Recursive reminding: Effects of repetition, printed frequency, connectivity, and set size on recognition and judgments of frequency

SERENA L. FISHER and DOUGLAS L. NELSON
University of South Florida, Tampa, Florida

In the present experiments, predictions of common path and recursive-reminding models of recognition (RG) and judgments of frequency (JOFs) were contrasted. The results indicated that each task is affected by study frequency, printed frequency, and associative connectivity. However, effect size analyses indicated that study frequency and item attributes show a double dissociation over tasks. Study frequency has a greater effect on JOFs than on RG, whereas printed frequency and associative connectivity have greater effects on RG than on JOFs. The recursive-reminding model predicts differential effects of study frequency, because it assumes that although both tasks are influenced by familiarity, JOF is more likely to be affected by recollective reminding as a procedure for encoding event frequency. Associative set size effects were absent in each task, suggesting that competitors play no role in either task.

A case can be made for the conclusion that recognition (RG) and judgments of frequency (JOFs) are based on a single process. First, the study trial experience can be identical in the two tasks, and they differ primarily in the information required during testing. In RG testing, both studied and nonstudied items are presented, with participants being asked to judge the list membership of each item. In JOF experiments, participants judge the frequency of occurrence of each item. Second, the same variables affect performance in each task. Presentation frequency (Hintzman, 2004), stimulus duration (Hintzman, 2004), level of processing (e.g., Greene, 1988), distribution of practice (e.g., Hintzman, 1969), and printed word frequency (e.g., Greene & Thapar, 1994) affect both JOFs and RG. Finally, single-process global memory models provide a good account of the findings associated with both tasks (e.g., Hintzman, 1988, 2001; Shiffrin, 2003). Because of the single-process assumption, Hintzman (2004) referred to these models as *common path models*. Variables that benefit performance in one task predict performance benefits in the other because they rely on a common mechanism, identified as familiarity or strength.

Common path models predict that the effect sizes associated with these variables will be comparable in the two tasks. If a variable explains $X\%$ of the variance in RG, it will explain about $X\%$ of the variance in JOFs. Allowances for task differences can soften the exactitude of this prediction, but common path models cannot predict crossover interactions. Of two variables that ostensibly increase familiarity, one cannot have a larger effect in one task and a smaller effect in the other, with the other variable having the opposite effect. However, Hintzman (2004) recently reported such a dissociation. Study frequency explained more variance in the JOF task than did stimulus duration, whereas the effects were reversed in the RG task. In an effort to explain the crossover interaction, he suggested that RG depends primarily on familiarity, whereas JOF depends primarily on *recursive reminding*. Reminding is a recollection that occurs during study when a repeated stimulus reminds the learner of its earlier occurrence. The reminding is recursive because the second presentation of the target word reminds the learner of the first, and the third presentation reminds the learner that the second presentation reminded him or her of the first, and so on. Recursive reminding is described as a subjective state of awareness involving spontaneous recalls that provide a distinctive frequency cue during testing.

The present investigation addresses the predictions of common path and recursive-reminding models in linked experiments. Study frequency was crossed with printed frequency in four experiments. Associative connectivity was also crossed with these variables in Experiments 1 and 2, which served as the JOF and the RG studies, respectively. Associative set size was crossed with these variables in Experiments 3 and 4 in the same way. The main difference between these experiments and Hintzman's (2004) lies in the manipulation of item attribute variables related to prior lexical knowledge, as opposed to manipulating only variables related to episodic knowledge (e.g., study frequency and duration). Theoretically, printed frequency effects are mediated by perceptual distinctiveness, whereas connectivity effects are mediated

by associative distinctiveness. Low-frequency words tend to have more distinctive letters and letter combinations, and when studied, low frequency has benefited RG (e.g., Estes & Maddox, 2002; Malmberg, Steyvers, Stephens, & Shiffrin, 2002), JOF (Greene & Thapar, 1994), and extralist cued recall (Nelson & Xu, 1995). Associative connectivity refers to the number of associate-to-associate links among the associates of a word. Higher levels of connectivity theoretically increase the activation strength of the studied word, making it more distinctive (Nelson, McKinney, Gee, & Janczura, 1998). A high level of connectivity is associated with improved recognition (e.g., Nelson, Zhang, & McKinney, 2001), as well as with a higher level of cued recall (Nelson & Zhang, 2000), but its effects on JOF are unknown.

The rationale underlying the manipulations of item attribute variables in the present experiments was driven by predictions made by the models. In Experiments 1 and 2, the common path model predicts that RG and JOFs will vary systematically with study frequency, printed frequency, and connectivity. Each of these variables theoretically affects the strength or familiarity of the target, and the model predicts that performance in each task will increase with increasing repetitions, increasing associative connectivity, and decreasing printed frequency. Importantly, the common path model predicts that each variable will show comparable effect sizes in each task.

In contrast, the recursive-reminding model predicts that study frequency will have a larger effect in the JOF task than in the RG task. Although RG will benefit from study frequency, the JOF task specifically requires frequency discrimination and is, therefore, more likely than the RG task to depend on the benefits of recursive reminding. As presently characterized, this model does not make clear predictions about the influence of item attribute variables in the JOF task, especially with reference to the magnitude of such effects. In the context of the model, any effects of item attributes must be attributed to either recursive reminding or familiarity. The problem is that the model does not specify how greater familiarity might benefit frequency estimation separately from recursive reminding. If item attributes affect JOFs, such results cannot be attributed to differential study frequency, because printed low-frequency words appear no more often than do high-frequency words during the task, and the same is true for words varying in associative connectivity.

