10-27-2004

Frequency Judgments and Recognition: Additional Evidence for Task Differences

Serena Lynn Fisher
University of South Florida

Frequency Judgments and Recognition:
Additional Evidence for Task Differences

by

Serena Lynn Fisher

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Arts
Department of Psychology
College of Arts and Sciences
University of South Florida

Major Professor: Douglas Nelson, Ph.D.
Michael Brannick, Ph.D.
Cathy McEvoy, Ph.D.

Date of Approval:
October 27, 2004

Keywords: memory, frequency estimation, printed frequency, connectivity,
set size

Table of Contents

List of Tables

List of Figures

Frequency Judgments and Recognition:
Additional Evidence for Task Differences

Serena Fisher

ABSTRACT

Four linked experiments were run in order to understand the relationship between frequency judgment and recognition discrimination tasks. The purpose of these studies was to contrast the common-path model and recursive reminding hypothesis as explanations for the underlying principles that drive these tasks. Item-attribute variables such as printed frequency, connectivity, and set size, and an episodic variable, study frequency were manipulated. Memory for recent episodes was evaluated using recognition and frequency judgment tasks. Although all of the variables, with the exception of set size, had significant effects in both tasks, an analysis of effect sizes revealed differences between the tasks in relation to the variables. Specifically, the item-attribute variables had larger effects in recognition than in JOF, and the effect size for study frequency was greater in the JOF task compared to recognition. The reliability of these differences was statistically established by a repeated measures analysis run on the correlations between each subject's mean and the variables. Although the effect size pattern is consistent with the reminding hypothesis, the effects of connectivity and printed frequency in the JOF task are not as they represent familiarity measures. Thus, this finding indicates that familiarity must be involved in making frequency judgments,

making the reminding hypothesis inadequate as an explanation as it does not take into account the effect of item-attribute variables and their contribution to familiarity with its subsequent effect on frequency estimates. Therefore, it is proposed that a dual-process approach that takes into account both the reminding and recollection at test in the JOF task, as well as attempting to explain the influence of an underlying construct such as familiarity that effects both tasks may be the most appropriate explanation for frequency estimation results.

Frequency Judgments and Recognition: Additional Evidence for Task Differences

Frequency judgment tasks (JOF's) involve participants seeing words at different frequencies during study, and then, at test, estimating how many times they think they saw a particular word during the study episode. Recognition discrimination tasks vary, but generally revolve around a participant seeing a list of items (words, patterns, pictures, etc.) and then later choosing the items seen previously, often from amongst items that were not studied. Logically, it would seem that these two tasks are related, because an item seemingly must be recognized before it can be judged for frequency. In fact, it has been argued that these two tasks might involve the same processes (Hintzman, 1984). In a recent article, Hintzman (2004) notes that Minerva2, REM (Shiffrin, 2003), and TODAM (Murdock, Smith, & Bai, 2001) all take for granted what he describes as a *common-path model* that assumes that one dimension underlies both recognition and JOF tasks, namely strength or familiarity. Hintzman (2004) cites several findings that support this idea. Among them are similar retrieval functions (Hintzman & Curran, 1994) and the fact that the mirror effect for printed frequency that can be found in recognition tasks can also be obtained in frequency discrimination tasks (Greene & Thapar, 1994). The common path model, demonstrated in Figure 1, predicts comparable effects for all variables in both JOF and recognition.

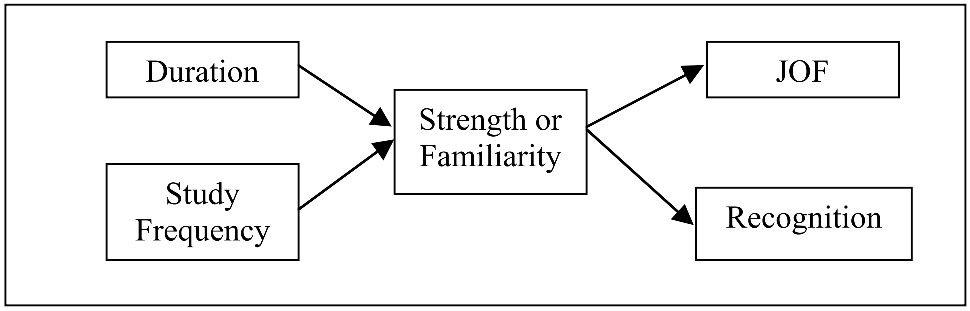

Figure 1. An adaptation from Hintzman (2004) demonstrating the common construct hypothesis

However, Hintzman (2004) also points out that there have been challenges to a common path model. Proctor (1977) finds differences in the accuracy of JOF's and recognition and in the ROC curves produced by the tasks as the motivation for his own attempt to discern differences between them. Also, Wells (1974) argued against a strength model for frequency in JOF tasks, claiming a more complex model was needed. Hintzman (2004) replicates Proctor's findings by crossing stimulus duration and study frequency resulting in differences of effect size and normalized memory operating characteristic (*z*MOC) curves between the two tasks. He then dismisses the common path hypothesis as inadequate, and introduces the idea of *recursive reminding* as an explanation for JOF's. The reminding hypothesis suggests that seeing a word a second, third, etc. time reminds the learner that they saw it before. Hintzman describes exposure to a word multiple times during study as "an accumulation of spontaneous acts of recognition," that can be recollected at test (Hintzman, 2004). However, perhaps both explanations play a role, allowing for the influence of both strength and recursive reminding in the JOF task. In order to assess the two models, the current article involves a series of experiments that manipulate variables that influence encoding strength in the two tasks, and then uses effect sizes to determine whether there are differences between JOF's and recognition.

In an effort to evaluate the common path and reminding hypotheses, study frequency, the number of times a word appears in the study list, was varied in both JOF and single-item recognition tasks. These manipulations were crossed with three item-attribute variables, including printed frequency, associative connectivity, and associative set size. Printed frequency and connectivity were varied to manipulate stimulus familiarity at non-semantic and semantic levels of distinctiveness, respectively. Printed frequency refers to estimates of word occurrence in the English language (Kucera & Francis, 1967). Low frequency words are likely to have more rare orthographic features in terms of distinctive letters and letter combinations (e.g., Malmberg, Steyvers, Stephens, & Shiffrin, 2002). Connectivity is illustrated in Figure 2, and as can be seen it refers to the interconnectivity between the associates of a word (Nelson, Bennett, Gee, Schreiber, & McKinney, 1993). High connectivity words have more pre-existing connections among their associates than low connectivity words. Theoretically, such inter-associate links increase the activation level of the studied word, which makes it more distinct (Nelson & Zhang, 2000).

Both printed frequency (e.g., Estes & Maddox , 2002) and connectivity (e.g., Nelson, McKinney, Gee, & Janczura, 1998) affect performance. Recognition is better when words are lower in printed frequency and when they are higher in connectivity. The effects of printed frequency and associative connectivity are ostensibly mediated by perceptual and associative distinctiveness, respectively. Low levels of printed frequency and high levels of connectivity are theoretically associated with higher levels of familiarity, that has a part in the common path model as an influence on recognition, but as Hintzman (2004) notes, familiarity's role in JOF's is still unknown. However, the presence of printed frequency and connectivity effects in the JOF task would suggest that familiarity may play a role in this task. Frequency judgments may depend to some extent on both familiarity and recursive reminding.

