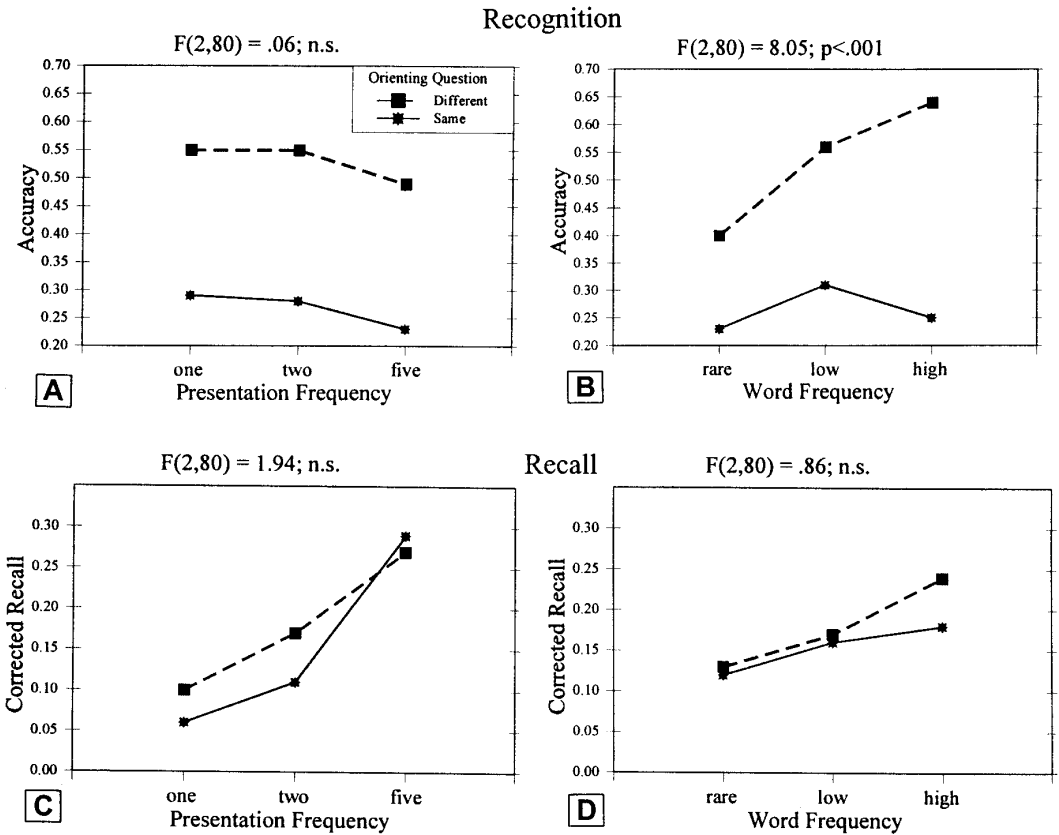


FIG. 1. Accuracy as a function of orienting question group during the recognition test of Experiment 1 (top panels) and during the free recall test of Experiment 2 (bottom panels). Group performance is shown as a function of presentation frequency (left panels) and word frequency (right panels). Dashed lines indicate the performance of the different-question groups, and solid lines show that of the same question groups. Panel B represents the only significant interaction.

hit and false alarm rates. Therefore the accuracy scores for this group (i.e., the difference between hits and false alarms) did not show an effect for word frequency. However, there was an overall main effect of word frequency on responding ($F(2,40) = 11.88, MS_e = .04, p < .001$) that resulted from both high ($M = .65$) and low ($M = .62$) frequency words being more likely to be endorsed than rare ($M = .53$) words, but not differing from one another. It is important to remember that this effect occurred regardless of whether the item is a target or distractor, and therefore represents a bias in favor of accepting high and low frequency words over rare words.

For the different-question group (right

panel, Fig. 2) word frequency significantly interacted with hit and false alarm rates, $F(2,40) = 18.69, MS_e = .03, p < .001$. The simple effects analysis of hit rates was significant, $F(2,40) = 23.43, MS_e = .04, p < .001$. Both high ($M = .87$) and low ($M = .82$) frequency items were more likely to be correctly endorsed than rare words ($M = .63$), but did not differ from one another. Thus, the pattern of hit rates mirrors the bias demonstrated in the same-question group above. However, the false alarm data for the different-question group does not follow this pattern. There was no simple effect for word frequency on the error rates for this group. Thus the different-question group was able to correctly reject

