Reflections of the mirror effect for item and associative recognition

WILLIAM E. HOCKLEY
Wilfrid Laurier University, Waterloo, Ontario, Canada

In five experiments, participants studied pairs of words and yes/no recognition memory for both item and associative information was tested. Two stimulus manipulations, nouns versus nonnouns and high versus low word concreteness, produced the mirror effect for both item and associative recognition. The mirror effect was reflected in both measures of accuracy and response latency. A word frequency manipulation, however, produced the mirror effect only for item recognition. Two additional experiments showed that the mirror effect could also be obtained between nouns and nonnouns and between high and low concrete words for associative recognition in a forced-choice recognition procedure. The results extend the generality of the mirror effect to measures of response latency and to associative recognition and also suggest that similar retrieval and decision processes underlie recognition of item and associative information.

The mirror effect has been shown to be an important feature of item recognition (Glanzer & Adams, 1985, 1990; Glanzer, Adams, & Iverson, 1991; Glanzer, Adams, Iverson, & Kim, 1993). The purpose of the present study was to compare and contrast the mirror effect for item and associative recognition memory.

Item information provides the basis for remembering that a particular event has previously been experienced, whereas associative information represents relationships between items or events (Humphreys, 1976; Murdock, 1974). Tests of item recognition involve discriminations between old (studied) and new (nonstudied) events. Associative recognition requires subjects to discriminate intact or studied pairs from rearranged or new pairs (see, e.g., Clark, 1992; Hockley, 1991, 1992; Humphreys, 1976, 1978; Murdock & Hockley, 1989). Rearranged test pairs consist of two previously studied words that have been presented in different study pairs. Thus, the discrimination of intact pairs from rearranged pairs cannot be based on the familiarity of the individual items in the test pairs but, rather, must be based on memory for the associations formed between items during the study phase.

Several findings provide strong empirical support for the theoretical distinction between these two types of information in human memory. For example, processes at encoding can differentially emphasize item or associa-

This work was supported by Natural Sciences and Engineering Research Council of Canada Grant OGPIN012. Portions of this study were presented at the XXV International Congress of Psychology, Brussels, in July 1992. I thank Murray Glanzer for providing the high/low word frequency pool; Deborah Connolly, Lyla Caudle, and Carolina Cristi for data collection; and Steve Clark, Michael Humphreys, Margaret Intons-Peterson, and Doug Hintzman for helpful comments on earlier versions of this paper. Correspondence can be addressed to the author, Department of Psychology, Wilfrid Laurier University, Waterloo, ON, Canada N2L 3C5 (e-mail: whockley@mach1.wlu.ca).

tive information (see, e.g., Bain & Humphreys, 1988; McGee, 1980). Dosher (1988) and Gronlund and Ratcliff (1989) examined the temporal dynamics of item and associative recognition using a response-signal procedure and found that the retrieval of item and associative information have different time courses. Finally, Hockley (1991, 1992) showed that item and associative recognition memory either have different rates of forgetting or differ in their susceptibility to the effects of interference.

The preceding studies demonstrate differences in the encoding, storage, and retrieval of item and associative information. In the present study, the mirror effect for item and associative recognition was examined in order to evaluate similarities or differences in the decision processes that underlie these two types of recognition memory.

The mirror effect can be characterized as follows: When there are two classes of stimuli, A and B, where class A is more accurately recognized than class B, then class A is more accurately recognized as old when old and class A is more accurately recognized as new when new. In other words, performance on new items mirrors performance on old items. Glanzer and Adams (1985), in a meta-analysis of 80 recognition experiments, showed that the mirror effect can be found for all types of recognition test procedures and a wide variety of stimulus comparisons. Glanzer and Adams (1990) and Glanzer et al. (1991) further extended the generality of the mirror effect and argued that the mirror effect has important implications for theories of the decision process underlying recognition memory.

Yet although it has been shown that the mirror effect is an important characteristic of item recognition, the effect has not been evaluated for associative recognition. An examination of the mirror effect for item and associative recognition should provide an additional basis for comparison of these two types of recognition memory, as well as a further test of the generality of the mirror effect itself.

Two different accounts of associative recognition have been proposed. According to one view, the discrimination of intact pairs from rearranged pairs is based on a recall process (e.g., Anderson, 1974; Clark, 1992; Humphreys, 1978; Mandler, 1980). That is, the discrimination between an intact test pair A–B and a rearranged pair test pair A–C would be made by recalling the item that was paired with A at study. Thus, as Clark (1992) has recently argued, there may be a strong similarity between recall and associative recognition. In this view, then, it is assumed that item and associative recognition responses are based on different retrieval and decision processes.

Global matching models of memory (e.g., the SAM model, Raaijmakers & Shiffrin, 1981; Gillund & Shiffrin, 1984; MINERVA 2, Hintzman, 1984; the Matrix model, Pike, 1984; and TODAM, Murdock, 1982; see Humphreys, Pike, Bain, & Tehan, 1989, for a comparison of the application of these models to associative recognition) provide an account of both item and associative recognition that does not include a recall process. In this class of models, both item and pair recognition are based on the strength of the match between the information encoded from the test probe and information represented in memory. Thus, the approach of global matching models is to assume that the same retrieval and decision processes underlie both item and associative recognition performance. Current versions of global matching models, however, do not predict a mirror effect. Moreover, Hintzman (1990) noted that the mirror effect "turns out to be a difficult phenomenon for global matching models to explain" (p. 119).

In the first five experiments of the present study, the mirror effect was examined with three different stimulus manipulations—nouns versus nonnouns, high versus low natural language word frequency, and high versus low word concreteness. In each experiment a study—test procedure was employed. Study lists consisted of random pairs of words wherein both words were from the same stimulus class. Test lists included both old and new single-word test probes (item recognition tests) and intact (old) and rearranged (new) test pairs (associative recognition tests). The rearranged test pairs consisted of two words (of the same stimulus class) that had been presented in two separate study pairs.

In addition to the examination of the mirror effect for item and associative recognition memory, a second purpose of the present study was to further evaluate the extent to which the mirror effect is reflected in the latency of the recognition decisions. Only a few studies that allow an assessment of the mirror effect have included the measurement of response latency. Hockley (1982) found a positive relationship between speed and accuracy in a comparison of item recognition for nouns and nonnouns; that is, the mirror effect was observed for both

measures of accuracy and response time. In contrast, Glanzer and Adams (1990, Experiments 1 and 3) found no relationship between speed and accuracy in their examination of the mirror effect for item recognition.