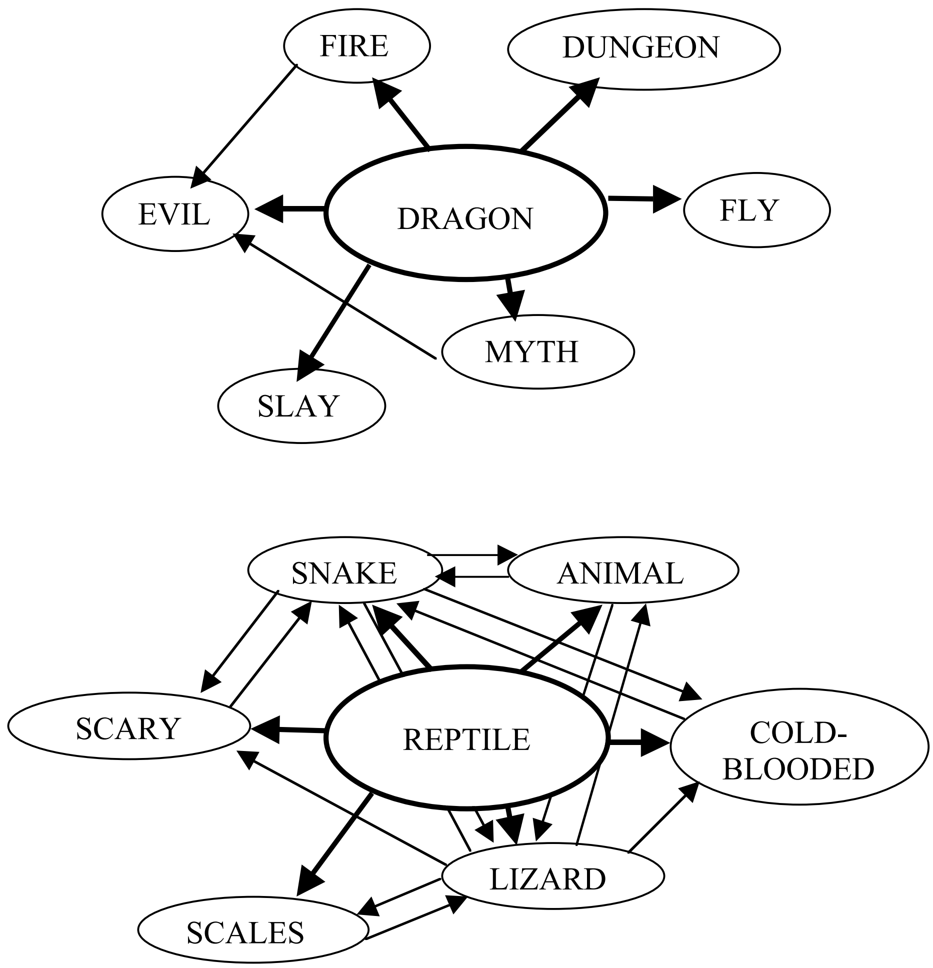

Figure 2. An example of words with low and high connectivity. Although REPTILE and DRAGON have the same number of associates, there are more connections among REPTILE's associates.

To provide a complete explanation for the findings common path theory predicts that the manipulations of both printed frequency and associative connectivity, as well as the manipulation of study frequency, will have the same effects on JOF and recognition. This prediction is based on the assumption that each of these variables has its effects because they influence a common mechanism, strength or familiarity. Such results would be consistent with predictions derived from Minerva2 and REM. Alternatively, if the reminding explanation is to provide a complete explanation of JOF, then item-attribute variables will affect recognition but not JOF and study frequency will have a larger effect on JOF than recognition. The absence of item-attribute effects on JOF estimates but not on recognition judgments would suggest that familiarity affects recognition but not JOF. In this case, the recursive reminding explanation would provide a complete explanation of JOF. Alternatively, if item-attributes affect frequency estimates, then a role for familiarity is implicated in this task even if the effects of these attributes are somewhat smaller than in recognition. Both familiarity and recursive reminding may influence frequency estimates.

Finally, in an additional test of similarities between frequency judgments and recognition, differences in set size were of interest over the two tasks. As can be seen in Figure 3, set size refers to the number of a word's associates. Set size has a negative effect in cued recall (Nelson & Freidrich, 1980), with words with a larger number of associates being recalled at lower rates than those with smaller numbers of associates, but typically set size has no effect on recognition discrimination (Nelson, Canas, & Bajo, 1987). However, its potential effects on JOF have never been evaluated.

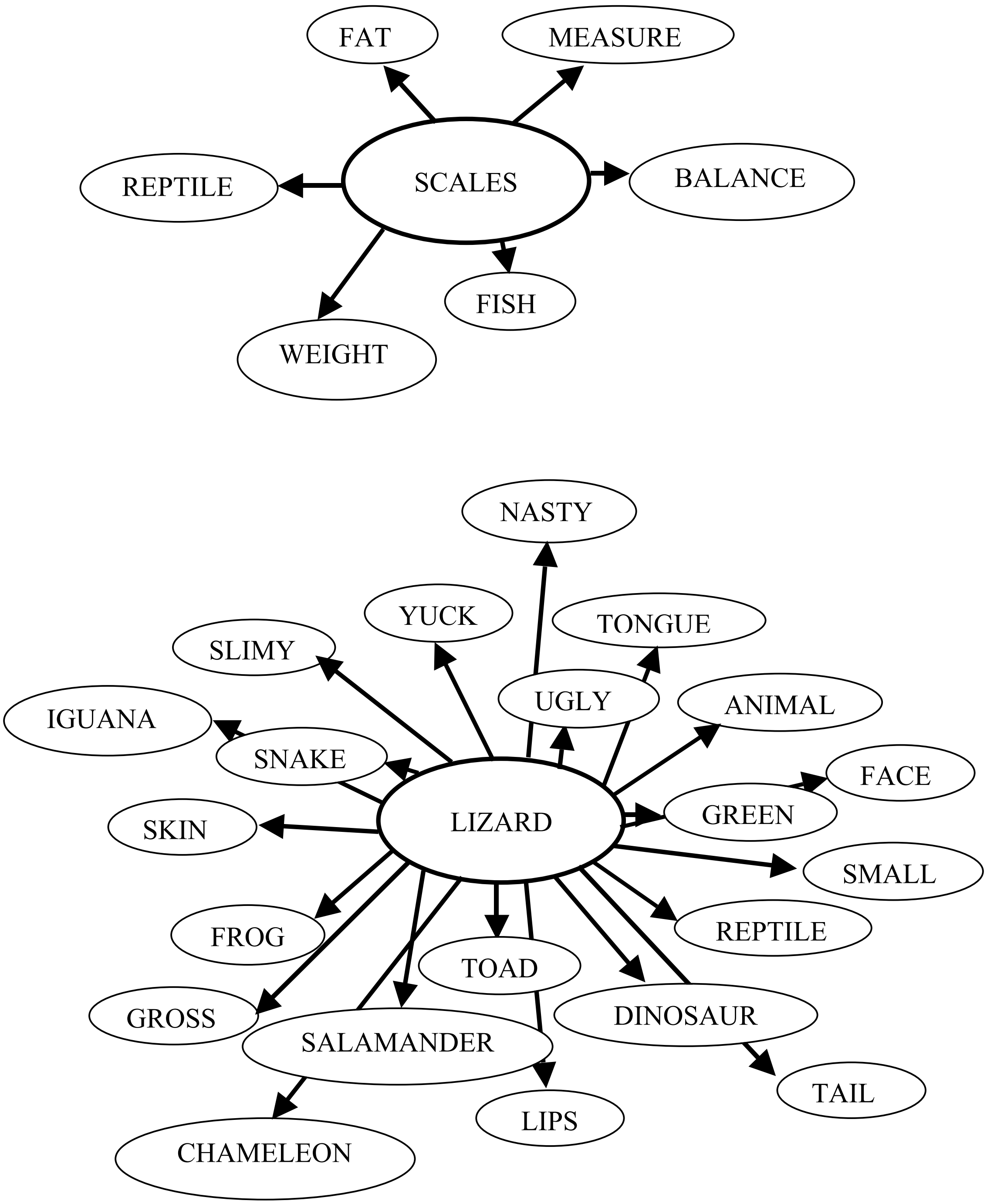

Figure 3. An example of words with small and large set sizes. The word LIZARD has a larger set of related associates than SCALES.