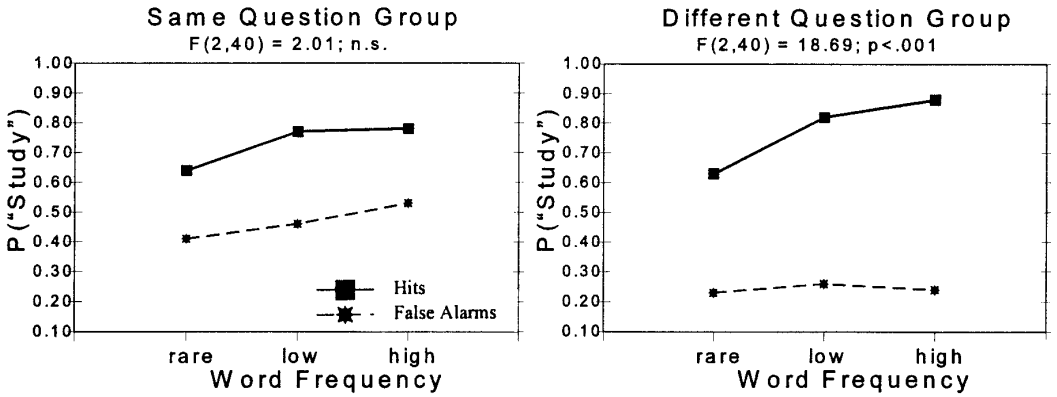


FIG. 2. Hit and false alarm data underlying the significant question group by word frequency interaction which occurred in the accuracy scores of Experiment 1.

distractors from all word frequency classes equally well.

Overall, this pattern of response data explains the orienting question group by word frequency interaction present in the accuracy scores and shown in Fig. 1B. In the case of word frequency, there is a bias toward accepting high and low frequency words over rare words. This tendency is present in the hit rate data for both groups, and in the false alarm data for the same-question group. However, there is no effect of word frequency on the false alarm data for the different-question group, which correctly rejects all distractors equally well. This suggests that this group is able to reject distractors based on categorical or qualitative information as a result of the shift in orienting question. The basis for this ability will be considered in the discussion.

Completely novel distractors. Under General Methods, we presented a brief argument against comparing discrimination accuracy between conditions which employ completely novel distractors to those in which the distractors have been seen in the experiment context one or more times, based on the possibility that the rejection of the two distractor types may rely on different cognitive processes. Comparison of the false alarm data to completely novel distractors with those seen one or more times in the experiment supports this argument. False alarm rates to completely

novel distractors are extremely low, did not differ across orienting question groups, and for five of the six cells, did not significantly differ from zero (see Table 3). This is not the case for distractors presented once or more during the experiment, for which error rates are greater than zero for all the cells presented in Table 3. Furthermore, simple inspection of the table shows that the increase in false alarm rate as one goes from completely novel distractors (zero) to those seen once prior to testing, is often several times larger than any of the other increases present, in some cases exceeding an order of magnitude. In sum, there is clear evidence suggesting that the relationship between completely novel and preexposed distractors is either qualitatively different or extremely nonlinear. This would confound any comparisons across conditions employing these different distractor types.

Discussion

The effects of presentation frequency on the same-question group replicate Tulving and Kroll (1995), with increased frequency impairing discrimination accuracy. However, the effect was considerably smaller than observed in their study. The current false alarm data suggest that this resulted because of large differences in the ability of subjects to correctly reject novel distractors versus their ability to make source discrimination judgments regard-

ing pre-exposed distractors; the latter being considerably more difficult. These effects were potentially confounded in Tulving and Kroll (1995) because their novel encoding condition employed completely novel distractors (see Table 1), which may have inflated the estimates of accuracy in comparison to the familiar condition in which distractors were preexposed.

Shifting the orienting question between pre-exposure and study lists greatly improved accuracy, primarily as a result of improving the different-question group's ability to correctly reject preexposed distractors. This improvement was independent of presentation frequency, but interacted with word frequency. If the processes underlying word and presentation frequency were identical (cf., Kinsbourne & George, 1974) such a difference should not have occurred. Closer inspection of the hit and false alarm data underlying this interaction suggested that this difference resulted from the shift in orienting question differentially affecting the correct rejection performance of the different-question group. In the case of the same-question group, there was a bias toward selecting high and low frequency items over rare items that occurred regardless of whether or not the item was actually studied (see also Mandler, 1982). However, this bias was only evident in the hit rate data of the different-question group, whose ability to correctly reject distractors did not differ across word frequencies (see Fig. 2). If we assume that the mechanism underlying this selection bias is present throughout testing, then the *effective* reduction in false alarm rates is actually greater for high and low frequency items compared to rare in the different-question group. The explanation of such an effect critically relies upon the process by which shifting orienting question improved the exclusion of distractors. The fundamental difference between targets and distractors following such a shift in the current paradigm is the nature of semantic processing performed upon them. Distractors are processed exclusively on the living/nonliving dimension, in contrast to targets in which two thirds of the cases are

processed on the living/nonliving, and all of which are judged for pleasantness. Given this, the most efficient way to exclude distractors would be for subjects to reject items that they were confident *had not* been judged for pleasantness. It appears that such a task is easier for items upon which subjects can be expected to make meaningful semantic judgments (i.e., high and low frequency) compared to those we expected to be relatively unfamiliar and poorly understood (i.e., rare items). However, it is important to keep in mind, that the subjects in the different-question condition are still considerably better at rejecting rare items than those in the same-question condition (see Fig. 2). Notably absent in the data is the standard word frequency effect for recognition, where low frequency items are recognized better than high frequency counterparts. However, these effects have typically been obtained in simple recognition tasks employing novel distractors, and not in the list discrimination format used here. Given the large differences in tasks, and in false alarm rates for novel versus familiarized distractors noted earlier, it is not clear whether word frequency effects can be expected to generalize to list discrimination recognition paradigms.

