Word frequency and list composition effects in associative recognition and recall

STEVEN E. CLARK and RICHARD E. R. BURCHETT
University of California, Riverside, California

The effects of list composition and word frequency on cued recall, associative recognition, and item recognition were examined in three experiments. For pure-frequency lists, cued recall and associative recognition show better performance on common high-frequency (HF) words than on rare low-frequency (LF) words. Item recognition, however, shows an advantage for LF words. In mixed lists, consisting of half HF and half LF words, the HF advantage in cued recall disappeared; however, the word frequency effects in item and associative recognition were unchanged. These results are inconsistent with explanations based on differential attention or co-rehearsal of HF and LF words. However, the results are consistent with list strength results which show that recognition is insensitive to strength-based list composition, but that recall is sensitive to list composition.

The word frequency effect (WFE) is a standard in research on human memory. Common high-frequency (HF) words are recalled better than rare low-frequency (LF) words. However, LF words are recognized better than HF words (see Gregg, 1976, for a review). This interaction between word frequency and type of test has important implications for theories of memory, particularly with respect to the relationship between recall and recognition. Most current theories of memory assume that recall and recognition are in some way fundamentally different, an assumption motivated in large part by the WFE.

This paper is concerned specifically with how WFEs may be modified by the composition of the study list. The HF word advantage in free recall is reduced or eliminated (Duncan, 1974; Gregg, 1976; Gregg, Montgomery, & Castano, 1980), and sometimes reversed (May, Cuddy, & Norton, 1979), when HF and LF words are mixed together within a single study list. The effect of mixing is less well documented for recognition. A sizeable LF advantage is shown in item recognition for both pure- and mixed-frequency lists (Garton & Allen, 1968; Schulman, 1967; Shepard, 1967). However, the effect of list composition on item recognition has been directly examined in only two studies. Gregg (1970) found a larger LF advantage for mixed lists than for pure lists, and Dorfman and Glanzer (1988) have shown that the size of the LF advantage increases as the proportion of LF words in a mixed list decreases.

In the experiments to be reported, subjects studied triples of words (denoted *ABC*, *DEF*, *GHI*, etc., where each letter denotes a word). Item recognition requires subjects to distinguish list items (A) from nonlist items (X). Associative recognition requires subjects to distinguish between intact study triples (*ABC*) and rearranged triples (*ABF*) in which one of the words (in this case, *F*) comes from a different study triple.

Clark (1992) has suggested that associative recognition may be something of a hybrid task. It is like item recognition, in that the subject need not recover or generate words from memory but need only make an *old/new* judgment about particular test items. However, it is like cued recall in that the critical information is in the association between words presented together rather than in item-specific information. Several recent experiments (Clark, 1992; Clark, Hori, & Callan, 1993; Clark & Shiffrin, 1992) have shown that associative recognition is better for HF words than for LF words, contrary to the usual LF advantage shown for item recognition. The HF word advantage is found typically in recall tasks, which supports the idea that associative recognition and recall have common characteristics.

Hockley (1992) has found a similar, but less dramatic, interaction: an LF word advantage for item recognition, but no effect of word frequency on associative recognition. The reasons for the differences between Clark's and Hockley's results are not yet apparent. However, it is important that all these experiments have shown different effects of word frequency on item and associative recognition, and that they are consistent with the proposal that associative recognition may share properties of both recognition and recall.

Moreover, these results have important implications for accounts of the WFE, and for current models of memory—

in particular, the SAM model (Gillund & Shiffrin, 1984) and Murdock's (1982) theory of distributed associative memory (TODAM). We will discuss the implications for models in more detail later, but we wish to emphasize here that the account of the WFE given within the framework of SAM cannot produce an HF advantage for associative recognition *and* an LF advantage for item recognition with the same parameters (see Clark, 1992, Appendix). One way in which SAM (or any other model) might produce the HF advantage in associative recognition is to build a recall component into the recognition model. Associative recognition may be based on retrieval processes like those which underlie cued recall, rather than, or in addition to, retrieval processes which underlie item recognition (Humphreys, 1978; Mandler, 1980). Given this possible connection between cued recall and associative recognition, it is important to examine how list composition might modify performance in these two tasks.

Predictions

Before presenting the experiments, we will briefly discuss some predictions based on explanations of mixing effects that have been offered for free recall and item recognition.

Attentional hypothesis. By definition, LF words are more novel than HF words. Their novelty may be particularly salient when they are mixed together with HF words in a single study list (May & Tryk, 1970). Thus, LF words may be given more attention in mixed lists than in pure lists. This would boost performance for LF words in mixed lists, which is consistent with results showing attenuation of the HF advantage in free recall, and magnification of the LF advantage in recognition (Dorfman & Glanzer, 1988; Gregg, 1970).