These issues were evaluated in two pairs of linked experiments. In Experiment 1, target connectivity and printed frequency were crossed with study frequency in a repeated measures factorial design. The purpose of Experiment 1 was to estimate effect sizes for study frequency, printed frequency, and connectivity in a frequency judgment task. Half of the list comprised words with high connectivity and the other half comprised words with low connectivity. Likewise, half of the high connectivity words were high or low in frequency, and half of the low connectivity words were high or low in frequency, allowing for all three variables to be tested within the confines of a single experiment. Study frequency was varied such that a word could appear 1, 2, 3, or 4 times depending on the given list, or 0 times, as a non-studied word. During testing, this task requires participants to estimate the number of times the word was seen in the study list, from zero to five times. Experiment 2 also varied study frequency, printed frequency, and connectivity. This experiment provided the recognition test comparison for Experiment 1 in order to determine whether these variables have comparable effect sizes in the two tasks. The same materials were used in both experiments.

Experiments 3 and 4 replicated the initial experiments, but replaced the connectivity manipulation with a set size manipulation. Studied words varied in set size and printed frequency and these variables were crossed with study frequency. Participants in Experiment 3 were asked to estimate frequency, whereas participants in Experiment 4 were given recognition instructions. The manipulation of target set size was motivated in part because earlier work done with set size and recognition (e.g., Nelson, et al, 1987) was not based on current knowledge about associative networks, so it is conceivable that the manipulation could have been confounded with uncontrolled variables. Hence, whether set size will affect recognition is an empirical question. Theoretically, set size effects are attributed to search effects directed toward recovering a studied word (e.g., LIZARD) when an associated, related word serves as a test cue (e.g., REPTILE). Set size effects ostensibly occur because cues and targets with larger sets of associates reduce the likelihood of selecting the target (Nelson, Schreiber, & McEvoy, 1992). Because searching the studied word's set is unlikely in recognition, such effects are also unlikely in a frequency judgment task. According to the common path hypothesis, an absence of set size effects in recognition predicts an absence of such effects in JOF's. The reminding hypothesis has no such constraint.

Theoretical interest is focused on the effect sizes produced by these variables in the two tasks. Subjects approach the two tasks similarly at study, encoding semantic, orthographic, and occurrence information, only to then rely on the information to varying degrees at test. Judgments of frequency and recognition may require different information or the same information at different strengths, a difference that would be apparent in effect size calculations. As for study frequency, we expect to see significant effects in both tasks. Study frequency should affect the context to word link, thereby improving frequency estimation and recognition with each additional presentation. We expect low frequency words to be recognized better than high frequency words because of their perceptual distinctiveness (Malmberg, et al., 2002). Similarly, high connectivity words will be recognized better than low connectivity words due to their associative distinctiveness (Nelson, et al, 1998). The interesting question is whether printed frequency and connectivity will have any effects on JOF.

General Method

Design and participants

The design for each of the following experiments formed a 2 x 4 x 2 repeated measures factorial. Printed frequency (high, low) and study frequency (the participant saw a word 1, 2, 3, or 4 times) will be manipulated in all experiments. In Experiments 1 and 2, connectivity (high, low) was varied with printed and study frequency. In Experiments 3 and 4, target set size (large, small) was manipulated with these two types of frequency. The odd numbered experiments required participants to estimate how often each studied word appeared during study. The even numbered experiments asked participants to recognize the studied words.

One hundred and sixty undergraduate students in the psychology and communication disorders programs at the University of South Florida served as participants, with forty in each experiment. Two equivalent study lists were created with 20 words for each within-subjects condition. Within each list, words were randomly assigned in a Latin Square design to either 1, 2, 3, or 4 presentations at study, resulting in four separate iterations of 200 words for each study list. The group of words assigned to appear once in List 1A would appear twice in List 1B, three times within List 1C, and four times within List 1D. The words appeared randomly without restriction at study. List 2 served as the zero presentation (non-studied) condition when the participant studied List 1, and List 1 served as the zero presentation condition when the participant studied List 2. Participation was randomly assigned to one of eight study lists. They received one point of extra credit for their participation.

Procedure

When participants arrived, they were seated in front of a Macintosh Quadra 800 computer and told to focus on the screen while the experimenter read the instructions for the experiment. Each was told that a long list of words would be presented following a short practice session that would provide experience on the presentation rate. The instructions indicated that they should try to remember as many words as possible and that some of the words would be seen more than once, but no information about the exact nature of the experiment was provided. Words were displayed at a rate of 1 every 3 seconds and were read aloud when shown. When the study phase was complete, the participant was told that additional words would be shown, some of which had just been studied. In Experiments 1 and 3, when the word appeared on the screen, they read it aloud and then gave a number from zero to five, corresponding with how many times participants believed they had seen the word in the studied list. The test was self-paced and the researcher entered the response with the keyboard number pad and the data were automatically stored onto the hard drive. In Experiments 2 and 4, when the word appeared on the screen, it was read aloud and followed by an "Old" or "New" decision, in which the participant made a judgment of "Old" if the word was from the original list or "New" if the word had not appeared in the earlier list. The researcher entered the response and the data were automatically stored.

Experiments 1 and 2 (Target Connectivity)

Materials

In these experiments, there were two independent lists with 20 words for each of the four printed frequency and connectivity conditions, for a total of 80 words per list (see Appendix A). In each list there were 20 words that were high in frequency and high in connectivity, 20 with high frequency and low connectivity, 20 with low frequency and high connectivity, and 20 with low frequency and low connectivity. The Kucera and Francis (1967) norms were used to index printed frequency, and a high printed frequency word was one that appeared in the range of 50-312 times per million words, whereas a low printed frequency word was one that appeared in the range of 0-10 times per million. The average occurrence rate for a high frequency word was 106.78 (SD= 58.34), while a low frequency word occurred an average of 4.50 (SD= 3.14) times per million words. When connectivity was high, each associate of the study word was connected to an average of 3.15 (SD= .32) other associates in its set (Nelson, McEvoy, & Schreiber, 1999). When it was low, each associate was linked to an average of .60 (SD= .18) other associates. Connectivity norms were determined using an analysis of free association data associated collected over the past twenty years (Nelson, et al, 1999).

Within each list, five of the 20 words in each condition were assigned to different study frequencies. A Latin square was used to randomly assign the words to study frequencies of 1, 2, 3, or 4, and thus four lists were created within List 1, allowing for cycling through study frequencies such that each word occurred at each different frequency (i.e., the group of words assigned to appear once in List 1A would appear twice in List 1B, three times within List 1C, and four times within List 1B). List 2 served as the zero presentation (non-studied) condition when the participant studied List 1, and List 1 served as the zero presentation condition when the participant studied List 2.