EXPERIMENTS 1A-1B Nouns Versus Nonnouns

Hockley (1982) compared nouns and nonnouns in a continuous yes/no item recognition paradigm. As discussed by Hockley and Murdock (1987), these results showed the mirror effect both in the analysis of hit and false alarm rates and in the analysis of the latency of correct responses. Murdock and Hockley (1989, Experiment 2) included a noun/nonnoun manipulation in their study of associative recognition performance and found that, overall, noun pairs were recognized more accurately than nonnoun pairs. These studies thus demonstrate that both item and associative recognition differ for nouns and nonnouns. Therefore, this stimulus manipulation should provide a good basis for comparing the mirror effect for item and associative recognition.

Method

Design. A $2\times2\times2$ within-subjects factorial design was used for each experiment. The three factors were type of stimuli, type of test (single item vs. pair recognition), and type of test probe (old vs. new).

Subjects. In Experiment 1A, 6 participants individually completed six 1-h sessions. The sessions were done on separate days, and the participants were paid. The participants in Experiment 1B and Experiment 2B were 66 students enrolled in an undergraduate cognitive laboratory course. Thirty-four students individually completed one session of Experiment 1B and 32 students completed one session of Experiment 2B as a class exercise. The students were not informed about the stimulus manipulations or about the mirror effect prior to testing.

Apparatus and Stimuli. List construction, presentation, and response recording were controlled by IBM-compatible personal computers in all experiments. The subjects used the "Z" key on the left side of the keyboard for "new" responses and the "/" key on the right side for "old" responses. The keyboards were fitted with opaque plastic covers that exposed only the response keys.

All words were selected from a 1,000-word version of the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982) consisting of common two-syllable words divided between nouns and nonnouns. The nonnouns consisted of adjectives, verbs, and adverbs (e.g., *ugly*, *discuss*, *greatly*). According to the norms provided by Friendly et al., which rate words on scales from 1 to 7, the nouns had a higher mean value of imagery (4.93 vs. 3.56) and concreteness (5.09 vs. 3.49) than did the nonnouns. The mean word length was similar for nouns (6.11) and nonnouns (6.07). The nonnouns had a higher mean natural language frequency (86.19 per million) than nouns (62.63), according to the Kučera and Francis (1967) counts.

Procedure. There were 10 study-test trials in every session. The participants initiated every trial. The study lists consisted of 32 critical pairs and 2 recency and 2 primacy buffer pairs, for a total list length of 36 pairs. The study pairs were constructed by pairing two nouns or two nonnouns selected at random without replacement during that session. A different random sequence was used for every experimental session. Order of presentation was partially random. There were 4 noun pairs and 4 nonnoun pairs in

each quarter of the study list (excluding primacy and buffer pairs). The study pairs were presented at the center of the monitor for 3 sec, with a 1-sec blank interval between presentations. A cue (?NEW/OLD?) presented for 3 sec signaled the end of the study presentation and was followed by the first test probe.

The test list consisted of a total of 32 presentations: 8 single old words, 8 single new words, 8 intact or same pairs, and 8 rearranged or new pairs. Half of each type of test probe were nouns and half were nonnouns. One exemplar of each type of test probe was obtained from each quarter of the study list (excluding buffer pairs). For single old test items, the right or left member of a study pair was randomly selected. New single-item test probes were randomly selected from the word pool. No word appeared in more than one trial in a given session.

The associative recognition intact (old) test pairs were presented in the same left/right order as were the study presentations. Rearranged (new) test pairs were formed by pairing the left member of one study pair with the right member of the immediately preceding study pair or the right member of the immediately following study pair. The particular combination was determined randomly for each test probe. Half the rearranged test pairs were composed of two nouns, and half were composed of two nonnouns. Rearranged test pairs were formed by pairing words from adjacent study presentations in order to equate as nearly as possible the relative recency of each old item in the rearranged pairs.

The order of the test presentations was random. The presentation rate was subject-paced, with a 1-sec blank interval between a response and the subsequent probe. Response latency was measured from the onset of the test probe to that of the response.

All subjects were given the same instructions. They were asked to indicate whether each single-item test probe and each pair test probe was new or old by pressing the appropriately labeled response key. They were instructed to learn the pairs by forming interactive visual images, or by generating sentences that included the two words from the study pair. The subjects were also encouraged to respond both as accurately and as quickly as possible, but the emphasis was placed on accuracy.

Results and Discussion

The .05 level of reliability was adopted for all statistical comparisons.

Experiment 1A. For item recognition, mean d' (based on subject means) was 1.53 for nouns and 1.15 for nonnouns. This difference was significant [t(5) = 3.31]. For pair recognition, the corresponding means were 1.78 and 1.02, and this difference was also reliable [t(5) = 4.50].

The proportion of "yes" responses (false alarms for new items and hits for old items) for nouns and nonnouns for both item and pair recognition tests are given in Table 1. The proportion of "yes" responses are arranged in Table 1 so that the mirror effect is evidenced by an increase in the means from top to bottom.

A 2×2 (stimulus type × test type) analysis of variance (ANOVA) was computed separately for hits and for false alarms. For hits, the main effect of stimulus type was highly significant $[F(1,5) = 22.75, MS_e = .002]$. The main effect of test type [F(1,5) < 1] and the stimulus type × test type interaction [F(1,5) < 1] did not approach significance. The same statistical outcome was obtained for false alarms—the main effect of stimulus type was highly reliable $[F(1,5) = 22.31, MS_e = .001]$, whereas the main effect of test type [F(1,5) < 1] and the stimulus × test type interaction $[F(1,5) = 2.39, MS_e = .003]$ were not significant.