Given this shortcoming, the results of manipulating item attributes in Experiments 1 and 2 would inform the recursive-reminding model, but they would not test it, because two alternative outcomes seem plausible. For example, if item attributes have larger effect sizes in JOF tasks than in RG tasks, such results would imply that item attributes exert their influence on recursive reminding. Perceptually and associatively more distinct words might serve as more effective cues for reminding learners about recent past occurrences than would words that are less distinctive along these dimensions. Such a finding would be interesting because it would call into question the need for

the familiarity process in the frequency estimation task. Alternatively, if printed frequency and associative connectivity produce reduced effect sizes in JOFs, the model could be used to attribute such effects to the reduced role of familiarity in JOF. Given this result, the problem for the model will be to explain how familiarity and recursive reminding are related and why item attributes have a greater effect on familiarity than on reminding.

Finally, Experiments 3 and 4 provided a replication of the effects of study frequency and printed frequency on RG and JOFs but focused on set size as the item attribute variable. Associative set size refers to the number of closely related associates, and rather than distinctiveness, set size effects indicate the presence of competition produced by related words. Set size has robust effects on extralist cued recall, with recall being consistently worse for words having larger sets of associates (e.g., Nelson & Zhang, 2000). The purpose of varying set size here was purely empirical, and it was to determine whether recursive reminding engages a target's competitors when its prior occurrence is cued. Normally, such interference does not occur in RG tasks, but competitors can influence RG when related words are present during testing (Nelson, Canas, & Bajo, 1987). In principle, competitors can influence RG when the testing conditions encourage recall processes, and it seemed plausible that such competitors could influence the success of the recursive-reminding process.

EXPERIMENTS 1–4

Method

Participants. One hundred sixty undergraduates in the psychology and communication disorders departments served as participants, with 40 in each experiment. The participants received an extra credit point for their participation.

Design. The design for each experiment formed a $5 \times 2 \times 2$ repeated measures factorial. In Experiments 1 and 2, study frequency (zero, one, two, three, or four), printed frequency (high or low) and connectivity (high or low) were varied. In Experiments 3 and 4, study frequency, printed frequency, and target set size (large or small) were varied. Odd- and even-numbered experiments involved JOFs and RG, respectively.

Procedure. Upon arrival, the participants were asked to focus on the screen while the experimenter read the instructions. Each was told that after a short practice session to familiarize them with the presentation rate, they would see a long list of words. The instructions indicated that they should try to remember as many words as possible and that some would be seen more than once, but no information about the retention test was provided. The words were displayed at a 3-sec rate and were read aloud when shown. When the study phase was complete, the participants were told that words would appear during testing, some of which had just been studied. In all the experiments, testing was self-paced, and the researcher entered the response into the computer. In JOF Experiments 1 and 3, when the word appeared on the screen, the participants read it aloud and provided a number from zero to five, corresponding with how many times they had seen the word in the studied list. In RG Experiments 2 and 4, when the word appeared on the screen, the participants read it aloud and provided an *old* or a *new* decision.

Materials in Experiments 1 and 2. There were two independent lists of 80 words (see Appendix A). In each list, 20 words represented each of the four factorial combinations of printed frequency

and connectivity. High-frequency words appeared an average of 107 ($SD = 58$) times per million, and low-frequency words occurred an average of 4 ($SD = 3$) times per million (Kučera & Francis, 1967). When connectivity was high, each associate of the study word was connected to an average of 3.15 ($SD = 0.32$) other associates in its set (Nelson, McEvoy, & Schreiber, 2004). When connectivity was low, each associate was linked to an average of 0.60 ($SD = 0.18$) other associates. Connectivity was determined by using free association data to index the presence of a connection between each pair of associates in the studied words' set (Nelson et al., 2004).

A Latin square was used to randomly assign the words to study frequencies of one to four, so that each word occurred equally often at each frequency across four sublists. Words assigned to appear once in one sublist appeared twice in another, three times in a third, and four times in the fourth. Each sublist was then randomized without restrictions for each participant. List 2 served as the zero study condition when the participant studied List 1, and List 1 served as the zero study condition for List 2.

Materials in Experiments 3 and 4. Two independent 80-word lists were also created for these experiments (see Appendix B), and each of the four combinations of printed frequency and set size were represented by 20 words. The average occurrence rate per million words for high-frequency words was 107 ($SD = 55$), whereas the average occurrence rate for low-frequency words was 4 ($SD = 4$). Hence, printed frequency was varied over the same range as that used in the initial experiments, but the words were different. The set size parameters were established using free association norms (Nelson et al., 2004). Word set size was determined by collecting free association responses and counting the number of different responses produced by 2 or more participants. Words with large set sizes had an average of 20.28 ($SD = 2.26$) associates, whereas words with small set sizes had an average of 6.60 ($SD = 1.49$) associates. The assignment of words to study frequency and the randomization were implemented as in Experiments 1 and 2. Finally, mean target connectivity was held constant in these experiments at the lower end of the continuum, averaging 1.53 links ($SD = 0.75$) per associate.

Results

In the main analyses, separate $4 \times 2 \times 2$ within-subjects ANOVAs were run on mean frequency judgments from Experiments 1 and 3 and on RG d', as well as hits, for Experiments 2 and 4. The complete patterns of d', hits, and false alarms are listed in Appendices C and D. ANOVA results are presented for each experiment separately, followed by analyses of effect sizes for each linked-experiment pair, in order to compare the JOF and RG findings directly.

Experiment 1. As can be seen in Figure 1, JOFs systematically increased with increasing study frequency. In addition, although the effects appear to be relatively small, JOFs increased when printed frequency was lower and when associative connectivity was higher. Study frequency [$F(3,117) = 246.53$, $MS_e = 107.05$], printed frequency [$F(1,39) = 15.93$, $MS_e = 4.56$], and connectivity [$F(1,39) = 7.70$, $MS_e = 1.76$] were significant sources of variance. Frequency estimates increased systematically for Study Frequencies 1–4, and they were 1.08, 1.82, 2.53, and 2.94, respectively. Higher frequency estimates were provided for low-frequency (2.17) than for high-frequency (2.01) words and for high-connectivity (2.14) than for low-connectivity (2.04) words. None of the interactions among these variables was significant.