Experiments 3 and 4 (Target Set Size)

Materials

In these experiments, there were two independent lists with 20 words for each printed frequency, set size condition, for a total of 160 words (see Appendix B). That is, there were 20 words that had high frequency and large target set sizes, 20 words with high frequency and small target set sizes, and so on. A high printed frequency word appeared in the range of 50-257 times per million words, whereas a low printed frequency word appeared in the range of 0-10 times per million. The average occurrence rate for a high frequency word was 107.23 (SD= 54.71), whereas a low frequency word occurred an average of 4.43 (SD= 3.56) times per million words. The set size parameters were established using norms (Nelson, et al, 1999). A word with a large set size could have anywhere from 18 to 26 other associates in its set, whereas a word with a small set size could have anywhere from 2 to 8 other associates in its set. Words with large set sizes had an average of 20.28 (SD= 2.26) associates, while words with small set sizes had an average of 6.60 (SD= 1.49) associates. Set size was defined by collecting free association responses and counting the number of responses produced by two or more participants to a particular word. Study frequency was established in the same manner as in Experiments 1 and 2.

Results

Analyses of Variance

Four analyses of variance (ANOVA's) were conducted on the data collected. First, a 2x4x2 within-subjects ANOVA was run on mean frequency judgments from Experiment 1, with printed frequency (high, low), study frequency (word presented 1, 2, 3, or 4 times), and connectivity (high, low) as factors. The same ANOVA design was run on the recognition d' scores collected in Experiment 2 and the complete pattern of hits, false alarms, and d' scores are listed in Appendix C. The design was the same for Experiments 3 and 4, but set size (large, small) replaced connectivity as a factor. All of the hits, false alarms, and d' scores for Experiment 4 are listed in Appendix D. At test, words from the non-studied list are presented as the zero frequency condition.

Experiment 1

The results of Experiment 1 are shown in Figure 4. As can be seen JOF's were higher when printed frequency was lower and when connectivity was higher. Furthermore, JOF's systematically increased with increases in study frequency. Printed frequency, $F(1,39) = 15.93$, $MSe = 4.56$, study frequency, $F(3,117) = 246.53$, $MSe = 107.05$, and connectivity, $F(1,39) = 7.70$, $MSe = 1.76$ were each significant sources of variance. Higher frequency estimates were provided for low (2.17) than high (2.01) frequency words, and for high connectivity (2.14) compared to low connectivity (2.04) words. Frequency estimates increased systematically for study frequencies 1-4, and they were, respectively, 1.08, 1.82, 2.53, and 2.94. None of the interactions were significant. Finally, when the analysis included the zero presentation JOF's (0.15), mirror effects were apparent because there was a significant printed frequency by study frequency interaction, F(4, 156)= 3.33, MSe= .664. The connectivity by study frequency interaction showed a similar pattern but only approached significance, F(4, 156)= 2.20, *p*= .07, MSe= .419.

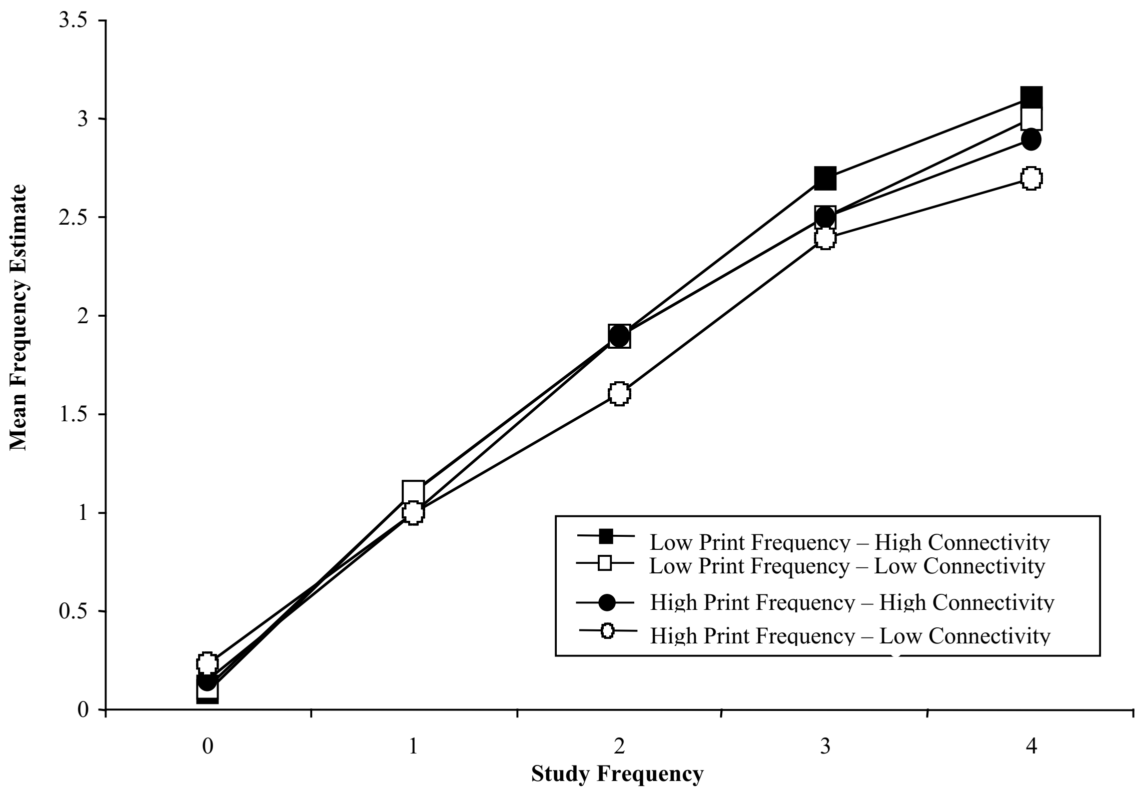

Figure 4. Mean ratings of frequency as a function of printed frequency, connectivity, and study frequency in Experiment 1.

Experiment 2

The results of Experiment 2 are shown in Figure 5. Recognition d' tended to be higher when printed frequency was low and when connectivity was high. Recognition increases systematically in each of these conditions with study frequency. Printed frequency, connectivity, and study frequency were each significant sources of variance, $F(1,39) = 31.07$, $MSe = 15.97$, $F(1,39) = 15.58$, $MSe = 25.76$, $F(3,117) = 39.15$, $MSe = 40.38$, 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). As in the prior experiment, d' scores increased systematically for study frequencies 1-4, and they were, respectively, 2.67, 3.30, 3.55, and 3.85. Also, there was a significant magnitude interaction between printed frequency and connectivity, $F(1,39) = 12.79$, $MSe = 8.99$, such that connectivity effects were more apparent for low frequency words. When frequency was low, high connectivity words (3.82) were recognized better than low connectivity words (3.18). At high frequency levels high connectivity words (3.27) were still recognized more accurately than low connectivity words (3.10), but the effect was smaller. Fisher's two-tailed least significance difference (LSD) was .18. There were no interactions between study frequency and printed frequency or between study frequency and connectivity. The magnitude of the printed frequency and connectivity effects was similar at each level of study frequency. No other interactions were significant.

