4

THE ROLE OF INTENTION ON FREQUENCY MEMORY

by

JEANETTE L. GONG-ARTIS

A dissertation submitted to the Graduate Faculty in Experimental Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1999

UMI Number: 9924811

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

This manuscript has been read and accepted for the Graduate Faculty in Experimental Psychology: Cognition, Learning and Perception in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

April 21, 1999
Date

Chair of Examining Committee

April 26, 1999
Date

Executive Officer

Dr. Neil Macmillan

Dr. Arthur Reber

Dr. Hollis Scarborough

Dr. Caren Rotello

Supervisory Committee

The City University of New York

ABSTRACT

THE ROLE OF INTENTION ON FREQUENCY MEMORY

by

Jeanette L. Gong-Artis

Adviser: Professor Don L. Scarborough

One debate in cognitive psychology concerns whether frequency memory is affected by intention, that is, whether or not a person intentionally pays attention to the occurrences of stimulus items and their frequencies. Hasher & Zacks (1984), Howell (1973) and others believe that frequency judgments are not significantly different between incidental and intentional groups. In contrast, Greene (1986, 1989), Naveh-Benjamin & Jonides (1986), and Brown (1995, 1997) believe that frequency judgments are better for intentional than for incidental groups. The goal of the new research described here was to examine further the role of intention on frequency processing. These experiments differ from earlier experiments in two ways: 1) the same subject is in both the incidental and intentional conditions and 2) signal detection analysis is utilized to study the effect of intention on frequency memory in a two alternative forced choice frequency judgment task by using a within subject design. By using signal detection analysis, the characteristics of frequency memory under incidental and intentional conditions can be explored. The different models of frequency memory such as the Strength theory and Multiple Trace theory are also reviewed and contrasted. Results show that intention improves frequency processing. It appears that subjects who intentionally process words are more likely to engage in different encoding strategies and therefore, are able to better retrieve from memory the frequencies of these words than those processed incidentally. In addition, the results from the signal detection analysis supports a uni-dimensional representation of frequency memory.

Acknowledgments

This dissertation is only possible from the support and guidance I have received over the years. I want to thank my mentor, Dr. Don Scarborough and my committee: Dr. Neil Macmillan and Dr. Arthur Reber for their wisdom, patience and advice. In addition, I want to thank my outside readers, Dr. Hollis Scarborough and Dr. Caren Rotello for their time, suggestions and support. I want to also thank my husband, Henry Artis, my sister, Ann Marie Gong, and my family and friends for all their patience and consideration through the years. I also want to extend my appreciation to my colleagues at the Infant and Child Learning Center and at SUNY HSCB-Downstate who have offered constant encouragement. Lastly, I dedicate this dissertation to both my parents, Doris Moy Lee Gong and Chuck Chung Lee Gong who are an inspiration to me.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

Ebbinghaus stated that "repetition is one of the most powerful variables affecting memory" (Hintzman, 1976), and knowledge of how often an event has occurred is important in perception and learning. Can you remember how many times you heard the word "impeachment" in the last six months or how many left turns you've made while driving the car in the last year? The answers to these questions tap our memory and touch on the role of frequency processing for both declarative and procedural knowledge. Frequency information about events in our lives can shape our expectations about the future. A person's ability to keep track of frequency information can shape what behavior he/she chooses to exhibit in different situations. For example, if we come in contact with a neighbor who is consistently rude and aggressive, we may choose not to say "hello" to this person because we expect this person to not return the greeting. In another example, it can be assumed that common sense is related to the ability to keep track of repeating events and it's assumed that someone with good common sense, is someone who learns quickly from those experiences. In summary, frequency information about events in our lives can shape our expectations, our behavior, and our thinking and is basic to how we interact in our world.

There has been a debate in the memory literature concerning whether frequency processing is influenced by intention. In their 1984 paper, Hasher and Zacks set several criteria as evidence of the automatic processing of frequency information. They believed that certain aspects of experience are continually registered in memory regardless of age, ability, education or motivation. They have two assumptions (these assumptions and criteria are derived directly from Hasher & Zacks' 1984 paper, p. 1373):

"1) memory consists of a collection of attributes and

2) an individual's momentary capacity for cognitive ability is limited:

a) people differ in systematic and predictable ways in the capacity they have available;

b) mental processes vary in the amount of capacity they require, with automatic processes requiring less and effortful processes requiring more;

c) automatic processes function at optimal levels, continuously & independently of intention for all intact humans, and these processes require only that an event be attended to."

These assumptions led to the development of six criteria that are required to conclude that an attribute is automatically encoded:

"1) People are sensitive to information without necessarily intending to be;

2) Information encoded in this way is no different than when intention is activated;

3) Training at processing such information does not improve encoding and neither does explicit feedback;

4) People differ very little in their ability to encode this information-neither education, social class or culture will influence the ability to process such information;

5) Encoding will be invariant across a wide age range; and

6) Disruptions due to arousal, stress, and/or additional simultaneous processing demands will have no impact on the processing of such information."

In this paper, I am primarily concerned with first and second criteria. Experiments 1 and 2 are designed to test whether intention changes the way frequency information is encoded. According to Hasher and Zacks (1984), there is "considerable evidence that the intention to encode frequency information does not improve the quality of stored knowledge of frequency." What is intention? For some, intentional processing might be the opposite of incidental processing. According to Reber (1995, p. 366), intention is "any desire, plan, purpose, aim or belief that is oriented toward some goal, some end state. Used by most with the connotation that such striving is *conscious*."

Several authors have presented evidence supporting Hasher and Zacks (1984) claim that intention has no effect on frequency processing while others dispute it. Howell (1973) found that intention did not affect frequency judgments of episodic events for subjects given either free recall or frequency estimation instructions. In Howell's experiment, all groups of subjects were presented words that occurred either 0, 1, 2, 4, 6, or 10 times. Subjects were presented a list with either 46, 115, 230 or 368 stimulus items constructed with either 10, 25, 50 or 80 unique words respectively. For example, a list of 46 words had two words each presented at 1, 2, 4, 6, & 10 repetitions (memory load of 10 unique words) while a list of 368 words had 16 words each presented at 1, 2, 4, 6, & 10 repetitions (memory load of 80 unique words). Howell (1973) introduced this condition to see whether the use of a counting strategy for the frequency estimation task would decrease as the number of different words in a list increased.

The words were presented auditorily (12/minute) and shown at the same time on index cards. Following each word, subjects then had 5 seconds to write down the word. Their written response was then covered so that subjects could not review their earlier responses. Half of the subjects were instructed that they would receive a free recall task after the list was presented and the other half were told to expect a frequency judgment task. In actuality, half of the free recall instructed subjects received a frequency estimation task after the stimuli were presented (the **MF** group: expect memory-get frequency; the incidental condition) and half received the free recall task (the **MM** group: expect memory-get memory). Subjects in the free recall condition were given 15 minutes to list all the words they could remember. Likewise for the subjects who were told to expect a frequency estimation task, half of them were actually tested for frequency estimation (the **FF** group: expect frequency-get frequency; the intentional group) and the other half received a free recall task (the **FM**: expect frequency-get memory). For the frequency estimation task, subjects saw the words one at a time and were instructed to write next to each word how often they thought that word occurred in the previous list (including "0"). Howell (1973) calculated the mean frequency estimate for items presented 0, 1, 2, 4, 6, & 10 times and across the memory loads. Howell's 1973 results are shown below.

Table 1.

Mean Frequency Estimates under all 48 Frequency Conditions (Howell, 1973):

		Memory Loads:				
Group	Word Freq.	10	25	50	80	**Mean**
Memory-Frequency	0	0.00	.07	.15	.15	0.09
(M-F)	1	2.08	2.72	1.98	2.28	2.27
	2	2.88	3.15	2.99	3.47	3.12
	4	5.04	5.73	4.50	4.81	5.02
	6	6.04	6.32	5.37	5.91	5.91
	10	7.92	9.08	7.18	6.90	7.77
	Mean	3.99	4.51	3.70	3.90	**4.03**

Mean Frequency Estimates under all 48 Frequency Conditions (Howell, 1973):

		Memory Loads:				
Group	Word Freq.	10	25	50	80	**Mean**
Frequency-Frequency	0	0.00	.15	.08	.02	.06
(F-F)	1	2.00	1.60	1.99	2.54	2.03
	2	2.33	2.63	3.23	3.58	2.94
	4	3.71	4.13	4.61	5.56	4.50
	6	5.50	4.92	6.20	7.14	5.94
	10	9.21	7.43	8.50	8.64	8.45
	Mean	3.79	3.48	4.10	3.58	**3.99**

Howell (1973) found that the free recall instructed group performed better at the free recall task (**MM** (mean %=56.6) > **FM** (mean %=48.8), F(1,88) = 5.98, p < .025) while both the free recall instructed group and the frequency instructed group performed equally well on frequency estimation (**MF**: mean freq. estimates=3.99 = **FF**: mean freq. estimates = 4.03; the difference was not significant). There was also no effect of memory load on frequency performance. Howell (1973) found only a significant effect of repetition frequency, F (4, 352) = 289.53, p < .01. None of the interactions was significant.

The following is a breakdown of the mean percent recall (Howell, 1973):

Table 2.
Percent recall across Frequency (Howell, 1973).

	Frequency of Occurence					
FREE RECALL (% of recall):	1	2	4	6	10	overall mean
Frequency-Memory (**FM**) means:	25	41	51	58	69	48.8
Memory-Memory (**MM**) means:	31	47	61	72	72	56.6

For Howell, intention refers only to frequency coding per se. But, both of his groups can be said to be in the intentional condition in that they both are told to study the stimulus words for a later memory test.

Hasher and Cromiak (1977) also believed that frequency processing is automatic and that informed subjects perform no differently than uninformed subjects. Half of their subjects were instructed to judge the frequency of words presented and half were given "nonspecific memory instructions" (subjects were told that they should remember the words so they could answer questions about them later; Hasher and Cromiak, 1977, p. 174-175). There were 40 subjects each from 2nd grade, 4th grade, 6th grade, and college (age ranges from 7 years old to early twenties). The presentation frequencies were 0, 1, 2, 3, and 4. The design was grade (4) X instruction (2) X frequency of occurrence (5). Each item was exposed for 4 seconds during presentation of the stimulus list. Following the list, subjects were asked to estimate how often an item occurred and were not given any upper limit on their frequency estimates. Hasher and Cromiak (1977) calculated two different dependent measures: 1) the absolute number of items whose frequency estimates were correct; and 2) the unsigned difference between each estimate and the actual frequency of occurrence. Hasher and Cromiak (1977) preferred the deviation method and calculated the mean unsigned (absolute) deviation for each subject at each frequency.

They found that the estimates increased with actual frequency, ($F(4, 608) = 624.37$, $MSe = .44$, $p < .05$). Subjects produced more accurate estimates for infrequent items than for frequent items, ($F(4, 608) = 128.57$, $MSe = .19$) and informed subjects produced larger frequency estimates than uninformed subjects (grand mean of the median frequency estimates: informed: 1.93 versus uninformed: 1.75, $F(1,151) = 7.82$, $MSe=.80$). These data go against the notion that intention does not make a difference in frequency judgments because informed subjects produced larger frequency estimates. The mean deviations are listed below across frequency. The informed group displayed increased accuracy (smaller deviation values) but only at the higher frequencies, $F(4, 608) = 9.49$, $p < .01$. No interaction was found for grade and frequency estimates, $F(12, 608) = 1.12$ or for grade, frequency estimates, and instruction, $F < 1$. Here, grade did not matter in terms of frequency performance.

Table 3.
Means of the Mean Unsigned Deviations (Hasher & Cromiak, 1977):

Freq.	0	1	2	3	4
Informed Condition	.18	.56	.78	.80	.95
Uninformed Condition	.14	.53	.91	1.13	1.39

Meanwhile, Zacks, Hasher, and Sanft, (1982) wanted to explore three variables that were expected to influence recall but not frequency judgments:

1) accuracy of expectation about memory tests

2) competing demands of capacity

3) individual differences

The authors wanted to look at the effect of informing subjects with accurate information about the forthcoming test versus providing them with misleading information. One third of the subjects was informed about the frequency test, another third was informed about the free recall test, and the last third was informed about both tests. Each of the three groups was further divided with half receiving a frequency test and half receiving a recall test. All subjects saw lists of words with frequency of occurence ranging from 0 to 7. Each word was presented for 4 seconds. Individual differences were examined by comparing the performance of students from two universities:

University A: small, selective, private: mean SAT verbal= 610

University B: state university, liberal admission policy: mean SAT verbal= 471

They found that, in terms of free recall results, there was a significant effect for university type ($F(1,171) = 11.56$, $p < .01$) and instructions ($F(2, 171) = 11.61$, $p < .01$). Closer examination revealed that University A students recalled more than University B subjects in the two conditions in which a recall test was expected ($F(1,171) > 10.89$, $p < .01$); but when a frequency test was expected, University A students performed no better than University B students (mean proportion correct on a forced choice test where subjects chose the most frequent item of a pair was .85 for both University A (SD.=.10) and University B (SD.=.08); $F<1$). In terms of the performance on a forced-choice frequency test, frequency accuracy was invariant across all instructional conditions as well as across the two populations. Neither test expectations, competing demands on capacity (asking subjects to prepare for both memory tests), nor individual differences influenced the encoding of frequency information. For contrast, for performance on free recall,the variables of test instruction, competing demands on capacity and individual differences (university) did play a role. Zacks, Hasher and Sanft (1982) believe that frequency encoding requires "minimal capacity and is independent of the use of conscious, effortful strategies." Below is a summary of the means for each instruction and test condition across both universities (the standard deviations are in parentheses, Zacks, Hasher, and Sanft, (1982)).

Table 4.
Mean proportion correct for free recall and forced choice frequency tasks (Zacks, Hasher, & Sanft, 1982)

RECALL CONDITION - mean proportion recalled

Instruction:	Recall (SD)	Frequency (SD)	Both (SD)
University A:	.74 (.25)	.50 (.23)	.65 (.18)
University B:	.60 (.11)	.55 (.13)	.50 (.13)

FORCED CHOICE FREQUENCY CONDITION - mean proportion correct for selecting the most frequent member in a forced-choice task

Instruction:	Recall (SD)	Frequency (SD)	Both (SD)
University A:	.82 (.10)	.85 (.10)	.87 (.12)
University B:	.81 (.14)	.85 (.08)	.86 (.11)

In contrast, Greene (1986,1989) found that intention played a significant role in how well frequency information is remembered. In Greene (1986), subjects were told that they would be tested on their memory for strings of five digits presented during the experiment (1 second per string). As a distractor task, subjects were required to repeat a word out loud after each string of digits. For example, during the experiment, a subject saw a string of five digits (e.g., " 2, 5, 1, 7, 3") and then a common word. The subject was required to say the word out loud every time an asterisk flashed on the screen (either once or twice in a 10 second rehearsal period). After that, the subject was required to recall the digits. After 96 trials, the subject counted backwards from 200 for one minute and then was tested for frequency memory of the words. To test for frequency, subjects were presented with 32 word pairs and asked to chose the member of each pair that was spoken twice. There were three different groups of subjects instructed in three different ways: the ***incidental*** group was not informed of the frequency test; the ***weak-intentional*** group was instructed that the digit recall task was more important than the frequency task; and the ***strong-intentional*** group was told that the frequency task was more important than the digit task. Greene found that the mean proportion correct for the frequency task was .85 for strong-intentional group, .72 for the weak intentional group, and .69 for the incidental group. Intention mattered for the the strong-intentional group but not for the weak-intentional group or the incidental group.

In another Greene experiment (1989), subjects viewed a list containing words that occurred either once or twice. Each word was presented for 10 seconds. While viewing the list of words, subjects were asked to find the rule that determined when a noun or adjective occurred on the list. Half the subjects were warned to memorize the words (intentional condition) and half were not warned (incidental condition). After the list was presented, subjects were given a relative frequency discrimination task. Each test pair contained one word that occurred once and one word that occurred twice. The subjects judged which member of a pair occurred more often. Greene found that the intentional group was more accurate than the incidental group (mean proportion correct for frequency discrimination: intentional = .73; incidental = .60). Therefore, Greene concluded that intention did affect how well subjects processed frequency information. Below are the mean proportion correct data for frequency discrimination responses for both incidental and intentional conditions as a function of the spacing between repeated words (Greene, 1989):

Table 5.
Mean proportion correct for both incidental and intentional conditions across different spacings (Greene, 1989).

Spacing-->	0	1	2	4	8	16	overall mean
Incidental:	.59	.60	.59	.60	.60	.60	.596
Intentional:	.62	.72	.72	.76	.77	.80	.732

Greene (1989) found a significant effect of intention, $F(1,46) = 16.24$, $MSe = 6.25$ and a significant interaction for spacing in the intentional group only, $F(5,115) = 8.34$, $MSe = 1.65$. Performance of the incidental and intentional groups at 0 lag is almost identical. In this paper, Greene (1989) proposed the ***Deficient-Processing Subtheory*** which he believes explains spacing effects in recognition tasks and frequency judgments. When a subject is presented with a second or repeated presentation of an item, the subject assigns less processing or rehearsal time compared to items presented only once because the item is familiar. As the spacing between repeated presentations increases, so does the amount of "contextual" information processed and memorized between the repetitions. For example, the contextual information may be the elements that are different since the last presentation of the word or it may be the items presented prior to the repeated word or the features of the word the subject decides to process (such as the meaning or typography). The more context information the subject decides to process while memorizing the word, the more features the subject can use to access that word later. Therefore, the longer the lag between repetitions, the more context information will be stored and thus, better retrieval (Study-Phase Retrieval Subtheory, Greene, 1989). Since there is no spacing between repetitions at 0 lag, subjects may process very little contextual information because they may think that they know the word very well and therefore, put in very little effort in processing the second presentation of a word.

Greene stated also that by manipulating intention, one manipulates the type of strategies subjects use. Greene believed that the question here is whether the strategies that subjects follow at the time of encoding affect later performance on frequency estimation. In Experiment 3, Greene (1986) attempted to manipulate encoding strategies directly to see how they affect frequency estimation. All subjects were informed that they were in an experiment to see how the use of strategies affects memory for frequency. He instructed subjects in three different ways:

a) To form a sentence mentally for each word (semantic processing)

*It's not clear whether subjects made up a new sentence for each repetition or could use the same sentence over;

b) To try to keep count of the frequencies of each item (counting);

c) To repeat each word silently (repetition).

Subjects saw a list of 100 words (40 one-syllable words presented either once, twice, three, or four times). After the list was shown (each word presented for 10 seconds), subjects were given a list of the words and asked to estimate how often each word had occurred. Below is a table of the data from this experiment (Greene, 1986, Experiment 3):

Table 6.
Mean Frequency Estimates across actual frequency and group type (Greene, 1986).

		TRUE FREQUENCY		
Group	1	2	3	4
Sentence	1.11	2.00	2.79	3.74
Counting	1.33	1.94	2.53	3.27
Repetition	1.24	1.82	2.64	3.19

There was a significant effect of frequency ($F(3,135) = 169.68$, $MSe = 22.94$) and no main effect of group. That shows that subjects' estimates increased with frequency of occurence. In addition, there was an interaction between actual frequency and group, F (6, 135) = 3.29, MSe = 6.98. Newman-Keuls tests show that subjects in the *sentence* group (mean accuracy = 0 .70, SD = 0.34) were significantly more accurate than the *repetition* (mean accuracy = 0.53, SD = 0.58) and *counting* (mean accuracy = 0.54, SD = 0.59) groups and that the latter two groups did not differ significantly from each other. The ineffectiveness of counting as a mnemonic strategy agrees with the results of Zacks et al. (1982). Below is a table of mean accuracy in the frequency estimation task from Greene's 3rd experiment (1986). The "Correct" column, specifies the proportion of items given the correct frequency. The "Deviation" column is the mean unsigned deviation between frequency estimate and true frequency, and the "Correlation" column is the mean correlation between true and estimated frequency. These data support Greene's hypothesis that the type of strategies that subjects use influences how well they perform on frequency estimation tasks. Greene believes that intention improves frequency processing because "it affects the probability that subjects will engage in meaningful processing of the items" (Greene, 1986, p. 494). Greene (1986, p. 492) also points out that "intentional learning leads to an increase in semantic processing."

Table 7.
Mean proportion correct, mean deviation, and mean correlation across group type.
(Greene, 1986).

	MEASURE OF ACCURACY		
Group	Correct	Deviation	Correlation
Sentence	.70	0.34	.84
Counting	.54	0.59	.70
Repetition	.53	0.58	.70

There was a significant effect found for instructional group and number correct, $F(2,45) = 8.24$, $MSe = 25.41$; unsigned deviations, $F(2,45) = 9.81$, $MSe = 51.13$, and correlations between actual and estimated frequency, $F(2,45) = 7.93$, $MSe = 0.01$ (Greene, 1986). Greene (1986) reports that Newman-Keuls analyses show that subjects in the sentence group were more accurate than those who used repetition or counting.

In another experiment on the effects of intention on frequency judgments by Naveh-Benjamin and Jonides (1986), subjects in the intentional condition were told to expect a frequency test after the list while subjects in the incidental condition were "told to pay attention to the words in order to answer some questions about them later." Naveh-Benjamin and Jonides (1986) believed that this type of instruction allowed them to more closely replicate the "usual" incidental conditions in previous experiments. A third of the subjects saw each word presented for 1 second, another third for 2 seconds and the last third for 4 seconds. Subjects saw a list of 42 two-syllable high frequency Hebrew words. There were seven words selected to appear at each frequency of occurence from one to six. Along with additional words to counteract primacy and recency effects, subjects saw a list with 161 words in all. After the stimulus list was presented, subjects were given a sheet with a list of these words mixed with 42 new words. Subjects were informed that the frequency test included words that were not on the stimulus list. They were instructed to write a frequency from 0-6 next to each word.