The primary goal for introducing the shift in orienting questions was to improve recognition via an improvement in recollection performance. This was based on the assumption that a large portion of what subjects remember when they see a test item is the nature of their previous encounter with it. If the semantic orienting task remains fixed between preexposure and study lists, this information will be rather limited in its contribution to list discrimination; shifting orienting tasks should alleviate this. However, our previous analysis and discussion has been largely focused on the effects of such a shift on items encoded before the manipulation, namely, preexposure list distractors. This is important because the term "recollection" is often taken to mean the retrieval of qualitative information regarding a study event. Such a definition makes considerable sense in the typical remember/know paradigm where remembering is based on infor-

mation regarding the simple *occurrence* of items within the experiment because the distractor pool is usually completely novel to the experiment context. From this, it becomes natural to equate remembering solely with the retrieval of information of target occurrence. This may be problematic for at least two reasons. First, it leaves unspecified what aspect of the previous occurrence the subject is accessing; that is, what is the nature of the information by which subjects make a judgment of "I remember?" Second, this emphasis overlooks the processes contributing to the correct rejection of distractors. Are these items dismissed on the basis of recollection, familiarity, or both? The current design closely resembles the process dissociation procedure of Jacoby and colleagues (e.g., Jacoby, 1991; Jacoby et al., 1994) and as such equates recollection with the use of qualitative information in controlled responding. The results of Experiment 1 suggest that the degree to which retrieved information can be held to be exclusive, or source distinctive, is a determining factor in the contribution of recollection to performance and that this applies not only to targets, but to potential distractors as well. This point will be explored further under General Discussion.

EXPERIMENT 2: RECALL

Given the results of Experiment 1, we were interested in observing the effects of the same manipulations on free recall performance. Both the dual process models of Mandler (1980) and Jacoby (1991) hold that discrete retrieval processes contribute to recognition performance. However, the model of Mandler assumes that this process is comparable to that which occurs during free recall, provided the study conditions and materials are similar. If this is the case, we would expect that shifting the orienting question would improve recall performance. However, the results of Experiment 1 demonstrated that this improvement was due in large part to facilitation of the correct rejection of preexposed distractors and therefore the shift reduced the subject's susceptibility to familiarity induced false alarms.

If this latter process is largely absent during free recall, shifting the orienting question may not greatly benefit performance. Failure to see an improvement would also suggest that the retrieval or selection processes differed across the two paradigms. Thus Experiment 2 replicates Experiment 1 with the only difference in procedure being a new group of 42 subjects were given a 5-min free recall test during which they entered items into the keyboard. The experimenter corrected any spelling errors and the program then automatically scored the subject's performance after each response, adding 15¢ for each correct recall and subtracting 5¢ for each intrusion. The total number of correct recalls, the total number of intrusions, and the total amount of money earned was shown on the monitor after each response.

All other conditions and materials were identical to the recognition experiment. All *post hoc* comparisons were conducted using Tukey's HSD tests, with an α level of .05 unless otherwise noted.

Results

Accuracy. Typically, statistical analysis for recall is conducted on the proportion of correct items produced, without correcting for intrusion errors prior to analysis. This is to be expected, since most researchers have no a priori reason to expect intrusions from a particular source. However, in the current study we anticipated that preexposure list intrusions would constitute the majority of errors, and indeed, greater than 70% of the errors were from that source. Therefore, we corrected recall scores, as in recognition, by subtracting the number of preexposure list intrusions from the proportion of correctly recalled items. Table 4 lists both the correct recall and the intrusion rates across the variables examined. As in Experiment 1, following accuracy analysis, effects of interest were broken down into correct recall and error rates to determine if either predominantly carried the effect.

A $2 \times 3 \times 3$ mixed-design ANOVA with independent variables of orienting question (same or different), word frequency (high,

TABLE 4
Correct Recall and Intrusion Rates for Experiment 2

Word frequency	Same-question condition			Different-question condition			
	Presentation frequency						
	One	Two	Five	One	Two	Five	
High	.11	.18	.41	P(Correct recall)	.20	.25	.40
	.02	.06	.07	P(Intrusion)	.03	.02	.07
Low	.10	.14	.33		.07	.19	.28
	.02	.02	.06		.01	.01	.02
Rare	.01	.10	.32		.06	.11	.22
	.00	.01	.06		.00	.00	.01

low, and rare), and presentation frequency (one, two, or five exposures prior to test) was conducted on corrected recall (items recalled minus intrusion errors). Results revealed no main effect of orienting question, with significant main effects of word frequency ($F(2,80) = 10.45$, $MS_e = .03$, $p < .001$) and presentation frequency, $F(2,80) = 41.89$, $MS_e = .03$, $p < .001$). There were no significant interactions. Recall was greater for high frequency words ($M = .22$) than low frequency ($M = .16$) or rare words ($M = .12$). Additionally, recall accuracy increased as a function of presentation frequency. Items seen five times ($M = .28$) were recalled better than those seen twice ($M = .15$), which in turn were recalled better than those seen only once ($M = .08$) prior to test. These accuracy effects are illustrated in the bottom row of Fig. 1.