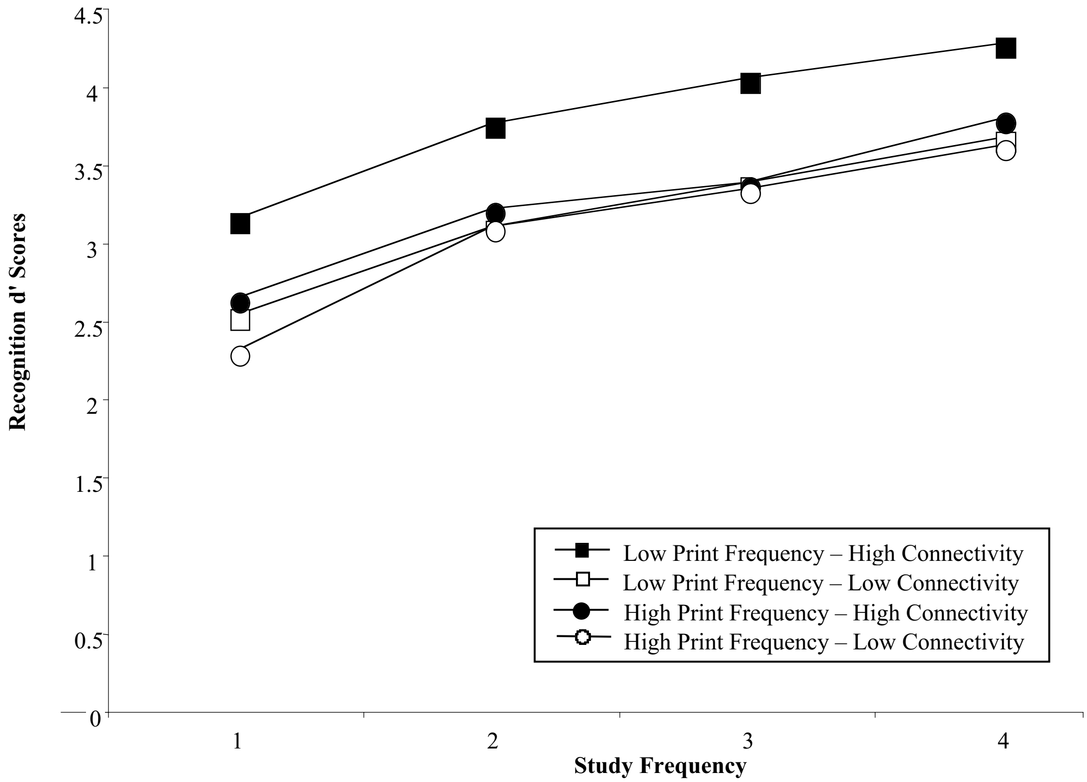

Figure 5. Recognition d' scores as a function of printed frequency, connectivity, and study frequency in Experiment 2.

Effect Size. All three variables affected performance on the tasks and in order to determine task differences, effect size tests were conducted.

	Study Freq	Printed Freq	Connectivity
Recognition d'	.22	.05	.06
Mean JOF	.80	.02	.01

Table 1. Proportion of Variance Explained by Study Frequency, Printed Frequency, and Connectivity (w^2)

Although the ANOVA's indicate significant effects for each of the variables included in Experiments 1 and 2, effect size analyses were run on study frequency, printed frequency, and associative connectivity in order to illustrate possible additional differences between frequency judgment and recognition tasks. Using Myers and Well's (1995) partial omega squared to estimate proportion of variance explained, the results shown in Table 1 were obtained. In keeping with the reminding hypothesis, the item-attribute variables had larger effects in recognition than in JOF. Also, consistent with Hintzman's (2004) conclusion that the frequency estimation task in more sensitive to study frequency, the effect size for study frequency was greater in the JOF task compared to recognition.

In order to provide statistical support for the hypothesis that there will be different effect sizes of the variables depending on the task, the results were subjected to a correlational analysis (Hintzman, 2004). Correlations were computed on a subject-by-subject basis for each of the three variables in the two tasks (connectivity, printed frequency, and study frequency for both the JOF and recognition tasks). An example can be seen in Appendix E. The *r*'s were then transformed into Fisher *Z* scores and evaluated in a repeated measures analysis. Figure 6 depicts the inverse-Fisher transformed mean *r*'s for the variables in Experiments 1 and 2. Lending support to the recursive reminding hypothesis and Hintzman's conclusions, there is a significant crossover interaction, $F(2,156) = 53.59$, $MSe = 4.31$. Study frequency has a larger effect size in the JOF task than in the recognition task, whereas item-attribute effects tend to have larger effect sizes in the recognition task and smaller effects sizes in the JOF task. Most importantly, although the item-attribute variables had a smaller effect on JOF, they do affect performance in this task.

In Experiments 1 and 2 there were significant effects of connectivity, printed frequency, and study frequency in both the JOF and recognition tasks. Effect size analyses revealed a pattern of greater effect size for connectivity and printed frequency in recognition, as well as a greater effect size for study frequency in the JOF task. Although this effect size pattern is consistent with the reminding hypothesis, the effects of connectivity and printed frequency in the JOF task are not. The number of study presentations should influence JOF's because each repetition should add its context information to the final recollection at test. In contrast, connectivity and printed frequency effects on frequency judgments represent an illusion. Low frequency and high connectivity words make people think that they have seen these words more often when they have seen them just as often as high frequency and low connectivity words. This finding indicates that familiarity must be involved in making frequency judgments. Therefore, a dual-process approach that takes into account both the reminding and recollection at test in the JOF task, as well as attempting to explain the influence of an underlying construct such as familiarity that effects both tasks may be the most appropriate explanation for frequency estimation results.

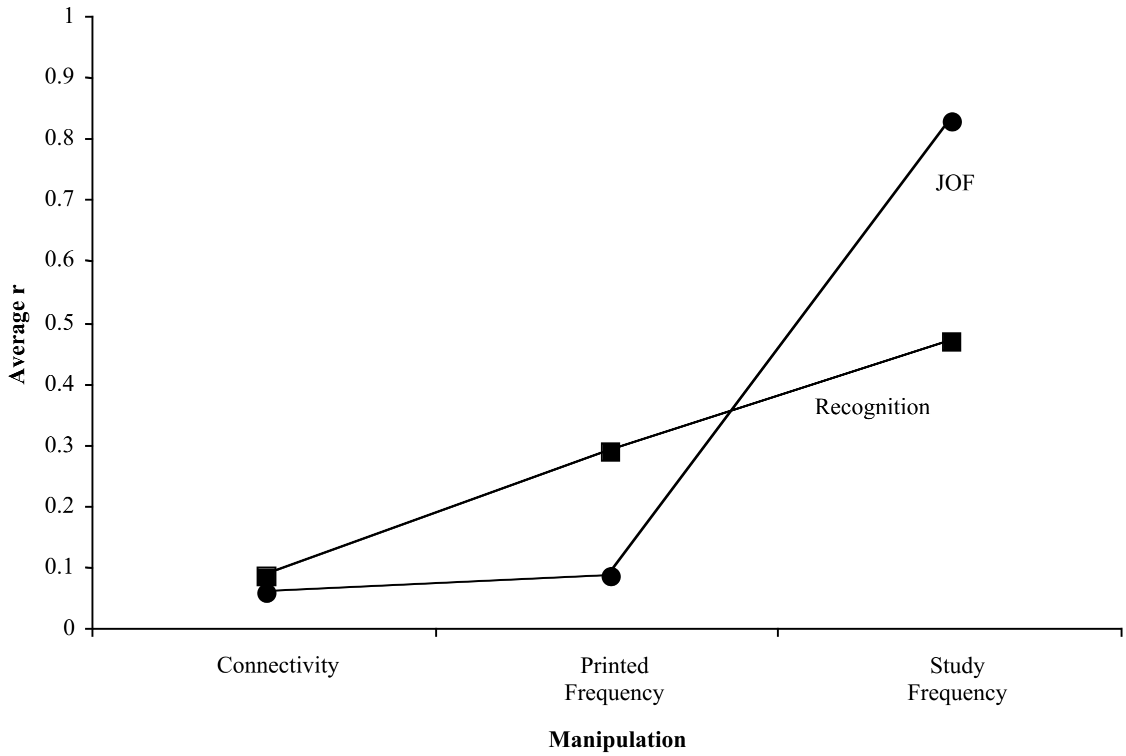

Figure 6. Correlational analysis of effects of connectivity, printed frequency, and study frequency on judgments of frequency (JOF) and recognition d'. For averaging, *r* values were Fisher transformed.