Naveh-Benjamin and Jonides (1986) analyzed the data by calculating the slope relating the frequency estimates to the actual frequency of occurence. A slope estimate was calculated for each subject individually and then these were averaged across subjects for each condition. According to the Naveh-Benjamin and Jonides (1986, p. 379), "this slope is an indicator of sensitivity to variations in frequency. As the slope becomes closer and closer to unity, one can conclude that subjects are more and more sensitive to variations in presented frequency." Naveh-Benjamin and Jonides (1986) performed a linear regression analysis on the estimates. The second measure they calculated was the amount of deviation between each frequency estimate and the mean judgment. The larger the deviation, the less consistent and reliable are the judgments.

The authors found that the slope estimates (comparing actual versus estimated frequency) were significantly higher (closer to 1) for the intentional condition than for the incidental condition for the 1 second and 2 second exposures but not for the 4 second exposures ($t(18) = 2.23$, $p < .05$ for 1 second.; $t(18) = 2.20$, $p < .05$ for 2 seconds; $t(18) = .58$, not significant for 4 seconds). Unfortunately, the authors do not report any significant instruction x exposure time interaction for the mean slope values. Below is a summary of the mean slope values across exposure times for both the incidental and intentional conditions (standard deviations are within the parentheses, Naveh-Benjamin and Jonides, 1986):

Table 8.

Mean Slope Values as a Function of Learning Instructions and Exposure Time (Naveh-Benjamin & Jonides, 1986):

Exposure times-->	1 second mean (SD)	2 seconds mean (SD)	4 seconds mean (SD)
Incidental:	.40 (.17)	.45 (.12)	.54 (.13)
Intentional:	.59 (.19)	.58 (.13)	.56 (.11)

The authors then examined the mean frequency estimates and found that subjects in the intentional condition tended to give more consistent estimates (less occurence of underestimation or overestimation of frequency) than those in the incidental condition, especially for words exposed for 1 and 2 seconds. To calculate the mean frequency estimates, "estimates were averaged across each of the actual frequencies to yield mean absolute estimates that were then averaged across subjects" in each exposure condition (Naveh-Benjamin and Jonides, 1986, p. 380). There was a significant interaction found between the instruction condition and actual frequency for 1 and 2 second exposure times (1 second: $F(5,90) = 2.52$, $p < .05$; 2 seconds: $F(5, 90) = 2.43$, $p < .05$). At the 4 second exposure time, there was no effect of instruction nor was there a significant interaction between instruction and frequency.

Naveh-Benjamin and Jonides (1986) also calculated the mean variability of estimates by measuring each subject's standard deviation of the estimates for each frequency. These were then averaged across actual frequency to get a a mean standard deviation for each subject. Below is a short summary of the ***mean standard deviations*** of the frequency judgments across instruction conditions and the three exposure times (standard deviations in parentheses, Naveh-Benjamin and Jonides, 1986):

Table 9.
Mean standard deviations across instruction and exposure times (Naveh-Benjamin & Jonides, 1986):

Exposure Times--->	1 second mean(SD)	2 seconds mean(SD)	4 seconds mean(SD)
Incidental:	1.41 (.28)	1.20 (.26)	1.16 (.24)
Intentional:	1.13 (.23)	1.08 (.31)	1.22 (.30)

Again, Naveh-Benjamin and Jonides (1986) report no significant instruction x exposure time ANOVA for mean standard deviations. There was no effect of intentional instruction at the longer exposure times, but there is an effect at the shortest exposure time ($t(18) = 2.24$, $p < .05$) and Naveh-Benjamin and Jonides (1986, page 384) took their cue from Fisk and Schneider (1984) who proposed that automatic coding of frequency may be an "early asymptote" controlled process and concluded that "modification of long term memory for frequency information may occur in the first few seconds of processing (before it asymptotes)...therefore, before it reaches the asymptote in performance, intention to code frequency can still impact on later judgments. This can be mediated either by the amount of effort expended or by the type of strategy used, or both."

Naveh-Benjamin and Jonides (1986) thus believe that frequency processing is not completely automatic because the data from their experiments show that strategy and intention can influence it, contradicting the criteria for automaticity listed by Hasher and Zacks (1984). The authors further suggest two points:

a) both automatic and non-automatic processes are not discrete entities but perhaps are part of the same continuum. Frequency processing falls somewhere along that continuum;

b) both automatic and non-automatic processes are at work when frequency encoding occurs and that sometimes experimental manipulations affect those variables (i.e., levels of processing; competing task demands) that have a direct influence on frequency encoding.

To summarize, Howell (1973), Hasher and Zacks (1984), and Zacks, Hasher, and Sanft (1982) found that performance for frequency judgments was not significantly different between the incidental and intentional groups. On the other hand, Greene (1986, 1989) and Naveh-Benjamin and Jonides (1986) found that performance for frequency judgments was better for intentional groups than for incidental groups. Thus, the role of intention in frequency processing is still unclear.

Why are there differences in results between these two camps? After reviewing the Naveh-Benjamin and Jonides (1986) experiment, the presentation times of the other experiments discussed here were reexamined. The presentation times ranged from 4 to 10 seconds. Below is a short summary listing the presentation times of those experiments.

<u>No Effect of Intention:</u>

Howell (1973) - 5 seconds

Zacks, Hasher & Sanft (1982) - 4 seconds

<u>Effect of Intention:</u>

Greene (1986, 1989) - 10 seconds

Naveh-Benjamin & Jonides (1986) - 1, 2, & 4 seconds

Because Naveh-Benjamin and Jonides (1986) ceased to find significant effects of intention at four seconds, one might expect Greene (1986, 1989) with a presentation time of ten seconds not to have a significant effect of intention, but he does. Therefore, it is not clear that processing times explain the difference in the results for these experiments.

Another factor to consider is that the instructions utilized in each experiment may differently affect the processing of frequency. In Howell's experiment (1973), subjects were told to prepare for a memory test or for a frequency test. In retrospect, both these types of instruction motivate subjects to prepare for a memory task which does not produce differences in the type of processing for subjects in Howell's experiment. Greene (1986) suggests that "subjects do not necessarily use a different strategy or process when trying to learn the frequency of items than when they are trying to identify the items. The crucial comparison should be between one group of subjects who know they will be tested on the items and another group who learn the items incidentally." Greene's third experiment (1986) shows that strategy does matter in frequency processing. Subjects' performance can be influenced by manipulating strategies (such as semantic processing or counting) at the time of encoding. Therefore, in experiments (Howell, 1973; Hasher and Zacks, 1984) that found no effect of intention, we might conclude that the instructions produced <u>no</u> change in the way subjects processed frequency information under both incidental and intentional conditions. Perhaps it is not really intention but different strategies used under incidental and intentional conditions that produce different results in frequency judgments.

Other work by Naveh-Benjamin and Jonides (1986) also showed that subjects who process items intentionally used more strategies or greater depth of processing (i.e., semantically versus phonetically) and also produced more accurate frequency judgments. Intention may alter the quality of the knowledge represented in memory on which subjects base their frequency judgments. Or as Brown (1995, p. 1539) so aptly states in his paper, "Participants use multiple strategies when estimating event frequencies; that strategy selection is determined, in part, by the contents of memory; and that the magnitude and accuracy of participants' frequency estimates are related to the strategy they select. Strategy selection can have a strong effect on estimation performance. This means that performance differences found in frequency estimation tasks can be very difficult to interpret. They may reflect differences in strategy use, or they may reflect both strategic and representational differences."

Theories of Memory Representation

Cognitive psychologists have proposed many kinds of memory such as episodic, semantic, procedural, declarative, implicit and explicit memory. If there is more than one form of memory and therefore, more than one way frequency information might be represented, the differences in the literature may reflect the involvement of these distinct memory forms. Therefore, this section outlines some ideas about these memory systems.

Recent work supports the idea that there are both explicit and implicit memory systems. According to The Penguin Dictionary of Psychology (Reber, 1995, p. 259), **explicit memory** is "characterizing that which is consciously known; explicit cognitive processes are those the person is aware of using." In contrast, **implicit memory** (Reber, 1995, p. 347) is defined as "covert, tacit and on occasion unconscious." Others such as Richardson-Klavehn and Bjork (1988) and Schacter (1987) note that the term **explicit** can also refer to memory systems and tasks that are ***declarative, autobiographical, direct,*** *episodic, and intentional.* On the other hand, the term **implicit** has been used to define memory systems that are *indirect, nonconscious, unintentional, nondeclarative, and incidental.*

Anderson (1990) states that explicit memory and declarative memory are synonymous because both terms represent memories that are consciously processed and known. Reber (1995, p.385) states that **declarative** memory is "knowledge about the world that can be represented as consciously known, factual knowledge." In contrast, procedural memory and implicit memory are similar because they involve "knowing things we cannot describe" or having "memories we are not conscious that we have". Reber (1995, p. 385) states that **procedural** memory is "knowledge that lies outside an individual's realm of consciousness; lies behind complex actions and is resistant to attempts to make it conscious." This becomes evident in real life when we start to verbally describe what we actually do when we drive a car or type a letter. We gain so much expertise in skills that we can perform them without any conscious awareness of the steps involved.

Examples of implicit memory can also be seen in patients suffering from Korsakoff's Syndrome because they are people who suffer from amnesia but retain and gain implicit memories that they consciously cannot recall (Graf, Squire, and Mandler, 1984). Graf, Squire, and Mandler (1984) found that amnesic patients performed worse than normal subjects for a previously seen word list on a free recall task (an explicit memory task) but performed as well as normal subjects on a word completion task (an implicit memory task). A word completion task involves presenting subjects the first three letters of a word and asking them to fill in the other letters to make an English word (for example, **bal** _ _ _ _ --> **bal** l o o n). In this experiment, the amnesics exhibited implicit memory for the word list in the word completion task.

Another example that supports the distinction between explicit and implicit memory is that of the patient H.M. who had parts of his temporal lobes removed as a curative for epilepsy. Because of this surgical procedure, he experienced the inability to learn new declarative information (anterograde amnesia) but he showed improvement in procedural performance on perceptual motor tasks and various cognitive tasks such as the Tower of Hanoi problem (Anderson, 1990). In the Tower of Hanoi problem, subjects are shown three pegs with a group of five different sized disks situated on the first peg with the largest disk on the bottom and the smallest disk on top. Subjects are required to move all the disks from the first peg to the third peg with the restriction that only the top disk on a peg can be moved and a larger disk cannot be placed on top of a smaller disk. When presented with the Tower of Hanoi problem, H.M. was able to learn as well as the normal subjects and could continue to solve the problem for a year after the initial training. Each day, H.M. did not recall seeing this puzzle but nonetheless displayed improved performance while solving it.

As the previous examples show, there is a difference in performance by subjects dependent on whether the task is an implicit or explicit one. Schacter (1987) and Richardson-Klavehn and Bjork (1988) have discussed the dissociation between performances on implicit and explicit memory tasks and that some variables appear to have different effects on explicit and implicit memory tasks. For example, Jacoby (1983) presented words such as "woman" in one of three ways: a) the word "woman" alone (no context); b) as part of an antonym word pair such as "man-woman" (semantic context); or c) a word such as "man" was presented and the subject generated the antonym "woman" (generative processing). Following this, half the subjects were asked to select the old words shown in the experiment from a list of old and new words (explicit task). The other half were presented words quickly (40 milliseconds) and asked to identify them (implicit task). Subjects in the semantic context and generative processing conditions performed best on the recognition task while subjects in the no context condition performed best on the word identification task. Please see the table below:

Table 10.
Mean proportion correct across learning conditions and testing conditions.

Ability to recognize a word in a memory test versus ability to identify it in a perceptual test as a function of how the word was originally studied (Jacoby, 1983):

	Experimental Condition		
	No Context	Context	Generate
Perceptual Identification:	.82	.75	.64
Recognition Judgment:	.55	.64	.70

In fact, 'no context' subjects were not able to discriminate between old and new words (proportion correct = .55 almost chance) but could identify the words in the word perception task (proportion correct = .82). On the other hand, subjects in the generative processing conditions did better in the old/new word recognition task but did more poorly in the tachistoscopic word identification task. Jacoby (1983) concluded that in the no context condition, subjects may not have been able to recognize that they had seen a word before but they were helped (even though they were presumably unaware of it) by the prior exposure of a word in the perceptual identification task. For Jacoby (1983), the subjects' performance on the no context condition showed a dissociation of explicit and implicit memory.

Hasher and Zacks (1984) claimed that frequency memory is unaffected by intention. As noted above, implicit memory has often been described as being unaffected by intention, and therefore, might be a basis for frequency memory. However, there is evidence that intention, per se, may not be important for explicit/declarative memory as well.

What role does intention play in explicit and implicit memory tasks? Reber (1995, p.

366) defines intention as "generally, any desire, plan, purpose, aim, or belief that is oriented toward some goal, some end state." The effects of intention can be measured from subjects' performance across different experimental conditions. Intention can be used as a manipulation when the experimenter instructs half the subjects on what information they will be directly tested on. In other words, intention can affect how subjects process the material to be remembered and how they attend to that material. Intuitively, when adult subjects (such as Introduction to Psychology subjects) are told to memorize a list of words because they will be tested on it, a majority of them will hopefully be motivated to use methods to process the material as deeply as possible. With greater depth of processing, or what Anderson (1990) terms "elaborateness of processing", the more fully subjects process the material, the better the memory.

According to Anderson (1990), "intention to learn is irrelevant to the amount learned. What is relevant is the way in which the information is processed." Anderson (1990) cites Postman's (1964) belief that what is really crucial is how the person processes the stimuli when they are presented during the experiment. Postman (1964) believes that if subjects utilize the same memorial processes when trying to memorize something, whether intentionally or incidentally, performance will be the same. For example, in Hyde and Jenkins (1973), subjects saw a list of 24 words with each word shown for 3 seconds. One group was told to check each word for the letter "e" or "g" while another group was asked to rate the pleasantness of each word (deeper processing). In addition, half the subjects were told they should learn the words because they would be tested on them later (intentional) and half were kept in the dark (incidental). As can be seen in the data below, Hyde and Jenkins (1973) found that subjects exhibited better recall in the pleasantness rating condition regardless of whether they were in the intentional or the incidental condition (68% versus 69%). Here, intention caused no significant difference in performance because subjects processed the material in the same way under both incidental and intentional conditions.

Table 11.

Percent correct across incidental and intentional conditions (Hyde and Jenkins, 1973).

	Orienting Task	
Condition	Pleasantness	Check Letters
Incidental	68%	39%
Intentional	69%	43%

Perhaps such processing differences can explain the opposing results of those who find no effect of intention on frequency memory (Howell, 1973; Hasher and Cromiak, 1977; Zacks, Hasher and Cromiak, 1982) and those who do find an effect (Greene, 1986, 1989; Naveh-Benjamin and Jonides, 1986). These experiments may differ in how subjects processed the stimuli under incidental and intentional conditions.

The **levels of processing** theory (Craik and Lockhart, 1972; Lachman, Lachman, and Butterfield, 1979) describes the representation of memory not by a multistore model (i.e., STM and LTM) but by what types of processing are done on the information as it is being memorized. As the information is initially being worked on, it goes through a perceptual-conceptual analysis which includes the sensory and semantic analysis of the stimuli. The type of the perceptual-conceptual analysis that is performed is influenced by the subject's intentions or motivations. Words that subjects are told to attend to may be processed at deeper levels. Craik and Lockhart (1972) believe that if a subject wants to remember information well, the level of processing must be deeper than if the subject wants to just remember what color the word is printed in. The result of the perceptual-conceptual analysis is coded stimulus information. For every type of analysis the stimuli have gone through, a different code results. According to Lachman, Lachman and Butterfield (1979, p.274), the subject's memory depends on the "nature of the codes, not on the properties of the memory stores." The different codes produced from the perceptual-conceptual analysis have different "intensities" or permanence. For example, the sensory codes (e.g., the color the word is printed in) have less permanence than the semantic codes (the meanings of the word). The longevity of a code can be increased through "recirculation" or updating it by, for example, repeating the stimulus word (repetition). In comparison, processing a word into a sentence or using a mnemonic is still a better way to retain a word in memory than just repetition. Recirculation or repetition is not as effective.

Craik and Tulving (1975) performed an experiment on incidental learning. They told subjects that they were participating in an experiment about perception and reaction time when in reality they were to be tested on their memory. Subjects were presented a question that they were required to answer about a following word. These questions were used to manipulate the levels of processing of the stimulus word (structural, phonemic, category, and sentence type questions). After each word was presented, the subjects were required to answer a yes-no question by pressing a key. Craik and Tulving (1975) measured how accurate subjects were and how long it took them to answer the questions. After they finished seeing all the words, subjects were given either a free recall or recognition task. Craik and Tulving (1975) found that the words that were deeply processed were recalled more. Some researchers proposed the "total time hypothesis" (Lachman, Lachman and Butterfield, 1979, p. 276) which "holds that retention increases with processing time: the longer a person processes an item, the better he should remember it." Craik and Tulving (1975) did not agree with this because there are types of shallow processing that can take longer to do than some higher level processing. It's not how long you process the material but how complexly the material is processed.

Later research showed that the levels of processing theory generally applies to verbal learning experiments (Lachman, Lachman and Butterfield, 1979). That is, the deep levels of processing help subjects do well on recall and recognition tests but researchers began to realize that retrieval factors also influence performance. Shallow processing may actually help subjects perform better on some tasks. Recall that subjects in the "no context" condition in Jacoby's (1983) experiment performed better than the subjects in the "semantic context" and "generative processing" conditions in the tachistoscopic word identification task. Bransford, Franks, Morris, and Stein (1978) put forth the idea of **transfer-appropriate processing.** The idea is that there is no level of processing better than another (shallow processing versus deeper processing) but that, different levels of processing are better for different types of memory retrieval tasks.

Explicit and implicit memory are also related to episodic and semantic memory respectively. Lachman, Lachman, and Butterfield (1979) summarize the differences between the two types of memory as follows. **Episodic** memories are often referred to as memories that are personal and autobiographical and people often can remember the context (i.e., time, location) of when the experience happened. **Semantic** memory is often referred to as a person's knowledge of words, rules, symbols, meanings, grammar, and word relations. Often, people cannot remember precisely how or when they learned this information.

Where does frequency memory fall in this division of memory? Most frequency estimation tasks are episodic in nature because subjects are asked to judge the frequency of stimuli in a particular experimental context. On the other hand, semantic memory comes into play because the subjects in frequency estimation tasks must be able to read the words that are being presented to them. Therefore, a majority of frequency estimation experiments involve an interaction of both episodic and semantic memories. With this background, I will next review theories of how frequency information is represented in memory.

<u>How is frequency information represented in memory?</u>

There are essentially three theories put forth to explain how frequency is represented (Hintzman, 1976; Howell, 1973). The first, **Strength theory**, is influenced by learning theory. Here, repeated presentations are believed to strengthen the memory trace in only a quantitative and not qualitative fashion. Also, it is assumed that the subject can assess the strength of the trace if asked to make a frequency judgment. With Strength theory, the representations of individual traces or episodes are not available and therefore, frequency is mainly coded in terms of cumulative strength (a sum total).

The second theory asserts that frequency can be represented by multiple traces. The **Multiple Trace theory**, assumes that each presentation may create its own memory trace and that all traces co-habitate with one another in the cognitive system. Since each presentation may form a new memory trace, each trace may contain what Hintzman (1976) called different "attributes". The most obvious difference between the Strength and Multiple Trace theories is that the latter maintains the individuality of the memory traces.

How does someone make a frequency judgment under the Multiple Trace theory? In a multiple trace simulation model such as MINERVA 2 (Hintzman, 1988), a probe activates all memory traces and they respond simultaneously producing a single, composite echo. The contribution of each echo is determined by its similarity to the memory traces. The intensity of the echo signifies how much the traces in memory match the probe. This echo intensity acts as a similarity signal and can be used when making frequency judgments.

Propositional Encoding is the third frequency theory. Here, frequency is represented in a propositional form. There are two possible ways frequency can be encoded. The first way assumes that each time the same item is presented, a new encoding is created. This mimics the Multiple Trace theory. The second way assumes that there is only one proposition that may represent frequency, and this may be updated with each new presentation. Also, Propositional Encoding assumes that with each new repetition or presentation, the subject encodes frequency information and context information. That is, related associations (eg., see dog and think of cat), the modality the item is presented in (auditorially or visually), and/or the information from previous repetitions. Since this is a more interactive theory, the presentation aspects the subject chooses to attend to may play an integral role on what type of information is processed and remembered. Hintzman (1976) noted that the Propositional Encoding system may operate a "counter" to keep track of frequency (like the Strength theory) but may encode individual presentations with their context information (like the Multiple Trace theory). These propositions can be hierarchically structured and therefore, quite complex. Propositions may be embedded within other propositions, and therefore, frequency information may be embedded within other frequency information. In this type of representation, factors such as intention and processing strategies should influence the structure of the information processed.