Although not germane to the main point of this article, a separate ANOVA including the zero-frequency condition was run for the sake of completeness. Mean JOF in the zero-frequency condition was .15, and when this condition was included in the analysis, mirror effects were apparent. The study frequency $\times$ printed frequency interaction was significant [$F(4,156) = 3.33$, $MS_e = .664$], and the study frequency $\times$ connectivity interaction showed a similar pattern but only approached significance [$F(4,156) = 2.20$, $MS_e = 0.42$, $p = .07$]. In both cases, low-frequency and high-connectivity words tended to show more hits and fewer false alarms than did their contrasting word types.

The results of Experiment 1 suggest that JOFs are substantially influenced by study frequency and that the
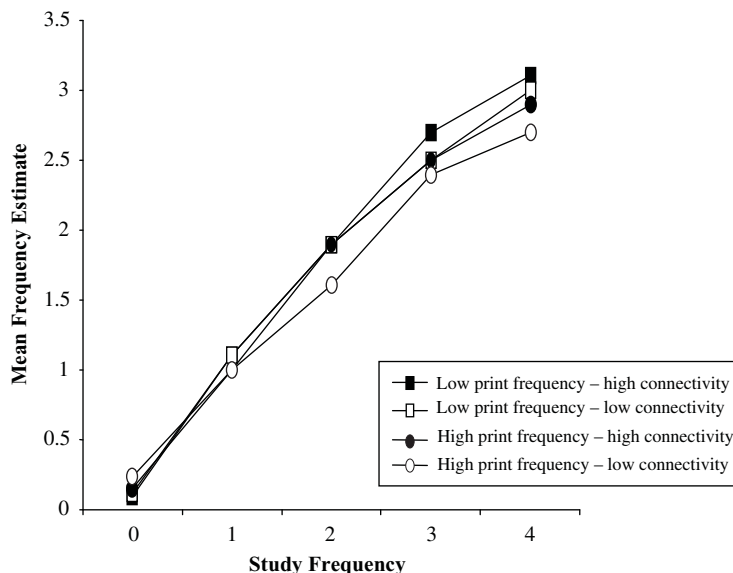

Figure 1. Judgments of frequency as a function of study frequency, printed frequency, and associative connectivity in Experiment 1.

item attribute variables of printed frequency and connectivity affect JOFs. Low-frequency and high-connectivity words were presented as often as high-frequency and low-connectivity words, but they were given higher frequency-of-occurrence estimates. The critical question concerns the magnitudes of the effect sizes of the item attribute variables in the JOF and RG tasks.

Experiment 2. At a gross level of analysis, the results of the RG and JOF experiments were similar. Figure 2 indicates that recognition d' systematically increased with increasing study frequency and tended to be higher when printed frequency was low and when connectivity was high. Study frequency, printed frequency, and associative connectivity significantly affected d' recognition [$F(1,39) = 31.07$, $MS_e = 15.97$, $F(1,39) = 15.58$, $MS_e = 25.76$, and $F(3,117) = 39.15$, $MS_e = 40.38$, respectively]. As in the JOF experiment, d' increased systematically with increasing frequency, and for frequencies of one to four, they were 2.67, 3.30, 3.55, and 3.85, respectively. Low-frequency words (3.50) were recognized more accurately than high-frequency words (3.18), and high-connectivity words (3.54) were recognized more accurately than low-connectivity words (3.14). In addition, there was a significant magnitude interaction between printed frequency and connectivity [$F(1,39) = 12.79$, $MS_e = 8.99$], so that frequency effects were more apparent for high-connectivity words. Fisher's two-tailed least significant difference (LSD) was .18. When connectivity was high, the d's for low- and high-frequency words were 3.82 and 3.27, respectively. When connectivity was low, these values were 3.18 and 3.10. No other interactions were significant.

The results of Experiment 2 confirm prior findings in showing that single-item recognition varies systematically with number of study presentations (e.g., Hintzman,

2004), printed frequency (e.g., Malmberg et al., 2002), and connectivity (e.g., Nelson et al., 2001). This experiment serves as a manipulation check on the materials and procedures, but more important, it provides the data for direct comparisons of the relative effects of these variables in the two tasks.

Effect Sizes in Experiments 1 and 2. In each experiment, effect size analyses were computed on study frequency, printed frequency, and associative connectivity, in order to determine whether comparable effect sizes were obtained. The common path model predicts that effect sizes for each variable will be equivalent for the JOF and the RG tasks. In contrast, the recursive-reminding model predicts that study frequency will have a greater effect size in JOF tasks than in RG tasks, but it makes no clear predictions about the relative effect sizes of item attribute variables in the two tasks.