Experiments 3 and 4

As can be seen, the JOF findings in Figure 7 indicate that printed frequency and study frequency affected the estimates, but set size had no apparent effects. The effects of printed frequency were significant, $F(1,39) = 19.25$, $MSe = 7.58$, with low frequency words (1.82) being given higher study frequency judgments than high frequency words (1.63). There was an effect of study frequency, $F(4,156) = 367.80$, $MSe = 211.10$, with frequency estimates increasing systematically for study frequencies 1-4, and they were, respectively, 1.03, 1.91, 2.55, and 2.99. When the zero condition was omitted, there were no reliable interactions. However when the zero condition (0.15) was included in the analysis, there was a significant printed frequency by study frequency interaction, F(4,156) = 13.19, MSe = 2.54, due to the mirror effect seen at the zero presentations condition.

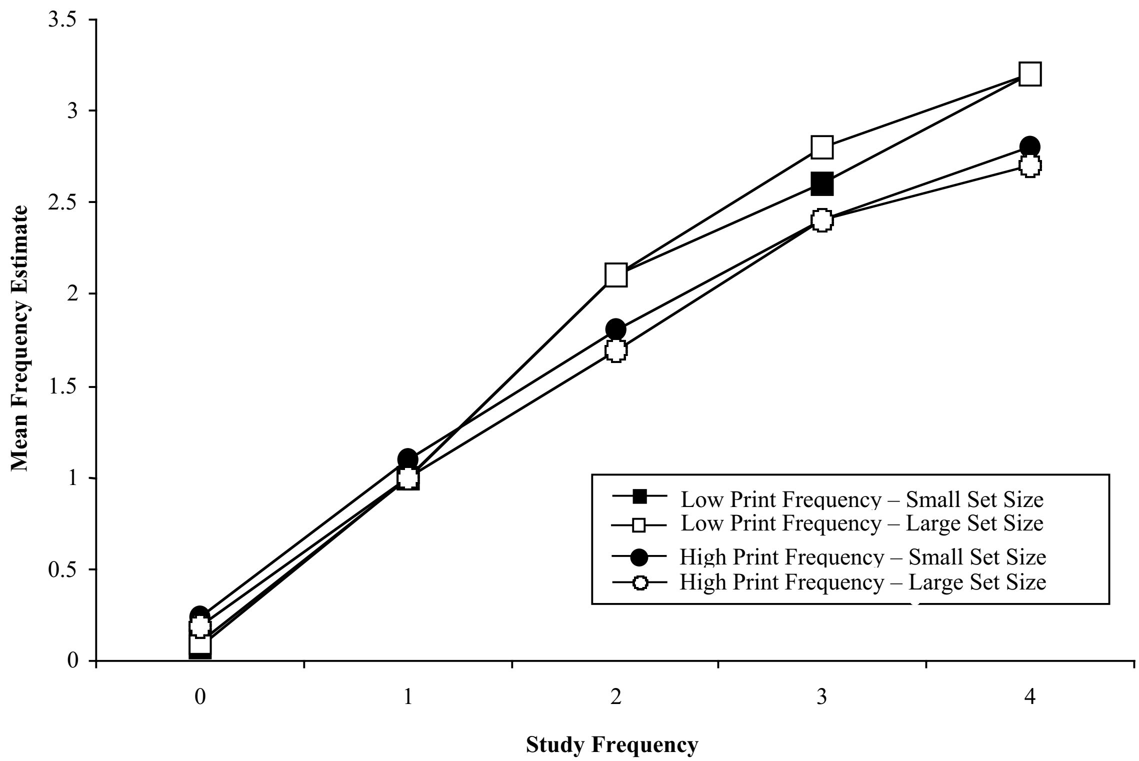

Figure 7. Mean ratings of frequency as a function of printed frequency, target set size, and study frequency in Experiment 3.

For Experiment 4, the recognition d' data reveal that there were significant main effects for printed frequency, F(1,39) = 95.41, MSe= 137.26, with low frequency words (3.62) being recognized more accurately than high frequency words (2.70). Set size had no effect. In addition to printed frequency, there was a main effect of study frequency, F(3,117) = 43.89, MSe = 36.31, with recognition accuracy increasing as study frequency increased. The means for study frequencies of 1-4 were 2.53, 3.10, 3.38, and 3.64, respectively. These effects can be seen in Figure 8. No interactions were significant.

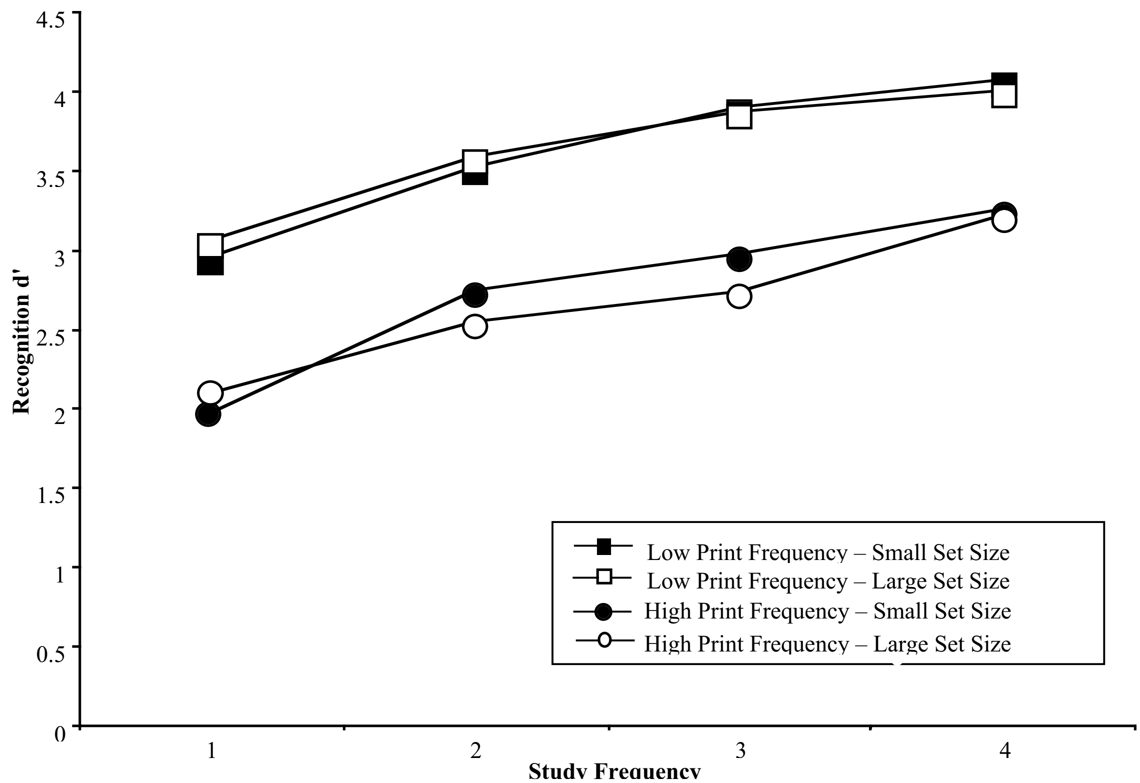

Figure 8. Recognition d' scores as a function of printed frequency, target set size, and study frequency in Experiment 4.