Correct recall and intrusion errors. Inspection of Fig. 3 shows the behavior of correct recall and intrusion errors as a function of presentation and word frequency. Based on this figure, and the data in Table 4, it is clear that the changes in accuracy noted above primarily reflect modulations in the correct recall of information, and not intrusion errors.²

Discussion

Overall, the recall results are consistent with early research demonstrating that correct

recall increases more rapidly than proactive interference as a function of preexposure (Anderson & Bower, 1972), that high frequency words are recalled more often than low frequency words (Gregg, 1976), and that intrusion or list differentiation errors during free recall tasks are relatively rare (for brief review, see Winograd, 1968). What may be somewhat surprising is that shifting the orienting task did not benefit free recall performance, especially since subjects had to perform distinctive encoding in order to satisfy the requirements of the new semantic orienting task. However, the study lists were unstructured and distinctive encoding of individual items, which does not result in relational or increased elaborative encoding, may not yield increases in free recall performance (e.g., Hunt & Smith, 1996; Tulving, 1983). This point will be further addressed under General Discussion.

Overall, these differences in free recall versus the recognition performance in Experiment 1 resulted in a double dissociation of the two types of task (see Fig. 1). Whereas mild impairments with increasing presentation frequency were observed for recognition, this manipulation resulted in increased free recall performance. The opposite pattern of results was observed when the orienting question was shifted between distractor and target lists; shifting the question greatly improved recognition performance, but left free recall unaffected. This finding, that the effects of distinc-

²The uniformly low error rates prevented statistical analysis due to violations of the distributional assumptions of ANOVA.

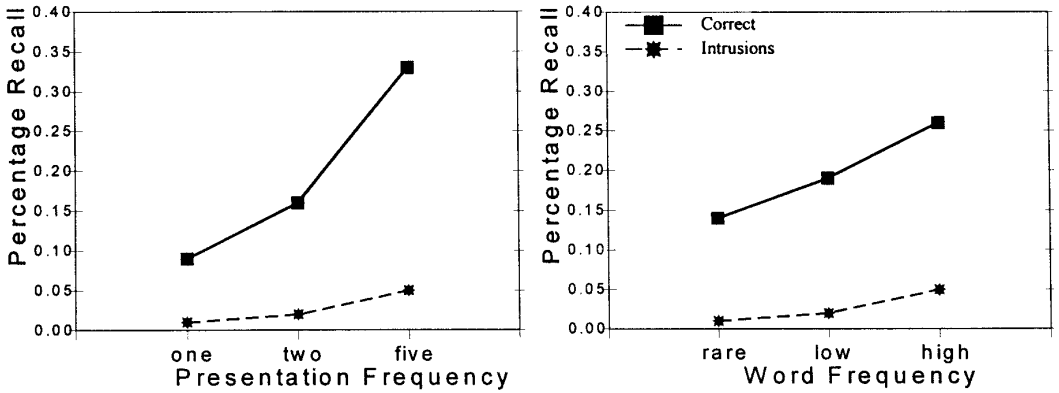


FIG. 3. Correct recall versus intrusion recall as a function of presentation (left panel) and word (right panel) frequencies. Scores are averaged over same- and different-question groups.

tive encoding operate differently during recognition and recall, suggests that the retrieval processes embodied in each task must differ (cf. Mandler, 1980). The possible nature of relationship between these retrieval processes is considered under General Discussion.

EXPERIMENT 3: TOTAL RECALL

The previous results demonstrated a marked dissociation between the recognition and recall performance in this task. One of the primary reasons for this difference is that the deleterious effects of distractor preexposure during free recall are small compared to those encountered during recognition. Presumably this pattern occurs because the subjects who are performing the recognition task can be lured into selecting distractors based on their familiarity, in absence of recollection of information which definitively identifies item source. The goal of the third experiment was to help determine why preexposure of distractors does not lead to a greater number of intrusion errors in free recall. For purposes of continuity, we will continue to use the terms “distractor” and “target” despite the fact that the items are not physically present during recall testing. It is possible that the low error rate during recall indicates that subjects are simply unable to retrieve distractors on which to err. Alternatively, it may be that they are able to retrieve these candidates, but that these

items do not lead to intrusion errors. We explored these two possibilities by repeating the earlier recall experiment with a different group of subjects—with one critical change. In this experiment, just prior to testing, subjects were informed that the previous instructions to focus only on the target list were inaccurate. Instead, they should report all items that came to mind, regardless of list origin. We reasoned that if subjects could provide considerably more distractor items than in the previous experiment, this would provide some evidence that distractor candidates are available, but do not induce responses. If the totals were not different, it would suggest that they can not retrieve these items in the first place, thus intrusion errors are rare because distractors are not retrievable.

Results

Table 5 shows the recall proportions for the study and preexposure lists, across the variables.