One way to examine the retrieval of frequency information is to look at response latency as a dependent variable (Howell, 1973). According to the Strength theory, strength may affect latency . The effect of strength on latency should depend on how subjects translate strength into a frequency estimate. Or for the Propositional Encoding theory, if frequency is encoded at the time of presentation and the subject retrieves the output of a pre-existing count judgment, then the response latency plotted against frequency should produce a flat, horizontal function.

For the Multiple trace theory, each presentation is represented by a memory trace. What type of function would result if the subject counts memory traces? One could assume a linear function with reaction times that increased with increasing frequency. One problem with this notion is how does the subject determine at what point all memory traces for a particular event or item have been found. What is the criterion for terminating trace retrieval? The function is not linear according to research performed by Voss, Vereb, and Bisanz (1975) who found that the response latency resembled an inverted "U" curve because reaction times increased as frequency increased up to eight presentations and decreased at higher frequencies.

Hockley (1984) examined latencies and the lag times between repetitions (range: 0-20) to discover what type of information subjects utilized in making frequency judgments. As subjects viewed repeating words, they were required to press keys noting whether this was a first, second, or third presentation of the word (P1, P2, or P3). Hockley believed that there was a difference between the accessibility and sources of information pertaining to frequency as lag time increased between two presentations. The assumption was that during short lag times, subjects could use two sources of information to make judgments (item information and response information about the last repetition from STM) and during long lags only one source of information (item information retrieved from LTM). Results showed that as lag time got bigger between presentations, the mean latencies for correct judgments increased and the proportion of correct responses decreased as presentations rose (percentage correct: P1=97%; P2=85%; P3=74%). Proportion of correct responses also decreased as a function of test lag.

With regard to the Strength theory, Hockley (1984, p. 230) believed that a Strength model that predicted a mean and a variance that "increased proportionally with the number of repetitions" would adequately describe his data. Each presentation of an item would increase the strength of the distribution and the size of that increase would get smaller as the frequency got bigger. This would establish a series of overlapping strength distributions that influenced frequency judgments. Since the size of the increase in the strength distribution gets smaller as actual frequencies increase, one can assume that there is a great deal of overlapping for the strength distributions at the higher frequencies. For Hockley

(1984, p. 241) latencies were a "function of the distance from a frequency boundary-the closer an item strength is to the boundary, the slower the response."

Another way to view this problem is to ask whether frequency is evaluated in a direct or indirect manner in memory. Jonides and Naveh-Benjamin (1987) have described two contrasting models: one assumes that frequency is represented by a specific attribute in memory such as a counter (direct, as in Propositional or/and Multiple Trace models) and the other model assumes that frequency information is culled from other information stored in memory (indirect, as in Strength or/and Multiple Trace models). The main difference between the two models is that Jonides and Naveh-Benjamin (1987) assumed that the direct method depends on intention and active processing on part of the subject in order to increase the counter for frequency. More importantly Begg, Maxwell, Mitterer and Harris (1986) pointed out that the direct method assumes frequency is processed during presentation while the indirect model assumes that frequency is processed later when requested. It is not clear whether frequency information is represented by either a direct or indirect model. Both representations can be true.

There is a way to determine which model is most adequate: the effect of intention and the effect of levels of processing on judgments. The notion concerning intention is that the informed subjects are more sensitive and accurate in their frequency estimates than those uninformed if subjects must actively encode frequency at the time of presentation. Recall that Naveh-Benjamin and Jonides (1986) discovered that intention influenced frequency estimates only when there was a short time to code because the advantage disappeared when time conditions were relaxed.

On the other hand, examining the effects of level of processing appears more promising because if frequency estimates are generated from information in memory traces, then the variables that affect the coding of these traces should then affect frequency estimation. Jonides and Naveh-Benjamin (1987) found in their first experiment that semantic processing produced estimates of frequency that were more accurate, consistent, and of higher magnitude than those items processed acoustically. In their experiment, subjects that were shown a list of words were asked to write down a different acoustic (rhyming) or semantic association for each presentation. There were five words at each presentation frequency from one to six. There was a main effect of coding strategy: $F(1,58)=11.5$, $MSe=.92$, $p<.05$ and a significant frequency by coding strategy interaction, $F(5,290)=2.80$, $MSe=.20$, $p<.05$. The interaction showed that the slope relating judged and actual frequency was significantly higher for semantic than acoustic processing (semantic slope=.73, acoustic=.64, $t(58)=2.01$, $p<.05$). To measure reliability and consistency, Jonides et al. (1987) measured the standard deviations of the estimates for each frequency and then averaged them across actual frequency for each subject to get a mean standard deviation. Semantic subjects were found to produced more reliable judgments (semantic=.92, acoustic=1.10; $t(58)=2.20$, $p<.05$). It must be noted that Jonides, et al. (1987) also found that subjects produced more varied associations for the semantic than the acoustic condition ($t(58)=2.19$, $p<.05$) and therefore, suggested that the number of associations available rather than whether they are semantic or acoustic may influence frequency judgments. The more associations available, the greater the estimates. To examine whether this was an artifact, subjects were run in a self paced version of the experiment to rule out that subjects performed worst in the acoustic condition because they produced fewer acoustic associations. This allowed subjects to spend greater time in the acoustic condition because acoustic associations were harder to generate than semantic ones. The pattern of results in the self-paced experiment was similar to the previous experiment and therefore, the time artifact was ruled out.

According to Brown (1995, p. 1539), frequency estimation can be either an enumeration or a nonenumeration process. Brown (1995) states that "enumeration occurs when individual items or events are retrieved and counted and when the count arrived at serves as the basis for an estimate." There are two categories of enumeration. For simple enumeration, the value of the estimate is equal to the the number of relevant episodes retrieved from memory; and for enumeration-and-extrapolation, the value of the estimate is greater than the number of episodes retrieved from memory (Brown, 1995, p. 1539).

For Brown (1995, p. 1540), nonenumeration processes can be divided into either direct retrieval and memory assessment strategies. For the direct retrieval process, it is assumed that facts about event or item frequency are directly stored in memory as described in Strength theory or Propositional Encoding theory. This information is retrieved from memory during the estimation process.

For the memory assessment process, it is assumed "that some aspect of memory performance is evaluated during the estimation process and that the outcome of this evaluation serves as the basis for a frequency judgment" (1995, p. 1540). Brown points out in his paper, that historically, there are other theories that describe some form of memory assessment strategy. These theories include Tversky and Kahneman's (1973) availability heuristic where the frequency estimate is determined by how easy it is for the subject to retrieve relevant information from memory. The easier it is, the higher the frequency estimate and the harder it is, the lower the frequency estimate. According to Brown (1995, p. 1540), another example of a memory assessment process is when the "similarity between the probe item and the contents of episodic memory is used as an index of item frequency. Frequency estimates are high when the target item closely resembles the items stored in memory and low when it does not. Memory assessment produces a relative or qualitative evaluation of event frequency." One can imagine that some form of memory assessment process occurs when frequency memory is stored in multiple traces.

As discussed by Jones and Heit (1993), this notion of similarity has been utilized in computer simulation programs such as Hintzman's MINERVA 2 (Hintzman, 1988) and and Gillund and Shiffrin's SAM (Search of Associative Memory model; Gillund and

Shiffrin, 1984). These computer models assume that each stimulus presentation produces a new memory trace. When a stimulus probe is presented, this stimulus is compared with each memory trace producing a similarity measure. These measures for each trace are then summed together to produce a total similarity score. The frequency judgments for the probes will increase with the total similarity score which depends on the number of similar memory traces retrieved. Thus, the more similar memory traces that are retrieved, the higher the frequency judgment (Jones and Heit, 1993).

In Brown's series of experiments described in his 1995 paper, subjects were required to study a list of paired words and then asked to make absolute frequency judgments. Brown (1995) also collected the subjects' response times to the frequency judgments. The presentation frequencies were 2, 4, 8, 12 and 16. Each pair consisted of a target word (which was a category label) and a context word (which was an example of that category). In a different-context word pair condition, subjects might see CITY-Boston, CITY-Cleveland, CITY-London (a different context each time the category of city was presented). In a same-context word pair condition, subjects might see CITY-Boston, CITY-Boston, CITY-Boston (the same exact pair each time it was presented).

Brown (1995, p. 1551-52) reached the following general results in his three experiments:

1) Response times increase steeply with presentation frequency in the different context condition but not in the same-context condition.

2) Different context subjects tend to underestimate event frequencies, and same-context subjects tend to overestimate frequencies. Brown (1995, p. 1551-52) concluded that:

1) Subjects use multiple strategies to estimate event frequencies.

2) Strategy selection is related to event properties. Subjects favor enumeration-based strategies when event instances are distinctive and rely on nonenumeration strategies when they are not.

3) Different context subjects retrieved and counted category exemplars and same-context participants did not.

4) Strategy selection can affect the magnitude of subjects' frequency judgments. Underestimation appears to be a necessary consequence of enumeration, especially when presentation frequency is high.

Therefore, Brown (1997, p. 898) believes that subjects utilize a multiple-strategy perspective where event properties and encoding factors determine how frequency is represented, and that the nature of the frequency representation influences the strategy utilized. In Brown (1997), he utilized lists of target-context word pairs similar in structure to those described in his 1995 paper. To evaluate the relation between context memory and strategy selection Brown manipulated variables such as study-phase instructions, study time, and target-context relatedness. Response times, frequency estimates and cued-recall data were measured as dependent variables. Subjects were assigned to three groups: a context memory group; a frequency group; and a general memory group. The context memory group was told they would be tested in a cued-recall task and were instructed to remember each word pair. The frequency group was informed of the frequency test and instructed to pay close attention to each target word. In the general memory group, subjects were told to pay close attention to the pairs and that their memory would be tested for them.

Brown (1997, p. 910-914) concluded the following:

1) Subjects are more likely to enumerate when event instances are readily available than when they are not. There is a threshold for enumeration and when event memory falls below this threshold, there is no enumeration.

2) Knowledge of event frequency can be affected by encoding factors and instructional manipulations which in turn affects estimation accuracy. One estimation strategy is not

necessarily better than another and each strategy (enumeration or nonenumeration) is capable of producing an accurate assessment of relative frequency.

3) Divided attention during the study phase and delayed testing both reduce the accuracy of frequency estimates. "Full attention facilitates the encoding of frequency -relevant information, regardless of its form, and forgetting caused by the delay makes this information less reliable, accessible, or predictive," (Brown, 1997, p. 912).

In summary, the studies reviewed support a multiple representations of frequency (Naveh-Benjamin and Jonides, 1986; Hintzman, 1976, Brown, 1995; Brown, 1997). The more complexly frequency information is processed, the better the retrieval of this information. This has implications for the automaticity hypothesis of Hasher and Zacks (1984) for frequency processing because subjects in the intentional condition (Greene, 1986, 1989 and Naveh-Benjamin and Jonides, 1986) exhibited better frequency performance than those subjects in the incidental condition. Subjects in the intentional condition processed frequency more complexly and this resulted in better frequency performance at test. This contradicts the theory that frequency processing is automatic because under both the incidental and intentional conditions, performance should be the same regardless of the differences in processing. Furthermore, recent studies that examined the effects of simultaneous processing, levels of processing, changes in context, and strategy all found they affected how well subjects performed on frequency tasks (Greene, 1986; Naveh-Benjamin and Jonides, 1986; Brown, 1995; Brown, 1997). If one accepts the notion that with each occurrence of an event, a trace may be created, then one can assume that any factors that affect the creation of these traces will later affect the frequency judgment of those traces. Frequency does not hold a special status in the multiple trace model; it is inferred from other information.

EXPERIMENT ONE

The basic aim of the following experiments is to examine and help resolve the role of intention in frequency processing. The experiments differ from earlier experiments in three ways. First, the design in the majority of these previous experiments is *between* subjects; usually one set of subjects is informed and another set of subjects is uninformed. Unlike past research, the design in the following experiments is *within* subject; the same subject is in both the intentional and incidental conditions. Second, the following experiments provide an opportunity to use signal detection analysis to study the effect of intention on frequency memory within a two-alternative-forced choice (2AFC) frequency judgment task. Third, because of the within-subject design and the use of signal detection analysis, the characteristics of frequency memory under incidental and intentional conditions can be examined (Squire, 1987).

Research in neurobiology (ex., amnesia) shows that there are differences between declarative and procedural knowledge and researchers such as Schacter (1987) have linked together declarative memory with explicit/intentional memory and procedural memory with implicit/incidental memory. Squire (1987, p. 162) suggests that "these two types of memories differ in their biological organization: differences exist in what kind of information is stored, how it is used, and what neural systems are required." Thus, the factors that mediate incidental frequency processing may not be the same factors that mediate intentional frequency processing.

In the first experiment, described below, subjects were presented with a list of paired words (one in capital letters, one in lowercase letters). They were instructed to pay attention to the capitalized word of each pair because afterwards they would be given a two-alternative forced choice (2AFC) test in which they circled the word of each test pair that was most frequent. In reality, after the presentation of the list, subjects were tested on both the capitalized words (intentional condition) and the words printed in lower case letters (incidental condition).

METHOD

Procedure

To examine the role of intention on frequency processing, subjects were presented with a list of paired words. These word pairs were presented one pair at a time by a slide projector. Subjects were instructed to attend to the target word of each pair because they would be tested for their memory of the target words only (the cover task). The target word was in capitalized letters and the other word, the context word, was in lowercase letters. Subjects were instructed that the context words would help them remember the target words better because they were semantically related as illustrated by the sample stimulus pairs shown below:

BANK	congress	ADMIT	reason
river	ACT	hospital	ACCOUNT

In the list, half of the pairs appeared once and half the pairs appeared twice. Although subjects were told to **focus** on the targets (informed), they were also later tested on the frequency of the context words (uninformed). This produced a within subject design because the subject responded in both the incidental and intentional conditions. The intentional condition consisted of frequency judgments for target words about which they were warned. The incidental condition consisted of frequency judgments for context words.

The subjects were given the following instructions:

"You will view a list of paired words one pair at a time on this screen. You will be tested on your memory for the words printed in capital letters. These capitalized words are ambiguous and therefore, the second word in the lowercase letters will help you understand the word better. For example, in the pair "VOLUME-space", the capitalized word or the word you should remember is "VOLUME". The lowercase word is "space" and that tells you that the word "VOLUME" refers to a dimension of size but not a degree of loudness. Another example is "alphabet-LETTER". Here the capitalized word to be remembered is positioned below the other word. Therefore, the capitalized word will sometimes appear on top and it will sometime appear on the bottom. It is also important for you to know that the words will only appear for several seconds, so make sure you read the pair of words carefully when it appears on the screen. Do you have any questions?"

A practice list was given prior to the experiment to orient the subject to the task and to focus the subject's attention on the target words. The practice list consisted of 18 pairs (6 presented once; 6 presented twice). After the practice, subjects were were given practice booklets and asked to make frequency judgments on the target words only. The experimenter checked each subject's booklet to see if the task was done correctly.

Stimulus list construction

The word pairs were taken from Perfetti, Lindsey, and Garson's (1971) *Association and Uncertainty: Norms of Association to Ambiguous Words.* Perfetti, et al. (1971) examined how many senses, that is, how many different meanings a word has. The more senses a word has, the more ambiguous the word. For each target word, they collected the number of senses (or associations) generated by subjects and for each sense, the percentage of responses it received. To prevent subjects in the present experiment from remembering a context word because of its strong association to the target word, the context word was chosen by selecting an item that represented the second or third sense.

For example the target word shift had 3 senses: the sense 1 response was "move" with 83% of the responses; the sense 2 response was "dress" with 9% of the responses; and the sense 3 response was "work" with 7% of the responses. In my experiments, the word pair was "SHIFT-work". Using this method of creating word pairs produced a total of 109 word pairs and from these, 80 word pairs were randomly selected for the stimulus list.

The word frequencies (how common the word is) of the items were culled from the *Word Frequency Book* by Carrol, Davies, and Richman (1971). The mean log frequency for target words was 2.529 (per million; log SD = .570) and the mean log frequency for context words was 2.568 (per million; log SD = .695).

Each word pair was presented either once (N=40 pairs) or twice (N=40 pairs), for a total stimulus list of 120 word pairs. In addition, there were 9 buffer pairs before and after the stimulus list to cover primacy and recency effects. Each word pair was presented for 5 seconds following Greene's (1989, 1990) presentation rate for word pairs. Each repeated pair had at least 5 intervening pairs between the first and second presentations. Past research (Madigan, 1969) showed that subjects displayed no significant difference in performance for spacing intervals equal to or greater than 5 items. In addition, Greene's (1989) data show no significant difference in performance for spacings of 4 or greater. There were eight pairs at each of the following spacings: 5, 6, 7, 8, and 9. The order of presentation of the word pairs was randomized.

All words were counterbalanced across the target/context "roles" and across frequency conditions. Therefore, a word that was a target on one list became a context word on another list. A word pair that occurred once on one list occurred twice on another list. A word that appeared on top of one pair appeared on the bottom in another list. Thus, the word "ring" would appear as a target and a context word and would occur once and twice in different stimulus lists. This produced eight counterbalanced stimulus lists. For example:

RING(target)	bell(context)	BELL(target)	ring(context)
bell(context)	RING(target)	ring(context)	BELL(target)
shown once (list 1)	shown once (list 2)	shown once (list 5)	shown once (list 6)
shown twice (list 3)	shown twice (list 4)	shown twice (list 7)	shown twice (list 8)

Within a list, the target word occurred equally often in the top and bottom positions; both with pairs presented once and pairs presented twice. When a pair was presented twice, the position of the target word remained the same for both presentations. There were no more than three consecutive word pairs presented with the targets positioned on top or bottom solely. This prevented subjects from developing a strategy and looking for target words only on the top or bottom of the slides during the experiment.

Relative Frequency Test

After a subject viewed the stimulus list, he/she was given a booklet containing test pairs, one pair per page. The words of each pair were printed on one line in capital letters. The subject was instructed to circle the word in each pair that had occurred most often. If the subject was unsure, he/she was asked to make the best guess. To insure that all subjects spent the same amount of time on each pair and to insure that subjects finished the relative frequency test at the same time (at the end of an eight second period), subjects were told to go onto the next page at the sound of a bell (actually a door chime).

The test pairs in the booklets were constructed in the following way. The test pairs in t_2-t_1 category contained a target word presented twice and a target word presented once. The test pairs in c_2-c_1 category contained a context word presented twice and a context word presented once. (The number subscripts that follow a word label represent the number of times the word occurred in the stimulus list.) To control for response bias for circling a word on either the left or right side, the most frequent word was presented on either side 50% of the time. This was true for all test pairs. The data from these categories showed whether subjects can make relative frequency judgments because the words within these test pairs do not differ in terms of intention but in how often they occurred. The data comparing the two categories also showed whether the instructions were effective and whether intention played a role in the accuracy of their frequency judgments.

1a) target1-target2	ex.**SEASON-HABIT**	
1b) target2-target1	ex.**FALL-BANK**	**t2-t1 category**
2a) context1-context2	ex.**TASTE-HAIR**	
2b) context2-context1	ex.**NUN-RIVER**	**c2-c1 category**
3a) target2-context2	ex.**FALL-NUN**	**t2-c2 category**
3b) context2-target2	ex.**KING-HABIT**	
4a) target1-context1	ex.**PUPIL-DANCE**	**t1-c1 category**
4b) context1-target1	ex.**RIVER-SEASON**	

The test pairs in t2-c2 category contained a target word presented twice and a context word presented twice. The test pairs in t1-c1 category contained a target word presented once and a context word presented once. These test pairs were constructed to see if intention played any role in how well subjects remembered these words because they differ not in frequency but in instruction. Since the frequency of occurrence for each word is the same for these pairs, the words differ in the amount of "attention" the subject spent on processing the target versus the context words. One would assume that their memory would be better for targets because they were warned they will be tested on their memory for these words. Subjects are predicted to select the target words in the t2-c2 and t1-c1 categories.

After subjects circled the most frequent word of each pair, they were asked to rate how confident they were of their response. Confidence ratings were added to the experiment to see how well subjects could assess their frequency performance. Do subjects feel more confident under the intentional or incidental conditions? Do subjects feel more confident when presented particular test pairs rather than others? Is there a dissociation between confidence ratings and actual frequency performance? Subjects were shown how to make confidence ratings during the practice trials. For example, each subject saw the following on a single page in the test booklet:

COMPANY - SPRING

Please rate how confident you feel of your response.

100% sure left 1-----2-----3-----4-----5-----6-----7 100% sure right
0 %
Guessing

Memory Test and Questionnaire

To determine whether subjects really processed the target words in terms of their “context”, a memory test and questionnaire were administered after subjects completed their test booklets. This task was included to see if subjects could remember the *context* in which a *target* word was presented. The memory test consisted of twenty sets of two-word pairs printed on a sheet of paper. For example, during the experiment subjects saw the word pair "**PUPIL-school**". During the memory test, subjects were presented with "PUPIL-teacher" and "PUPIL-eye" and asked to circle the pair that was most similar in meaning to the one that appeared in the stimulus list.

The memory test consisted of 10 target words presented once and 10 target words presented twice. Only target words were used for this memory test. The word sets were randomly selected from the stimulus list and randomly presented on the memory test. The correct pair of a set was located 10 times on the right and 10 times on the left. The correct answer was not located more than three times consecutively on either the right or left side.

After the memory test, subjects had to answer several questions regarding the experiment. These questions were:

1) What do you think the experiment is about?
2) Did you use any strategy to memorize the capitalized words? Please describe the strategy.
3) Did you find it difficult to rate how confident you were of your responses?
4) Did you find the task in the experiment easy to do?