The prediction for cued recall and associative recognition is straightforward. The increased attention given to LF words in mixed lists should improve performance for LF words across the board for any memory task. Thus, the HF advantage for cued recall and associative recognition would be diminished in a mixed list.

Co-rehearsal hypothesis. Gregg (1970) suggested that the attenuation of the HF advantage in free recall with mixed lists is due to joint rehearsal of HF and LF words. In many experiments, the study list is presented as a series of single words, and it is likely that subjects rehearse HF and LF words together, which, of course, cannot occur in pure lists.

It is precisely this mechanism that produces mixing effects for free recall in SAM. The model assumes that HF words are better *cues* than LF words; thus, in pure lists, HF words are recalled better than LF words. However, if HF and LF words have been rehearsed together in a mixed list, an HF cue will be as likely to access an LF word as it is to access an HF word in memory. For cued recall, the model produces an HF advantage in pure-frequency lists, again because HF cues are stronger than LF cues. However, in a mixed list, if the items within a group are of the same frequency, thus minimizing rehear-

sal between HF and LF words, there will be no change in the HF advantage relative to the pure list comparison.[1]

What about item recognition? SAM produces an LF word advantage by assuming that preexperimental connections between LF words are weaker than preexperimental connections between HF words. This allows easier discrimination of list words from nonlist words when LF words are tested than when HF words are tested. As in recall, recognition performance is assumed to be due to the frequency of the cues used to probe memory (i.e., the test items), *not* the items in memory. Thus, the cue-to-image strength between an HF cue and an HF word in memory is the same as the cue-to-image strength between an HF word and an LF word in memory, and likewise for LF cues. Thus, a word of a given frequency used as a cue will have the same overall familiarity, irrespective of the frequency of the words in memory. Consequently, list composition should have no effect on WFEs in item recognition.

Sensitivity to list composition. In another domain, it has been shown that recall and recognition are differentially sensitive to list composition. Specifically, for recall, memory performance for a given word in the list decreases as other words on the list are strengthened (by increased study time). Ratcliff, Clark, and Shiffrin (1990) have termed this the *list strength effect*. This effect is quite large for free recall; it is somewhat smaller for cued recall; and it does not occur for item or associative recognition (Hockley, 1992; Murnane & Shiffrin, 1991). Thus, free and cued recall are sensitive to the mixing of strong and weak items within a list, but recognition is insensitive to list composition (mixing of strong and weak items). Given this, one might expect that, in general, recall (both cued and free) is more sensitive to other items on the list than is recognition (both item and associative). One might therefore expect larger frequency-mixing effects for recall than for recognition.

SAM is able to account for the differential sensitivity of recall and recognition to list composition because the model assumes that quite different retrieval processes are involved. Specifically, recall is assumed to operate on the basis of a serial search, whereas recognition is based on global matching. Global matching does not require the retrieval of particular items, but involves access to the list as a whole. Independent of models, recall requires that information for specific items be assembled and output, whereas recognition does not. The processes that are used to gain access to and output specific items may be sensitive to list composition, but the processes used to gain access to the list as a whole are not.

Implications for Current Models

Later we will discuss how these issues are related to current mathematical models of memory, in particular SAM and TODAM. At the outset it should be noted that these hypotheses have differing statuses with respect to the models. The attentional hypothesis is something of a theoretical "tack-on" in that it can be applied to any the-

ory. It allows one to have different sets of parameters for mixed and pure lists, which allows the model to produce any pattern of results.

The joint rehearsal hypothesis is much less of a tack-on and is a natural consequence of fundamental properties of the SAM model. The demonstration of a mixing effect for cued recall or associative recognition would be problematic for SAM.

The differential sensitivity hypothesis is not driven by any particular model. In fact, most current models have trouble with the list strength results (see Shiffrin, Ratcliff, & Clark, 1990), although SAM can be modified to handle them. The pattern of word frequency results predicted on the basis of the differential sensitivity hypothesis does not favor any particular model.