A 2 × 2 mixed ANOVA with independent variables of experiment group (selective recall-Experiment 2 or total recall-Experiment 3) and item-type (target and distractor) was conducted on recall proportions. Results yielded main effects for group ($F(1,40) = 17.29, MS_e = .001, p < .001$) and item-type ($F(1,40) = 445.30, MS_e = .001, p < .001$) and a group by item-type interaction, $F(1,40)$

TABLE 5

Proportion Recalled from Study and Preexposure-Only Lists for Experiment 3

Word frequency	Same-question condition			Presentation frequency	Different-question condition		
	One	Two	Five		One	Two	Five
	High	.14	.24		.42	P(Study list)	.15
Low	.06	.10	.23	P(Pre-Exposure-Only)	.05	.10	.17
	.09	.16	.34		.13	.18	.28
Rare	.02	.06	.12		.02	.03	.19
	.05	.08	.33		.04	.10	.31
	.01	.02	.14		.02	.04	.06

= 5.56, $MS_e = .001$, $p < .05$. For the group by item-type interaction, simple effects analysis revealed that the groups did not differ in the number of study list items produced, however, subjects instructed to recall all possible items generated more pre-exposure items ($M = .08$) than subjects instructed to provide target items only ($M = .03$), $F(1,40) = 35.35$, $MS_e = .001$, $p < .001$. This suggests that a small portion of distractor items (5%) were indeed available to subjects for recall in Experiment 2, but that they did not trigger intrusions. We reasoned that the preexposure items provided in this experiment represented an upper bound of possible intrusion errors that could have occurred. That is, subjects would be unable to make intrusion errors at a higher rate than this, because this is the maximum number of distractor items they are able to intentionally provide. However, even when these responses are treated as true intrusions and recall rates are "corrected," the data still show an increase in accuracy as a function of increasing experimental frequency, $F(2,80) = 33.98$, $MS_e = .03$, $p < .001$. Again, this is opposite to the pattern observed during recognition. As in Experiment 2, this increase occurs because correct recall rates of target items rise more quickly than those of non-target items as items are familiarized—even when subjects are trying to recall non-target items.

Discussion

Overall, the results suggest that a small proportion (approximately 5%) of distractors are

in fact available for retrieval, but do not trigger errors as in recognition. Furthermore, even when subjects are instructed to produce these items, recall for the originally intended targets still increases more rapidly as a function of presentation and word frequency. The fact that even when instructed to do so, subjects could only provide a small percentage of the distractor items shows that a large proportion of these items are unavailable by the time testing begins. The reason for this is unclear. One possibility is that the instruction for subjects to focus exclusively on the target list for remembering results in a disproportionate degree of elaborative rehearsal or maintenance processes between the two lists. This may result in fewer effective retrieval cues for the first list and make intentional retrieval difficult (e.g., directed forgetting). Alternatively, subjects may have simply forgotten the preexposure list items over the course of the study list learning. In any case, it is clear that during free recall in this paradigm, subjects are not spontaneously generating large numbers of distractor items and then subsequently rejecting them based on source considerations. Given this, the processes which facilitated the correct rejection of familiar distractors in Experiment 1 appear to be largely unnecessary. Because overall the error rates were low in both recall experiments, one might infer that the null effect of shifting the orienting question was the result of floor effects on intrusion rates. However, "intrusion" rates for the

highest presentation frequency in Table 5 are well above floor, and yet there is no consistent reduction in intrusions as a result of shifting the orienting question. Nevertheless, it remains to be seen whether such a null effect would be preserved in the face of substantially higher intrusion rates.

GENERAL DISCUSSION

The current study raises critical questions regarding both the processes underlying recognition and the relationship of these processes to those occurring during free recall. In the case of recognition, it was determined that shifting the orienting question improved recollection, benefiting target and distractor discrimination. Closer inspection of the data also revealed that this benefit was not carried in the hit rates, but in the improved ability of the different-question group to correctly reject familiar distractors. One approach to explaining this phenomenon is to employ the construct of distinctiveness (Hunt & McDaniel, 1993). From this perspective, recollected information will only contribute to performance when it can be used to unambiguously determine item source.

In the typical process dissociation paradigm (Jacoby, 1991), considerations of source distinctiveness are not germane because the dimension on which subjects must include or exclude items is directly related to qualitative processing differences between the encoded items during study (e.g., male vs female voice). Recollection is estimated by the ability of subjects to exclude items processed in a particular manner, despite the fact that their exposure within the experimental context will make them somewhat familiar. In this case, if a subject remembers the task performed with the item, uncertainty is completely eliminated during responding. This high degree of exclusivity or source distinctiveness is largely absent in the preexposure paradigm of Tulving and Kroll (1995) and its variant used here (same-question condition, Experiment 1). Given that the study pool is partially drawn from the items shown during preexposure, remembering that an item occurred during the

preexposure list is not useful for identifying targets. Furthermore, processing task information is likewise ineffective because the same semantic orienting task is used across lists. Based on these considerations, we expected the contribution of recollection to performance to be relatively low, and conversely, the influence of familiarity processes during responding to be relatively high. Inspection of the recognition performance of the same-question group (top panels Fig. 1) shows this to be the case. This group has considerable difficulty discriminating targets from distractors, which was exacerbated by both being seen an equal number of times within the experiment. Shifting the orienting question served to improve the distinctiveness of the processing between target and distractor lists, and enabled subjects to more efficiently reject familiar distractors. It is important to remember that distractors were encoded identically in the two groups, and therefore the increased ability of the different-question group to reject distractors was the result of processes which took place *after* encoding. This highlights a point that is sometimes obscured in dual process research. Namely, that estimates of recollection do not directly reflect the amount of qualitative information retrieved during item presentation, but rather the usefulness of this information in terms of selective responding (Jacoby, 1991). Indeed, given that distractors were encoded identically in the two orienting question groups, it is likely that the information directly retrieved is similar for each. However, the information is not equivalent in terms of its distinctiveness or task utility. It is only when subjects could reject items because they have not been encoded on a second, and different semantic dimension that performance improved.