Table 1
Mean Proportion of "Yes" Responses, Mean
Correct Response Times (RT, in Milliseconds), and
Standard Errors for Experiment 1A

| | It | em Re | cognitio | n | Associative Recognition | | | | |
|------------|------------|-------|----------|-----|-------------------------|-----|-------|-----|--|
| | Proportion | | RT | | Proportion | | RT | | |
| Test Probe | M | SE | М | SE | M | SE | M | SE | |
| New | | | | | | | | | |
| Nouns | .203 | .06 | 1,363 | 106 | .190 | .07 | 2,058 | 96 | |
| Nonnouns | .229 | .06 | 1,433 | 113 | .286 | .06 | 2,419 | 169 | |
| Old | | | | | | | | | |
| Nonnouns | .618 | .05 | 1,479 | 126 | .650 | .06 | 1,957 | 140 | |
| Nouns | .700 | .04 | 1,383 | 138 | .736 | .06 | 1,670 | 105 | |

To assess the reliability of the mirror effect for item and associative recognition, separate comparisons were made. One-tailed t tests were used for these planned comparisons. For item recognition, the difference in hit rates between nouns and nonnouns was reliable [t(5) = 2.76], but the difference in false alarm rates did not reach significance [t(5) = 1.27]. For associative recognition, the differences in both hit rates [t(5) = 4.08] and false alarm rates [t(5) = 3.11] were statistically significant.

Mean response latencies for correct rejections and hits are also shown in Table 1. A 2 (stimulus type) \times 2 (old vs. new test probes) ANOVA was computed separately for item and pair recognition tests. (Response latency was not compared between item and pair recognition tests because, since the number of items in each probe set differed, it was expected on the basis of reading time alone that latencies for pair tests would be slower than latencies for item tests.) The pattern of the mean latencies for item recognition was consistent with the mirror effect, but the main effect of type of stimulus $[F(1,5) = 1.35, MS_e = 30,540]$, the difference between positive and negative test probes [F(1,5) < 1], and the interaction between these variables [F(1,5) < 1] did not approach statistical significance.

For pair recognition, the difference in mean response latency between nouns and nonnouns $[F(1,5) = 12.69, MS_e = 49,530]$ and the difference between positive and negative test probes $[F(1,5) = 19.33, MS_e = 56,073]$ were both reliable. The stimulus by probe type interaction did not approach significance [F(1,5) < 1]. The mirror effect was reflected in the latencies of the correct responses—the difference in mean response time between nouns and nonnouns was reliable for both hits [t(5) = 4.68] and correct rejections [t(5) = 2.45].

Experiment 1B. For item recognition, mean d' was 2.43 for nouns and 1.84 for nonnouns. The difference was reliable [t(33) = 3.32]. For associative recognition, mean d' was 2.57 for noun pairs and 1.60 for nonnoun pairs, and this difference was also statistically significant [t(33) = 6.95].

Mean hit and false alarm rates and the mean latencies of correct responses for each stimulus and test type are presented in Table 2. For item recognition, the hit rate for nouns was significantly greater than the hit rate for non-

Table 2
Mean Proportion of "Yes" Responses, Mean Correct
Response Latencies (RT, in Milliseconds), and
Standard Errors for Experiment 1B

| | lt | em Re | cognitio | n | Associative Recognition | | | | |
|------------|------------|-------|----------|----|-------------------------|-----|-------|-----|--|
| | Proportion | | RT | | Proportion | | RT | | |
| Test Probe | M | SE | М | SE | М | SE | М | SE | |
| New | | | | | | | | | |
| Nouns | .118 | .02 | 1,470 | 68 | .140 | .02 | 2,565 | 157 | |
| Nonnouns | .150 | .02 | 1,544 | 74 | .278 | .03 | 2,801 | 145 | |
| Old | | | | | | | | | |
| Nonnouns | .724 | .02 | 1,571 | 85 | .795 | .02 | 2,066 | 113 | |
| Nouns | .795 | .02 | 1,368 | 71 | .848 | .02 | 1,655 | 80 | |

nouns [t(33) = 4.27] and the false alarm rate for nouns was reliably lower than the false alarm rate for nonnouns [t(33) = 2.13]. The same pattern was obtained for associative recognition—the hit rate was reliably greater [t(33) = 4.18], and the false alarm rate was significantly lower [t(33) = 7.07], for noun pairs compared to nonnoun pairs.

The latency of correct responses largely mirrored the pattern of hit and false alarm rates. For item recognition, the mean latency of hits was significantly faster for nouns than for nonnouns [t(33) = 2.94]. The mean latency of correct rejections was also faster for nouns; this difference approached significance [t(33) = 1.63, p = .06]. For associative recognition, the mean latencies of hits [t(33) = 5.64] and correct rejections [t(33) = 2.52] were both reliably faster for noun pairs than for nonnoun pairs.

The pattern of results of Experiments 1A and 1B showed that a noun versus nonnoun stimulus manipulation produces the mirror effect for both item and associative recognition. Furthermore, the mirror effect was observed in both the analysis of accuracy and the analysis of the latency of correct responses.

In the following experiments, different stimulus comparisons were made in order to examine the generality of the mirror effect for associative recognition and for response latency. Natural language word frequency was manipulated in Experiments 2A and 2B. Glanzer and Adams (1990) and Glanzer et al. (1991) have demonstrated that this stimulus variable reliably produces the mirror effect in item recognition. Furthermore, Clark (1992) and Clark and Shiffrin (1992) have found that word frequency produces opposite effects for item and associative recognition.

EXPERIMENTS 2A-2B Word Frequency

It is a well-established finding that low-frequency words are recognized more accurately than high-frequency words, whereas high-frequency words are generally recalled better than low-frequency words (see Gillund & Shiffrin, 1984, for a review). The effect of word frequency on associative recognition has not been determined as clearly. Gronlund and Ratcliff (1989, Ex-

periment 3) and Hockley (1991, Experiment 2) failed to find an effect of word frequency on associative recognition. In contrast, Clark (1992) and Clark and Shiffrin (1992) reported a high-frequency advantage for associative recognition.

Gronlund and Ratcliff (1989) and Hockley (1991) manipulated word frequency within lists, whereas Clark (1992) and Clark and Shiffrin (1992) used a between-list comparison. The advantage for low-frequency words in tests of item recognition is usually obtained in both within-list and between-list comparisons, but the advantage for high-frequency words in tests of item recall is typically only observed in between-list comparisons (cf. Gillund & Shiffrin, 1984). Clark (1992) has taken his finding of a high-frequency advantage in associative recognition as support for the view that associative recognition involves a recall process. If this interpretation is correct, it follows that the effects of word frequency on associative recognition would largely be found in between-list, rather than within-list, comparisons. In Experiments 2A and 2B, the effects of a between-list word frequency manipulation were evaluated for both item and associative recognition memory.