The first question was to determine if subjects knew what the experiment was about prior to the debriefing. The second question was to find out what kinds of strategies subjects used. The third and fourth questions were to determine if the frequency and confidence ratings tasks were difficult for Introduction to Psychology students to do within the prescribed time. The entire experiment including the debriefing took about 45 minutes.

Subjects

To determine how many subjects should be run in the experiment, the size of the effect of intention on relative frequency discrimination tasks was estimated. How large should the sample be to get adequate statistical power? One way is to see how large an effect Greene (1989) had in his Experiment 4 by looking at the ANOVA. Recall that for his relative frequency discrimination task, Greene (1989) presented subjects with test pairs. Each test pair contained one word that had occurred once and one word that occurred twice. The subject was required to select the word that occurred most frequently.

Greene had a significant effect of intention, $F(1,46) = 16.24$, $MSe = 6.25$. To calculate the effect size using an F score, I utilized the following formula (Rosenthal and Rosnow, 1984) :

$$\text{eta} = \sqrt{\frac{(F)*(\text{df effect})}{((F)*(\text{df effect})) + (\text{df error})}}$$

F (1, 46) = 16.24, df effect = 1, df error = 46

$$\text{eta} = \sqrt{\frac{(16.24)\ (1)}{((16.24)\ (1)) + (46)}} = \sqrt{\frac{16.24}{62.24}}$$

$$\text{eta} = \sqrt{.2609} = .5108$$

How large is an effect of .51? According to Rosenthal and Rosnow (1984, Table 24-2), an effect size of .51 is considered a large effect.

Although Greene (1989) had a large effect in his experiment, a conservative approach was taken and it was assumed that our manipulation might not be as large as Greene's. To determine the size of the sample required for adequate statistical power, the number of subjects that was needed to detect a "medium" effect was calculated. According to Rosenthal and Rosnow (1984, Table 24-2), at least 40 subjects are needed.

Forty-two subjects participated in the first experiment. Subjects were recruited from an Intro Psych pool for laboratory credit. All subjects were required to bring their prescription glasses if they needed them. Each subject was called the night before to remind him/her of the scheduled session. All subjects were run in small groups of 5 to 7 between 6:00 to 9:00 P.M. on weeknights.

Of the 42 subjects, 25 or 60% were female and 17 or 40% were male. The ages ranged from 18 to 42 with the mean age being 22.6 years.

Apparatus

Word pairs in the stimulus list were printed in the Helvetica font by a Macintosh Laser printer on white paper. Each word pair was then photographed with Kodak tungsten slide film using a 35mm camera. The tungsten slide film was developed commercially at Baboo Color Labs in Manhattan.

The stimuli were presented on a Kodak Ektagraphic slide projector and projected onto the wall under low ambient illumination (the front lights of the room were turned off). The projector was 185 inches from the wall and the size of the slide image on the wall was 32 inches high and 46 inches wide. The size of capitalized letters were 2 1/4 inches high and lower case letters were 1 5/8 to 2 1/4 inches high. The visual angle of a letter was approximately 0.62 degrees.

One Hunter timer was connected to the slide projector and was used to pace the presentation of the word pairs at a 5 second rate. An Archer Electronic door chime (model 273-071) attached to another Hunter timer was used during the test period. This enabled the experimenter to pace the subjects' responses to the test pairs at an 8 second per stimulus rate.

DATA ANALYSIS

Predictions

PREDICTION 1: Hasher and Zacks (1984) and Howell (1973) would predict that performance on all test pairs should not be significantly different because of intention. Therefore, the estimated relative frequency judgments would look like this:

$$\mathbf{d}'_{t2\text{-}t1} = \mathbf{d}'_{c2\text{-}c1} \text{ and } \mathbf{d}'_{t2\text{-}c2} = \mathbf{d}'_{t1\text{-}c1} = \text{chance}$$

For this prediction, the assumption is that frequency of occurrence would make the only difference in performance.

PREDICTION 2: According to Greene (1986, 1989), the mean proportion correct for target words will be higher than for context words. For example, the mean proportion correct from Experiment 4 from Greene (1989) was .60 for the incidental condition and .77 for the intentional condition.

The d' represents how sensitive the subject is to relative frequency in judging the pairs. When a subject produces a d' = 0, it tells us that the subject is unable to tell the difference between a word that occurred once versus a word that occurred twice. On the other hand, a subject with a d' = 1 is a more sensitive subject. The larger the d', the more sensitive the subject. To determine how sensitive Greene's subjects were to frequency of occurrence and to see how intention affects frequency judgments, the d' values for his conditions were calculated. To calculate d' using the proportion correct from Greene's data (1989), the following formula from Macmillan and Creelman was utilized (1991, Table A5.2 gives d' directly for various 2AFC proportions):

$$P(C)_{max,\ 2AFC} = \Phi\,(d'/\sqrt{2})$$

$$d' = (z)(\sqrt{2}) \quad \Phi = \text{area under the curve using } p(c)$$

By looking up the table in Macmillan and Creelman (1991) for the area under the normal curve that corresponds to Greene's proportion correct of 0.77 (intentional condition = .77), we find $z = .739$. The calculated d' is:

$$d'_{intentional} = (.739)(\sqrt{2}) = (.739)(1.414) = 1.045$$

Similarly, for the area under the curve for the proportion correct of 0.60 (incidental condition = .60), we find $z = .253$. The calculated d' here is:

$$d'_{incidental} = (.253)(\sqrt{2}) = (.253)(1.414) = .3577$$

The d' values from Greene's experiment show that subjects' sensitivity for relative frequency judgments is higher in the intentional condition than in the incidental condition. Subjects are more discriminating in the intentional condition.

Calculated ANOVA and Newman-Keuls tests:

Proportion correct was calculated for each subject according to the four test pair categories: $target_2$-$target_1$ pairs (target only, t_2-t_1), $context_2$-$context_1$ pairs (context only, c_2-c_1), $target_1$-$context_1$ pairs, and $target_2$-$context_2$ pairs.

For test pairs t_2-t_1 and c_2-c_1, the correct answer referred to the member that occurred most often. For test pairs t_1-c_1 and t_2-c_2, the "correct" answer referred to the target member of each pair since frequency of occurrence is the same. Subjects gave significantly more correct responses for $target_2$-$target_1$ pairs than for the other three categories. In fact, performance was as follows:

$$\begin{array}{ccccccc} 0.56 & & 0.60 & & 0.63 & & 0.65 \\ context_2\text{-}context_1 & < & target_1\text{-}context_1 & < & target_2\text{-}context_2 & < & target_2\text{-}target_1 \end{array}$$

Arc sin transforms were also applied to the proportion correct and the specific information can be found in Appendix C. Significant results were replicated. To examine if there were stimulus list effects because of the different log word frequencies, a t-test was performed comparing data from lists 1, 2, 3, and 4 to data from lists 5, 6, 7, and 8. There were no significant differences in subjects' performance.

Signal Detection Curves :

For each subject, a d' was calculated for each of the four test pair categories. The mean d' s for each category were as follows:

Table 12.

Mean d' values and mean percent correct across test pair category.

Test pair category:	Mean d'	Mean % Correct
$target_2$-$target_1$:	.60	.65
$target_2$-$context_2$:	.53	.63
$target_1$-$context_1$:	.38	.60
$context_2$-$context_1$:	.25	.56

*t2-t1 versus c2-c1 pairs are significantly different.

The subjects in this experiment displayed more sensitivity for target2-target1 pairs than for the other three types. (For individual d' values for Experiment 1, please see Appendix C.) From these d' scores, it can be seen that intention played a role in how sensitive or discriminating subjects were in this experiment. They exhibited the least amount of sensitivity for context only pairs (incidental condition) and the most sensitivity for target only pairs (intentional condition). An ANOVA was performed on the d' values, and a significant difference was found among the four pair categories ($F(3,123) = 3.23$, $p < .025$). Newman-Keuls post hoc tests were performed to determine any significant differences among the mean d primes. There was a significant difference found between t2-t1 versus c2-c1 pairs, $q_4 = 4.03$, $p < .05$.

A look at the top subjects

Although the effect of intention was significant, it was not as great as that found in Greene's experiment (1989). The range of performance for subjects in Experiment 1 was wide (mean proportion correct ranged from 0.35 to 0.90) and many subjects performed at about chance levels. Therefore, many of these subjects added noise to the data. Subjects' performance may have been affected because the experiment was run during Finals week at the end of the Spring semester. Top subjects were selected by their percentage correct on the t_2-t_1 pairs because overall data showed that performance was best for this pair category. The cutoff point was 70% correct (as in Greene's 1989 data). A closer look at the top subjects' performance (those who scored at least 70% correct in the intention condition = 19) revealed higher proportion correct across all four conditions. The data changed in that mean proportion correct performance on c2-c1 pairs went from .56 to .63; t2-c2 performance from .63 to .66 and performance on t2-t1 pairs from .65 to .79. Performance on t1-c1 pairs was about the same (.60 versus .61) The following is the proportion correct for both the total 42 subjects and the top 19 subjects across the four pair categories:

Table 13.

Comparing the mean proportion correct between all subjects and the top 19 subjects.

Pair Category	All 42 subjects	Top 19 subjects
t2-t1:	.65	.79
t2-c2:	.63	.66
t1-c1:	.60	.61
c2-c1:	.56	.63

The following is a comparison of **d '** between all 42 subjects and the top 19 subjects:

all 42: 0.25 0.38 0.53 0.63

context2-context1<target1-context1<target2-context2<target2-target1

top 19: 0.40 0.47 0.58 1.14

target1-context1<context2-context1<target2-context2<target2-target1

An analysis of variance performed on d' values for all four pair types for the top 19 was significant, $F(3,54) = 7.05$, $p < .001$. The top 19 subjects' performance was significantly better for the t2-t1 than for the c2-c1 word pairs (Newman-Keuls: $q3 = 5.43$, $p < .01$) and therefore, intention played a factor in these subjects' performance. A Newman-Keuls analysis of the d' values also showed significant differences between t2-t1 versus t1-c1 pairs ($q4 = 5.80$, $p < .01$) and between t2-t1 and t2-c2 pairs ($q2 = 3.89$, $p < .05$). No significant Newman-Keuls values were found for comparisons between t2-c2 versus t1-c1; between t2-c2 versus c2-c1; and between c2-c1 versus t1-c1 pairs. In conclusion, subjects made more correct frequency judgments for t2-t1 test pairs than for c2-c1 or t1-c1 pairs and therefore, intention plays a role in frequency judgments..

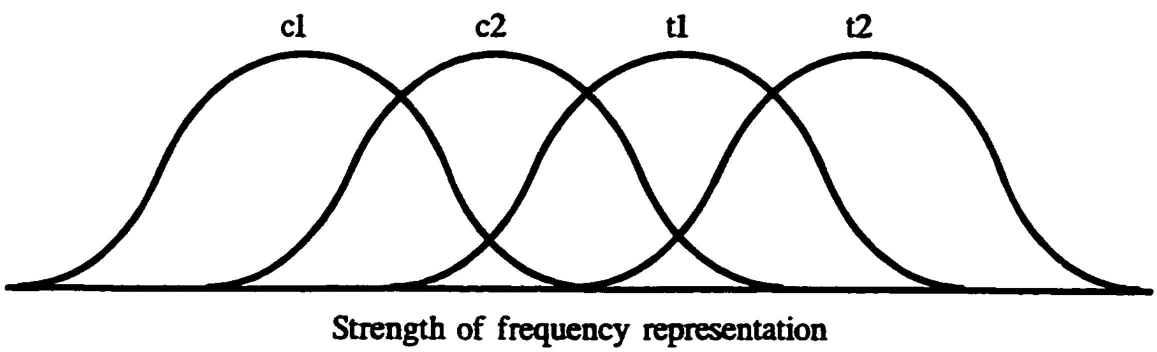
The Representation of Frequency using Signal Detection Analysis

Earlier I discussed several theories of memory as well as several ideas about how frequency information might be represented in memory. This discussion raised the idea that there may be several different types of memory, and several different ways in which frequency can be represented. This leads to the possibility that subjects in frequency judgment experiments might use more than one form of frequency memory. The current experiment provides a test of this possibility.

As a starting point, assume that subjects rely on a Strength representation of frequency memory. In this case, a subject should chose the stimulus with the higher strength. From this perspective, the mean strength for each type of stimulus item, i.e. c1, c2, t1 and t2, can be represented as a point on a line representing memory strength, and the d' scores reflect the average difference in the strength of the various stimulus types. This is illustrated in Figure 1, where the normal curves represent the distributions of the strengths of the individual items around the mean.

Figure 1.

An illustration of strength value distributions for different word types



With this representation, if we know the distances separating c1 and c2, c2 and t1, and t1 and t2, we can calculate the distance between any other pair of items. For example, the distance from c1 to t2 is just the sum of the distances for (c1-c2) + (c2-t1) +(t1-t2). Because the d' scores represent these distances in z-scores units, we can apply this analysis to this experiment as shown below where c1, c2, t1, and t2 correspond to the means of the distributions in the figure below depicting the d' distances corresponding to the stimulus pairs in the experiment.

Figure 2.

Diagram of the means of the distributions for different test pairs.

```
          |<--d'c1c2 = 0.25-->|<------------d'c2t2 = 0.53----------->|
Strength---c1------------------c2------------------t1------------------t2-
          |<---------------d'c1t1 = 0.38---------->|<--d't1t2 = 0.63->|
```

This diagram leads to the following prediction:

$$d'_{c1c2} + d'_{c2t2} = d'_{c1t1} + d'_{t1t2} \qquad (1)$$

Using the d' values shown in Figure 2, $d'_{c1c2} + d'_{c2t2} = 0.78$, which differs from $d'_{c1t1} + d'_{t1t2} = 1.01$. If frequency is represented in terms of a single strength-like dimension, these two values should be equal. This prediction can also be expressed as follows:

$$d'_{c1c2} + d'_{c2t2} - d'_{c1t1} - d'_{t1t2} = 0 \qquad (2)$$

$$0.25 + 0.53 - 0.38 - 0.63 = -0.23$$

Using the d' values from Figure 2, the actual result is -0.23.

According to this analysis, if we calculate $(d'_{c1c2}) + (d'_{c2t2}) - (d'_{c1t1}) - (d'_{t1t2})$ for each subject, the result should be equal to 0 except for estimation error.

On the other hand, if frequency information can be represented in more than one way, as suggested by Brown (1995) and Hintzman (1988), we might not expect to obtain the result just described. To illustrate this, assume that there are two components or dimensions to the representation of frequency, e.g. an implicit memory representation (possibly based on Strength) and an explicit memory representation (possibly

Propositional). Then the representation of the frequency of any item could be described by a point in two-dimensional space as illustrated below.

Figure 3.
An illustration of two components that represent frequency in 2-dimensional space.

```
implicit     |-            c2
frequency    |
dimension    |-                      t2
             |
             |
             |- c1
             |-   t1
             |
             |__|__|_______|_____|_____________
             explicit frequency dimension
```

Here, each point (t1,t2,etc.) represents the mean of a two - dimensional Gaussian density function that rises out of the two - dimensional plane defined by the implicit and explicit dimensions.

Given this representation, how might subjects make their 2AFC frequency judgments? In fact, there are a number of possibilities that depend on how subjects might use the information from each dimension.

One possibility is that frequency judgments might be based on the Euclidean distance of each stimulus from the origin. That is, the farther a point is from the origin, the greater its subjective frequency. Given two stimuli, a subject might compute the distance of each point from the origin and the point that was farthest from the origin would be judged the most frequent.

Although it may not be immediately obvious, this decision algorithm effectively makes the representation equivalent to the one - dimensional representation. That is, the only variable that matters is the distance from the origin, and it does not matter whether a particular stimulus lies closer to one or the other of the two axes. In this case, the prediction that $d'_{c1c2} + d'_{c2t2} - d'_{c1t1} - d'_{t1t2} = 0$ should still hold and obtaining this result would not tell us whether the underlying frequency representation was uni- or multi-dimensional.

On the other hand, if subjects have a multi-dimensional representation, it is possible that they might not be able to integrate the dimensions so as to calculate the Euclidean distance of each stimulus item from the origin. If so, there are several possible alternative decision algorithms:

Alternative 1: Subjects might make their judgments of the frequency of a stimulus based on the representation of the stimulus on only one of two dimensions (e.g. only on the Explicit dimension). Which dimension was used for the judgment might depend on the type of stimulus involved, e.g. context stimuli might be judged using only information from the implicit dimension. Or a subject might be predisposed to use a particular dimension regardless of the stimulus involved, perhaps because that dimension was in some way more salient or because it generally provides higher acuity in making frequency decisions.

Alternative 2: Another possibility is that the information provided by each dimension is combined in some non-Euclidean form. This might be done in several ways:

a) When two stimuli are presented, a subject evaluates each stimulus with respect to its representation on each dimension. If one stimulus is higher than the other on both dimensions, the subject reports this as the more frequent stimulus (e.g. t2 versus c1 in the diagram above). On the other hand, if one stimulus is higher on one dimension but lower on the other dimension (e.g. c2 versus t1 in Figure 3 above), the appropriate response is less obvious. One possibility is that in the case of conflicting information of this sort, one of the dimensions (e.g. Explicit) might be the preferred basis for decision. A second possibility is that if subjects can equate the scales of the two dimensions, the response might be based on the dimension that encodes the highest frequency, that is, on the maximum of the values of the stimuli on each dimension. For example, in comparing a c2 and a t2 stimulus, if the location of t2 on the Explicit axis is greater than the location of c2 on the Implicit axis, t2 would be chosen as the more frequent stimulus.

b) The possibility just mentioned, i.e., that subjects can compare the scalar distances along each axis, leads to the possibility that subjects might make their judgments by combining the information provided by each dimension, say by using a city-block measure of the distance of each stimulus from the origin.

If subjects followed either alternative 1 or 2 we would not expect the contrast d'c1c2 + d'c2t2 - d'c1t1 - d't1t2 = 0 to hold. To test this idea d'c1c2 + d'c2t2 - d'c1t1 - d't1t2 was calculated for each subject. The mean of the 42 subjects was 0.181, with a Standard Error of 0.119, for a non-significant $t = 1.52$, $p > .05$.

As noted earlier, overall performance in this experiment was fairly low, with some subjects essentially at chance performance. Including such subjects in any analysis adds considerable noise. Therefore, the analysis just described was repeated using only the 32 subjects who were above chance on the t1-t2 stimulus pairs. For these subjects, mean value of the contrast was 0.292 with a standard error of 0.138, for a non-significant $t = 2.119$, $p > .05$.

In summary, the results of this analysis are consistent with either the idea that subjects have a single underlying representation of frequency, or with the idea that if there are multiple representations, these representations can be combined into a single integrated representation in order to make 2AFC judgments of the sort required in this task. It is worth noting, however, that the range of d's observed was relatively small because of the low level of performance. A larger range of d's would provide a stronger test.

The Memory Test

Recall that a memory test was administered after the frequency test because I wanted to see if subjects remembered the context that the target words were presented in. I wanted to be sure that subjects did not ignore the context (lowercase) words and did not just attend to the target words (uppercase words) when they were presented during the stimulus list. If subjects only attended to the target words, there would be no incidental condition in the experiment.

The subjects were required to circle the word pair that had the most similar context to the word pair presented in the stimulus list (ex., the original pair on the stimulus list is PUPIL-school: is it PUPIL-teacher or PUPIL-eye?). Subjects, on the average, circled the correct pair 13 times out of 20 or 65%. The mean arc sine transform was 1.9 and the standard deviation was .347.

A t-test was performed to determine whether subjects performed significantly above chance on the memory test. Arc sine transformations on the actual memory scores were compared to those of chance performance and a significant difference was found (t_{41} = 6.17, $p < .001$). Therefore, subjects in the first experiment performed significantly above chance when it came to remembering what context the target word was presented in. Therefore, subjects did attend to the context words and did not ignore them.

To determine if there was any relationship between the memory test and the subjects' performance in the four test pair categories, Pearson r correlations were computed. Before the correlations were computed, the proportion correct for each subject was changed to arc sine transformations. The following includes the Pearson r correlations for the overall mean performance and for each test pair condition:

Table 14.

Comparing the performance between the memory test and the test pairs using Pearson r.

Comparison type:	Pearson r:
a) memory test vs overall mean performance =	+.15
b) memory test vs t2-t1 test pairs =	+.18
c) memory test vs c2-c1 test pairs =	-.12
d) memory test vs t1-c1 test pairs =	-.01
e) memory test vs t2-c2 test pairs =	-.04

As you can see, the greatest correlation was for the t2-t1 pairs. None of the correlations was found to be significant. This may mean that the factors that are used for frequency estimation are not the same factors used for memory tests.

Questionnaire Summary

A questionnaire was administered to see if subjects knew what the real motivation for the experiment was prior to the debriefing and to determine if the frequency task was too difficult. Often when people are asked to estimate how often something happened, people report that they find it difficult to give a frequency estimation and they often report that they are not too confident about their response.