Although we have argued that source distinctiveness critically affects the usefulness and, hence, expression of recollection, direct estimates of recollection have not been derived. One method to obtain such estimates (outlined in the Appendix) is to fit the observed receiver operating characteristics (ROCs) of each group with a mathematical

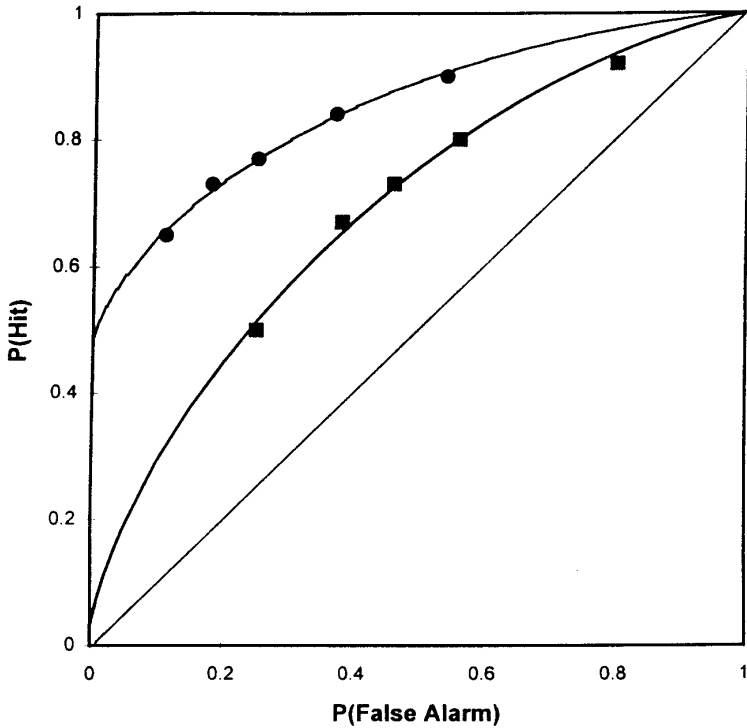


FIG. 4. Observed receiver operating characteristics (ROCs) for the same (squares) and different (circles) orienting question groups. Solid lines indicate the fit of the dual process, signal detection model.

model that assumes the independent contributions of a threshold recollection and signal detection based familiarity process to performance. This dual process, signal detection model has received considerable use in the recognition literature (e.g., Yonelinas, 1994; Yonelinas, Regehr, & Jacoby, 1995) and provides estimates of familiarity in terms of d' and recollection in terms of a simple probability. The observed ROCs for the two question groups in Experiment 1 are plotted in Fig. 4 along with the functions generated by the model. Table 6 shows the parameter estimates of recollection and familiarity for each group, along with the errors of prediction. The results are consistent with our contention that the shift in orienting question primarily benefited recollection, and that the contribution of recollection in the same-question condition was necessarily quite low. In the case of the different-question condition the recollection estimate is .47 which represents greater than a

15-fold increase over the same-question condition, where recollection is only .03. Importantly, this does not mean that subjects in the former condition remembered 15 times more information than those in the latter, merely that what was remembered was decidedly more useful for discriminating when orienting questions shifted. The estimate of familiarity also increased from same to different-question

TABLE 6

Parameter Estimates and Sum of Squared Errors of Prediction for Dual Process Equal Variance Signal Detection Model

Orienting question	<i>SSE</i>	d'	<i>R</i>
Same	0.0004	0.67	.03
Different	0.0002	0.86	.47

Note. *SSE*, Sum of squared errors of prediction; *R*, recollection.

conditions, although the increase was relatively modest compared to that observed in recollection.

The model underlying the ROC analysis assumes that recollection and familiarity are independent, and this is consistent with the finding that recollection can be used to reject the familiar. That is, if the two processes are independent, then there will be some proportion of items that elicit both recollection and familiarity. To the extent that the recollected information is distinctive, subjects can use it as a basis for the control of behavior (cf., Jacoby, 1991) and more efficiently reject lures despite their familiarity. The fact that the distinctiveness of recollected information affects the behavioral expression of familiarity neither requires, nor suggests that the two are related at the process level. Analogously, the notion that observers may adopt a strict response criterion, which will necessarily reduce observed hit and false alarm rates, does not indicate that decision bias and discrimination processes are related. In fact within signal detection theory (SDT), the two are explicitly assumed to be independent.