The purpose of Experiment 1 was to test these hypotheses by directly examining WFEs in cued recall and item and associative recognition with pure- and mixed-frequency lists. We examined cued rather than free recall for several reasons: (1) A central aim of these experiments was to investigate the possible role of recall-like retrieval processes in associative recognition. If such processes do operate in associative recognition, they would likely behave in a manner similar to cued, rather than free, recall. For example, following the study of pairs *AB* and *CD*, a rearranged distractor (*AD*) may be correctly rejected by using one of the test items (*A*) to retrieve its original pair member (*B*) (Humphreys, 1978; Mandler, 1980). (2) Free recall involves various sequential factors, such as the use of recalled items as cues for subsequent recall attempts. These factors may be important in producing list composition effects. If so, list composition effects should not appear in cued recall. (3) In cued recall, both the cues and the items to be recalled are specified by the experimenter, whereas in free recall, the cues are relatively unknown, and the items to be recalled are more vaguely defined.

EXPERIMENT 1

Experiment 1 will be reported as three experiments, 1A, 1B, and 1C. The study materials for the experiments were identical. Experiment 1A tested with cued recall, Experiment 1B tested with associative recognition, and Experiment 1C tested with item recognition.

For cued recall, subjects were presented with one or two words from a triple as cues and were to type in the remaining word(s) from that study triple. For cued recall, the number of cues may be important in producing mixing effects. As cues are added, the cued recall task takes on properties of the associative recognition task. In particular, two cues may configure to provide higher order associative information *in the cue*, which may serve to focus retrieval on the third item in the triple. Such higher order associative information is not available when one cue is given. Thus, any associative information must be recalled, because none is given in the cue set. This added dependence on recall processes may increase the likeli-

hood of a mixing effect in cued recall. Also, free recall involves sequential processes in which information recalled may be used to recall additional items. This kind of sequential recall should not occur when two cues are used to recall one item, but it may come into play when one cue is used to recall two items. If sequential factors play a role in list composition effects, the number of cues used in cued recall may be important.

For associative recognition, subjects distinguished between intact and rearranged test triples; the rearranged triples consisted of two items from one study triple, and the third item from a different study triple. For item recognition, subjects distinguished between old and new single items. Although we are primarily concerned with differences in overall levels of performance, a secondary issue regards how performance differences are produced. Glanzer and Adams (1985, 1990) have noted that in recognition tasks that involve different classes of items, class *A* and class *B*, where performance is better for class *A* items, the performance advantage is shown with both old and new items. Old items are better recognized as old (higher hit rates), and new items are better recognized as new (lower false alarm rates). This symmetry of facilitation for hit and false alarm rates is termed the *mirror effect*.

Typically, the LF advantage in item recognition shows the mirror pattern. However, in our previous experiments with triples (Clark, 1992; Clark & Shiffrin, 1992), the LF item recognition advantage and the HF associative recognition advantage were produced primarily by differences in false alarm rates. The mechanisms underlying mirror effects are not completely understood, and our investigation of them here is quite exploratory.

Method

Subjects. Forty-nine subjects participated in Experiment 1A (cued recall), and 38 participated in Experiment 1B (associative recognition), in both cases as part of a requirement for introductory psychology courses. Thirty-two subjects participated in Experiment 1C (item recognition), either to fulfill the introductory psychology requirement, or for cash.

Materials and Procedure. A study-test procedure was used. The study list materials were identical for Experiments 1A, 1B, and 1C. All materials were selected from published word frequency norms (Thorndike & Lorge, 1944, and Kučera & Francis, 1967). HF words were defined as occurring at least 50 times per million words, and LF words as those with 4 or fewer occurrences per million in both norms.

The subjects were presented with three lists, a pure list of 40 HF word triples, a pure list of 40 LF word triples, and a mixed list of 20 HF triples and 20 LF triples. Here, the word triples will be denoted *ABC, DEF, GHI*, and so forth, where each letter denotes a word. Each word triple was presented alone on a computer screen for 7 sec in Experiments 1A and 1B, and for 4 sec in Experiment 1C (item recognition).[2] The subjects were fully instructed regarding the test procedures that would follow each list, and they were given a series of practice trials. All subjects were encouraged to remember the word triples by forming mental images, or creating sentences with the words in a triple.

Following each list was a mental arithmetic task that required the subjects to add singly presented digits for 30 sec. Mental arith-

metic was followed by cued recall (Experiment 1A), associative recognition (Experiment 1B), or item recognition (Experiment 1C). The details of each of these are described below.

Cued recall (Experiment 1A). The first and last two study triples were excluded from testing, and memory was tested for the middle 36 triples. On half the trials, one word from a study triple was presented, and the subject was required to type in the remaining two words. On the other trials, two words from a triple were presented, and the subject was required to type the remaining word. The positions of the cues were counterbalanced across position within the triple.

The cued recall test was computer controlled and self-paced. The order of test trials was random for each subject, as was the assignment of words to study and test conditions.