When subjects were asked what they thought this experiment was about, 86% thought it was a memory experiment. When subjects were asked what strategy they used to memorize the target words, 26% used repetition (repeating words over and over) or counting (keeping an account of how often a word occurred); 17% concentrated on only the repeated words; another 17% reported no strategy at all; 12% related words to their own generated associates; 10% made words into stories or sentences; 10% concentrated only on the capitalized words; 7% related words from one pair to another pair; and 2% categorized the words into grammatical categories. A majority of the subjects used counting or repetition as a strategy to memorize the words and that may explain why the best mean correct frequency performance was only at 65%. Recall that from Greene's experiment (1989), that those who made sentences from the words (mean accuracy = 70%) performed better than those who just utilized counting or repetition (mean accuracy = 53-54%). When subjects were asked if they found it difficult to score confidence ratings, 55% found it difficult and 33% found it easy. Lastly, when subjects were asked if they found the task in the experiment easy, 48% found it easy; 36% found it difficult and 7% found the directions easy but remembering hard.

Discussion

The data support the idea that intention affects relative frequency judgments. Subjects performed significantly better for t2-t1 pairs than for c2-c1 pairs (d': .60 versus .25, t_{41} = 2.20, $p < .05$). This finding is contrary to Hasher and Zacks (1984) who stated that frequency processing is automatic with intention having no effect on frequency processing. This is not to say that frequency processing does not occur without intention because in Experiment 1, subjects performed slightly above chance for the c2-c1 word pairs (mean proportion correct = .56). A d' of 0 would show that subjects cannot discriminate frequency but d' performance on c2-c1 pairs is .25. With intention, subjects' d' performance was .60 which shows more sensitivity to frequency information for t2-t1 pairs than for c2-c1 pairs (d': .63 versus .25).

The t1-c1 and t2-c2 pairs are "funny" in that subjects are required to circle the word that occurred most often when in actuality both words were presented at the same frequency. Both words in each pair have the same frequency of occurrence but presumably differ in the amount of intention used in processing the words. Although the difference in performance between t2-c2 and c2-c1 pairs was not significant, subjects displayed a greater tendency to circle the target word in a t2-c2 pair than to circle the more frequent context word in a c2-c1 pair (mean percentage correct: .63 versus .56). These two pairs were compared because in the t2-c2 pair, the words have the same frequency but different intentions while in the c2-c1 pairs, the words have the same intentions but different frequencies. One would assume performance may be better for c2-c1 pairs because the words only differ in terms of frequency, while one would predict performance for t2-c2 be worse because both words have the same frequency. In addition, no significant difference was found between t1-c1 and t2-c2 pairs.

Word pair designs that include c2-t1 and t2-c1 combinations are needed to get a better picture of how frequency memory is represented under incidental and intentional conditions. In Experiment 2, these two new frequency word pairs will be added to see if they will help to explain how frequency memory is represented.

There are two issues of concern in the first experiment. First, how do we know that the subjects looked at both the context and target words of each stimulus pair? Subjects recalled the correct context in which the target word was presented 65% of the time but 10% of the 42 subjects mentioned that they only looked at the target words on the stimulus list. The second experiment included a task that insured that subjects looked at both context and target words in each pair. Second, to more accurately measure the characteristics of frequency memory, new test pairs were constructed. These test pairs were comprised of t2-c1 word combinations and c2-t1 word combinations. With the addition of these new test pairs, the hypothesis that frequency memory falls along a single dimension under both incidental and intentional conditions can be better tested.

EXPERIMENT TWO

In the second experiment, we will add two additional categories to the test pairs because it appeared in Experiment 1 that not enough test pair categories were used to explain fully how frequency memory is represented. For the second experiment, the same methodology and stimuli were used except for the following three modifications:

A) One change from Experiment 1 is that subjects were given a semantic memory test both after the practice list and after the actual stimulus list. This memory test is different in style from the memory test in Experiment 1 as seen below:

Experiment 1:
Please circle the the pair that was most similar in meaning to the
one that appeared in the stimulus list:
a) PUPIL - teacher
b) PUPIL - eye

In Experiment 1, the semantic memory test was presented only after the stimulus list, but in this experiment, we really wanted to reinforce the subjects for looking at both the target and context words. This memory test had different definitions of the stimulus word. For example, the stimulus pair "**PUPIL-school**" was presented during the experiment. During the memory test, subjects were to select the definition that best describes what the subject saw. For example, the subject was asked the following:

Experiment 2
Which best describes the meaning of the word "PUPIL" in the experiment?
Please circle one of the following:
a) eye
b) teacher

The memory test presented after the practice stimulus list consisted of six test items similar in format to the one above. This informed the subjects that they were expected to look at the context words because they would be tested on their memory for the definitions of the targets. The memory test presented after the stimulus list contained 26 test items. Unlike the memory test in the first experiment, half of the items tested subjects' memory for target words and half tested subjects' memory for context words. Ten of the correct answers were located at "a" and ten were located at "b" with the constraint that no more than three consecutive correct responses be located at either a or b. In addition, half of the target and context words were words presented once and half were words presented twice. All words were counterbalanced across frequency of presentation and target/context category. For example, how a word used in the memory test is counterbalanced across target/context and frequency of occurence:

Stim. List:	Word Pair:	Description:
1 & 2	NUT-bolt	target - context x 1
3 & 4	NUT-bolt	target - context x 2
5 & 6	nut-BOLT	context - target x 1
7 & 8	nut-BOLT	context - target x 2

To counterbalance for lexical frequencies across responses at "a and b", each item was checked for lexical log frequencies (of occurrences in language) in Carrol, Davies, and Richmond's The American Heritage Word Frequency Book (1971). All responses were checked to ensure that half of the right answers had higher log frequencies and half of the wrong answers had higher log frequencies. The mean log frequency for right answers was 1.904 (SD=.8369) and the mean log frequency for wrong answers was 2.069 (SD=.6267). An illustration of the counterbalancing is shown in Appendix B.

B) Secondly, the test pair categories were modified and new test pairs constructed (12 pairs per category). The test pair categories were changed to the following:

1a) target1-target2	ex.**SEASON-HABIT**	**t2-t1 category**
1b) target2-target1	ex.**FALL-BANK**	
2a) context1-context2	ex.**TASTE-HAIR**	**c2-c1 category**
2b) context2-context1	ex.**NUN-RIVER**	

3a) target2-context2	ex.**FALL-NUN**	**t2-c2 category**
3b) context2-target2	ex.**KING-HABIT**	
4a) target1-context1	ex.**PUPIL-DANCE**	**t1-c1 category**
4b) context1-target1	ex.**RIVER-SEASON**	
*5a) target2-context1	ex.**FALL-DANCE**	**t2-c1 category**
*5b) context1-target2	ex.**RIVER-HABIT**	
*6a) target1-context2	ex.**PUPIL-NUN**	**c2-t1 category**
*6b) context2-target1	ex.**KING-SEASON**	

Note. * = these are the new word pairs.

The t2-t1, c2-c1, t1-c1 and t2-c2 pairs are just like the pairs in Experiment 1, but two new pairs were added: t2-c1 and c2-t1. With data from these six test pair categories, a better picture can be developed concerning intention and frequency memory. It is predicted that subjects will again display better performance for t2-t1 pairs than for c2-c1 pairs. In addition, it is predicted that they will show better performance for t2-c1 pairs than for c2-t1 pairs. This is with the theory that for the t2-c1 pairs, target words processed twice intentionally are easier to remember than context words processed once incidentally. The frequency judgments for c2-t1 pairs were expected to be difficult to make because the memory for context words processed twice incidentally might not be different from a target word processed once intentionally.

With the addition of t2-c1 and c2-t1 pairs, we may more carefully measure the characteristics of memory under incidental and intentional conditions using signal detection analysis. In Experiment 1, when subjects were presented with t1-c1 and t2-c2 pairs, they could not base their judgments on frequency. Subjects may have used other characteristics of the words to make their judgments or based their judgments on different types of memory representations (i.e., explicit frequency memory versus implicit frequency memory). By adding t2-c1 and c2-t1 test pairs, an examination of whether incidental and intentional frequency memory falls along more than a single dimension can be more fairly ascertained.

C) Thirdly, the test booklets were further modified because the structure of the confidence ratings was changed to resemble more of a 2AFC design. The following is an example:

COMPANY - SPRING

Please rate how confident you feel of your response.

100% sure left 3---2---1---0---1---2---3 100% sure right
0 %
guessing

This design more closely parallels the 2AFC design of the test pairs. It is practical to ask questions comparing d' and confidence ratings. Subjects may be more or less accurate in frequency judgments than they believe they are. Will subjects display less confidence for context pairs (c2-c1) and more for target pairs (t2-t1)? The differences in subjects' confidence ratings for target and context words may not be consistent with actual performance.

To determine how many subjects to run, the following formula was utilized to calculate the effect size using the F score (from the arc sin transformations of the four test pair categories, $F(3,123) = 3.41$, $p < .025$) from the first experiment (Rosenthal and Rosnow, 1984, Table 24-2):

$$\text{eta} = \sqrt{\frac{(F)*(\text{df effect})}{((F)*(\text{df effect})) + (\text{df error})}}$$

F (3,123) = 3.41, df effect = 3, df error = 123

$$\text{eta} = \sqrt{\frac{(3.41)\ (3)}{((3.41)\ (3)) + (123}} = \sqrt{\frac{10.23}{133.23}}$$

$$\text{eta} = \sqrt{.08} = .2828$$

How large is an effect of .28? According to Rosenthal and Rosnow (1984), an effect size of .28 is considered a medium size effect. Therefore, the size of the sample required to get adequate statistical power needs to be determined. By looking at a table in Rosenthal and Rosnow (1984, Table 24-2), it is determined that 40 subjects minimally need to be run.

Subjects were run in the Spring semester during finals week. They were students who were fulfilling a subject pool requirement for an Introduction to Psychology course at Brooklyn College. Forty subjects were run in groups of 3 to 5 in a classroom using the same equipment and materials described in the methodology section of Experiment 1.

RESULTS

Analysis of Experiment 2:

Table 15.

Comparing the mean percentage correct from Experiments 1 and 2.

pair type:	t2-t1 t1-t2	c2-c1 c1-c2	t2-c2 c2-t2	t1-c1 c1-t1	t2-c1 c1-t2	c2-t1 t1-c2
Experiment 2						
mean %:	65.8	60.8	63.6	58.5	73.3	50.5
SD %:	15.36	12.78	20.15	15.26	14.24	16.50
Experiment 1 (no t2-c1 and c2-t1 pairs)						
mean %:	65.07	56.43	62.86	59.81	NA	NA
SD %:	15.75	13.22	14.48	12.82	NA	NA

Note. t=target, c=context, %=percentage correct, SD=standard deviation

The subjects in Experiment 2 displayed the best frequency discrimination memory for the t2-c1 pairs and the worst memory for the c2-t1 pairs (the exact 2 new pair categories). It appears that subjects' memory for targets presented twice was 'stronger' than context words presented once (mean correct %=73.3). Two factors contribute to the memory of a word in this experiment: 1) frequency of presentation; and 2) whether the item was intentionally (target) or incidentally processed (context). Therefore, when subjects are asked to chose the word in each pair that occurred most frequently, t2-c1 pairs were easier than c2-t1 pairs. Because subjects are asked to really concentrate on the target words in the experiment, memory for context words is not as good. Subjects' memory for c2 words must be similar to t1 words because their performance for c2-t1 pairs is at chance performance (mean correct %=50.5). From looking at the mean percentage correct in Experiment 2, subjects displayed a tendency to chose the target items (range of mean %

correct: 59-73%) except for the chance performance for the c2-t1 pair.

As in Experiment 1, circling the target item in the t2-c2 and t1-c1 test pairs was scored as correct in Experiment 2. For the t2-t1, c2-c1, t2-c1, and c2-t1 test pairs, the most frequent item was scored as correct. Subjects' performance for the t2-t1 and the t2-c2 pairs was very similar for both Experiments 1 and 2 (mean % correct t2-t1: 65.8 vs. 65.07; mean % correct t2-c2: 63.6 vs. 62.86). Also, it should be noted that in the t2-c2 pair category, subjects selected the t2 item approximately 64% over the c2 item even though the subjects saw both items equally often. Frequency discrimination memory is most different for the c2-c1 pair in the two experiments: Experiment 2 = 60.8 versus Experiment 1 = 56.43.

Table 16.

Significant Newman-Keuls for Experiment 2.

Significant Newman-Keuls for d' (percent correct and d' in parenthesis):

t2-c1 (73%, 1.03) versus c2-c1 (61%, .42) : $q4 = 5.68$ at $p < .01$ level
t2-c1 (73%, 1.03) versus t2-t1 (66%, .636) : $q2 = 3.66$ at $p < .05$ level
c2-c1 (61%, .419) versus c2-t1 (51%, .026) : $q3 = 4.14$ at $p < .05$ level

---

The d' values were calculated for each of the six test pair categories as shown below. The mean d' is derived from each score for each subject in each category using the formula and the z proportions from Macmillan and Kaplan (1985) : $d' = (z)(\sqrt{2})$. (Note: subjects' response booklets were checked to see if there were any preference for selecting either a word on the right or left only and no such preference was found).

Table 17.
The d' values, standard deviations and means percentage, for test pairs from smallest to greatest value.

least sensitive<-->most sensitive

	c2-t1	t1-c1	c2-c1	t2-c2	t2-t1	t2-c1
mean d':	.026	.340	.419	.586	.636	1.03
SD d':	.629	.626	.517	.843	.650	.846
mean %:	50.5	58.5	60.8	63.6	65.8	73.3

*Note: Please see Appendix D for individual d' values.

According to Table 17, the highest mean d' or the greatest sensitivity was found for the t2-c1 pairs and the smallest mean d' or least sensitivity was found for the c2-t1 pairs. An ANOVA of the d' values shows that the pair types differed significantly, $F(5, 195) = 10.63$, $p < .001$. Newman-Keuls tests were performed to test possible differences among the mean d' primes. The following were significant d primes at least at the .05 level:

a) t2-c1 versus c2-c1 (1.03 vs .419): Here we are comparing subjects' performance for selecting the most frequent word for pairs with context words presented once versus target words presented twice or context words presented once versus context words presented twice. Subjects' were significantly better and more sensitive at discriminating t2-c1 pairs than c2-c1 pairs, $q_4 = 5.68$, $p < .01$ level.

b) t2-c1 versus t2-t1 (1.03 vs .636): Subjects were more sensitive for t2-c1 pairs than for t2-t1 pairs, $q_2 = 3.66$, $p < .05$. Memory strength for c1 words is not as strong as for t1 words and therefore, the contrast is greater between t2-c1 words than for t2-t1 words.

c) c2-c1 versus c2-t1 (.419 vs .026): Subjects are better at discriminating c2-c1 pairs than c2-t1 pairs, $q_3 = 4.14$, $p < .05$. Although both c1 and t1 words are presented once, they differ in "intention". The memory trace of t1 words may be more stronger than c1 words, therefore subjects exhibit chance performance for c2-t1 pairs.

<u>The Representation of Frequency using Signal Detection Analysis</u>

As I described in discussing the results of Experiment 1, if we assume that frequency information is represented by a single underlying type of memory, such as is proposed by the Strength theory of frequency representation, then the mean strength of each of the four types of items, c1, c2, t1 and t2 can be represented by a point along a single line and the 2AFC d' values are measurements of the distance separating the means. The current experiment provides an additional opportunity to test this hypothesis.

Because the second experiment included the same test pairs, we can again test the idea that the representation of frequency falls along a single underlying dimension as represented by the diagram below, which was introduced earlier in the discussion of Experiment 1.

Figure 4.

<u>Strength diagram for Experiment 1.</u>

```
                 |<-----d'c1c2------>|<----------------d'c2t2---------------->|
Strength---------c1-------------------c2-------------------t1-------------------t2-
                 |<----------------d'c1t1----------------->|<-----d't1t2----->|
```

As noted earlier, this leads us to predict that:

$$d'c1c2 + d'c2t2 - d'c1t1 - d't1t2 = 0 \qquad (2)$$

Because Experiment 2 included additional types of test stimuli, there are other possible contrasts that can be formed as illustrated in the next diagram.

Figure 5.

<u>Strength diagram for Experiment 2.</u>

```
                                                (0.42 + 0.03 + 0.64) = 1.09
              |<-d'c1c2 = 0.42-->|<-d'c2t1=0.03-->|<-d't1t2=0.64--->|
Strength---------c1----------------c2----------------t1----------------t2-
              |<-------------------d'c1t2 = 1.03------------------->|
```

This diagram illustrates the prediction that:

$$d'_{c1c2} + d'_{c2t1} + d'_{t1t2} = d'_{c1t2} \qquad (3)$$

Using the d' values shown in Figure 5, we find that $d'_{c1c2} + d'_{c2t1} + d'_{t1t2} = 1.09$ which is fairly close to $d'_{c1t2} = 1.03$. This prediction can also be expressed as follows:

$$d'_{c1c2} + d'_{c2t1} + d'_{t1t2} - d'_{c1t2} = 0 \qquad (4)$$

The data of the second experiment were tested against the predictions of equations 2 and 4. For contrast 2, the mean value was -0.029 while for contrast 4, the mean value was 0.029, neither of which was significantly different from 0.

As with Experiment 1, this analysis was repeated using only subjects for whom $d'_{t1t2} > 0$. For these 32 subjects, the two values were 0.22 and 0.10, neither of which was significantly different from 0.

In conclusion, based on the results of Experiments 1 and 2, the data support the hypothesis that there is a single one-dimensional representation of frequency underlying performance in this task.

<u>Semantic Memory Test Data Analysis :</u>

In terms of the semantic memory test given to subjects after the frequency discrimination task, performance was better for Experiment 2 than Experiment 1:

Table 18.
Comparing the mean percent correct from the memory test for both Experiments.

Experiment 2:	Target mean % = 75.75	Context mean % = 72.25
	% SD = 12.99	% SD = 15.61
Experiment 1:	Target mean % = 65.24	Context mean % = NA
	% SD = 15.58	% SD = NA

The proportions correct were changed using the arc sine transformation. Pearson r correlations were computed comparing the memory score with the results from the relative frequency judgments. Data analysis on memory items showed that overall mean percentage correct was 75%; mean percentage correct for target words was 76% and mean percentage correct for context words was 72%. A correlated t-test was done to see if the performance on target and context words were significantly different and it was not ($t(39) = 1.095$). Therefore, subjects' performance showed that subjects remembered target and context words about equally well in terms of their semantic "flavor" in the stimulus list.

Correlations were done to see the relation between memory test performance (arc sine transforms) and the correct arc sine from the frequency judgment test pairs. The following is a summary of the correlations for Experiment 1 and 2 (the Pearson r from Experiment 1 were not significant)

Table 19.

Comparing the Pearson r from Experiments 1 and 2.

Category	Exp. 2 r:	Exp 1 r:
memory vs. t2-t1	-.2360	+.18
memory vs. c2-c1*	.3126	-.12
memory vs. t2-c1*	.3125	n/a
memory vs. c2-t1	-.0770	n/a
memory vs. t2-c2	.2188	-.01
memory vs. t1-c1	.1214	-.04

Using the significance test in the Bruning and Kintz's Computational Handbook of Statistics (1968; pg. 155), none of the correlations was found significant, but the two categories with asterisks came close. The following is the formula used to determine significance: $z = r\sqrt{n-1}$. If the z score reaches +/- 1.96, it's significant at the .05 level. Although the correlations are not significant, it appears coincidental that in Experiment 1, the performance on the memory test (65%) is very similar to the best performance test pair, t2-t1 (mean percentage correct = 65.07%). In Experiment 2, performance on the memory test is 75% and this is very similar to the best performance test pair, t2-c1 (mean percentage correct = 73.3%).

Median Confidence Ratings Analysis for Experiment 1 and Experiment 2

Confidence ratings were asked to assess whether subjects could evaluate how well they made frequency judgments. To evaluate the relationship between actual performance and how confident subjects were in their performance. It seems plausible to assume that confidence would be related to the source of information in a subject's memory. Another assumption was that subjects would exhibit higher confidence ratings for items intentionally processed (target words) compared to those incidentally processed (context words).

In Experiment 2, subjects were asked to circle how confident they were of their frequency recognition response to a pair of words. Each subject was asked to circle the word he/she thought was the most frequent word in the pair (APPLE - HOUSE). Then the subject was required to circle how confident he/she was of his/her response. For example,

APPLE - HOUSE

3 --------- 2 --------- 1 --------- 0 ---------- 1 ---------- 2 --------- 3
100% 0% 100%
sure left Guessing sure right

In Exp 1, subjects were given a confidence rating scale that looks like this:

APPLE - HOUSE

1 --------- 2 --------- 3 --------- 4 ---------- 5 ---------- 6 --------- 7
100% 0% 100%
sure left Guessing sure right

To make the confidence ratings more synchronous with Exp 2, the ratings in Exp 1 were translated to the following corresponding numbers:

Ratings in Exp 1		Ratings in Exp 2
1, 7	became	3
2, 6	became	2
3, 5	became	1
4	became	0

In recording confidence ratings, they were grouped by test pair category (i.e., t2-t1, c2-c1, t2-c1, c2-t1, t2-c2, and t1-c1; t = target, c = context; 1 = presented once, 2 = presented twice). It's important to note here that in Experiment 1, there were only four test pair categories: t2-t1, c2-c1, t1-c1, and t2-c2. They were further divided into correct and wrong confidence ratings within a test pair category. For each subject in Experiment 2, there were 12 confidence rating responses collected for each of the six test pair categories. For each subject in Experiment 1, there were 20 confidence rating responses for each of the four test pair categories. The absolute ratings were recorded and whether it was a "3" on the left or right was not a factor in the analysis.