With respect free recall, the results of Experiment 2 were in accord with previous research on the effects of word frequency, presentation frequency, and distinctive encoding on recall performance. Most notably, shifting the orienting task did *not* improve recall performance. The lack of recall improvement with distinctive encoding may appear somewhat surprising, however, given that the new orienting task did not require relational encoding among the study list items above that required when the orienting task remained fixed, there would be no reason to expect an increase in recall performance. Hunt and McDaniel (1993) argued for at least two types of distinctive encoding processes during recall paradigms. The first can be said to be item specific, in which the encoding task forces the subjects to consider the individual attributes of each item (e.g., living/non-living judgments). The second also requires consideration of the individual attributes, however, these are then used in a comparison process along a particular dimension common to all list members (e.g.,

sorting items by size). We will refer to this latter type of distinctive encoding as relational. Distinctive encoding which is item specific provides relatively poor free recall retrieval cues compared to relational encoding, which serves to increase intralist associations. Both tasks employed in the current study focused only on item specific information, thus switching tasks would not serve to increase relational processing and improve recall. The current study extends these notions by demonstrating that this is clearly not the case for recognition, which benefited from the listwise distinctiveness of item specific encoding, despite its non-relational and non-organizational nature. In this instance, distinctive encoding relative to that performed on the distractor pool, greatly improved performance by making recollected information useful for responding.

Examination of the results of Experiments 1 and 2 show that *all three* of the critical variables dissociated across recognition and recall (see Fig. 1). In the case of presentation frequency, it mildly impaired recognition while improving free recall. Increases in word frequency served to improve recall. This also improved recognition, but only in the case where the orienting question shifted between lists. Finally, a shift in orienting questions greatly elevated recognition performance, but did not benefit free recall. The reason underlying these differences revolves primarily around the fact that recognition performance is highly susceptible to the influences of familiarity during item selection, especially when either recollection is low, or its usefulness limited. Such a pattern has also been demonstrated in organic amnesia (e.g., Huppert & Piercy, 1976, 1978). Indeed, both the shape of the ROC, and the estimates of recollection and familiarity observed in the same-question group (see Fig. 3 and Table 5), are consistent with the performance of subjects we have tested who suffer from memory impairment as a result of cerebral vascular accident (Yonelinas, Kroll, Dobbins, Lazzara, & Knight, in press).

This strong influence of familiarity is decidedly absent in recall performance in the current study, and even when instructed to do so, subjects cannot intentionally provide enough

'distractor' items to mimic recognition effects. At present, it is unclear why it is so difficult to induce intrusions during free recall. One suggestion, offered by Koriat and Goldsmith (1994), is based on the notion that subjects typically withhold responses about which they are unsure during free report conditions. Again, however, given the results of Experiment 3, this appears not to be a major factor in the current study because subjects are clearly not holding back large numbers of incorrect responses during free recall. A second factor identified by Koriat and Goldsmith (1994) involves the distinction between selection and production processes in memory report. In the case of selection (i.e., recognition), we have employed dual process models as explanatory constructs. With regard to the production processes underlying free recall, we have little to add to early research (see Winograd, 1968) suggesting that the free recall of items during list discrimination paradigms is typically accompanied by accurate identification of source; especially at short retention intervals.

The current study, and research employing the process dissociation paradigm (Jacoby, 1991, 1994), highlights the importance of recollection processes in the control of responding. This becomes readily apparent when the functional distractors are preexposed to the subject within the experiment context. In the absence of recollective control, subjects are quite vulnerable to errors based on familiarity. Such potentially competitive processes may often not play an appreciable role during free recall. Consequently, it would appear to be a mistake to directly relate recollection estimates with free recall performance (cf., Mandler, 1980), or to assume that recognition and free recall rely entirely on the same neurocognitive system.

APPENDIX

Using the Dual Process Signal Detection Model to Estimate the Contribution of Recollection and Familiarity to Recognition

Extensive reviews of the logic and equations underlying the dual process signal detec-

tion model of recognition can be found in Jacoby, Toth, and Yonelinas (1993); Yonelinas (1994); Yonelinas, Regehr, and Jacoby (1995); and Yonelinas, Dobbins, Szymanski, Dhaliwal, and King (1996). In this Appendix we provide a brief description, along with the equations necessary for the empirical curve fitting discussed in the paper.

The dual process signal detection model is based on the assumption that recognition performance is based on the independent contribution of recollection and familiarity. Familiarity reflects an equal-variance signal detection process of the type that underlies standard d' tables. Recollection, on the other hand, reflects the retrieval of qualitative information about the study event. Because the two processes are assumed to be independent, the probability of accepting an old item as old (i.e., a hit), is equal to the probability that the item is recollected (R), or the probability that the item is not recollected ($1 - R$), but exceeds the subject's response criterion for endorsing familiar items ($F_0 > c$):

$$P(\text{'old'} \mid \text{old}) = R + (1 - R)(F_0 > c) \quad (1)$$

Given that the subject adopts a certain familiarity criterion, there will be some new items that exceed the criterion and these items will be incorrectly accepted as old. Thus the probability of accepting a new item as old (i.e., a false alarm) will be equal to the probability that the familiarity of the new item exceeds response criterion:

$$P(\text{'old'} \mid \text{new}) = (F_n > c) \quad (2)$$

Because familiarity is assumed to reflect an equal variance signal detection process, the familiarity components can be replaced by formulae representing the portion of old and new item normal distributions which exceed the observer's criterion, ($F_0 > c$) = $\Phi(d'/2 - c)$ and ($F_n > c$) = $\Phi(-d'/2 - c)$ (see Macmillan & Creelman, 1991, for a detailed discussion of signal detection theory). Substituting these terms into Eqs. (1) and (2) and sub-

tracting Eq. (2) from Eq. (1) leads to the following equation:

$$P(\text{"old"} \mid \text{old}) = P(\text{"old"} \mid \text{new}) + R + (1 - R)\Phi(d'/2 - c) - \Phi(-d'/2 - c) \quad (3)$$