Associative recognition (Experiment 1B). Again the middle 36 triples were tested. Twelve intact targets (*ABC*) and 12 rearranged distractors (*ABF*) were tested in random order. The position of the "odd" word in rearranged triples was counterbalanced. Word position within a triple was maintained for both intact and rearranged triples; that is, test triples such as *ACB* and *AFB* were not tested. A given study triple contributed to only one test trial. Thus, if *ABC* was tested, *ABF* would not also be tested.

Item recognition (Experiment 1C). On each trial, either a list word (i.e., *A*) or a nonlist word (i.e., *X*) was presented for an *old/new* judgment. Eighteen list and 18 nonlist words were tested. The position of the test word (left, middle, right) was counterbalanced for both list and nonlist trials. The list items for the recognition test were randomly sampled from the middle 36 triples, with the constraint that each study triple contributed to only one test trial. Thus if *A* was tested, *B* and *C* would not be tested.

Apparatus. Stimulus presentation and response collection were controlled by ATARI 1040ST computers, running independently for each subject.

Design. Word frequency and list composition were manipulated within subjects for each experiment; in Experiment 1A, the number of cues for a given cued recall trial was also manipulated within subjects.

Results

Experiment 1A: Cued recall. Proportion of correct items recalled was calculated separately for one-cue and two-cue conditions, for HF and LF words in mixed and pure frequency lists. Recall proportions were calculated on the basis of the number of possible items for each test trial (thus, the denominator is twice as large for the two-cue condition as for the one-cue condition). These recall proportions, shown in Table 1, were subjected to a $2 \times 2 \times 2$ analysis of variance (ANOVA), with word frequency, list composition, and number of cues as factors.

Proportion of items recalled was better with two cues than with one [$F(1,43) = 50.41, MS_e = .01, p < .0001$]; overall, HF words were recalled better than LF words [$F(1,43) = 21.78, MS_e = .01, p < .0001$], and there was no main effect for list composition.

There were two significant interactions. Word frequency interacted with number of cues [$F(1,43) = 12.08, p <$

Table 1
Recall Probabilities for Experiment 1A

Word Frequency	Pure			Mixed		
	1 Cue	2 Cues	*M*	1 Cue	2 Cues	*M*
High	.12	.22	.17	.09	.19	.14
Low	.07	.10	.09	.09	.17	.13

.001]. This would indicate that the improvement in performance due to adding an additional cue was larger for HF words than for LF words. More critical to the current study was the finding of a significant interaction between list composition and word frequency [$F(1,43) = 12.48, p < .001$]. The breakdown of this interaction is straightforward. In pure-frequency lists, cued recall performance was better for HF words than for LF words, for both one-cue [$t(43) = 3.34, p < .01$] and two-cue [$t(43) = 6.14, p < .001$] conditions. However, there was no effect of word frequency in mixed lists, for one [$t(43) = .16$] or two [$t(43) = 1.21, p > .20$] cues.

The three-way interaction (number of cues, word frequency, and list composition) was not significant.

Experiment 1B: Associative recognition. Table 2 shows hit rates, false alarm rates, and recognition d' for HF and LF words in mixed- and pure-frequency lists. A 2×2 ANOVA computed on the d's for each condition[3] showed that HF triples were recognized better than LF triples [$F(1,37) = 9.74, MS_e = .55, p < .005$]. List composition showed no main effect and did not interact with word frequency (both $Fs < 1$).

For pure lists, the HF advantage showed a mirror effect: hit rates were higher [$t(37) = 1.88, p < .07$], and false alarm rates lower [$t(37) = 2.48, p < .02$], for HF words. It is somewhat puzzling that the hit rate advantage was not statistically reliable, even though the difference was larger than for false alarm rates. This was due to greater variability for hit rates. For mixed lists, hit rates did not differ for HF and LF words [$t(37) = .34$], but false alarm rates were lower for HF words [$t(37) = 2.69, p < .02$].

Experiment 1C: Item recognition. Hit and false alarm rates are presented in Table 3, along with recognition d' averaged across subjects. A 2×2 ANOVA performed on d' showed that LF words were recognized better than HF words [$F(1,25) = 21.89, MS_e = .45, p < .001$]. List composition did not produce a main effect, nor did it interact with word frequency (both $Fs < 1$).

In the pure lists, the LF advantage was produced primarily by lower false alarm rates [$t(31) = 2.79, p < .01$], while hit rates were only slightly higher for LF words [$t(31) = 1.07, p > .20$]. For mixed lists, hit rates were higher for LF words [$t(31) = 2.57, p < .02$], and false alarm rates were .06 lower; however, the false alarm rate difference was not statistically reliable [$t(31) = 1.41, p > .15$].