Method

Subjects. In Experiment 2A, 6 participants individually completed 10 experimental sessions on separate days. In Experiment 2B, 32 participants (described in Experiment 1B) individually completed one experimental session.

Stimuli. All words were randomly selected without replacement for each subject session from Glanzer and Adams's (1990) pool of 248 high-frequency words (mean frequency of 177.3 per million, according to Kučera & Francis, 1967) and 248 low-frequency words (mean Kučera-Francis frequency of 11.2).

Procedure. There were six study-test trials per session. The construction of the study and test lists was identical to that of Experiments 1A and 1B, with the exception of the stimulus manipulation. In Experiments 2A and 2B, half the lists were composed wholly of high-frequency word pairs and half were composed of low-frequency word pairs. The order of lists was random in each session.

Results and Discussion

Experiment 2A. For item recognition, mean d' was 1.73 for low-frequency words and 1.32 for high-frequency words. This difference was reliable [t(5) = 5.58], and it constitutes a replication of the item recognition advantage for low-frequency words. Mean d' for associative recognition was 1.46 for low-frequency word pairs and 1.40 for high-frequency word pairs. This difference was not significant [t(5) = 0.48].

The hit and false alarm rates for low- and high-frequency item and pair recognition tests are presented in Table 3. For item recognition the hit rate was significantly greater for low- than for high-frequency words [t(5) = 2.12], and the false alarm rate was reliably lower for low- than for high-frequency words [t(5) = 3.62]. In contrast, for associative recognition there was no reliable difference in the hit rate [t(5) = 0.53] or in the false alarm rate [t(5) = 0.99] between low- and high-frequency word pairs.

Table 3
Mean Proportion of "Yes" Responses, Mean Correct
Response Times (RT, in Milliseconds), and Standard Errors
for High- and Low-Frequency Tests for Experiment 2A

| Test | It | em Re | cognition | | Associative Recognition | | | | |
|-------|------------|-------|-----------|----|-------------------------|-----|-------|-----|--|
| | Proportion | | RT | | Proportion | | RT | | |
| Probe | M | SE | M | SE | М | SE | M | SE | |
| New | | | | | | | | | |
| Low | .203 | .07 | 1,005 | 45 | .246 | .07 | 1,674 | 152 | |
| High | .265 | .08 | 1,133 | 79 | .230 | .06 | 1,708 | 144 | |
| Old | | | | | | | | | |
| High | .726 | .04 | 1,051 | 95 | .699 | .05 | 1,330 | 68 | |
| Low | .780 | .04 | 913 | 56 | .710 | .06 | 1,275 | 50 | |

The mean response latencies for hits and correct rejections for each stimulus and test type are also presented in Table 3. For item recognition the mean latency of hits [t(5) = 2.52] and the mean latency of correct rejections [t(5) = 2.38] were both significantly faster for low-than for high-frequency words. For associative recognition, neither the mean latency of hits nor the mean latency of correct rejections differed reliably between low- and high-frequency word pairs.

Experiment 2B. The results of Experiment 2B replicated the pattern of findings found in Experiment 2A. For item recognition, mean d' was 2.81 for low-frequency words and 2.10 for high-frequency words, and the advantage for low-frequency words was reliable [t(31) = 3.40]. The corresponding means for associative recognition were 2.57 and 2.74. These means did not differ statistically [t(31) = 0.93].

The hit and false alarm rates and mean latencies of correct responses are presented in Table 4. For item recognition, the hit rate was significantly greater [t(31) = 2.19], and the false alarm rate was significantly lower [t(31) = 3.30], for low-frequency words, and the mean latency of hits [t(31) = 2.23] and correct rejections [t(31) = 2.24] were significantly faster for low-frequency words than for high-frequency words. For associative recognition, there were no reliable differences in hit rates [t(31) = 0.92], false alarm rates [t(31) = 1.15], mean latency of hits [t(31) = 1.25], or mean latency of correct rejections [t(31) = 0.08] between low- and high-frequency word pairs.

Table 4
Mean Proportion of "Yes" Responses, Mean Correct
Response Latencies (RT, in Milliseconds), and Standard
Errors for High- and Low-Frequency Tests
for Experiment 2B

| | It | em Re | cognition | 1 | Associative Recognition | | | | |
|-------|------------|-------|-----------|-----|-------------------------|-----|-------|-----|--|
| Test | Proportion | | RT | | Proportion | | RT | | |
| Probe | M | SE | M | SE | M | SE | М | SE | |
| New | | | | | | | | | |
| Low | .084 | .01 | 1,476 | 83 | .133 | .02 | 2,610 | 166 | |
| High | .157 | .02 | 1,772 | 186 | .115 | .03 | 2,600 | 201 | |
| Old | | | | | | | | | |
| High | .794 | .03 | 1,426 | 78 | .839 | .02 | 1,731 | 93 | |
| Low | .839 | .02 | 1,301 | 66 | .852 | .02 | 1,852 | 111 | |

The word frequency manipulation used in Experiments 2A and 2B produced the mirror effect for item recognition, both in measures of accuracy and response latency. Thus, these results replicate those reported by Glanzer et al. (1991), and they provide another demonstration that the mirror effect is also reflected in the latency of correct responses.

In contrast, word frequency had no discernible effect on associative recognition. Although this result is consistent with the findings of Gronlund and Ratcliff (1989, Experiment 3) and Hockley (1991, Experiment 2), it constitutes a failure to replicate the between-list high-frequency advantage in associative recognition reported by Clark (1992; Clark & Shiffrin, 1992). The reason for this discrepancy is not clear. It is possible that Clark used a more extreme manipulation of word frequency than that used in the present study. If so, the fact that a significant effect of word frequency was still found for item recognition would suggest that item recognition is more sensitive to the effects of word frequency than is associative recognition.

Experiment 3 was designed to evaluate the generality of the mirror effect for associative recognition further, by using a third stimulus manipulation.

EXPERIMENT 3Word Concreteness

Method

Subjects. Thirty introductory psychology students completed one experimental session for course credit.

Stimuli. The word pool consisted of 240 high concrete nouns and 240 low concrete nouns selected from Paivio, Yuille, and Madigan (1968). According to the norms provided by Paivio et al., where ratings are based on a scale of 1 to 7, the high concrete nouns had a mean concreteness value of 6.71, a mean imagery value of 6.27, and a mean meaningfulness value of 6.67. The corresponding means for the low concrete nouns were 2.46, 3.32, and 5.01.