The confidence ratings were divided by test categories to see if subjects' ratings were different across test pair categories. The confidence ratings were also divided into correct and wrong responses to see if there is a difference across accuracy. Finally, for each subject and test pair category, the median rating response was found for both the correct and incorrect responses. For example, for subject 17 in category one (t2-t1, t1-t2) pairs, there were seven correct ("r") responses and five incorrect responses ("w").

Figure 6.

Illustration of how the confidence rating median was calculated.

	Correct median:	Wrong median:
	r1 = 3	w1 = 3
Total Median = 3	r2 = 3	w2 = 2
[3, 3, 3, 3, 3, **3**, **3**,	r3 = 3	**w3 = 3**
3, 3, 3, 2, 2]	**r4 = 3**	w4 = 3
3 + 3 = 6/2 = 3	r5 = 3	w5 = 3
	r6 = 2	
	r 7 = 3	
	---------	-----------
	r median: 3	w median: 3

Note. r = correct response, w = wrong response.

The correct medians and wrong medians were systematically done for every subject in each of the six test pair categories. To calculate median confidence ratings, the scores were ordered from lowest to highest value and the number at the 50th percentile was selected as the median. If there was an even number of scores, the two middle numbers were averaged. If there were an odd number of scores, the one in the middle was selected as the median.

Experiment 2: An analysis of variance of the median confidence ratings for all responses across subjects and across the six test pair categories was significant, $F(5, 195) = 9.55$ ($p < .001$). Therefore, subjects' confidence ratings were significantly different across the six test pair categories. An analysis of variance for correct confidence ratings was also significant, $F(5, 195) = 9.86$ ($p < .001$). The analysis of variance for wrong confidence ratings across test pair categories was not significant.

Experiment 1: An analysis of variance for total median confidence ratings across the 4 test pair categories was significant, $F(3, 123)=10.98$ ($p < .001$). In addition, the ANOVAS for the correct and wrong confidence ratings were also significant: correct-$F(3,123)=10.82$ ($p < .001$); wrong-$F(3, 123)=2.95$, ($p < .05$).

Below is a listing of the means of the median confidence ratings for both experiments across the test pair categories for total, correct and wrong responses from the highest median to the lowest median. To compare confidence ratings with performance, subjects' d' values are also listed.

Table 20.

The mean confidence ratings for all medians, correct medians and wrong medians from Experiments 1 and 2 from highest to lowest rating.

Highest						Lowest
Experiment 2 - Mean total median confidence ratings						
	t2-c2 >	**t2-c1 >**	**t2-t1 >**	**c2-t1 >**	**t1-c1 >**	**c2-c1**
mean:	2.30	2.25	2.23	2.08	1.91	1.75
	t2-c1 >	**t2-t1 >**	**t2-c2 >**	**c2-c1 >**	**t1-c1 >**	**c2-t1**
mean d':	1.03	.636	.586	.419	.340	.026
Experiment 1 - Mean total median confidence ratings						
	t2-t1 >	**t2-c2 >**	**t1-c1 =**	**c2-c1**		
mean:	2.25	2.18	1.71	1.71		
	t2-t1 >	**t2-c2 >**	**t1-c1 >**	**c2-c1**		
mean d':	.603	.533	.384	.249		
Experiment 2 - Mean correct median confidence ratings						
	t2-c1 >	**t2-t1 >**	**t2-c2 >**	**c2-t1 >**	**t1-c1 >**	**c2-c1**
mean:	2.40	2.37	2.33	1.91	1.90	1.83
	t2-c1 >	**t2-t1 >**	**t2-c2 >**	**c2-c1 >**	**t1-c1 >**	**c2-t1**
mean d':	1.03	.636	.586	.419	.340	.026
Experiment 1 - Mean correct median confidence ratings						
	t2-t1 >	**t2-c2 >**	**t1-c1 >**	**c2-c1**		
mean:	2.31	2.27	1.93	1.67		
	t2-t1 >	**t2-c2 >**	**t1-c1 >**	**c2-c1**		
mean d':	.603	.533	.384	.249		
Experiment 2 - Mean wrong median confidence rating						
	c2-t1 >	**t2-c2 >**	**t2-t1 >**	**t1-c1 >**	**t2-c1 =**	**c2-c1**
mean:	2.00	1.91	1.88	1.81	1.64	1.64
	t2-c1 >	**t2-t1 >**	**t2-c2 >**	**c2-c1 >**	**t1-c1 >**	**c2-t1**
mean d':	1.03	.636	.586	.419	.340	.026
Experiment 1 - Mean wrong median confidence ratings						
	t2-t1>	**t2-c2 >**	**t1-c1 >**	**c2-c1**		
mean:	1.82	1.81	1.52	1.44		
	t2-t1 >	**t2-c2 >**	**t1-c1 >**	**c2-c1**		
mean d':	.603	.533	.384	.249		

Looking at the mean of the median confidence ratings for trials where the subjects were correct, confidence and performance are clearly highly correlated. For example, in Experiment 1, the order of the test pair categories from the highest to lowest for the mean of the median confidence ratings and the d' values is the same (c1-c2 < c1-t1 < c2-t2 < t1-t2). This is what would be expected if frequency memory is explicit for both the intentional and incidental conditions. The observed correlation is also consistent with the idea that frequency memory is one-dimensional. However, this correlation is based on means over subjects, and there may be large individual differences in how subjects use confidence ratings. Thus, the confidence ratings may incorporate response biases that differ from subject to subject. If this is so, then using the mean of the confidence ratings over subjects may provide a misleading picture of the relationship between confidence and performance.

Subjects may differ in their use of the confidence ratings. For example, one might have a pessimistic subject with high d' values, but low confidence ratings. Or, conversely, one might have an optimistic subject with high confidence ratings and low d' values. That is, the confidence ratings may incorporate response biases that differ from subject to subject. A possible solution is to normalize the confidence ratings for each subject. For example, find the mean confidence rating for a subject collapsed across all conditions. Then, for each condition for that subject, use the difference between the median for that condition and the "grand mean" divided by the standard deviation. The values are listed in Table 21 below:

Table 21.
Means of normalized confidence ratings for both experiments.

Experiment 1

Category	t2-t1	c2-c1	t1-c1	t2-c2
Total mean	.407	-.328	-.354	.268
SD	.565	.787	.650	.630
Correct mean	.356	-.391	-.221	.257
SD	.587	.736	.769	.559

Wrong mean	.043	-.288	-.108	.296
SD	.824	.829	.807	.741
d' values:	.603	.249	.384	.533

Total normalized confidence rating: t2-t1 > t2-c2 > c2-c1 > t1-c1

Correct normalized confidence rating: t2-t1 >t2-c2 >t1-c1 > c2-c1

Wrong normalized confidence rating: t2-c2 >t2-t1 >t1-c1 >c2-c1

Experiment 2

Category	t2-t1	c2-c1	t2-c1	c2-t1	t2-c2	t1-c1
Total mean	.281	-.618	.230	-.037	.368	-.393
SD	.758	.677	.848	.751	.625	.837
Correct mean	.291	-.528	.492	-.236	.301	-.293
SD	.745	.734	.632	.980	.579	.811
Wrong mean	.083	-.242	-.175	.231	.141	-.025
SD	.951	.652	.863	.622	.939	.975
d' values:	..636	.419	1.03	.026	.586	.340

Total normalized confidence rating: t2-c2 >t2-t1 >t2-c1 >c2-t1 >t1-c1 >c2-c1

Correct normalized confidence rating: t2-c1 >t2-c2 >t2-t1 >c2-t1 >t1-c1 >c2-c1

Wrong normalized confidence rating: c2-t1 >t2-c2 >t2-t1 >t1-c1 >t2-c1 >c2-c1

The trends in the confidence ratings remain the same. Subjects typically score higher for t2-t1, t2-c2 and t2-c1 pairs for the normalized confidence ratings. They scored very low for t1-c1 and c2-c1 test pairs. Subjects show better confidence ratings for intentionally processed items than for incidentally processed items. Also, the medians are higher for total and right confidence ratings than for wrong confidence ratings.

Correlations were done to see the relationship between d' values and the normalized median confidence ratings in both experiments. Does a subject's confidence change in the same direction with performance as measured by the d' values? Below are the Pearson rs for Experiment 1 for t2-t1; c2-c1; t1-c1; and t2-c2 categories.

Table 22.

Pearson r values comparing d' and median confidence ratings for Experiment 1 across subjects.

	d':	CR:	r:	
t1-c1:	.38	-.354	+.4495*	a = 1.3659; b = 0.9075; sx = 0.5106; sy = 1.0308
t2-t1:	.60	.407	+.2611	a = 2.0580; b = 0.3184; sx = 0.6546; sy = 0.7982
c2-c1:	.25	-.328	+.2017	a = 1.6178; b = 0.3919; sx = 0.5151; sy = 1.0008
t2-c2:	.53	.268	+.1691	a = 2.0565; b = 0.2289; sx = 0.6361; sy = 0.8612

Note. d' = mean d'; CR = mean of the median total confidence rating; r = Pearson r

Having normalized the confidence ratings for each subject makes it possible to compare subjects to see if variations in performance are correlated with variations in confidence. Pearson r correlations were calculated across subjects for each of test-pair category. In general, the correlations are positive although only one correlation reaches significance. (There are two negative correlations, both in Experiment 2, neither significant, In one of these cases (c2-t1), the correlation is close to 0 (-0.07) as is the d' (0.03). More curious however, is the t2-t1 test pair where the correlation is -0.17 and the d' = 0.64. However, since there was a positive correlation for this pair in Experiment one, the most likely explanation of this non-significant correlation is noise.) The following is the Pearson r for Experiment 2 (193) for d' and median confidence ratings.

Table 23.

Pearson r comparing d' values and normalized median confidence ratings for Experiment 2.

	d'	CR:	r:	
c2-c1:	0.42	-.618	+.2830	a = 1.5761; b = 0.4155; sx = 0.5175; sy = 0.7596
t1-c1:	0.34	-.393	+.2055	a = 1.8249; b = 0.2573; sx = 0.6258; sy = 0.7835
t2-c2:	0.59	.368	+.1854	a = 2.2252; b = 0.1276; sx = 0.8826; sy = 0.6076
t2-t1:	0.64	.281	-.1737	a = 2.3354; b = -0.1736; sx = 0.6501; sy = .6499
t2-c1:	1.03	.230	+.1729	a = 2.1013; b = 0.1446; sx = 0.8456; sy = 0.7071
c2-t1:	0.03	-.037	-.0695	a = 2.0732; b = -0.0695; sx = 0.6287; sy = 0.6845

Note: d' = mean d', CR=mean of the median total confidence rating; r = Pearson r

None of the above correlations was significantly different from r = 0 and therefore, subjects in Experiment 2 were not good at judging how accurate their frequency judgments were in the six categories. Even the significant performance in the t1-c1 category from Experiment 1 did not replicate in Experiment 2. In conclusion, although the confidence ratings track d' performance over the various test pairs, the correlations of the confidence ratings and d' across subjects for each test pair type are relatively weak.

Questionnaire Data:

Of the 40 subjects in Experiment 2, 37.5% were freshmen, 32.5% were sophomores, 20% were juniors, 7.5% were seniors and 2.5% were unknown. When asked what they thought this experiment was about; 67.5% thought it was about memory; 22.5% thought it was about obtaining information and remembering it; and the remaining 10% thought it ranged from attention span to word association to *exploiting* the mind. When asked if they utilized a strategy; 75% said yes and 25% said no. When asked if they thought the confidence ratings were difficult to do; 50% said yes and 50% said no. When asked if they thought the frequency task was difficult; 70% said it was easy and 30% said it was difficult. The following is a list of strategies use in the order of preference: relating words together; repetition; concentrating on the words; made up stories using the words; and counting.

In hindsight, the questionnaire needs to be constructed better to answer the following questions: 1) Do subjects process the target words differently than the context words?; 2) Are subjects aware that they are being tested on the context words?; 3) What role did they think the context words played?, and 4) Did they know that some pairs contained words that had equal frequency of occurence (i.e., t1-c1, t2-c2) and how did they make frequency judgments for these pairs?

EXPERIMENT THREE

The first two experiments directly measured the effect of intentional processing via a relative frequency task. The results were that intentional processing improved frequency discrimination. But as discussed earlier, some investigators have found incidental processing to be superior in some tasks (e.g. Jacoby, 1983). What about measuring the memory for context words by using a method that is sensitive to words processed incidentally? As mentioned earlier, much discussion has taken place over the definition of implicit and explicit memory systems. In the literature, **implicit** and the term "indirect" and **explicit** and the term "direct" are used interchangeably (Richardson-Klavehn and Bjork, 1988; Schacter, 1992 and 1987) to describe both memorial systems and test measures. Richardson-Klavehn and Bjork (1988) describe an **explicit** task as one that makes obvious reference to a particular event or episode in the subject's personal history (e.g., the experiment). The term explicit refers to both memory systems and tasks that are *declarative, autobiographical, direct, episodic, or intentional* (Richardson-Klavehn and

Bjork, 1988; Schacter, 1987, 1992; Zola-Morgan and Squire, 1990). According to these authors, explicit tasks such as recall and recognition necessitate that a subject utilize conscious recollection or awareness when retrieving information from memory. For Richardson-Klavehn and Bjork (1988), subjects in an experiment are made "aware" via instructions (prior to the presentation of the stimuli) that they will be tested on their memory for those stimuli.

In contrast to the term explicit, **implicit** memory systems or tasks have been defined as *nondeclarative, indirect, nonconscious, unintentional, or incidental* (Richardson-Klavehn and Bjork, 1988; Schacter, 1987, 1992; Zola-Morgan and Squire, 1990). Essential to implicit tasks is that at the time of testing, no explicit reference is made to subjects' specific events in their personal history (e.g., experimental situation) even though subjects' performance may be influenced by those events. Implicit effects are indicated by improved accuracy or faster response time in performance due to prior exposure to the experimental stimuli. Typical implicit memory tasks include stem completion or fragment completion tasks (e.g., TRA___ or B_SH_L). When the test stimuli are semantically, graphemically, phonemically, or morphologically related to the experimental stimuli, increased accuracy or decreased latency in performance is generally interpreted as reflecting the effects of implicit memory. Implicit or indirect measures can also include tests of procedural knowledge, tests of conceptual, lexical, or perceptual knowledge, and measures of behavioral change such as neurophysiological responses and conditioning measures (Richardson-Klavehn and Bjork, 1988).

Schacter (1992) and Zola-Morgan and Squire (1990) point out that explicit and implicit tests have produced dissociations in the effects of some experimental factors on subjects' performance. Generally, factors such as depth of processing will affect the outcome of an explicit measure but not an implicit one. On the other hand, factors such as change in modality (auditory versus visual) and typography at study and at test will affect the outcome of an implicit measure but not an explicit one. In addition, research on amnesia patients such as H.M. has shown that while these subjects exhibit little explicit memory they can show improvement in an implicit, procedural task, (e.g., decreased latency in completing a puzzle or maze without conscious awareness of ever doing the puzzle or maze before, Zola-Morgan and Squire, 1990).

Researchers like Merikle and Reingold (1991), Blaxton (1989), Schacter (1992), Neill, Beck, Bottalico, and Molloy (1990) and Richardson-Klavehn and Bjork (1988) all discussed these dissociations between explicit and implicit memory measures. Some like Schacter (1992, 1987), Zola-Morgan and Squire (1990) and Graf and Ryan (1990) suggest that there are separate memory systems such as semantic and episodic memories or declarative and nondeclarative memories to explain these dissociations. In contrast, others like Blaxton (1990) believe that different processing is called into play when using implicit and explicit measures. Taking a cue from Morris, Branford, and Franks (1977), Blaxton (1989) suggests that how much one remembers is determined by the amount of overlap in the processes utilized at both the time of study and at test (transfer-appropriate processing). For example, for explicit tasks like recall and recognition, deliberate or conscious recollection of a previous episode is necessary while for implicit tasks such as word completion, this is not required. According to Reber, Allen and Reber (1999), recent neuroanatomical work shows that "conscious, explicit learning and memory seem to depend on specific brain structures, while implicit learning is likely subserved by different and rather diffuse brain structures."

What Schacter (1992) does caution is that the field of cognition not create a new memory system whenever dissociations occur. Schacter points out that there are dissociations between recall and recognition performance; does this mean new memory systems must be created for dissociations within explicit memory? He warns of "theoretical chaos" and recommends that the field look for external validity for existing dissociations in implicit and explicit memory in other fields such as cognitive neuroscience (Schacter, 1992).

Merikle and Reingold (1991) raise several objections to using the term explicit and implicit because the terms have been used to describe both memory tasks and memory processes. The authors believe that this "confuses the distinction between theoretical constructs and empirical measures" (Merikle and Reingold, 1991). They prefer to exchange the term indirect for implicit and direct for explicit for these experimental tasks because there is no need to specify underlying memory processes or structures or any need to describe phenomenal experiences such as awareness or intention.

Results from the first two experiments show that frequency processing is affected by intentional instructions, but this does not preclude frequency processing that occurs incidentally. Recall that Hasher and Zacks (1984) state that one criterion for automaticity is that people are sensitive to frequency information without necessarily intending to be. Naveh-Benjamin and Jonides (1986) suggested that both "direct and indirect" factors may affect frequency processing. In Experiments 1 and 2, the frequency discrimination task taps intentional frequency memory. The task in experiment 1 and 2 tested explicit conscious processes by using an old/new discrimination recognition task (a direct measure). Perhaps in order to measure automatic, incidental frequency processing, an implicit task should be administered to tap this type of memory. The goal of the third experiment is to see whether the difference between intentional and incidental processing found in the first two experiments will disappear in an implicit memory task. Work by Greene (1986), Neill, Beck, Bottalico, and Molloy (1990) and Schacter (1987) suggests that the word fragment or stem completion task is one method for measuring implicit memory.

Experiment 3 Design:

The design of this experiment is similar to the design of Experiment 2 except that after the presentation of the stimulus list, subjects did not make judgments of the relative frequency of test pairs but instead performed a stem and fragment completion task. The task entails presenting word stems (ex., tab_ _) and asking subjects to complete them. The question is whether subjects' performance is equally good (percentage correct) for the stem completion task under both the incidental and intentional conditions in contrast with the results of the frequency judgment tasks (percentage correct) under incidental and intentional conditions in Experiment 2. Again, the design is *within* subject: each subject responds in both the incidental and intentional conditions.

Subjects:

Subjects were recruited from the Introduction to Psychology Subject Pool. Twenty-two subjects were run in small groups.

Stimulus List and the Stem/Fragment Completion Task:

Subjects viewed the same practice stimulus list and experimental stimulus lists used in Experiment 2. All words presented on the experimental stimulus list were counterbalanced across subjects for the target/context condition and for 1 versus 2 presentations. An example of a stimulus pair and the corresponding fragment completion items follow:

Stimulus pairs: BELL - ring X 1
BELL - ring X 2
RING - bell X 1
RING - bell X 2

Stem/Fragment Completion Task: b e __ __
r __ n __

Memory Test Example:

Which describes the meaning of "bank" in the experiment?
____money

____river

Again, since all the stimulus items were counterbalanced across frequency conditions (once or twice presented) and type (target or context), the same words were tested under both incidental and intentional conditions across subjects. Each test booklet had one fragment completion item per page. There were a total of 40 fragment completion items for both the incidental and intentional conditions (with four buffer items each at the beginning and end of the list). Each fragment completion item was chosen with the restriction that there would be no other match for it from the stimulus list. The experiments by Blaxton (1989) show that typography does not affect explicit memory performance but does affect implicit memory performance. This effect can be tested by presenting the intentional and incidental words in both upper and lowercase letters. Therefore, the fragment completion item was printed in the same typeface as the stimuli on the list (Helvetica) but half were printed in uppercase and half were printed in lowercase letters.

Stimulus list:

Target = Intentional condition = uppercase
Context = Incidental condition = lowercase

Test list:

Intentional words = upper and lowercase
Incidental words = upper and lowercase

Experiment 3 Procedure:

1-Instructions and present Practice List (18 pairs)

2-Administer Practice Memory Test

3-Review instructions to see subjects understand experiment

4- Present Stimulus List

5-Administer Fragment Completion Booklets

6-Collect Test Booklets and Hand out Memory Tests and Questionnaires

7-Collect Memory Tests and Debrief Subjects

Subjects were instructed that they would be tested on their memory for target words and would be given a practice stimulus list and then a practice memory test (like the one in Experiment 2; see above). The practice memory test contained only questions referring to the target items from the practice stimulus list and therefore, helped to ensure that subjects attended to the target items. The practice memory test was completed by subjects and was then reviewed by the group. Subjects were then asked if they understood the task and if there were any further questions. Once all questions were answered, the stimulus list was presented. After the stimulus list was presented, subjects were given the fragment completion task.

Subjects were instructed "that each fragment completion item is part of an English word. Some words will be in uppercase letters and some will be in lowercase letters. We would like to examine whether the typography of the words will affect memory for those words. So when you see the letters, please write the first word that comes to mind because you will have a limited time to fill in each word. No proper names are allowed." These instructions are similar to those used by Graf, Shimamura, and Squire (1985). Subject were given 20 seconds to complete each item. This is the amount of time Blaxton (1989) gave each subject for the word fragment completion. A subject's response was considered correct only if the word appeared in the stimulus list in the same form - "no plurals, changes in tense or adjectival forms" (Graf, Shimamura, and Squire, 1985). No confidence ratings were administered in this experiment. After the fragment completion task was completed, subjects took the memory test (like the one Experiment 2) to see if they remembered the context of the target words presented during the stimulus list. After the memory test, subjects filled out a questionnaire asking them the same questions used in Experiments 1 and 2.