As was expected, the main effects of printed frequency and study frequency conformed to previous research findings (Estes & Maddox, 2002; Hintzman, 2004). In keeping with expectations, set size was not affected by task type. This finding may be seen as support for the idea that neither recognition tasks nor judgment of frequency tasks appear to involve a search process like that engendered in the extralist cued recall task.

Effect Size. Additionally, as can be seen in Table 2, the effect sizes for printed frequency and study frequency were calculated (Myers and Well's, 1995). A similar pattern to that of Experiments 1 and 2 was obtained. Printed frequency had a greater effect size in the recognition task than in the JOF task, whereas study frequency had a greater effect size in the JOF task compared to the recognition task.

	Study Freq	Printed Freq	Set Size
Recognition d'	.24	.31	Indeterminate
Mean JOF	.76	.05	Indeterminate

Table 2. Proportion of Variance Explained by Study Frequency, Printed Frequency, and Set Size (w^2)

Statistical support was determined by obtaining *r* values, transforming them into Fisher *Z* scores, and then running a repeated measures analysis. Target set size replaced connectivity as a correlated variable from the 80 participants in Experiments 3 and 4. Figure 9 depicts the mean correlations for Experiments 3 and 4. As in the previous experiments the significant crossover interaction, $F(2,156) = 74.28$, $MsE = 6.49$, showed differential effects for the variables depending on task. Study frequency has a larger effect size in the JOF task than in the recognition task, whereas item-attribute effects tended to be larger in the recognition task and smaller in the JOF task.

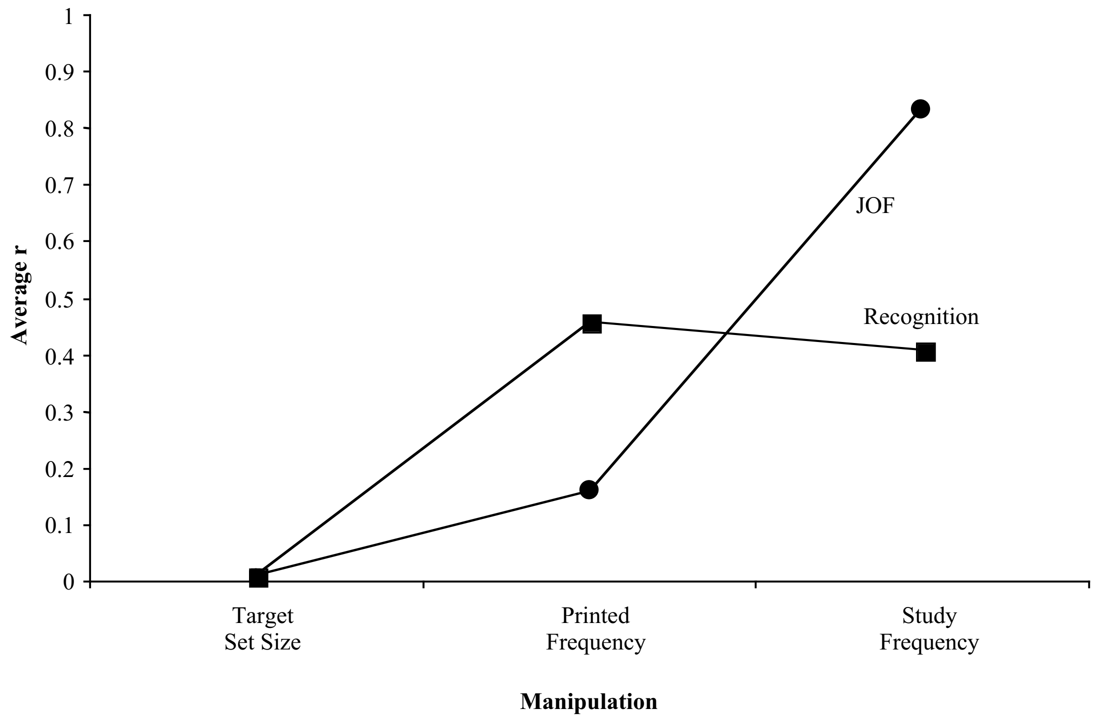

Figure 9. Correlational analysis of effects of target set size, printed frequency, and study frequency on judgments of frequency and recognition d'. For averaging, *r* values were Fisher transformed.

Discussion

The purpose of these experiments was to determine whether JOF's are determined by familiarity, reminding, or some combination of these processes. The results show that item-attribute variables, such as printed frequency and connectivity, affect memory performance in JOF and recognition tasks, but they have stronger effects in the recognition task. We also confirmed and extended Hintzman's (2004) findings in showing that study frequency has a greater effect size on JOF's than on recognition d'. However, it may still be premature to conclude that we have confirmed his conclusion that the reminding hypothesis is the solution to the differences between JOF's and recognition, as familiarity does appear to play a role in the JOF task.

In his 2004 article, Hintzman postulates "JOF may be strongly tied to recollection." That is, if the reminding hypothesis is correct, JOF's involve "recollection that occurs during a study trial, " implying that a search occurs for a recent presentation of the word upon seeing it an additional time. However, when Nelson, Canas, and Bajo (1987) initiated a search process within a recognition task, they were able to find a significant effect of set size. Yet, in the JOF tasks present in this study, no set size effects were statistically significant. Therefore, either the basic assumptions of the reminding hypothesis are incorrect and JOF's are not closely tied to recollection, but can instead be explained through some other means, or researchers need to take a second look at the common path model.

PIER2 depends on the interaction of information previously acquired in everyday life, or implicit memory, and information acquired within an episode, explicit memory, to explain many different phenomena (Nelson, et al, 1998; Nelson & McEvoy, in press). The model assumes that two types of representation are created for a familiar word during study. There is an automatically activated implicit representation of the word and its associates, and an explicit representation resulting directly from the nature of the encoding operations applied to the word while studied (e.g., rehearsal, rating for pleasantness, and so on). In PIER2 strength is not unidimensional as suggested by the common path model. Variables such as printed frequency and influence implicit strength, whereas study frequency affects explicit strength. Participants may encode both implicit and explicit information at study and then differentially rely primarily on one type of information or another in different types of retention test. Perhaps a better description of the common construct model would have strength divided in order to account for implicit and explicit memory strength as seen in Figure 10, to account for item-attribute variables and study frequency, to account for why set size has no effect in recognition or JOF's. Future research could focus on trying to discern the differing levels of strength and their influences on tasks.