The partial omega squared procedure was used to estimate the proportions of variance explained for each variable (Myers & Well, 1995). In keeping with the reminding model, study frequency accounted for more variance in the JOF task than in the RG task, accounting for 80%, as opposed to 22%, of the total variance. This pattern replicates Hintzman's (2004) findings, despite the use of *old–new* judgments, as compared with confidence judgments, in the RG task. In contrast, item attribute variables had smaller effects on JOFs than on RG. Printed frequency accounted for 2% and 5% of the variance in JOFs and RG, and connectivity accounted for 1% and 6% of the variance in these tasks, respectively. These differences were evaluated statistically by computing correlations on a subject-by-subject basis for each of the three variables in the two tasks (see Hintzman, 2004). The rs were transformed into Fisher z scores and evaluated in a repeated measures
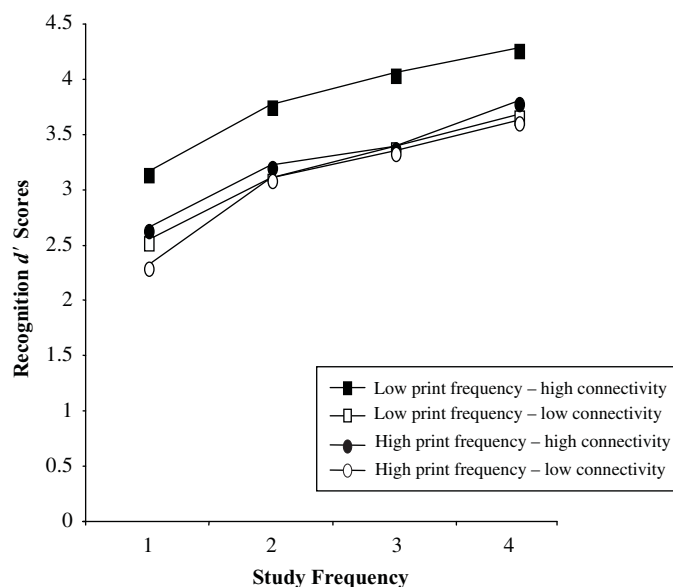

Figure 2. Recognition d' as a function of study frequency, printed frequency, and associative connectivity in Experiment 2.

analysis. Figure 3 depicts the inverse-Fisher transformed mean *r*s for the variables in Experiments 1 and 2. There was a significant crossover interaction [$F(2,156) = 53.59$, $MS_e = 4.31$]. Study frequency had a larger effect size in the JOF task than in the RG task, whereas item attribute effects tended to have larger effect sizes in the RG task and smaller effect sizes in the JOF task. Although the item attribute variables had a much smaller effect on JOFs, they did affect performance in this task. Finally, we note that the crossover pattern also was apparent for hits alone, indicating that it was not due to the d' correction for false alarms.

In Experiments 1 and 2, there were significant effects of study frequency, printed frequency, and connectivity in both JOF and RG tasks. Effect size analyses revealed a pattern of greater effect size for printed frequency and connectivity in the RG task, as well as a greater effect size for study frequency in the JOF task. The differential effect sizes for study frequency are consistent with the recurrent-reminding model, and the differential effects of item attributes in these tasks can be used to inform this model. Item attributes failed to have larger effect sizes in the JOF than in the RG task. This outcome suggests that low-frequency words and high-connectivity words do not serve as more effective reminding cues than do high-frequency words and low-connectivity words. This result is consistent with cued recall results showing that cue frequency has no observed effects on extralist cued recall (Nelson & McEvoy, 2000). Similarly, unpublished cued recall findings from our lab indicate that cue connectivity has no effects after correcting for mediators. Hence, the effects of item attributes on JOF do not appear to be caused by processes associated with recurrent reminding but, by default, by strength or familiarity. The effects of low-frequency and high-connectivity words in RG have been attributed to their relatively greater distinctiveness

Figure 3. Average correlations between task (judgment of frequency [JOF] and recognition) and associative connectivity, printed frequency, and study frequency in Experiments 1 and 2.

(e.g., Nelson et al., 2001; Shiffrin & Steyvers, 1997), and it appears that these variables affect JOF in a similar manner. However, as is indicated by the crossover interaction in the effect size analysis, their effects are smaller in JOF than in RG. Although the effects are smaller, judging how often a word occurred is influenced by its familiarity or strength, and this finding must be incorporated into any explanation for the effects of repeated presentations on JOF.

Experiment 3. The findings portrayed in Figure 4 indicate that both study frequency and printed frequency affected JOFs. The effect of study frequency [$F(4,156) = 367.80$, $MS_e = 211.10$] was significant, with frequency estimates increasing systematically for study frequencies of one to four, 1-4, and they were 1.03, 1.91, 2.55, and 2.99, respectively. The effects of printed frequency also were significant [$F(1,39) = 19.25$, $MS_e = 7.58$], with low-frequency words (1.82) being afforded higher JOFs than were high-frequency words (1.63). Set size had no apparent effects ($F < 1$), and there were no reliable interactions.

For the sake of completeness, when the JOFs for the zero condition (0.15) were included in the analysis, there was a significant study frequency × printed frequency interaction [$F(4,156) = 13.19$, $MS_e = 2.54$], due to the mirror effect seen in the zero-presentation condition. The JOF mirror effect was present in Experiments 1 and 3, and its presence replicates findings reported by Greene and Thapar (1994).

Experiment 4. The recognition d' data in Figure 5 indicate that, as in Experiment 2, RG performance varied with study frequency and printed frequency. The mean d' values for study frequencies of one to four were 2.53, 3.10, 3.38, and 3.64, respectively, and mean d' scores for low and high printed frequencies were 3.62 and 2.70. Both study frequency [$F(3,117) = 43.89$, $MS_e = 36.31$] and printed frequency [$F(1,39) = 95.41$, $MS_e = 137.26$] had significant effects. Set size had no effect ($F < 1$), and no interactions were significant.

Effect Sizes in Experiments 3 and 4. Effect sizes for study frequency, printed frequency, and set size were calculated (Myers & Well, 1995). Effect size for the set size variable was indeterminate and could not be calculated, but study and printed frequency showed patterns similar to those obtained in Experiments 1 and 2. Study frequency explained 76% of the JOF variance but only 24% of the RG d' variance, whereas printed frequency explained 31% of the RG d' variance and 5% of the JOF variance. The crossover interaction pattern was also evident when hits alone were used as the dependent measure. With either measure, study frequency had a greater effect size in the JOF task than in the RG task, whereas printed frequency had a greater effect size in the RG task than in the JOF task.