Experiment 3 Data Analysis:

Each response was scored correct if it matched exactly a word in the stimulus list. Each subject's booklet was scored for the number of correct completed word fragments for target words (intentional/explicit processing condition) and for context words (incidental/implicit processing condition). Each subject's proportion correct was arc sine transformed as in the first two experiments.

Analysis of Variance and t-tests were applied to arc sine values to see if performance was significantly different for target words and for context words and to see whether typography was a factor. It was hypothesized that typography would not affect fragment completion of target words, but that subjects would display a higher percentage correct for context words tested in the lowercase type than the uppercase type. This is because the context words were originally studied in the lowercase type and according to Blaxton (1989), subjects perform best on stem completion tasks when the typography is the same at study and at test.

Percent correct was calculated in two ways: 1) with blank responses included in the denominator and 2) with blank responses excluded from the denominator. For example, suppose that a subject produces 5 "matching" responses, 2 blank responses, and 3 "unmatch" responses. If blanks are included, the subject has 50% (5/10) correct response and if blanks are not included, the subject has 62.5% (5/8) correct response.

Table 24.

Percentage correct across item type and frequency.

Fragment Completion Data

Blanks Included in the Denominator:

Item Type	Frequency 1	2	Mean
Target	16.8%	19.1%	18.0%
Context	24.1%	15.9%	20.0%
Mean	20.5%	17.5%	

Blanks Excluded in the Denominator:

Item Type	Frequency 1	2	Mean
Target	18.7%	21.2%	20.0%
Context	28.3%	16.7%	22.5%
Mean	23.5%	19.0%	

The trend in the performance was that context words presented once had the best performance compared to the other 3 conditions (c1 > t2 > t1 > c2). But an analysis of variance using the arc sine transformed percent correct values showed no significant difference between the four conditions (blanks included arc sine: $F(3,63) = 1.49$; blanks deleted arc sine: $F(3, 63) = 1.97$; all not significant).

The data were further categorized according to typography, frequency and target/context conditions. Arc sine values can be found in Appendix F. Below is a summary of means and standard deviations for percentage correct:

Table 25.
Percentage correct across item type and typography in Experiment 3.

Blanks Included in the Denominator:

Mean (SD)	Target 1 UU	Target 1 UL	Target 2 UU	Target 2 UL
% :	21.83 (23.02)	12.73 (15.79)	20.00 (17.46)	18.18 (17.36)

Mean (SD)	Context 1 LU	Context 1 LL	Context 2 LU	Context 2 LL
%:	19.09 (14.44)	29.09 (19.25)	17.27 (15.49)	14.55 (14.05)

Blanks deleted from the Denominator:

Mean (SD)	Target 1 UU	Target 1 UL	Target 2 UU	Target 2 UL
% :	23.18 (25.66)	13.64 (16.34)	24.09 (20.85)	19.09 (18.17)

Mean (SD)	Context 1 LU	Context 1 LL	Context 2 LU	Context 2 LL
%:	22.88 (18.25)	33.78 (21.37)	18.41 (17.14)	15.00 (14.31)

*Note: U = uppercase; L = lowercase

Correlated t-tests performed on the blanks deleted data showed no significant difference in fragment completion between c2 and t2 words. While there was a significant difference between c1 and t1 words ($t(21) = 1.97$, $p < .05$) for data with the blanks deleted, there was no significant difference found for the blanks included data, ($t(21) = 1.65$). In addition, the significant result is marginal and may be a type I error.

Memory Test Results:

The memory test was administered directly after the fragment completion task. It was identical to the one used in Experiment 2 when subjects were presented with a stimulus word with different meanings and asked to chose the definition that best defined the flavor of the word in the experiment. For example,

Which describes the meaning of "bank" in the experiment?

____money
____river

Subjects displayed a mean percent correct of 73.64 for target words and 75.91 for context words . More specifically, comparing the target performance on the fragment completion task versus the memory task produced a r = 0.07 while a Pearson r comparing performance on context words on the fragment completion task versus the memory task is r = .43.

The previous results show general performance on the memory test but unlike Experiment 2, not all the items tested in the memory test were items in the stem completion task. Because of the restrictions of creating items for the stem completion task, not all the words that appeared on the stimulus list were used, unlike in the relative frequency test in Experiment 2. Therefore, there were six words tested in both the fragment completion task and memory test. Performance on these six words is shown below:

Table 26.

Mean percent correct for memory test and stem completion task.

Semantic Memory Test:

Mean (SD):	Target (SD):	Context (SD):	Total (SD):
%	74.23 (32.51)	75.77 (34.44)	75.00 (23.44)

Fragment Completion Task:

Mean (SD):	Target (SD):	Context (SD):	Total (SD):
%	7.50 (14.15)	10.50 (15.73)	9.09 (9.90)

A correlated t-test performed on the semantic memory test arc sine values for target and context words showed no significant difference in subjects' performance ($t(21) = .10$, $p > .05$). Pearson r calculations were done to compare subjects' performance on the semantic memory test and the fragment completion task for the six words:

Table 27.

Pearson r comparing stem completion versus memory test.

Category:	Pearson r:
target memory vs. target fragment	+.24
context memory vs. context fragment	-.26
total memory vs. total fragment	-.37

Calculations were computed to determine if any of these correlations was significantly different from a $r = 0$ using the following formula from Bruning and Kintz (1968): $z = r\sqrt{n-1}$. If the z score is +/- 1.96, it is significant at the .05 level. Not one of the correlations was found significant. A Pearson r was calculated to determine the relationship between context and target words and this also was not significant ($r = .1621$).

Questionnaire Results:

When asked what they thought the experiment was about, 43% thought it was a memory experiment, 25% thought it was how one remembers words, 14.5% thought it was about understanding meanings, and the remaining 17.5% thought the experiment was about miscellaneous topics from vocabulary to cognitive skills. Subjects were then asked whether they thought the fragment completion task was difficult and 50% said it was hard to do and 45% said it was easy (5% did not answer this question). When asked if they were given enough time to do the fragment completion task, 77% said there was enough time and 23% said there was not enough time. When then asked if they processed target words differently than context words; 68% said they processed them the same while 32% said they processed them differently. The following is the list of the most popular strategies in order of preference: related both words; repetition; made up stories with words; concentrated only on repeated words; and counted words.

In summary, subjects' performance on the stem completion task was very low. Percentage correct ranged from 15.9 to 24.1% for all subjects across the four conditions. Although the trend in the data was : c1 > t2 > t1 > c2, performance was not significantly different across the four conditions. To improve the methodology of future experiments, I would suggest that 1) more stimulus items be used, and 2) a greater range of frequency of occurence be utilized. Using just frequencies of one and two may not be enough to test incidental and intentional processing in a stem completion task.

GENERAL DISCUSSION

The results of this research lead to two main conclusions. First, intention did affect frequency processing. Second, frequency judgments in these experiments were based on a one-dimensional memory representation that was common to both the intentional (target) and incidental (context) items. These two conclusions are discussed more fully below.

Intention and Frequency Processing

The results of Experiments 1 and 2 show significant effects of intention. The d' values show better frequency discrimination for words that were intentionally processed than for words incidentally processed. This contradicts the second criterion set forth by Hasher and Zacks (1984) that states that frequency processing is automatic because "information encoded in this way is no way different than when intention is activated."

Specifically, in Experiment 1, d' for t2-t1 word pairs was significantly better than for c2-c1 word pairs. These two types of pairs are similar in that both compare a word presented twice with a word presented once, but they differ in "intention". Therefore, subjects are more discriminating when judging the relative frequency of target word pairs (intentional) compared to context word pairs (incidental).

In Experiment 2, subjects were best with the t2-c1 pairs and worst with the c2-t1 pairs, where performance was close to chance. The d' values for t2-t1 test pairs versus c2-t1 test pairs also differed significantly. In these two test pairs, both t2 and c2 words have the same frequency of occurrence but were under different "intention" conditions. Here, it is easier for subjects to discriminate between t2 and t1 words than to discriminate between c2 and t1 words. Another difference was found for t2-c1 versus c2-c1 pairs where t2-c1 performance was better. These results lead to the conclusion that the subjects in Experiments 1 and 2 show better frequency memory for the target words than for the context words because of the effects of intention.

Why did intention improve frequency memory for target words in the present experiments, in contrast to the results reported by Hasher and Zacks and others? One obvious possibility is that intention, per se, has no direct effect on memory but rather affects processing. For example, as noted above, Hyde and Jenkins (1973) also found no effect of intention in a free recall task but found a big effect of the task that subjects performed during the learning phase. This raises the obvious possibility that under some conditions, variations in intention might lead to variations in processing. As has been shown by Greene (1986) and Brown (1995, 1997), subjects who are instructed to pay particular attention to words are more likely to engage in different strategies to process the words. In the present experiments, telling subjects to remember target words based on information provided by the context word may have produced deeper or more complex processing of the target words.

For example, in an experiment reported by Greene and Thapar (1994), subjects were asked to deeply process words (at the time of presentation, subjects were asked questions like "How easy is it to form an image of the word?") and subjects were asked to shallowly process words (subjects were asked questions like "How many ascending letters are in the word?"). Half of these subjects were informed that some of the words were deeply and shallowly processed and the other half were uninformed. Greene and Thapar (1994) found that for deeply processed pairs in both informed and uninformed conditions, frequency judgment performance was about the same (.80 versus .81). For shallowly processed words across informed and uninformed subjects, performance was almost the same (.58 versus .64). For the repeated deeply processed word-once presented shallowly processed word test pair (.92 versus .92) and for the repeated shallowly processed word-once presented deeply processed word test pair (.35 versus .25), performance was similar under informed and uninformed conditions. In fact, frequency judgment performance was found to be not significantly different across the informed and uninformed conditions. The reason Greene and Thapar (1994) did not find any difference for the informed and uninformed groups was that the level of processing was the same across both groups. That is why, it is crucial to account for levels of processing when asking questions about intention in frequency processing.

In contrast, in the same paper, Greene and Thapar (1994) presented subjects with a list of words and non-words presented once or twice. The test pairs were constructed as follows: word-word (mean proportion correct =.89), non-word-non-word (mean proportion correct =.65); repeated word-once presented non-word (mean proportion correct =.77) and repeated non-word and once presented word (mean proportion correct =.81). Subjects differed in the levels of semantic processing for the words and non-words but the results did not show a uni-dimensional representation in memory. It is not clear how levels of processing of different kinds of stimuli (such as words and non-words or pictures) may translate into effects on frequency performance.

Unfortunately, the present experiments and the post-experimental questionnaire were not designed to assess the possibility of different types of processing for target and context words, which should clearly be an objective for further work in this area. In any case, to the extent that it is the type of processing rather than intention that is important, the inconsistency in the results of experiments that have focused on intention without controlling processing (including the present experiment) is perhaps not altogether surprising.

So where does this leave the question of automaticity and frequency processing? Naveh-Benjamin (1986) suggested that both automatic and non-automatic processes are at work in frequency memory tasks, and that experimental manipulations could alter the relative influence of these processes. But this idea appears to be inconsistent with the results of the d' analysis of the present experiments, as discussed next.

The Representation of Frequency Information

The Naveh-Benjamin and Jonides (1986) suggestion that both automatic and non-automatic mechanisms might be involved in processing frequency information raises that possibility that there could be more than one way for frequency information to be represented in memory. This possibility was discussed earlier in the sections on "Theories of Memory Representation" and "The Representation of Frequency Information." Current conceptions of memory support the idea that there are different representations associated with explicit versus implicit learning conditions, and neurological evidence strongly supports the idea that there are different brain mechanisms for perceptual and procedural learning as opposed to declarative and episodic learning. Further, even within one type of learning, information may be represented in more than one way. For example, frequency information in episodic memory might be represented in terms of trace strength, multiple traces, or propositionally by explicitly counting stimuli. And these possibilities need not be exclusive. All might be employed.

From this rich theoretical perspective, it is perhaps surprising that the results of the present experiments provide an argument for a single representational mechanism. This argument is based on the d' analysis, particularly of Experiment 2. This analysis found that the d' values for the various test pairs could be represented as falling on a single underlying strength-like representation of frequency regardless of whether the pairs involved context or target items. This is perhaps surprising because as just discussed above, the superior memory for the frequencies of the target items could have arisen because target items were processed differently, e.g. more deeply. And as also discussed above, Brown (1995, 1997) and others have suggested various ways in which intention can lead to different processing and consequently different representations of frequency information. Therefore, it is somewhat puzzling that although the intention variable did lead to better frequency memory, this variable apparently did not lead to different representations of frequency information for target and context items. One possible solution to this puzzle is to recognize that subjects might be able to combine frequency information from various memory representations so as to make their 2AFC judgments. To the extent that subjects can do this, the combination of information in the judgment process will conceal the multiple representations that might underlie that judgment.

In fact, Jonides and Jones (1992) suggest that both a direct code (such as a frequency tag) and an indirect code (such as deriving frequency information from the number of traces) are involved in the encoding of frequency information and that subjects use both when making frequency judgments. One way in which these two sources of frequency information may work together is that the "direct code might provide a ballpark estimate of frequency for an item, and then this anchor would be finely tuned on the basis of the multiplicity of traces stored for that item," (Jonides and Jones, 1992, p. 378).

On the other hand, subjects were asked to make confidence judgments, and it seems plausible to assume that confidence would be related to the source of the information in a subject's memory. The decision to include confidence ratings in these experiments was based, in part, on the intuition that subjects might report greater confidence for explicitly learned information than for implicitly learned information. That is, it seemed possible that subjects might show comparable performance for judgments of target (intentional) and context (incidental) items, but be more confident with the target items. However, this was not the case. Confidence judgments showed close correspondence with the d' values, which is again consistent with the idea that subjects were basing their judgments on a single underlying representation of frequency.

If there is only a single representation of frequency information in these tasks, what was it? Unfortunately, the 2AFC results can be potentially described by any of several representations, e.g. trace strength, multiple traces, or a propositional "counter" representation. In fact, even a simple two-state threshold model in which each item in a 2AFC test pair was judged simply as either familiar or unfamiliar might be able to predict the present results. To discriminate among these possibilities, a richer set of data would be needed such as might be obtained by having more than two levels of presentation frequency.

Experiment 3

The 2 AFC frequency task in Experiments 1 and 2 taps intentionally and incidentally processed frequency information using an explicit or direct retrieval task. Perhaps some conflicting research on the role of automaticity of processing frequency information is due to the type of task used to retrieve this information. In order to measure incidental frequency processing, an implicit task should be administered to tap this type of memory.

Experiment 3 was run to see if we would find stronger memory for the incidentally processed context words using the fragment completion task.

Performance in Experiment 3 was poor, with percent of items completed ranging from 16% to 24% across conditions, and the differences across the various types of items, c1, c2, t1, t2, were not significant. Performance was low in part because many subjects showed problems with spelling their responses correctly. Although the power of this analysis is weak because of low performance and the number of subjects, the lack of any significant differences across the item types provides some support for the idea that the fragment completion task is not affected by intention, in contrast to the clear effects of intention on the frequency judgment task found in Experiments 1 and 2. In addition, it would have been desirable to include control items in the test, that is, items that had not appeared in the stimulus list.

In terms of performance on the semantic memory test, the subjects in Experiment 3 were as accurate as the subjects in Experiment 2, with a mean percent correct of 74% for target words and 76% for context words. On the other hand, the correlation between semantic memory test performance and the fragment completion task was not significant. This may reflect the difference in the two types of task, or may be due to the relatively low power of this analysis because of the small number of items in the semantic memory task and the relatively small number of subjects. These results would be more convincing if based on more trials and more subjects with better spelling abilities. In summary, although the idea behind Experiment 3 was interesting, these results contribute little to our understanding of the frequency judgment experiments.

Summary

First, by using a within subject design (the subject was in both the incidental and intentional conditions) and a 2AFC frequency judgment task, the data from Experiments 1 and 2 show that intention affects frequency processing. This contradicts one of the criterion that Hasher and Zacks (1984) utilized to qualify frequency processing as automatic. The results of the two experiments also suggest that subjects processed target words differently than they processed the context words.

Secondly, the d' data from Experiment 2 show that the d' values for the different test pairs could be represented by a single underlying strength-like representation regardless of whether the pairs involved context or target words. This goes against the notion that intentional words are processed differently or more deeply than context words because research suggests that different processing produces different representations. However, subjects may be able to combine frequency information from these various representations to make their 2 AFC judgments.

In future experiments, it would be interesting to see what results we would get with increased frequency of occurence because Experiments 1 and 2 used only frequencies of one and two. A better test of the uni-dimensional representational model of frequency memory would be to utilize a greater range of event frequencies. It might also be interesting to vary exposure time as Naveh-Benjamin & Jonides (1986) did to see if that would affect the type of direct and indirect codes being utilized at encoding. It might also be interesting to manipulate the types of orienting tasks (and therefore processing) while still maintaining a within subject design. For example, subjects may be told to create a sentence for target words (the uppercase words) but record whether context words (the lowercase words) begin with a vowel or a consonant. This may serve to accentuate the different levels of processing for the intentional and incidental words.

In addition, a questionnaire should be constructed to evaluate more fully the following:

1) whether subjects utilized different strategies for processing target versus context words;

2) what type of strategies did they used;

3) whether they recognized that they were being tested on context items;

4) what role did they think the context items played;

5) did they know that some test pairs contained items with the same frequency of occurrence; and

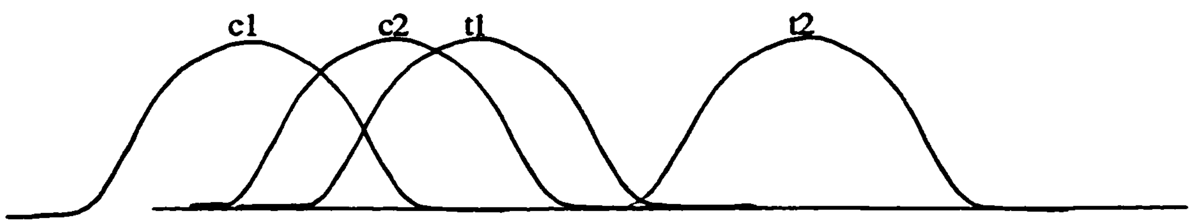
6) how did they make a decision about which item to chose in those pairs with equal frequencies.

By including more specific questions into the questionnaire, we can better ascertain what subjects are doing in regards to the the target and context words.

Appendix A

Determination of the formula used to calculate the predicted t2c2 distribution curves.

The d' values tell us the distance between specific distributions. For example, the distance between c1 and c2 distributions is 0.25 (C). The distance between c1 and t1 distributions is 0.38 (A) and the distance between t1 and t2 distributions is 0.60 (B).



c1 c2 t1 t2

C =0.25 X B = 0.60

|-----------|-------|---------------------------|

| A = 0.38 |

D = ?

|-----------------------------------|

c2 t2

What is the distance between t2 and c2 distributions? What is t2-c2 (D)?

t2-c2 (D) = t2-t1 (B) + c2-t1 (X) but c2-t1 (X) is unknown and therefore, the term c2-t1 (X) will be replaced with t1-c1 (A) minus c2-c1 (C).