Discussion

The results of Experiments 1A, 1B, and 1C are straightforward. For cued recall, mixing HF and LF words into a single list eliminates the advantage for HF words shown in pure-frequency lists. This result is consistent with previous results for free recall (Duncan, 1974; Gregg, 1970, 1976; May et al., 1979). Associative recognition also showed an HF word advantage, replicating previous results (Clark, 1992; Clark et al., 1993; Clark & Shiffrin, 1992). However, the magnitude of the HF advantage was

Table 2
Hit Rate, False Alarm Rate, and d' for Experiment 1B

Word Frequency	List Type					
	Pure			Mixed		
	HR	FAR	d'	HR	FAR	d'
High	.70	.25	1.40	.66	.22	1.55
Low	.63	30	1.02	.65	.31	1.14
High–low	.07	−.05	0.38	.01	−.09	0.41

Note—HR, hit rate; FAR, false alarm rate.

Table 3
Hit Rate, False Alarm Rate, and d' for Experiment 1C

Word Frequency	List Type					
	Pure			Mixed		
	HR	FAR	d'	HR	FAR	d'
High	.72	.30	1.20	.64	.29	1.09
Low	.75	.22	1.64	.75	.23	1.68
High–low	−.03	.08	−0.46	−.11	.06	−0.59

Note—HR, hit rate; FAR, false alarm rate.

unaffected by list composition. For item recognition, the usual LF word advantage was shown, and it was unaffected by list composition.

That associative recognition shows an HF word advantage (like cued recall) and an insensitivity to list composition (like item recognition) is consistent with the idea that associative recognition is a hybrid task that contains components of both recall and recognition. Like cued recall, associative recognition shows an HF word advantage, but like item recognition, the WFE is unaffected by list composition.

One might argue that the differences in sensitivity to list composition were due to study differences, rather than to the operation of different retrieval processes at test. There is evidence that subjects alter their study strategies when they expect recall or item recognition tests, following lists of single words (Balota & Neely, 1980). However, there are two reasons to suspect that this is not the sole source of the differences among test conditions. First, all subjects were given the same instructions, to form interactive images, or connect all the words in a triple in a single sentence. Second, the task requirements, at least for cued recall and associative recognition, were similar in that performance depended on associative information.

The mirror effect analyses are difficult to interpret, because the sources of performance differences were quite inconsistent. Sometimes, clear mirror patterns were shown (associative recognition, pure lists), sometimes clearly not shown (associative recognition, mixed lists), and sometimes weakly shown (item recognition). One might argue for combining the data over list composition, since list composition did not produce main effects or interactions. Although it is not clear that collapsing in this way is warranted (which is why we did not do it), it does simplify the mirror effect issue by producing clear mirror effect patterns for both item and associative recognition. The LF advantage in item recognition is then given by .07 hit

rate and .07 false alarm rate differences, and the HF advantage is produced by .04 hit rate and .07 false alarm rate differences.

To return to the main issue, the overall results have important implications for accounts of the WFE and for accounts of mixing effects specifically, as well as important implications for models. However, before addressing these issues, we will present Experiments 2 and 3. In both experiments, list composition was manipulated between rather than within subjects. Experiment 2 is a replication of the cued recall results of Experiment 1A, and Experiment 3 is a replication of the associative recognition results of Experiment 1B.

EXPERIMENT 2

Experiment 2 was a replication and extension of Experiment 1A for cued recall. In Experiment 1A, recall performance for LF words was quite low, above .10 in only one condition. In Experiment 2, subjects studied word pairs, rather than triples. We reasoned that it would be easier for subjects to form and remember associations for pairs than for triples. List composition was varied between, rather than within, subjects.

Method

Subjects. Forty-nine different subjects from the same pool as in Experiment 1 participated, 24 in the mixed-list condition, and 25 in the pure-list condition.

Materials and Procedure. Subjects studied four lists of word pairs, each list consisting of 40 pairs, presented at a rate of 4 sec/pair. Subjects in the pure-list group studied two HF and two LF lists, in random order; subjects in the mixed-list group studied four mixed lists, each consisting of 20 HF and 20 LF pairs, randomly mixed. The HF and LF words were taken from the same pool as for Experiment 1.

Following each list, the subjects engaged in mental addition for 30 sec and then were given a cued recall test, testing the middle 36 test pairs from the preceding list. For 18 of the test trials, the left-hand word was presented as a cue to recall the right-hand word, and vice versa for the other 18 trials. The ordering of the test trials was randomly determined for each subject.