Finally, note that the effect size for printed frequency was much smaller in Experiment 2 than in Experiment 4. The former experiment was based on materials that varied systematically in both printed frequency and connectivity, whereas in the latter experiment low-connectivity words were used. Given Experiment 2's results showing that printed frequency effects were smaller for low-connectivity words,
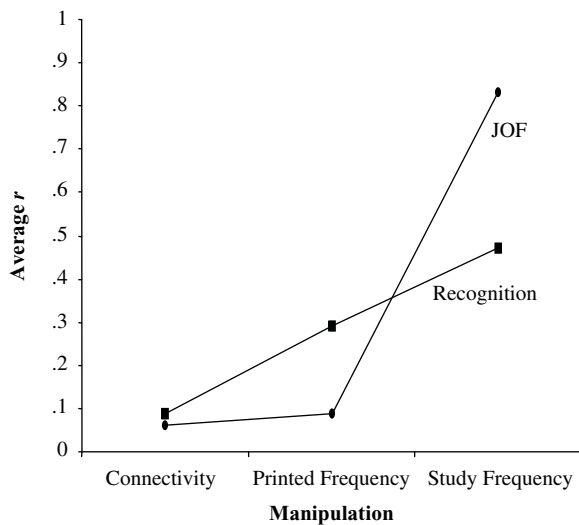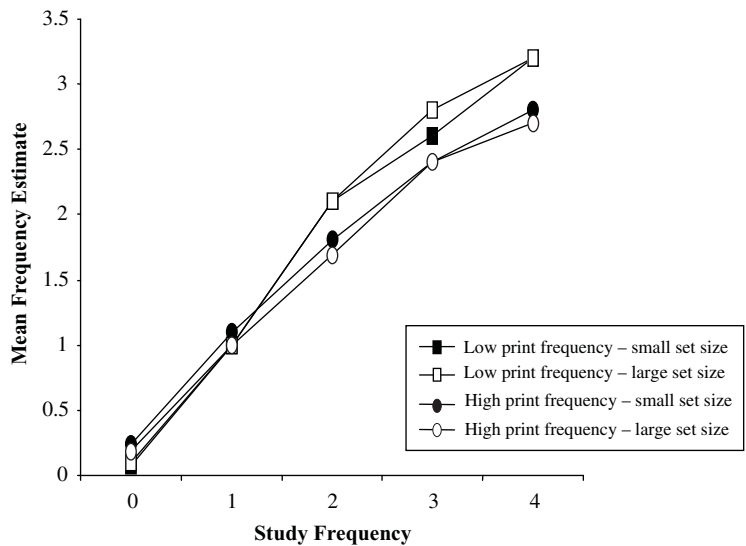
Figure 4. Judgments of frequency as a function of study frequency, printed frequency, and associative set size in Experiment 3.

the difference in effect size goes in the opposite direction of what might have been expected. If the magnitude of the printed frequency effects was determined by connectivity as such, these effects should have been relatively small in Experiment 4, because the majority of the words were low in connectivity. The larger effect size in Experiment 4 cannot be attributed to the level of connectivity of the items by itself. Apparently, the participants in Experiment 4 were forced to rely more on the perceptual distinctiveness of the words for making their recognition decisions. In contrast, those in Experiment 2 could rely on either perceptual or associative distinctiveness or on both attributes for making these decisions. In decisions about list membership,

the availability of distinctive associative information may reduce reliance on distinctive perceptual information.

As in the earlier experiments, correlations for the individual participants were calculated, transformed into Fisher z scores, and then subjected to a repeated measures analysis. Figure 6 depicts the mean correlations for Experiments 3 and 4, and as in the previous experiments, the crossover interaction was significant [$F(2,156) = 74.28$, $MS_e = 6.49$], showing differential effects for the variables, depending on task. Study frequency had a larger effect size in the JOF task than in the RG task, whereas printed frequency effects tended to be larger in the RG task and smaller in the JOF task.
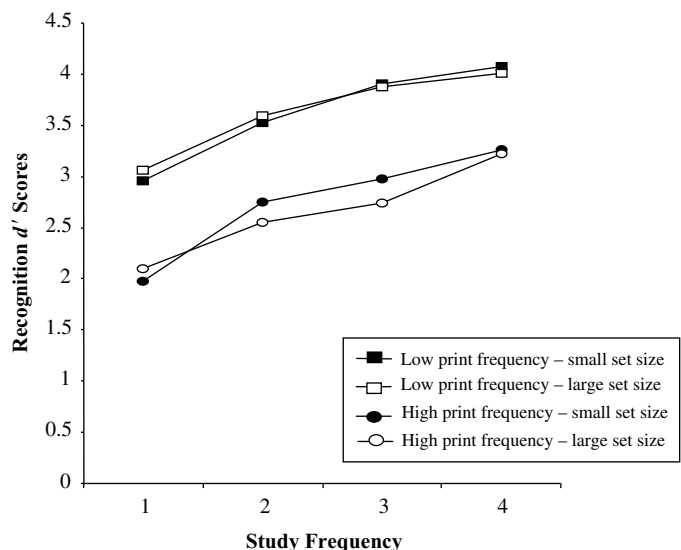

Figure 5. Recognition d' as a function of study frequency, printed frequency, and associative set size in Experiment 4.
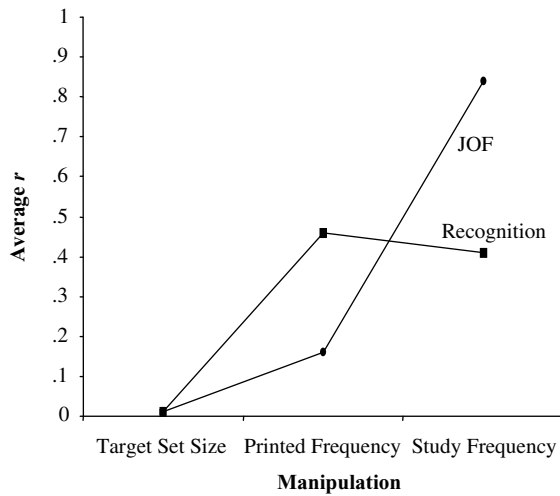
Figure 6. Average correlations between task (judgment of frequency [JOF] and recognition) and target set size, printed frequency, and study frequency in Experiments 3 and 4.