Procedure. There were six study—test trials in a session. Except for the change in the stimulus manipulation, list construction, list presentation, and testing were identical to those in Experiment 1B.

Results and Discussion

For item recognition, mean d' was 2.53 for high concrete words and 1.73 for low concrete words. This difference was reliable [t(29) = 6.84]. The corresponding means for associative recognition were 2.70 and 1.19, and this difference was also highly significant [t(29) = 8.29].

The proportion of "yes" responses (false alarms for new tests and hits for old tests) for high and low concrete items and pairs are given in Table 5. Separate comparisons showed that for item recognition the difference between high concrete and low concrete items was reliable for both hits [t(29) = 5.16] and false alarms [t(29) = 4.33]. The differences between high and low concrete hits [t(29) = 7.77] and false alarms [t(29) = 6.96] were also significant for associative recognition.

Mean latencies for correct responses for both item and associative recognition are also shown in Table 5. For item recognition, the difference in mean latency between high and low concrete items approached signifi-

Table 5
Mean Proportion of "Yes" Responses, Mean Correct
Response Times (RT, in Milliseconds), and Standard Errors
for High and Low Concrete Words for Experiment 3

| | It | em Re | cognition | l | Associative Recognition | | | | |
|-------|------------|-------|-----------|-----|-------------------------|-----|-------|-----|--|
| Test | Proportion | | RT | | Proportion | | RT | | |
| Probe | М | SE | M | SE | M | SE | M | SE | |
| New | | | | | | | | | |
| High | .099 | .02 | 1,424 | 68 | .132 | .02 | 2,281 | 110 | |
| Low | .174 | .02 | 1,585 | 110 | .319 | .03 | 2,838 | 173 | |
| Old | | | | | | | | | |
| Low | .729 | .02 | 1,500 | 75 | .727 | .02 | 2,094 | 95 | |
| High | .824 | .03 | 1,327 | 78 | .876 | .03 | 1,641 | 73 | |

cance for hits [t(29) = 1.92, p = .07] and was reliable for correct rejections [t(29) = 2.21]. The difference between high and low concrete pairs was significant for both hits [t(29) = 5.64] and correct rejections [t(29) = 3.63] for associative recognition.

Experiment 3, like Experiments 1A and 1B, shows that the mirror effect can be produced for both item and associative recognition. In addition, Experiment 3 provides a further demonstration of the mirror effect in the pattern of response latencies.

A yes/no recognition test procedure was used in Experiments 1–3. Glanzer and Adams (1985) have shown that a mirror effect for item recognition can be obtained in both yes/no and forced-choice test procedures. Experiments 4A and 4B were designed to extend the mirror effect for associative recognition to the forced-choice procedure.

EXPERIMENTS 4A-4B Forced-Choice Recognition

Method

Subjects. Forty different undergraduate students participated in each experiment. The students either received course credit or were paid for their participation.

Stimuli. The noun/nonnoun word pool described in Experiment 1 was used for Experiment 4A and the high/low concreteness word pool described in Experiment 3 was used for Experiment 4B.

Procedure. A session consisted of 10 (Experiment 4A) or 6 (Experiment 4B) study—test trials followed by an end-of-session final test. In both Experiments 4A and 4B, every study list consisted of 32 critical pairs plus 2 primacy buffer pairs and 2 recency buffer pairs, for a total list length of 36 pairs. In Experiment 4A, half the pairs consisted of two nouns and half consisted of two nonnouns, and in Experiment 4B, half the pairs consisted of two high concrete words and half consisted of two low concrete words. All study pairs were constructed by randomly selecting (without replacement within a session) two words from the same class from the appropriate word pool. The pairs were presented in a random order with the constraint that there was an equal number of the two types of pairs in each critical half of the study lists. Every pair was presented for 1.5 sec, and there was a 1-sec blank interval between presentations.

An associative recognition test list followed every study list. The immediate test lists consisted of eight forced-choice discriminations. There were four different types of discriminations, and one of each type of discrimination was formed from study pre-

sentations in each half of the study lists. In Experiment 4A, the four types of discriminations were (1) intact noun pairs versus rearranged noun pairs, (2) intact noun pairs versus rearranged nonnoun pairs, (3) intact nonnoun pairs versus rearranged noun pairs, and (4) intact nonnoun pairs versus rearranged nonnoun pairs. In Experiment 4B, the corresponding discriminations were (1) intact high concrete pairs versus rearranged high concrete pairs, (2) intact high concrete pairs versus rearranged low concrete pairs, (3) intact low concrete pairs versus rearranged high concrete pairs, and (4) intact low concrete pairs versus rearranged low concrete pairs. The test pairs were presented to the left and right of the middle of the screen, and the left/right order of the two test pairs was randomly determined. Participants identified the intact test pair by pressing the corresponding left ("Z") or right ("/") response key. The presentation rate was subject paced.

The final recognition test list consisted of eight forced-choice discriminations from each of the 10 (Experiment 4A) or 6 (Experiment 4B) study lists. Thus, there were 80 final discriminations in Experiment 4A and 48 final discriminations in Experiment 4B. The intact test pairs were study presentations that had not been previously tested. The rearranged test pairs consisted of the complementary words from the study pairs used to form the rearranged pairs for the immediate tests. The order of the test discriminations was random, and the presentation rate was subject paced. The participants were not informed about the final recognition test until after they had completed the last study-test trial.

Results and Discussion

Experiment 4A. The proportions of correct forced-choice decisions for the combinations of intact and rearranged test pairs for both the immediate and the final test lists are presented in Table 6. For the mirror effect to be demonstrated, the discrimination between intact nonnoun versus rearranged nonnoun pairs (NN-NN) should be less than the discrimination of both intact nonnoun versus rearranged noun pairs (NN-N) and intact noun versus rearranged nonnoun pairs (N-NN). The discrimination of the mixed test pairs should be equal and less than the discrimination of intact noun versus rearranged noun test pairs (N-N).

Significant differences between conditions were determined by one-tailed *t* tests for predicted differences and two-tailed *t* tests otherwise. The obtained results are indicated in Table 6, where the symbols denote the means in each row that are not significantly different from each other. The pattern of differences between the conditions conforms to the mirror effect for the immediate tests. For the delayed test, there is an exception to the pattern predicted by the mirror effect—the accuracy of the NN-N condition is greater than the accuracy of the N-NN condition.