Results and Discussion

The proportions of items correctly recalled in the pure lists were .25 for HF words and .12 for LF words. In the mixed list, these proportions were .26 for HF words and .21 for LF words. A 2×2 ANOVA showed better recall for HF words than LF words [$F(1,47) = 25.42$, $MS_e = .01$, $p < .001$], no effect of list composition [$F(1,47) = 2.13, p > .15$], and a frequency × list composition interaction [$F(1,47) = 4.25$, $MS_e = .01$, $p < .05$]. The HF advantage was .13 for pure lists and .05 in the mixed lists. Separate analyses showed that the HF word advantage was significant for both pure [$t(24) = 4.77, p < .001$] and mixed [$t(23) = 2.25, p < .04$] lists.

The goal of Experiment 2 was to increase overall performance, while replicating the pattern of results of Experiment 1A. Overall performance was slightly higher, and the word frequency × list composition interaction was

replicated. The interaction was produced in similar fashion, as well: recall of LF words increased from .12 in pure lists to .21 in the mixed lists. Experiment 1A showed the same increase in performance for LF words.

Experiment 2 also extends the results of Experiment 1A by demonstrating the interaction with a between-subjects manipulation of list composition. Likewise, Experiment 3 was also designed to replicate the associative recognition results of Experiment 1B with a between-subjects manipulation of list composition.

EXPERIMENT 3

Method

Materials and Procedure. Subjects studied four lists each consisting of 40 word triples, presented at a rate of 4 sec/triple. For subjects in the mixed-list group, each list consisted of 20 HF and 20 LF word triples. For subjects in the pure-list group, two lists consisted of all HF words, and two lists consisted of all LF words. The test procedures were identical to those in Experiment 1B: following a 30-sec mental arithmetic task, the subjects were presented with an associative recognition test requiring discrimination between intact (ABC) and rearranged (ABF) triples.

Results and Discussion

Average hit rates, false alarm rates, and d' are shown in Table 4 for each test condition. The d' scores were submitted to a 2×2 ANOVA with word frequency and list composition as factors. Overall performance was better for subjects in the mixed-list group [$F(1,71) = 5.93$, $MS_e = .08$, $p < .05$], and HF words were recognized better than LF words [$F(1,71) = 50.87, MS_e = .02, p < .0001$]. The word frequency × list composition interaction was not significant.

Analyses of hit and false alarm rates showed that for pure lists, the HF advantage was produced by higher hit rates [$t(35) = 2.74, p < .01$] and lower false alarm rates [$t(35) = 3.58, p < .005$] for HF words. For mixed lists, however, the HF advantage was produced entirely by lower false alarm rates; hit rates did not differ significantly for HF and LF words ($t < 1$).

Thus, the results of Experiment 2, in which list composition was varied between subjects, replicated those of Experiment 1B, in which list composition was varied within subjects: mixing LF and HF words in each study list does not attenuate the advantage for HF words shown in associative recognition, relative to pure-frequency lists. In addition, the pattern for hit and false alarm rate differences also replicated those of Experiment 1B. The mir-

ror pattern was clearly shown for pure lists, but the HF advantage in mixed lists was due almost entirely to lower false alarm rates.

GENERAL DISCUSSION

The present results show that for pure-frequency lists, memory performance was better with HF words than with LF words in cued recall and associative recognition, but performance was better with LF words than with HF words in item recognition. When HF and LF words were mixed together at study, the HF advantage in cued recall was eliminated, but the HF word advantage in associative recognition and the LF word advantage in item recognition remained unchanged. The results are inconsistent with accounts of mixing effects based on differential attention or on co-rehearsal of HF and LF words.

The differential attention hypothesis proposes that LF words attract more attention in mixed lists than in pure lists of all LF words. This would decrease the HF word advantage in cued recall and associative recognition, and increase the LF word advantage in item recognition. Contrary to this prediction, the current results showed no effect of list composition on associative or item recognition. More complex versions of the attentional hypothesis may still be viable, but the simple version is clearly unsupported by the present results.

The co-rehearsal hypothesis proposes that the HF word advantage in recall is eliminated because HF and LF words are rehearsed together in a mixed list. Co-rehearsal is the backbone of the SAM model's account of mixing effects. The procedures used in the present experiments should minimize co-rehearsal of HF and LF words. Thus, without co-rehearsal, word frequency effects should be unchanged in mixed-frequency lists. The demonstration of a mixing effect for cued recall is therefore a problem for the SAM model account of the WFE.