The effects of study frequency and printed frequency in Experiments 3 and 4 replicated the results of the first pair of experiments with different materials and a manipulation of set size, instead of connectivity. The smaller effect of printed frequency in the JOF task once again is consistent with the conclusion that familiarity, in addition to recursive reminding, contributes to estimates of how often words appeared during study. The results of Experiments 3 and 4 indicate that associative set size had no effects on either JOFs or RG. The absence of set size effects in the RG task replicates past findings obtained many times, but with lists that were better controlled (e.g., Nelson et al., 1987). The absence of set size effects on JOFs is a new finding, and although the caveats for null effects apply, this result suggests that such judgments are unaffected by the number and strength of the competitors linked to a word when it is being used as a cue for retrieving its prior occurrences. Set size effects are robust in extralist cued recall, and they can occur in RG tasks when conditions encourage participants to rely on associative cuing processes to determine whether a word appeared in the study list (Nelson et al., 1987). However, judging the recency or frequency of a word just seen normally does not engender competition effects produced by related lexical knowledge. Using one word to cue the recall of another engenders such effects, but using a word as cue for itself generally does not. Hence, the nature of the recall involved in recursive reminding differs from associatively motivated recall, because of the differences in the nature of the cues. Recursive reminding may be very cue specific, and to the extent that this is the case, substituting a related word for the target in the presentation sequence should fail to add to an estimate of frequency. When *lion* is seen four times in the sequence, judgments of its frequency will be higher than when it is seen three times. In contrast, if *lion* is seen three times and then the strong associate *tiger* is seen later in the list, the estimate of *lion*'s frequency should be unaf-

fected. Depending on the degree of specificity, even small changes, such as modality switching or font switching, may eliminate the benefits of repetition.

Discussion

The present findings show that study frequency and the item attributes of printed frequency and associative connectivity affect both JOFs and RG. Such results might normally be used to support common path models, because these variables affect performance in both tasks (Hintzman, 2004). However, effect size findings show a crossover interaction, or double dissociation, between manipulations of study frequency and item attributes, and this dissociation indicates that the similarity in results is superficial. Study frequency has a significantly larger effect on JOFs than on RG, and although item attributes influence performance in both tasks, they have larger effects on RG than on JOFs. The relative advantages for low-frequency words and for high-connectivity words are more apparent when items are judged for list membership than when frequency of occurrence is estimated. The present findings also show that associative set size has no apparent effects on the performance of either task.

Differences between the JOF tasks and the RG tasks are unlikely to have been due to what was encoded, because the materials and study conditions were virtually identical. Familiarization and recursive reminding occurring during study should have been equally likely in each task, so performance differences between the tasks could have been caused only by the differential demands of the retention tests. RG requires discriminating studied from unstudied words, whereas JOF requires the recovery of event frequency. Different information is needed in these two tasks, and it must be recovered from the same encoding. How our processing systems implement these differences in test instructions remains as one of the most significant unsolved problems in cognitive science, and we will not attempt to resolve it here, because this issue goes so far beyond the goals set for the present experiments. Instead, we will focus, as was planned, on the implications of the findings for RG and for the differences between this task and JOF tasks.

Recognition. In RG, different kinds of variables have similar effects on performance. Despite qualitative differences, the effects of variables linked to the study episode and variables linked to item attributes have been successfully attributed to a common familiarity process (e.g., Shiffrin & Steyvers, 1997). Study frequency, stimulus duration, level of processing, and other episodically defined variables affect familiarity because they influence the number of representations of the same word (Hintzman, 1988) or because they influence encoding strength (Criss & Shiffrin, 2004). In either case, episodic experience renders studied items stronger and, therefore, more familiar than those that are not studied. Item attributes, such as printed frequency, associative connectivity, and rated concreteness, affect distinctiveness ostensibly because of knowledge acquired prior to the laboratory episode, and the influence of such attributes is often attributed to fa-

miliarity as well (Nelson et al., 2001; Shiffrin & Steyvers, 1997). Differences in familiarity can arise from perceptual, associative, and imaginal sources and can affect RG above and beyond that acquired through episodic study. Mirror effects are found for all three types of attributes. Regardless of whether familiarity is derived directly from the episode, from item attributes, or from both sources, greater familiarity affects the task of judging list membership. The typical absence of associative set size effects in RG reinforces the emphasis on familiarity as the primary process in RG. Interference from competing associates linked to a studied target provides a marker for a cued recall process, but in the standard single-item recognition task, competition from related words has not been evident in dozens of experiments, including the present one (Nelson et al., 1987).