Experiment 4B. The proportion of correct discriminations in each condition of Experiment 4B are also given in Table 6. The means that do not significantly differ in each row of the table are noted by the symbols.

In both Experiments 4A and 4B, the overall pattern of results is consistent with the pattern that defines the mirror effect. In both immediate and final tests, the best discrimination was made between pairs consisting of words from the more memorable stimulus class and the poorest discrimination was made between pairs composed of words from the less memorable class. Discriminations of

| Table 6 |
|---|
| Mean Proportion of Correct Discriminations Between Intact and Rearranged Test Pairs |
| for the Immediate and Final Tests of Experiments 4A and 4B |

| | Intact-Rearranged Test Pairs | | | | | | | | | | |
|-----------|------------------------------|---------|---------|-------|-------|---------|---------|-------|--|--|--|
| | | Experin | nent 4A | | | Experin | nent 4B | | | | |
| | NN-NN | NN-N | N-NN | N-N | LC-LC | LC-HC | HC-LC | HC-HC | | | |
| Immediate | .720 | .760* | .755* | .831 | .683 | .838*# | .802* | .863# | | | |
| Final | .610# | .705* | .628# | .726* | .567 | .758 | .700 | .794 | | | |

Note—NN, nonnoun-nonnoun pair; N, noun-noun pair; LC, low concrete word pair; HC, high concrete word pair. Symbols represent means in each row that are not significantly different from each other.

mixed tests involving pairs from both classes were largely intermediate.

For the mixed-pair tests, however, there was an advantage favoring the discrimination of intact nonnoun pairs versus rearranged noun pairs in Experiment 4A, and intact low concrete pairs versus rearranged high concrete pairs in Experiment 4B. This difference was more pronounced on the final test list. The reason for this asymmetry is not clear, but it could reflect a bias to choose the test pair composed of words from the less memorable class in situations in which neither test pair can be recognized. Such a bias is consistent with Glanzer et al.'s (1991) finding that when participants are presented with two new items in a forced-choice item recognition test (null comparisons), participants choose the new item from the less memorable stimulus class more often than the new item from the more memorable stimulus class.

GENERAL DISCUSSION

The present study provides a further demonstration of the regularity of the mirror effect in item recognition. A mirror effect was produced by the stimulus manipulations of nouns versus nonnouns (Experiments 1A and 1B), low versus high word frequency (Experiments 2A and 2B), and high versus low word concreteness (Experiment 3). In addition, the pattern of results for the latency of item recognition responses mirrored the pattern of results for accuracy. That is, there was a positive relationship between response latency and accuracy. This finding replicates the results of Hockley (1982) and indicates that the mirror effect can be seen in both accuracy and latency measures of recognition performance.

Glanzer and Adams (1990), however, found no relationship between response latency and accuracy. It should be noted, though, that Glanzer and Adams used a more complex procedure to measure response time. In their Experiment 1, participants pressed the space bar to end the timing period, then pressed either an "old" or a "new" response key, and finally pressed one of four keys to indicate confidence. In their Experiment 3, participants pressed one of eight keys to indicate both their recognition decision and their confidence about the decision. The more complex response procedures, and the fact that participants had to make both a recognition decision and a confidence judgment, may be the reason

why Glanzer and Adams did not find a positive relationship between response latency and accuracy. Nevertheless, it should be emphasized that both the finding of no relationship and the present finding of a positive relationship between speed and accuracy provide strong evidence against a speed—accuracy tradeoff interpretation of the mirror effect.

A strong and consistent mirror effect was also obtained for associative recognition in Experiments 1A, 1B, and 4A (nouns vs. nonnouns) and in Experiments 3 and 4B (high vs. low word concreteness). As in item recognition, the mirror effect was observed in both measures of accuracy and response latency and was obtained by using both yes/no and forced-choice test procedures. The manipulations of nouns versus nonnouns and word concreteness had a similar effect on both item recognition and associative recognition. Concrete words are typically recognized more accurately than abstract words (see, e.g., Anderson, Goetz, Pichert, & Halff, 1977; Glanzer & Adams, 1990) and, in paired-associate learning, concrete word pairs are remembered better than abstract word pairs when the members of the pairs have been related and a cue appropriate to the relationship is provided at test (Marschark & Hunt, 1989). Since nouns have higher concreteness and imagery values than nonnouns do, the memorial advantage of nouns over nonnouns is most probably an effect of word concreteness.

Marschark and Hunt (1989) have suggested that the effects of word concreteness are a joint function of the processing of relational and distinctive information (see also Begg, 1982). Imaginal encoding of concrete words enhances the discriminability of these items from other items in memory. For concrete word pairs, integrative imagery can provide associative links between the words (e.g., Morris & Stevens, 1974; Paivio, Clark, & Khan, 1988). Thus, for concrete word pairs, imagery "can serve a dual function of providing a relational link between words at the pair level while enhancing distinctiveness between pairs at the list level" (Marschark & Hunt, 1989, p. 717). The present findings of increased recognition for both concrete words and concrete word pairs are consistent with this view.

In contrast to the effects of the noun/nonnoun and word concreteness comparisons, the manipulation of word frequency in Experiments 2A and 2B did not affect associative recognition performance. As noted previously, this constitutes a failure to replicate the high-

frequency advantage reported by Clark (1992) and Clark and Shiffrin (1992). Thus, the present results do not provide support for Clark's (1992) suggestion that associative recognition involves a recall process. The present results are, though, consistent with Clark's, insofar as they show that word frequency has a differential effect on item and associative recognition.

Few studies have examined the effects of word frequency on the acquisition and retention of paired associates. Postman (1962, Experiment 1) varied the frequency of stimulus and response terms in paired-associate learning and found that while speed of acquisition increased with the frequency of the response terms, the number of trials to criterion was a nonmonotonic function of the frequency of the stimulus terms. Tests of retention showed no effect of the frequency of the response terms, but, like acquisition, retention was a nonmonotonic function of the frequency of the stimulus terms. Martin (1964) found a more consistent relationship: The learning rate of word pairs that were not normatively associated increased as a function of the frequency of the stimulus terms. Postman (1962, Experiment 2) and Martin (1964) both found that the strength of preexperimental associations affected paired-associate learning when the stimulus terms were of low frequency, but there was little effect of associative strength when the stimulus terms were of high frequency. These results indicate that it is more difficult to form associations between low-frequency words than between high-frequency words.