We should note that previous studies (Dorfman & Glanzer, 1988; Gregg, 1970) have shown a mixed-list increase in the magnitude of the LF word advantage for item recognition. Both attentional and co-rehearsal accounts seem unlikely, however, to apply to Dorfman and Glanzer's Experiment 2, in which subjects performed a lexical decision task during study. It is hard to explain why the lexical decision task would involve more processing of LF words in a set of mixed trials than in a set of pure-frequency trials. It also seems unlikely that any co-rehearsal would occur under those conditions. In the present experiments, words were studied in triples. It seems reasonable to assume that very little co-rehearsal of HF and LF words would occur under these conditions. Additional research is necessary for these experiments to be reconciled.

The current list composition results mimic the pattern shown with list strength manipulations, in which strength is manipulated by increasing study time. Ratcliff et al. (1990) showed that recall is sensitive to the strength of *other* items on a list, but that item recognition and associative recognition (Murnane & Shiffrin, 1991) are not. The

Table 4
Hit Rate, False Alarm Rate, and d' for Experiment 3

Word Frequency	Pure HR	Pure FAR	Pure d'	Mixed HR	Mixed FAR	Mixed d'
High	.66	.27	1.13	.65	.15	1.57
Low	.59	.35	0.66	.64	.29	1.02
High−low	.07	−.08	0.47	.01	−.14	0.55

Note—HR, hit rate; FAR, false alarm rate.

present results show a similar pattern of sensitivity. Cued recall was sensitive to the nominal frequency of other items on the list, but item and associative recognition were not.

It is important to note, however, that although the present results mimic list strength results in terms of sensitivity to list composition, the patterns of the results are different in the two cases. The list strength effect shows that differences in recall performance due to study time differences are magnified in a mixed list as opposed to a pure-list comparison. However, the mixing of HF and LF words produced the opposite pattern. The advantage of the better recalled (strong) HF words is *smaller* in the mixed-list than in the pure-list comparison.

Subjects may partition HF and LF words, as they would items in different taxonomic categories. Evidence for this comes from a release from proactive interference when subjects are switched to words of a different nominal frequency (HF to LF or LF to HF) after several trials in a Brown-Peterson recall task (Swanson & Wickens, 1970). However, this partitioning hypothesis would predict that overall performance would be better with mixed lists than with pure lists, for both HF and LF words, a result that was not observed.

Shiffrin et al. (1990) argued that the differential sensitivity of recall and recognition to strength-based list composition differences is due to different underlying retrieval processes for recall and recognition. Our results lead to a similar conclusion: that item and associative recognition are both based on a global matching process. Thus, the fact that associative recognition and recall both show better performance with HF words does not necessarily mean that recall processes are operating in associative recognition.

The important difference between item and associative recognition seems to be due to differences in the kinds of information that are relevant to the two tasks, rather than to differences in retrieval processes. Many theorists distinguish between two fundamentally different kinds of information that underlie the kinds of tasks used here. Item-specific information indicates that a particular event has occurred, whereas associative information indicates that two events have co-occurred (see Clark & Shiffrin, 1992; Humphreys, 1976, 1978; Hunt & Einstein, 1981; Murdock, 1974, 1982).

Item and associative recognition performance are based on different kinds of information: discrimination of list words from nonlist words (item recognition) is based on item-specific information, whereas associative recognition is based on the association formed between words at study. That item and associative recognition show different patterns of results for word frequency suggests that the strength of item and associative information may vary independently.

The independence of item and associative information is a fundamental property of TODAM. Associations are stored as higher order units, which can be accessed independently of the item-level units that they comprise. In this respect, TODAM differs from all of the other global

matching models, including SAM, MINERVA 2 (Hintzman, 1988), and the Matrix model (Pike, 1984). SAM can be modified to include higher order associative units as well (see Shiffrin, Murnane, Gronlund, & Roth, 1988), but the independence of item-specific and associative information is attained by assumption, rather than by design.

The present results are consistent with the proposal that item-specific and associative information are separate, and at least to some degree independent, which is consistent with TODAM. Our results did not provide any evidence that the HF advantage in associative recognition was due to recall-like retrieval processes. Using different procedures, others have come to similar conclusions regarding the operation of recall processes in associative recognition (see Dyne, Humphreys, Bain, & Pike, 1990; Gronlund & Ratcliff, 1989). We do not rule out the possibility that they operate at some level; however, firm evidence to support this proposal has not been given by the present results.

REFERENCES

Balota, D. A., & Neely, J. H. (1980). Test-expectancy and word-frequency effects in recall and recognition. *Journal of Experimental Psychology: Human Learning & Memory*, 6, 576-587.

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

Clark, S. E., Hori, A., & Callan, D. E. (1993). Forced-choice associative recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 19, 871-881.