This equation represents the predicted relationship between hits and false alarms, and can be used to estimate recollection (R) and familiarity (d') by fitting it to an observed ROC. The logic is similar to that underlying linear regression. In a linear regression, a linear equation is fit to a set of data points by minimizing the sum of squared errors (SSe) between the predicted line and the observed data points. The analysis produces parameters representing the slope and intercept of the linear equation. In the current analysis, we used the dual process equation (Eq. (3)), rather than a linear one, and estimated R and d' rather than a slope or intercept. The equation was fit to the observed ROC using a gradient descent search algorithm which minimized the SSe between the predicted function and observed ROC data points. Similar estimates can be obtained using a maximum log likelihood method.

REFERENCES

- Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, **79**, 97–123.
- Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In R. C. Atkinson, R. D. Luce, D. H. Krantz, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: I. Learning, memory and thinking*. San Francisco, CA: W. H. Freeman.
- Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition*, **18**, 23–30.
- Gregg, V. (1976). Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183–216). Bath, England: Wiley.
- Guttentag, R. E., & Carroll, D. (1994). Identifying the basis for the word frequency effect in recognition memory. *Memory*, **2**, 255–273.
- Haist, F., Shimamura, A. P., & Squire, L. R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 691–702.
- Hintzman, D. L., Curran, T., & Oppy, B. (1992). Effects of similarity and repetition on memory: Registration without learning? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 667–680.
- Hirshman, E., & Palij, M. (1992). Rehearsal and the word frequency effect in recognition memory. *Journal of Memory & Language*, **31**, 477–484.
- Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory & Language*, **32**, 421–445.
- Hunt, R. R., & Smith, R. E. (1996). Accessing the particular from the general—The power of distinctiveness in the context of organization. *Memory & Cognition*, **24**, 217–225.
- Huppert, F. A., & Piercy, M. (1976). Recognition memory in amnesic patients: Effects of temporal context and familiarity of material. *Cortex*, **12**, 3–20.
- Huppert, F. A., & Piercy, M. (1978). The role of trace strength in recency and frequency judgements by amnesic and control subjects. *Quarterly Journal of Experimental Psychology*, **30**, 347–354.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513–541.
- Jacoby, L. L. (1994). Measuring recollection: Strategic versus automatic influences of associative context. In M. Moscovitch & C. Umiltà (Eds.), *Attention and performance 15: Conscious and nonconscious information processing. Attention and performance series* (pp. 661–679). Cambridge MA: MIT Press.
- Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, **122**, 139–154.
- Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 13–48). Mahwah, NJ: Erlbaum.
- Kinoshita, S. (1995). The word frequency effect in recognition memory versus repetition priming. *Memory & Cognition*, **23**, 569–580.
- Kinsbourne, M., & George, J. (1974). The mechanism of the word-frequency effect on recognition memory. *Journal of Verbal Learning & Verbal Behavior*, **13**, 63–69.
- Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, **123**, 297–315.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence: Brown Univ. Press.
- MacLeod, C. M., & Kampe, K. E. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of Experimental Psy-*

- chology: Learning, Memory, & Cognition*, **22**, 132–142.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, **87**, 252–271.
- Mandler, G., Goodman, G. O., & Wilkes-Gibbs, D. L. (1982). The word-frequency paradox in recognition. *Memory & Cognition*, **10**, 33–42.
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 13515–13522.
- Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Columbia University.
- Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, **26**, 1–12.
- Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin & Review*, **2**, 387–390.
- Tulving, E., Markowitsch, H. J., Craik, F. E., Habib, R., & Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cerebral Cortex*, **6**, 71–79.
- Tulving, E., Markowitsch, H. J., Kapur, S., Habib, R., & Houle, S. (1994). Novelty encoding networks in the human brain: positron emission tomography data. *Neuroreport*, **5**, 2525–2528.
- Winograd, E. (1968). List differentiation as a function of frequency and retention interval. *Journal of Experimental Psychology Monographs*, **76**, (2 pt 2).
- Wixted, J. T. (1992). Subjective memorability and the mirror effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 681–690.
- Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1341–1354.
- Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition*, **5**, 418–441.
- Yonelinas, A. P., Kroll, N. E. A., Dobbins, I., Lazzara, M., & Knight, R. T. (in press). Recollection and familiarity deficits in amnesia: Convergence of remember/know, process dissociation and ROC data. *Neuropsychology*.
- Yonelinas, A. P., Regehr, G., & Jacoby, L. L. (1995). Incorporating response bias in a dual-process theory of memory. *Journal of Memory & Language*, **34**, 821–835.

(Received March 19, 1997)

(Revision received November 12, 1997)