McGovern (1964) suggested that paired-associate recall involves two independent stages: A response-recall phase that is dependent on contextual (item-to-list) associations, and an associative phase that is dependent on forward and backward associations between the stimulus and response terms. Similarly, associative recognition may involve both recognition of the individual items and retrieval of the associations between items. Noun pairs and high concrete word pairs are recognized better because high concrete words facilitate both item recognition and the encoding and retrieval of associations between items. In contrast, although low-frequency words are easier to recognize, the results of Postman (1962) and Martin (1964) indicate that it is more difficult to form associations between low-frequency words than between high-frequency words. The present experiments, therefore, may not have resulted in an effect of word frequency on associative recognition because the opposing effects of word frequency canceled each other out.1 Clark (1992) may have obtained a high-frequency advantage in associative recognition because the lowfrequency words he used were more difficult to associate than the low-frequency words used in the present experiments.

The present study shows that the mirror effect is a feature of both item and associative recognition. This suggests that similar retrieval and decision processes may underlie recognition of both types of information. Such a conclusion is consistent with the general approach of global matching models. The common assumption of

this class of models is that item and associative recognition decisions are both based on the familiarity of the test probe. This familiarity represents the strength of the match (or global activation) between the information encoded from the test probe and the information stored in memory. The match could be based on both item and associative information, as in Gillund and Shiffrin's (1984) SAM model, or item and associative information could be assessed independently, as in Murdock's (1982) TODAM model.

It should be emphasized, however, that a complete account of the mirror effect has not been provided within this class of models. For example, Hockley and Murdock (1987) showed that their recognition decision model could fit the pattern of results characterizing the mirror effect by assuming that the differences in accuracy and response latency for different classes of stimuli are due to differences in the underlying distributions representing the matching strength of the new and old items of each stimulus class. The same assumption could also be used to account for the mirror effect for associative recognition. The weakness of this approach, however, is that it does not provide a rationale for why the distribution for new items of the less memorable stimulus class should have a greater mean than the new item distribution for the more memorable stimulus class.

Glanzer and Adams (1990; Glanzer et al., 1991) have proposed attention-likelihood theory not only as an account of the mirror effect, but as an alternative approach to strength theories of recognition memory. In this model, a given recognition decision is based on a likelihood ratio derived from a comparison of the proportion of marked features an old item of this stimulus type is expected to have and the proportion of marked features a new item of this stimulus type is expected to have. Thus, in attention-likelihood theory, differences in the memorability of the stimuli as a class are incorporated into the recognition decision process. This approach could also be extended to associative recognition if appropriate assumptions concerning the representation of associative information are made.

One current limitation of attention-likelihood theory should be noted in light of the present results. Attention-likelihood theory has not, as yet, been extended to account for the latency of recognition decisions. The present results show that the mirror effect is reflected in measures of both accuracy and response latency. Thus, a complete description of the mirror effect must provide an account of both of these aspects of recognition performance.

It is not obvious how easily attention-likelihood theory could account for differences in the latency of recognition decisions. As described by Kim and Glanzer (1993), at test a subset of features is sampled from the test probe, and the number of marked features are counted. The number of marked features that an old item and a new item of that stimulus class are likely to have is then estimated. These estimates are used to compute the likelihood ratio for that item, and the recognition decision

is based on the likelihood ratio. This decision involves the processes prescribed by signal detection theory for making a response, but attention-likelihood theory differs from strength theories of recognition memory because the decision is based on likelihood ratios rather than on memory strength.

There does not seem to be any obvious reason why the counting of the marked features, or the calculation of the likelihood ratio, should take less time for items from the more memorable stimulus class than for items from the less memorable class. Thus, differences in response latency would, presumably, be due to the time taken to make the decision. Since the decision process is the same as in signal detection theory, the problem is to specify the transfer function that maps response latency onto distance from the decision criterion. Murdock (1985) has shown that such an approach can work; but it is not simple, and it would require a number of parameters and some ad hoc assumptions.

Hintzman (1994) has recently criticized the attentionlikelihood model on the grounds that this approach requires subjects to be able to estimate each old and new item distribution at test, but that it gives no explanation of how subjects know the true state of their memories. Hintzman suggested a simpler, item-specific method or strategy for rescaling the familiarity of the test item prior to the recognition decision to produce a mirror effect. In Hintzman's test-trial increment rescaling approach, the information needed for rescaling is generated by learning the test item on the test trial. The increment in familiarity produced by learning at test provides an estimate of the increased familiarity value that the item would have if it had been in the study list. Hintzman also noted that test-trial increment rescaling could be added to any model of memory (such as global matching models) that allows learning to occur on a test trial. It is not clear, though, how easily the test-trial increment rescaling approach, like the attention-likelihood model, could be developed to also provide an account of the latency of recognition decisions.

If the mirror effect is based on a postretrieval evaluation of the memorability of test items, either individually or as a class, the evaluation or rescaling process cannot be conscious, effortful, or time consuming. Hintzman, Caulton, and Curran (1994) examined item recognition for words of high and low frequency and high and low concreteness. A divided-attention procedure was used to constrain the opportunity for conscious evaluation at test, and a response-signal procedure was used to control recognition processing time. These retrieval manipulations did not eliminate or attenuate the mirror effect.

The findings of Hintzman et al. (1994) from the response-signal procedure are particularly noteworthy in that they indicate that the information that provides the basis for the mirror effect is available early in the retrieval/decision process. Hintzman et al. suggested that the retrieval dynamics of recognition, and of the mirror effect in particular, might be modeled by a random-walk

process in which evidence for a positive or negative decision accumulates over time (see, e.g., Ratcliff, 1978).

Hintzman et al.'s (1994) results, however, would seem to be more consistent with the recognition decision model proposed by Hockley and Murdock (1987, 1992). In this model, the input to the decision system is the output of the memory comparison process, which represents the strength of the match between the information encoded from the test probe and information stored in memory plus random noise. The temporal dynamics of recognition are modeled by a two-criterion decision process that evaluates the output from the memory comparison process over time in the face of fluctuating random noise. This model, then, does not assume that information accumulates over time. Rather, the result of the memory comparison process stays constant while extraneous noise varies over time. Thus, in this model, information would be available early in the decision process to provide a basis for the mirror effect.