Clark, S. E., & Shiffrin, R. M. (1992). Cuing effects and associative information in recognition memory. *Memory & Cognition*, 20, 580-598.

Dorfman, D., & Glanzer, M. (1988). List composition effects in lexical decision and recognition memory. *Journal of Memory & Language*, 27, 633-648.

Duncan, C. P. (1974). Retrieval of low-frequency words from mixed lists. *Bulletin of the Psychonomic Society*, 4, 137-138.

Dyne, A. M., Humphreys, M. S., Bain, J. D., & Pike, R. (1990). Associative interference effects in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16, 813-824.

Garton, R. F., & Allen, L. R. (1968). Familiarity and word recognition. *Quarterly Journal of Experimental Psychology*, 20, 385-389.

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

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, 13, 8-20.

Glanzer, M., & Adams, J.K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16, 5-16.

Gregg, V. H. (1970). *Word frequency effects in human memory*. Unpublished doctoral dissertation, Birbeck College, University of London.

Gregg, V. H. (1976). Word frequency, recognition, and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183-216). London: Wiley.

Gregg, V. H., Montgomery, D. C., & Castano, D. (1980). Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning & Verbal Behavior*, 19, 240-245.

Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 846-858.

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

HOCKLEY, W. E. (1992, November). *Tests of the mirror and list strength effects for item and associative recognition.* Paper presented at the 33rd Annual Meeting of the Psychonomic Society, St. Louis.

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 & Verbal Behavior,* **17,** 175-187.

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

KUČERA, H., & FRANCIS, W. N. (1967). *A computational analysis of present-day American English.* Providence, RI: Brown University Press.

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

MAY, R. B., CUDDY, L. J., & NORTON, J. M. (1979). Temporal contrast and the word frequency effect. *Canadian Journal of Psychology,* **33,** 141-147.

MAY, R. B., & TRYK, H. E. (1970). Word sequence, word frequency, and free recall. *Canadian Journal of Psychology,* **24,** 299-304.

MURDOCK, B. B., JR. (1974). *Human memory: Theory and data.* Hillsdale, NJ: Erlbaum.

MURDOCK, B. B., JR. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review,* **89,** 609-627.

MURDOCK, B. B., JR., & OGILVIE, J. C. (1968). Binomial variability in short-term memory. *Psychological Bulletin,* **70,** 256-260.

MURNANE, K., & SHIFFRIN, R. M. (1991). Word repetitions in sentence recognition. *Memory & Cognition,* **19,** 119-130.

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

RATCLIFF, R., CLARK, S. E., & SHIFFRIN, R. M. (1990). The list-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **16,** 163-178.

SCHULMAN, A. I. (1967). Word length and rarity in recognition memory. *Psychonomic Science,* **9,** 211-212.

SHEPARD, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning & Verbal Behavior,* **6,** 156-163.

SHIFFRIN, R. M., MURNANE, K., GRONLUND, S. D., & ROTH, M. (1988). On units of storage and retrieval. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane-Floweree Symposium on Cognition* (pp. 25-68). Hillsdale, NJ: Erlbaum.

SHIFFRIN, R. M., RATCLIFF, R., & CLARK, S. (1990). The list-strength effect: II. Theoretical mechanisms. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **16,** 179-195.

SWANSON, J. M., & WICKENS, D. D. (1970). Preprocessing on the basis of frequency of occurrence. *Quarterly Journal of Experimental Psychology,* **22,** 378-383.

THORNDIKE, E. L., & LORGE, I. (1944). *The teacher's wordbook of 30,000 words.* New York: Columbia University, Teachers College Press.

NOTES

1. The SAM model predictions for cued recall may seem to contradict Gillund and Shiffrin's (1984) earlier description of the model. They fit SAM to cued recall results from a mixed-list word frequency experiment. The good fit of the model may incorrectly be taken to indicate that the model is able to account for mixing effects. However, to demonstrate that the model can account for such list composition differences, it must be shown to simultaneously fit data from both mixed- and pure-frequency lists. In Gillund and Shiffrin's experiment, to which SAM was fit, there was no pure-list condition.

2. The differing presentation rates were due to pilot results, which showed floor effects in cued recall for short presentation times, and ceiling effects in item recognition for long presentation times.

3. For all the experiments, the data were analyzed using both d' and hit rate minus false alarm rate as dependent variables. Because response probabilities of zero and one are undefined in d' analyses, these probabilities must be estimated. We used the estimation procedure used by Murdock and Ogilvie (1968). The pattern of the results was the same for d' and hit rate minus false alarm rate analyses, and thus only the d' analyses are reported.