Recognition and judgments of frequency. The common path model predicted that both study frequency and item attributes would have the same effect sizes in JOF and RG tasks. In this model, a single process, familiarity or strength, underlies performance in both tasks. Contrary to this model, however, study frequency and item attributes show a double dissociation. Item attributes have a larger effect on RG, but study frequency has a larger effect on JOF, suggesting that something other than an item's individual familiarity mediates JOFs. Hintzman (2004) attributed the differential effects of study frequency to recursive reminding, on the assumption that the second presentation of the target word during study reminds the learner of the first, the third serves as a reminder that the second presentation served as a reminder of the first, and so on. An embedded, recall-based reminding process is at work during study, and during test, it provides a discriminative cue for frequency of occurrence. Such information would be just as available on an RG test, but this task does not call for frequency information. RG calls only for discriminating presence–absence information, and therefore, study frequency is not as likely to be used in making the discrimination. The recursive-reminding model tacitly assumes that the goal adopted by the learner differs in the two tasks and that this difference causes different information to be recovered from the same memory trace.

The recursive-reminding model is new and relatively undeveloped, so it could not predict how item attributes, such as printed frequency and associative connectivity, would affect performance in the two tasks. The findings indicate that these variables did not affect JOFs more than they affected RG. Such a result intuitively seems unlikely, but it was plausible at the outset, and it would have carried interesting implications for the recursive-reminding account of the effects of repeating exemplars in a sequence. Finding that study frequency, printed frequency, and associative connectivity each had larger effect sizes in JOFs than in RG could be interpreted as indicating that the frequency judgments are determined solely by a reminding process. Familiarity need play no role, because regardless of the source, distinctive cues are simply more effective as recursive-reminding cues. This effect, however, was not obtained, so this interpretation is ruled out by the data. One

conclusion that can be drawn is that these item attributes contribute to JOF through their influence on familiarity independently of recursive reminding. Low-frequency words and high-connectivity words seem more familiar, and so they are more likely to be correctly selected as list words and correctly rejected as new words (mirror effect). This conclusion makes recursive reminding a dual-process model, with performance in both JOF and RG tasks determined by both reminding and familiarity.

Such a dual-process model easily accommodates the present findings if we assume that both processes contribute to successful performance in both tasks, with the relative weights on each process determined by the task-defined goals. Study frequency explains more variance in the JOF task than in the RG task, because the JOF task explicitly requires reporting the number of times the target occurred, whereas in the RG task such information may be used more or less randomly. The information is available but is not used consistently. Similarly, item attributes explain more variance in the RG task than in the JOF task, because the RG task requires only a presence–absence discrimination, which can be made on perceptual and associative information that discriminates old from new words. In the RG task, the use of frequency information that goes beyond mere presence may occur more or less randomly. Each kind of information contributes independently, with the value of a given type determined by the goal of the retrieval process. The explicit assumption underlying this dual-process model is that learners often have information available to them that is not used in making judgments about what they know and that decisions are driven largely by goals that appear to define a limited purpose.

Additional points. Two additional points need to be made, and the first is methodological. Hintzman's (2004) use of explained variance to evaluate the dominant model of JOFs should remind researchers that such information is useful, even critical, in evaluating theory. Reporting the proportions of explained variance should be necessary in any research purporting to claim that a variable has the same effect in different tasks. In such research, it is no longer acceptable to maintain such claims by reporting only the results of significance tests. The second point is theoretical and is directed toward the concept of reminding. Hintzman (2004, p. 345) described a reminding event as a subjective state of conscious awareness that becomes encoded as an episodic memory. No data in the present experiments challenge this view, but alternative models that avoid this assumption need to be developed, because frequency information may be encoded consciously and automatically as well (e.g., Hasher & Zacks, 1984).

Developing such a model is beyond the scope of this article, but the recently modified REM model for recognition could provide a starting point (Criss & Shiffrin, 2004). In this version, both context and item attributes are encoded about targets, and both types of information introduce error during testing. With two additional assumptions, this model may be able to handle the double-dissociation findings. The first assumption is that context serves as a better

discriminative cue for frequency than does item information, and the second assumption is that test instructions differentially bias the recovery of encoded context and item features. JOF instructions increase the probability of recovering context features, and RG instructions increase the probability of recovering item features. Under these assumptions, JOFs will be affected more by presentation frequency than is RG, and in turn, RG will be affected more by variations in item attributes than are JOFs. The feasibility of these suppositions remains to be determined, but they suggest that the influence of these variables might be explained through a matching process that can be biased by test instructions, without having to make assumptions about the necessity of conscious processing.

REFERENCES

CRISS, A. H., & SHIFFRIN, R. M. (2004). Context noise and item noise jointly determine recognition memory: A comment on Dennis and Humphreys (2001). *Psychological Review*, **111**, 800-807.

ESTES, W. K., & MADDOX, W. T. (2002). On the processes underlying stimulus-familiarity effects in recognition of words and nonwords. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 1003-1018.

GREENE, R. L. (1988). Generation effects in frequency judgment. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 298-304.

GREENE, R. L., & THAPAR, A. (1994). Mirror effect in frequency discrimination. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 946-952.

HASHER, L., & ZACKS, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist*, **39**, 1372-1388.

HINTZMAN, D. L. (1969). Apparent frequency as a function of frequency and the spacing of repetitions. *Journal of Experimental Psychology*, **80**, 139-145.

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

HINTZMAN, D. L. (2001). Similarity, global matching, and judgments of frequency. *Memory & Cognition*, **29**, 547-556.

HINTZMAN, D. L. (2004). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory & Cognition*, **32**, 336-350.

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

MALMBERG, K. J., STEYVERS, M., STEPHENS, J. D., & SHIFFRIN, R. M. (2002). Feature frequency effects in recognition memory. *Memory & Cognition*, **30**, 607-613.

MYERS, J. L., & WELL, A. D. (1995). *Research design and statistical analysis*. Hillsdale, NJ: Erlbaum.

NELSON, D. L., CANAS, J., & BAJO, M.-T. (1987). The effects of natural category size on memory for episodic encodings. *Memory & Cognition*, **15**, 133-140.