Hintzman et al. (1994) concluded that the mirror effect "may best be attributed to the inherent nature of the retrieval and judgment processes that underlie recognition memory" (p. 286). The present study extends the generality of the mirror effect to encompass associative recognition and response latency. The fact that the mirror effect is a characteristic of both item and associative recognition suggests that similar retrieval and decision processes underlie both types of recognition performance. The finding that the mirror effect is also reflected in response latency further emphasizes the importance of considering the temporal dynamics of recognition in general, and of the mirror effect in particular.

REFERENCES

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, **6**, 451-474.

ANDERSON, R. C., GOETZ, E. T., PICHERT, H. M., & HALFF, H. M. (1977). Two faces of the conceptual peg hypothesis. *Journal of Experimental Psychology: Human Learning & Memory*, 3, 142-149.

BAIN, J. D., & HUMPHREYS, M. S. (1988). Relational context: Independent cues, meanings, or configurations? In G. M. Davies & D. M. Thompson (Eds.), *Memory in context: Context in memory* (pp. 97-137). London: Wiley.

Begg, I. (1982). Imagery, organization, and discriminative processes. Canadian Journal of Psychology, 36, 273-290.

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

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

DOSHER, B. A. (1988). Retrieval dynamics of item and associative information. Unpublished manuscript.

FRIENDLY, M., FRANKLIN, P. E., HOFFMAN, D., & RUBIN, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, 14, 375-399.

GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model for both recognition and recall. *Psycholological 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.
- GLANZER, M., ADAMS, J. K., & IVERSON, G. (1991). Forgetting and the mirror effect in recognition memory: Concentering of underlying distributions. *Journal of Experimental Psychology: Learning, Mem*ory, & Cognition, 17, 81-93.
- GLANZER, M., ADAMS, J. K., IVERSON, G. J., & KIM, K. (1993). The regularities of recognition memory. *Psychological Review*, 100, 546-567.
- 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. (1984). MINERVA 2: A simulation model of human memory. Behavior Research Methods, Instruments, & Computers, 16, 96-101
- HINTZMAN, D. L. (1990). Human learning and memory: Connections and dissociations. *Annual Review of Psychology*, **41**, 109-139.
- HINTZMAN, D. L. (1994). On explaining the mirror effect. Journal of Experimental Psychology: Learning, Memory, & Cognition, 20, 201-205.
- HINTZMAN, D. L., CAULTON, D. A., & CURRAN, T. (1994). Retrieval constraints and the mirror effect. *Journal of Experimental Psychol*ogy: Learning, Memory, & Cognition, 20, 275-289.
- HOCKLEY, W. E. (1982). Retrieval processes in continuous recognition. Journal of Experimental Psychology: Learning, Memory, & Cognition, 8, 497-512.
- HOCKLEY, W. E. (1991). Recognition memory for item and associative information: A comparison of forgetting rates. In W. E. Hockley & S. Lewandowsky (Eds.), Relating theory and data: Essays on human memory in honor of Bennet B. Murdock (pp. 227-248). Hillsdale, NJ: Erlbaum.
- HOCKLEY, W. E. (1992). Item versus associative information: Further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1321-1330.
- HOCKLEY, W. E., & MURDOCK, B. B., JR. (1987). A decision model for accuracy and response latency in recognition memory. *Psychologi*cal Review, 94, 341-358.
- HOCKLEY, W. E., & MURDOCK, B. B. (1992). Speed-accuracy trade-off and item recognition: A reply to Gronlund and Ratcliff. *Journal of Mathematical Psychology*, 36, 461-467.
- 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.
- HUMPHREYS, M. S., PIKE, R., BAIN, J. D., & TEHAN, G. (1989). Global matching: A comparison of SAM, MINERVA 2, Matrix, and TODAM models. *Journal of Mathematical Psychology*, 33, 36-37.
- KIM, K., & GLANZER, M. (1993). Speed versus accuracy instructions, study time, and the mirror effect. *Journal of Experimental Psychol*ogy: Learning, Memory, & Cognition, 19, 638-652.

- KUČERA, F., & FRANCIS, W. (1967). 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.
- MARSCHARK, M., & HUNT, R. R. (1989). A reexamination of the role of imagery in learning and memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 710-720.
- MARTIN, J. G. (1964). Associative strength and word frequency in paired-associate learning. *Journal of Verbal Learning & Verbal Be*havior, 3, 317-320.
- McGee, R. (1980). Imagery and recognition memory: The effects of relational organization. *Memory & Cognition*, **8**, 394-399.
- McGovern, J. B. (1964). Extinction of associations in four transfer paradigms. *Psychological Monographs*, 78 (16, Whole No. 593).
- MORRIS, P. E., & STEVENS, R. (1974). Linking images and free recall. Journal of Verbal Learning & Verbal Behavior, 13, 310-315.
- 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, 609-626.
- MURDOCK, B. B., JR. (1985). An analysis of the strength-latency relationship. *Memory & Cognition*, 13, 511-521.
- MURDOCK, B. B., & HOCKLEY, W. E. (1989). Short-term memory for associations. In G. H. Bower (Ed.), The psychology of learning and motivation: Vol. 24 (pp. 71-108). San Diego: Academic Press.
- Paivio, A., Clark, J. M., & Khan, M. (1988). Effects of concreteness and semantic relatedness on composite imagery ratings and cued recall. *Memory & Cognition*, 16, 422-430.
- PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monographs*, 76 (1, Pt. 2), 1-25.
- PIKE, R. (1984). Comparison of convolution and matrix distributed memory systems for associative recall and recognition. *Psychological Review*, 91, 281-294.
- Postman, L. (1962). The effects of language habits on the acquisition and retention of verbal associations. *Journal of Experimental Psychology*, **64**, 7-19.
- RAAIJMAKERS, J. G., & SHIFFRIN, R. M. (1981). Search of associative memory. Psychological Review, 88, 93-134.
- RATCLIFF, R. (1978). A theory of memory retrieval. Psychological Review, 85, 59-108.

NOTE

1. I thank Douglas Hintzman for suggesting this possibility.

(Manuscript received August 16, 1993; revision accepted for publication January 3, 1994.)