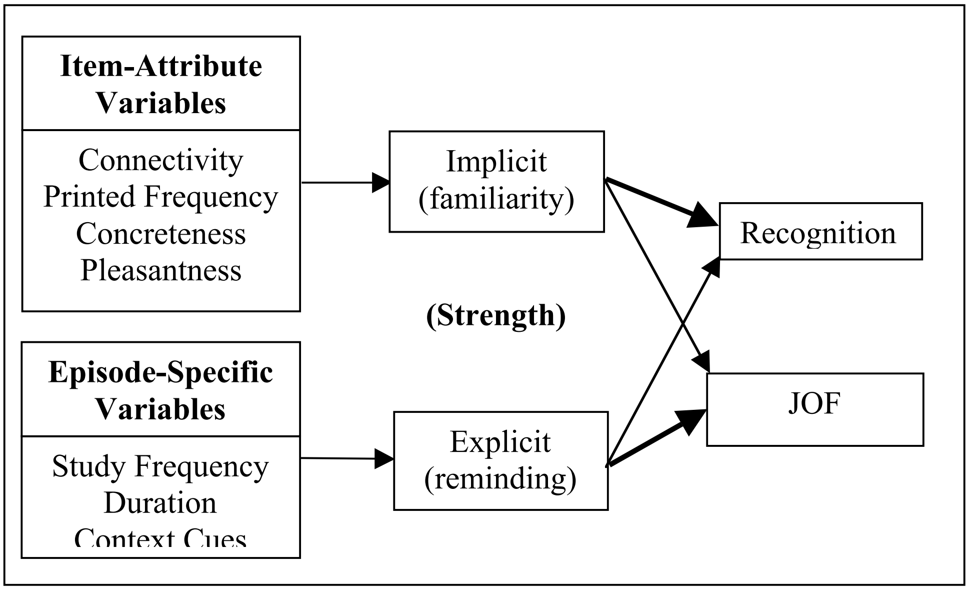

Figure 10. A possible expansion of the common path hypothesis. Thicker lines represent greater influence of certain information on different tasks.

This study can be used to evaluate the relevant significance and validity of several previously held assumptions. In addition to forcing Minerva2, REM, and TODAM to rethink the approach taken with regards to recognition and JOF's, the results obtained in this article also present a problem for the early version of the PIER2 model. In order the explain the printed frequency effect found in recognition, PIER2 relied on Jost's Law, which states that "when two associates are of equal strength, a repetition strengthens the older more than the younger" (Boring, 1957). However, Jost's Law would predict that printed and study frequencies should have interactive effects, with printed frequency effects diminishing with increasing presentations. The effects of item-attribute variables should diminish as the items are learned. Distinctive features at all levels will become less important as the item becomes more strongly encoded in the context. Jost's Law predicts that the effects of printed frequency and connectivity should diminish with increasing numbers of presentations. After a word is seen once in the experimental context, seeing it a second, third, and fourth time in the same context should add less and less to its activation strength. In Jost's terms, such an item would no longer be "old." However, the current studies show no such interactions, implying that Jost's Law is not sufficient to explain printed frequency effects and that PIER2 needs to incorporate a different explanation into the model, such as the associative distinctiveness explanation used here.

Future directions suggested by this work might include varying additional item-attribute variables within a frequency estimation task, such as forward and backward strength, to test their contributions to reminding. Different study instructions might be used in order to examine their influence on encoding and results at test. Researchers may also seek to understand strength as a divisible construct by determining what comprises strength and how it really relates to JOF and recognition. Or, if the reminding hypothesis can be redefined it might still be a valuable tool in discriminating between JOF and recognition. The relationship between familiarity and reminding needs to be further delineated so that researchers can know how the two processes operate in relation to one another. Are they parallel processes? Does one begin where the other ends? Can one operate without the other? Could familiarity be interpreted as weights within a reminding equation? There is still much to be done in understanding how the brain accomplishes frequency estimation and recognition.

References

Boring, E. G. (1957). *A History of Experimental Psychology* (2nd ed.). New York: Appleton-Century-Crofts, Inc.

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, and Cognition, 28*, 1003-1018.

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

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

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

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

Kucera, 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.

Murdock, B., Smith, D., & Bai, J. (2001). Judgments of frequency and recency in a distributed memory model. *Journal of Mathematical Psychology, 45,* 564-602.

Myers, J. L., & Well, A. D. (1995). *Research Design and Statistical Analysis.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Nelson, D. L., Bennett, D. J., Gee, N. R., Schreiber, T. A., & McKinney, V. (1993). Implicit memory: Effects of network size and interconnectivity on cued recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 747-764.

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., & Friedrich, M. A. (1980). Encoding and cuing sounds and senses. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 717-731.

Nelson, D.L, & McEvoy, C. L., (in press). Implicitly activated memories: The missing links of remembering.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). The University of South Florida word association, rhyme and word fragment norms. http://cyber.acomp.usf.edu/FreeAssociation/

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., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review, 99,* 322-348.

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.

Proctor, R. W. (1977). The relationship of frequency judgments to recognition: Facilitation of recognition and comparison to recognition-confidence judgments. *Journal of Experimental Psychology: Human Learning & Memory, 3,* 679-689.

Shiffrin, R. [M.] (2003). Modeling memory and perception. *Cognitive Science*, *29,* 341-378.

Wells, J. E. (1974). Strength theory and judgments of recency and frequency. *Journal of Verbal Learning and Verbal Behavior, 13,* 378-392.

Appendices

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

High Printed Frequency, High Connectivity

# of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
Mean	.08	.73	.85	.88	.96
d'	NA	2.65	3.22	3.39	3.80

High Printed Frequency, Low Connectivity

# of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
Mean	.06	.66	.83	.88	.93
d'	NA	2.31	3.11	3.35	3.63

Low Printed Frequency, High Connectivity

# of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
Mean	.03	.79	.92	.96	.99
d'	NA	3.16	3.77	4.06	4.28

Low Printed Frequency, Low Connectivity

# of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
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

High Printed Frequency, Large TSS

# of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
Mean	.08	.62	.75	.83	.88
d’	NA	2.10	2.55	2.75	3.22

High Printed Frequency, Small TSS

# of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
Mean	.08	.59	.77	.84	.89
d’	NA	1.99	2.75	3.00	3.26

Low Printed Frequency, Large TSS

# of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
Mean	.03	.75	.89	.93	.96
d’	NA	3.06	3.59	3.88	4.01

Low Printed Frequency, Small TSS

# of Study Presentations	0 (False Alarm)	1 (Hit)	2 (Hit)	3 (Hit)	4 (Hit)
Mean	.03	.74	.85	.93	.96
d’	NA	2.96	3.53	3.90	4.07

Appendix E
A Numerical Example of a Correlation Computation

For the 5 words that fall into the High Printed Frequency x High Connectivity x 1 Study Presentation condition, Participant 1 gave frequency estimations of 1, 0, 0, 1, and 1. Therefore, the person's mean for this condition 0.6. This mean and all of the other condition means for Participant 1 were entered into the analysis program as follows: The first column represents the dummy-coding of the printed frequency variable, such that means for the High Printed Frequency condition are associated with number 1 and means for the Low Printed Frequency condition are associated with number 2. A correlation computation is run on the two columns in order to obtain the correlations between Participant 1's scores and the frequency variable.

Variable	Participant's Score
1	0.6
1	2.0
1	3.4
1	3.2
1	0.6
1	1.4
1	2.6
1	3.2
2	1.6
2	1.8
2	1.8
2	2.6
2	1.0
2	2.0
2	3.0
2	3.0