NELSON, D. L., & McEVOY, C. L. (2000). What is this thing called frequency? *Memory & Cognition*, **28**, 509-522.

NELSON, D. L., McEVOY, C. L., & SCHREIBER, T. A. (2004). The University of South Florida free association, rhyme and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, **36**, 402-407.

NELSON, D. L., McKINNEY, V. M., GEE, N. R., & JANCZURA, G. A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological Review*, **105**, 299-324.

NELSON, D. L., & XU, J. (1995). Effects of implicit memory on explicit recall: Set size and word frequency effects. *Psychological Research*, **57**, 203-214.

NELSON, D. L., & ZHANG, N. (2000). The ties that bind what is known to the recall of what is new. *Psychonomic Bulletin & Review*, **7**, 604-617.

NELSON, D. L., ZHANG, N., & McKINNEY, V. M. (2001). The ties that bind what is known to the recognition of what is new. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 1147-1159.

SHIFFRIN, R. M. (2003). Modeling memory and perception. *Cognitive Science*, **27**, 341-378.

SHIFFRIN, R. M., & STEYVERS, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, **4**, 145-166.

APPENDIX A
Materials Used in Experiments 1 and 2

	High Frequency		Low Frequency	
	High Connectivity	Low Connectivity	High Connectivity	Low Connectivity
List 1	activity	again	amp	annual
	bottle	apart	amuse	beetle
	cent	assignment	bandage	boxer
	child	baseball	bruise	bucket
	community	Christian	burglar	cardboard
	cousin	desk	clinic	cricket
	crisis	dinner	dessert	doorbell
	depth	hear	fitness	flake
	different	law	freeway	hanger
	excellent	nice	grief	hose
	island	pay	marker	kinetic
	marine	philosophy	meteor	legion
	painting	pretty	pants	mask
	performance	radio	proton	owl
	plenty	rifle	scallop	paste
	poetry	send	sew	poison
	relation	stream	soccer	pumpkin
	scientific	stress	surgery	sack
	tiny	temperature	towel	shingle
	trial	writer	watt	toaster
List 2	attempt	beginning	ape	banner
	attention	board	bargain	booth
	blood	bridge	broil	braces
	bottom	center	clarinet	cable
	chain	claim	cushion	crater
	chief	column	diamond	crust
	direct	data	forbid	dragon
	dream	follow	fraud	gloves
	failure	force	Jupiter	helmet
	hair	Indian	microphone	husk
	key	line	notebook	kitten
	news	nose	raft	magnet
	platform	numbers	scar	nail
	post	phone	scuba	pane
	rose	question	shrub	peanut
	spread	rock	snack	poker
	stretch	round	sponge	packet
	symbol	rules	thaw	scout
	title	set	vest	snail
	win	supply	weird	vanish

APPENDIX B
Materials Used in Experiments 3 and 4

	High Frequency		Low Frequency	
	Large Target Set Size	Small Target Set Size	Large Target Set Size	Small Target Set Size
List 1	ability	addition	alley	addiction
	absence	apart	barley	ashtray
	block	bank	bead	brook
	clear	bible	blessing	cable
	decision	conclusion	bruise	chalk
	effort	correct	chunk	cork
	fresh	enter	defrost	despise
	island	everything	flute	faucet
	lack	fast	garlic	hammer
	mark	fingers	hiker	kite
	pass	front	insult	maple
	philosophy	job	lobster	noun
	please	king	mastery	paste
	reality	near	outlaw	puddle
	sharp	novel	poison	robin
	stage	remain	referee	shampoo
	train	seek	singer	socks
	uniform	similar	stain	toaster
	union	son	toy	vanish
	wind	winter	unequal	vent
List 2	advice	aid	ambulance	affection
	baseball	attempt	asphalt	banner
	chance	begin	biscuit	bouquet
	direction	circle	bribe	cardboard
	engage	difficulty	butterfly	comb
	friend	dinner	clamp	cradle
	issue	evening	dragon	drench
	leader	final	forbid	fracture
	local	frame	hallway	hornet
	mass	hour	hobby	knob
	park	library	jewel	noisy
	phase	orchestra	lace	pail
	poetry	question	mildew	pebble
	range	simple	napkin	pumpkin
	shoulder	six	plaid	scissors
	space	south	robe	shingle
	tradition	spend	scrap	timid
	view	thin	swamp	umbrella
	welfare	west	tunnel	wager
	wood	youth	valve	yoke

APPENDIX C
Hits, False Alarms, and d' Scores for Experiment 2

No. of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
High Printed Frequency, High Connectivity					
Mean	.08	.73	.85	.88	.96
d'	NA	2.65	3.22	3.39	3.80
High Printed Frequency, Low Connectivity					
Mean	.06	.66	.83	.88	.93
d'	NA	2.31	3.11	3.35	3.63
Low Printed Frequency, High Connectivity					
Mean	.03	.79	.92	.96	.99
d'	NA	3.16	3.77	4.06	4.28
Low Printed Frequency, Low Connectivity					
Mean	.04	.76	.87	.93	.99
d'	NA	2.94	3.45	3.80	4.14

APPENDIX D
Hits, False Alarms, and d' Scores for Experiment 4

No. of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
High Printed Frequency, Large TSS					
Mean	.08	.62	.75	.83	.88
d'	NA	2.10	2.55	2.75	3.22
High Printed Frequency, Small TSS					
Mean	.08	.59	.77	.84	.89
d'	NA	1.99	2.75	3.00	3.26
Low Printed Frequency, Large TSS					
Mean	.03	.75	.89	.93	.96
d'	NA	3.06	3.59	3.88	4.01
Low Printed Frequency, Small TSS					
Mean	.03	.74	.85	.93	.96
d'	NA	2.96	3.53	3.90	4.07