Predicted t2c2 (D) = t2-t1 (B) + c2-t1 (X)
predicted t2c2 (D) = t2-t1 (B) + (t1c1 (A) - c2-c1 (C))
predicted t2c2 (D) = 0.60 + (0.38 - 0.25)
predicted t2c2 (D) = 0.73

Appendix B

WORD PAIRS: (words in bold lettering are the ones used in memory test)

List 1 & 2: TARGET x 1
Nut-bolt
SEASON-taste
SENTENCE-term
FLY-zipper
MUG-cup

List 3 & 4: target x 2
List 5 & 6: context x 1
List 7 & 8: context x 2

List 1 & 2: Context x 1
FENCE-**steal**
SECOND-**minute**
DUCK-**stoop**
BRIDGE-**cards**
STABLE-**steady**

List 3 & 4: context x 2
List 5 & 6: target x 1
List 7 & 8: target x 2

List 1 & 2: TARGET x 2
FAIR-average
COLD-virus
HABIT-nun
ROW-line
BOWL-pin

List 3 & 4: target x 1
List 5 & 6: context x 2
List 7 & 8: context x 1

List 1 & 2: Context x 2
TEACH-**train**
POUND-**beat**
LIGHT-**heavy**
CHECK-**hat**
PEN-**pig**

List 3 & 4: context x 1
List 5 & 6: target x 2
List 7 & 8: target x 1

Appendix C

d' for Experiment 1

Subject	t2-t1	c2-c1	t1-c1	t2-c2
1	-.544	0	-.178	.358
2	-.178	-.178	-.544	.544
3	0	0	.953	.741
4	-.544	0	.741	.178
5	-.741	-.741	.178	.178
6	1.191	.358	1.813	.953
7	.953	.358	.544	1.813
8	0	.178	.358	-.178
9	.358	.178	.953	.358
10	.544	.178	1.191	.358
11	-.178	-.178	.358	-.544
12	1.191	-.178	0	0
13	.358	-.358	.544	.358
14	.741	-.178	.178	2.326
15	.106	-.544	.544	.358
16	.544	-.544	.544	1.813
17	.544	.544	0	0
18	1.191	.358	.544	.178
19	.953	.358	1.465	.178
20	.544	.544	.178	-.358
21	.358	-.358	-.544	1.465
22	1.191	.953	0	0
23	-.178	.178	.544	0
24	.544	0	.358	.178
25	1.191	.544	.544	.741
26	2.326	-.178	.358	1.465
27	.358	.358	.544	.741
28	1.191	1.465	.741	.358
29	-.358	-.178	.178	.544

Subject	**t2-t1**	**c2-c1**	**t1-c1**	**t2-c2**
30	.953	.741	0	.178
31	1.813	1.465	0	.741
32	.953	-.178	1.465	1.813
33	0	0	.741	1.465
34	.953	.544	-.178	.358
35	1.191	.544	.178	.178
36	.741	.358	-.178	.178
37	.544	.358	0	.741
38	.544	1.465	-.358	0
39	.358	.358	.544	0
40	1.465	.953	.358	.544
41	.953	.358	.106	.741
42	1.191	.544	.358	.358

ARC SINE VALUES:

The arc sine transform means for each category in Experiment 1 were as follows:

Pair Category:	arc sine transform:	mean % correct:
$target_2$-$target_1$ mean	1.895	.65
$context_2$-$context_1$ mean	1.706	.56
$target_1$-$context_1$ mean	1.777	.60
$target_2$-$context_2$ mean	1.850	.63

ANOVA: $F(3,123) = 3.41$, $p < .025$

Experiment 1 (no t2-c1 and c2-t1 pairs) [t=target, c=context, Ø=arc sin transform]

pair type:	t1-t2 t2-t1	c2-c1 c1-c2	t2-c2 c2-t2	t1-c1 c1-t1	t2-c1 c1-t2	c2-t1 t1-c2
mean Ø:	1.895	1.707	1.850	1.777	NA	NA
SD Ø:	.3450	.2800	.3303	.2761	NA	NA

Significant Newman-Keuls: t2-t1 versus c2-c1 = $q4 = 4.21$, $p < .05$

Analysis of Experiment 2: [t=target, c=context, Ø=arc sin transform]

Experiment 2

pair type:	t1-t2 t2-t1	c2-c1 c1-c2	t2-c2 c2-t2	t1-c1 c1-t1	t2-c1 c1-t2	c2-t1 t1-c2
mean Ø:	1.903	1.798	1.876	1.751	2.097	1.557
SD Ø:	.3507	.2760	.4563	.3327	.3653	.3447

Significant Newman-Keuls performed on the arc sine transforms from Experiment 2:

t2-c1 versus c2-t1: q6 = 9.87 at $p < .01$ level
t2-c1 versus t1-c1: q5 = 6.32 at $p < .01$ level
t2-c1 versus c2-c1: q4 = 5.46 at $p < .01$ level
t2-c1 versus t2-c2: q3 = 4.04 at $p < .05$ level
t2-c1 versus t2-t1: q2 = 3.56 at $p < .05$ level
t2-t1 versus c2-t1: q5 = 6.31 at $p < .01$ level
t2-c2 versus c2-t1: q4 = 5.83 at $p < .01$ level
c2-c1 versus c2-t1: q3 = 4.41 at $p < .05$ level
t1-c1 versus c2-t1: q2 = 3.55 at $p < .05$ level

Significant Newman-Keuls performed on the percentage correct from Experiment 2:

t2-c1 versus c2-t1: q6 = 9.90 at $p < .01$ level
t2-c1 versus t1-c1: q5 = 6.17 at $p < .01$ level
t2-c1 versus c2-c1: q4 = 5.19 at $p < .01$ level
t2-c1 versus t2-c2: q3 = 4.03 at $p < .05$ level
t2-c1 versus t2-t1: q2 = 3.12 at $p < .05$ level*
t2-t1 versus c2-t1: q5 = 6.79 at $p < .01$ level
t2-c2 versus c2-t1: q4 = 5.87 at $p < .01$ level
c2-c1 versus c2-t1: q3 = 4.71 at $p < .01$ level
t1-c1 versus c2-t1: q2 = 3.73 at $p < .05$ level

Appendix D

d' for Experiment 2

Subject	t2-t1	c2-c1	t2-c1	c2-t1	t2-c2	t1-c1
1	.953	.622	.953	-.622	.622	.286
2	.286	.286	.622	.953	.953	0
3	-.286	0	.622	.286	1.987	0
4	0	.953	1.987	.178	-.286	.286
5	.953	1.987	1.349	.622	.953	.286
6	-.622	.286	0	.953	0	-.178
7	0	.953	1.987	.622	1.987	.286
8	0	0	0	.286	0	-.622
9	.286	-.286	1.987	-.622	1.987	.622
10	.953	.622	-.286	.953	0	0
11	1.987	-.286	1.349	-.622	.286	.286
12	.286	.622	0	0	0	.286
13	1.349	.622	.953	-.953	-.286	.286
14	0	.622	.953	0	-.953	1.349
15	.622	.622	.953	0	-.286	.286
16	.622	.286	1.987	-1.349	1.349	.953
17	.286	.506	1.091	-.178	-1.349	.741
18	-.286	.286	.622	.286	.622	.286
19	.622	0	.622	.286	.622	.953
20	0	.286	.622	.286	.286	0
21	.953	0	1.987	-1.349	1.987	.622
22	1.987	1.349	.953	.622	.622	.286
23	.953	.953	1.987	.286	1.987	.953
24	.953	.953	4.369	-.622	.286	1.987
25	1.987	.286	.622	.286	1.349	0
26	.953	.622	.953	-.622	1.349	-.622
27	1.349	.622	.622	-.622	1.987	1.987
28	.286	-.286	.286	-.286	0	-.622
29	.622	-.286	1.987	-.286	1.987	-.286
30	.622	0	1.349	-.953	-.286	.953
31	.286	1.349	1.987	-.286	.286	.622
32	.953	0	.286	.286	-.953	0
33	.953	953	.286	.286	.953	0

d' for Experiment 2

Subject	**t2-t1**	**c2-c1**	**t2-c1**	**c2-t1**	**t2-c2**	**t1-c1**
34	.286	0	.622	-.953	.622	-.286
35	.953	.286	.622	.286	.286	.953
36	1.349	.286	1.349	.867	0	0
37	-.622	0	.622	0	.286	-.286
38	.286	.953	.286	0	.953	-.622
39	1.349	-.286	.953	-.286	.953	.622
40	.953	0	.622	.953	.286	.953

Appendix E

Results of d'(t1-c1) + d' (t2-t1) - d'(c2-c1) - d' (t2-c2):

Experiment 1:

Subj. #:	Contrast/Result:
1	-0.364
2	-0.356
3	1.694
4	0.375
5	-1.126
6	3.315
7	3.668
8	0.358
9	1.847
10	2.271
11	-0.542
12	1.013
13	0.902
14	3.067
15	0.464
16	2.357
17	1.088
18	2.271
19	2.954
20	0.908
21	0.921
22	2.144
23	0.544
24	1.080
25	3.020
26	3.971
27	2.001
28	3.755
29	0.186
30	1.872
31	4.019

Subj. #:	Contrast/Result:
32	4.053
33	2.206
34	1.677
35	2.091
36	1.099
37	1.643
38	1.651
39	1.260
40	3.320
41	2.158
42	2.451

mean: 1.745; SD: 1.316

Experiment 2:

Subj. #:	Contrast/Result:
1	2.483
2	1.525
3	1.701
4	0.953
5	4.179
6	-0.514
7	3.226
8	-0.622
9	2.609
10	1.575
11	2.273
12	1.194
13	1.971
14	1.018
15	1.244
16	3.210
17	0.184
18	0.908
19	2.197
20	0.572
21	3.562

Subj. #:	Contrast/Result:	
22	4.244	
23	4.846	
24	4.179	
25	3.622	
26	2.302	
27	5.945	
28	-0.622	
29	2.037	
30	1.289	
31	2.543	
32	0.000	
33	2.859	
34	0.622	
35	2.478	
36	1.635	
37	-0.622	
38	1.570	
39	2.638	
40	2.192	mean: 1.980; SD: 1.542

Appendix F

Mean Data from Experiment 3: Blanks included in total denominator
Correct responses - U = uppercase; L = lowercase fragment completion test

Subject #:	target1-U arc sine (%)	target1-L arc sine (%)	target2-U arc sine (%)	target2-L arc sine (%)
1	80 (2.21)	0 (0)	20 (.927)	0 (0)
2	0 (0)	0 (0)	0 (0)	40 (1.37)
3	0 (0)	0 (0)	20 (.927)	20 (.927)
4	20 (.927)	20 (.927)	20 (.927)	20 (.927)
5	20 (.927)	20 (.927)	20 (.927)	20 (.927)
6	20 (.927)	0 (0)	20 (.927)	0 (0)
7	0 (0)	0 (0)	0 (0)	20 (.927)
8	20 (.927)	0 (0)	40 (1.37)	0 (0)
9	0 (0)	0 (0)	20 (.927)	20 (.927)
10	0 (0)	20 (.927)	0 (0)	0 (0)
11	20 (.927)	0 (0)	0 (0)	0 (0)
12	0 (0)	20 (.927)	0 (0)	0 (0)
13	60 (1.77)	0 (0)	40 (1.37)	20 (.927)
14	0 (0)	20 (.927)	0 (0)	40 (1.37)
15	20 (.927)	20 (.927)	20 (.927)	40 (1.37)
16	40 (1.37)	20 (.927)	20 (.927)	60 (1.77)
17	40 (1.37)	0 (0)	40 (1.37)	40 (1.37)
18	20 (.927)	60 (1.77)	0 (0)	20 (.927)
19	0 (0)	20 (.927)	20 (.927)	0 (0)
20	20 (.927)	0 (0)	40 (1.37)	20 (.927)
21	60 (1.77)	40 (1.37)	40 (1.37)	0 (0)
22	40 (1.37)	20 (.927)	60 (1.77)	20 (.927)
mean %:	21.82	12.73	20.00	18.18
mean arc sine:	.785	.522	.771	.709

Subject #:	context1-U arc sine (%)	context1-L arc sine (%)	context2-U arc sine (%)	context2-L arc sine (%)
1	20 (.927)	40 (1.37)	40 (1.37)	20 (.927)
2	20 (.927)	20 (.927)	0 (0)	0 (0)
3	20 (.927)	20 (.927)	40 (1.37)	40 (1.37)
4	0 (0)	60 (1.77)	20 (.927)	20 (.927)

Subject #:	**context1-U arc sine (%)**	**context1-L arc sine (%)**	**context2-U arc sine (%)**	**context2-L arc sine (%)**
5	60 (1.77)	40 (1.37)	0 (0)	0 (0)
6	20 (.927)	20 (.927)	40 (1.37)	20 (.927)
7	20 (.927)	20 (.927)	20 (.927)	20 (.927)
8	40 (1.37)	60 (1.77)	0 (0)	0 (0)
9	20 (.927)	40 (1.37)	20 (.927)	40 (1.37)
10	20 (.927)	60 (1.77)	20 (.927)	20 (.927)
11	20 (.927)	40 (1.37)	0 (0)	20 (.927)
12	20 (.927)	40 (1.37)	20 (.927)	20 (.927)
13	20 (.927)	20 (.927)	0 (0)	20 (.927)
14	20 (.927)	0 (0)	0 (0)	20 (.927)
15	0 (0)	20 (.927)	40 (1.37)	0 (0)
16	40 (1.37)	40 (1.37)	0 (0)	0 (0)
17	0 (0)	40 (1.37)	0 (0)	0 (0)
18	20 (.927)	20 (.927)	20 (.927)	20 (.927)
19	0 (0)	0 (0)	20 (.927)	0 (0)
20	0 (0)	0 (0)	20 (.927)	40 (1.37)
21	20 (.927)	0 (0)	40 (1.37)	0 (0)
22	20 (.927)	40 (1.37)	20 (.927)	0 (0)
mean %:	19.09	29.09	17.27	14.55
mean arc sine:	.795	1.035	.691	.608

Mean Data from Experiment 3: Blanks deleted from total denominator
Correct responses - U = uppercase; L = lowercase

Subject #:	**target1-U arc sine (%)**	**target1-L arc sine (%)**	**target2-U arc sine (%)**	**target2-L arc sine (%)**
1	100 (3.14)	0 (0)	25 (1.05)	0 (0)
2	0 (0)	0 (0)	0 (0)	40 (1.37)
3	0 (0)	0 (0)	33.3 (1.23)	25 (1.05)
4	20 (.927)	25 (1.05)	25 (1.05)	20 (.927)
5	20 (.927)	25 (1.05)	25 (1.05)	25 (1.05)
6	20 (.927)	0 (0)	25 (1.05)	0 (0)
7	0 (0)	0 (0)	0 (0)	20 (.927)
8	20 (.927)	0 (0)	50 (1.57)	0 (0)
9	0 (0)	0 (0)	20 (.927)	20 (.927)

Subject #:	target1-U arc sine (%)	target1-L arc sine (%)	target2-U arc sine (%)	target2-L arc sine (%)
10	0 (0)	20 (.927)	0 (0)	0 (0)
11	25 (1.05)	0 (0)	0 (0)	0 (0)
12	0 (0)	20 (.927)	0 (0)	0 (0)
13	60 (1.77)	0 (0)	66.6 (1.91)	20 (.927)
14	0 (0)	25 (1.05)	0 (0)	50 (1.57)
15	20 (.927)	25 (1.05)	25 (1.05)	40 (1.37)
16	40 (1.37)	20 (.927)	25 (1.05)	60 (1.77)
17	40 (1.37)	0 (0)	50 (1.57)	40 (1.37)
18	20 (.927)	60 (1.77)	0 (0)	20 (.927)
19	0 (0)	20 (.927)	20 (.927)	0 (0)
20	25 (1.05)	0 (0)	40 (1.37)	20 (.927)
21	60 (1.77)	40 (1.37)	40 (1.37)	0 (0)
22	40 (1.37)	20 (.927)	60 (1.77)	20 (.927)
mean %:	23.18	13.64	24.09	19.09
mean arc sine:	.839	.544	.838	.729

Mean Data with Blanks deleted from the denominator:

Subject #:	context1-U arc sine (%)	context1-L arc sine (%)	context2-U arc sine (%)	context2-L arc sine (%)
1	20 (.927)	40 (1.37)	50 (1.57)	20 (.927)
2	20 (.927)	33.3 (1.23)	0 (0)	0 (0)
3	25 (1.05)	33.3 (1.23)	40 (1.37)	40 (1.37)
4	0 (0)	60 (1.77)	20 (.927)	20 (.927)
5	75 (2.09)	50 (1.57)	0 (0)	0 (0)
6	25 (1.05)	20 (.927)	40 (1.37)	25 (1.05)
7	20 (.927)	20 (.927)	20 (.927)	20 (.927)
8	50 (1.57)	60 (1.77)	0 (0)	0 (0)
9	25 (1.05)	40 (1.37)	20 (.927)	40 (1.37)
10	25 (1.05)	60 (1.77)	20 (.927)	20 (.927)
11	33.3 (1.23)	50 (1.57)	0 (0)	25 (1.05)
12	25 (1.05)	40 (1.37)	20 (.927)	20 (.927)
13	25 (1.05)	20 (.927)	0 (0)	20 (.927)
14	25 (1.05)	0 (0)	0 (0)	20 (.927)
15	0 (0)	25 (1.05)	40 (1.37)	0 (0)
16	50 (1.57)	66.6 (1.91)	0 (0)	0 (0)
17	0 (0)	50 (1.57)	0 (0)	0 (0)
18	20 (.927)	25 (1.05)	20 (.927)	20 (.927)

19	0 (0)	0 (0)	20 (.927)	0 (0)
20	0 (0)	0 (0)	20 (.927)	40 (1.37)
21	20 (.927)	0 (0)	50 (1.57)	0 (0)
22	20 (.927)	50 (1.57)	25 (1.05)	0 (0)
mean %:	22.88	33.78	18.41	15.00
mean arc sine:	.881	1.134	.714	.619

Mean Semantic Memory Test Data (general):

Subject #:	Correct Target arc sine (%)	Correct Context arc sine (%)	Correct Total arc sine (%)
1.	2.50 (90)	3.14 (100)	2.69 (95)
2.	2.50 (90)	2.50 (90)	2.50 (90)
3.	1.16 (30)	1.77 (60)	1.47 (45)
4.	2.50 (90)	2.21 (80)	2.35 (85)
5.	1.57 (50)	1.98 (70)	1.77 (60)
6.	2.50 (90)	2.50 (90)	2.50 (90)
7.	2.21 (80)	2.50 (90)	2.35 (85)
8.	2.21 (80)	1.98 (70)	2.09 (75)
9.	1.98 (70)	2.50 (90)	2.21 (80)
10.	2.50 (90)	2.50 (90)	2.50 (90)
11.	1.98 (70)	3.14 (100)	2.35 (85)
12.	1.98 (70)	1.98 (70)	1.98 (70)
13.	2.21 (80)	1.57 (50)	1.88 (65)
14.	1.77 (60)	1.98 (70)	1.88 (65)
15.	2.21 (80)	1.57 (50)	1.88 (65)
16.	1.77 (60)	2.21 (80)	1.98 (70)
17.	2.50 (90)	1.77 (60)	2.09 (75)
18.	2.21 (80)	2.50 (90)	2.35 (85)
19.	1.77 (60)	1.57 (50)	1.67 (55)
20.	1.77 (60)	2.21 (80)	1.98 (70)
21.	1.77 (60)	1.98 (70)	1.88 (65)
22.	2.50 (90)	1.98 (70)	2.21 (80)

Mean Semantic Memory Test (6 words only) Data:

Subject #:	Correct Target arc sine (%)	Correct Context arc sine (%)	Correct Total arc sine (%)
1	3.14 (100)	3.14 (100)	3.14 (100)
2.	3.14 (100)	3.14 (100)	3.14 (100)

Subject #:	Correct Target arc sine (%)	Correct Context arc sine (%)	Correct Total arc sine (%)
3.	1.22 (33)	1.92 (67)	1.57 (50)
4.	3.14 (100)	3.14 (100)	3.14 (100)
5.	1.22 (33)	1.22 (33)	1.22 (33)
6.	3.14 (100)	3.14 (100)	3.14 (100)
7.	1.92 (67)	3.14 (100)	2.29 (83)
8.	3.14 (100)	1.22 (33)	1.92 (67)
9.	3.14 (100)	3.14 (100)	3.14 (100)
10.	3.14 (100)	3.14 (100)	3.14 (100)
11.	0 (0)	3.14 (100)	1.57 (50)
12.	3.14 (100)	3.14 (100)	3.14 (100)
13.	3.14 (100)	1.22 (33)	1.97 (67)
14.	1.22 (33)	1.92 (67)	1.57 (50)
15.	3.14 (100)	0 (0)	1.57 (50)
16.	1.22 (33)	3.14 (100)	1.92 (67)
17.	3.14 (100)	3.14 (100)	3.14 (100)
18.	1.22 (33)	3.14 (100)	1.92 (67)
19.	1.92 (67)	0 (0)	1.22 (33)
20.	1.92 (67)	3.14 (100)	2.29 (83)
21.	1.92 (67)	1.92 (67)	1.92 (67)
22.	3.14 (100)	1.92 (67)	2.29 (83)

Mean Fragment Completion Task Data (6 words only):

Subject #:	Correct Target arc sine (%)	Correct Context arc sine (%)	Correct Total arc sine (%)
1.	1.22 (33)	0 (0)	.842 (16.7)
2.	1.22 (33)	0 (0)	.842 (16.7)
3.	0 (0)	1.22 (33)	.842 (16.7)
4.	0 (0)	0 (0)	0 (0)
5.	0 (0)	1.22 (33)	.842 (16.7)
6.	0 (0)	0 (0)	0 (0)
7.	0 (0)	0 (0)	0 (0)
8.	0 (0)	1.22 (33)	.842 (16.7)
9.	1.22 (33)	0 (0)	.842 (16.7)
10.	0 (0)	0 (0)	0 (0)

Subject #:	Correct Target arc sine (%)	Correct Context arc sine (%)	Correct Total arc sine (%)
11.	0 (0)	1.22 (33)	.842 (16.7)
12.	0 (0)	0 (0)	0 (0)
13.	0 (0)	0 (0)	0 (0)
14.	0 (0)	0 (0)	0 (0)
15.	1.22 (33)	0 (0)	.842 (16.7)
16.	0 (0)	1.22 (33)	.842 (16.7)
17.	0 (0)	0 (0)	0 (0)
18.	1.22 (33)	1.22 (33)	1.22 (33)
19.	0 (0)	1.22 (33)	.842 (1.22)
20.	0 (0)	0 (0)	0 (0)
21.	0 (0)	0 (0)	0 (0)
22.	0 (0)	0 (0)	0 (0)

Data Analysis of Experiment 3 but with the blanks deleted from the denominator:

Proportion Correct

Subject	Target 1	Target 2	Context 1	Context 2
1	50	12.5	30	33.3
2	0	20	25	0
3	0	28.6	28.6	40
4	22.2	22.2	33.3	20
5	22.2	22.2	55.5	0
6	10	11.1	22.2	33.3
7	0	10	20	20
8	10	22.2	55.5	0
9	0	20	33.3	30
10	10	0	44.4	20
11	14.3	0	42.9	11.1
12	10	0	33.3	20
13	33.3	37.5	22.2	10
14	14.3	25	14.3	11.1
15	22.2	33.3	12.5	22.2
16	30	44.4	57.1	0
17	20	44.4	25	0

18	40	10	22.2	20
19	10	10	0	10
20	11.1	30	0	30
21	50	22.2	11.1	25
22	30	40	33.3	11.1
-------	----------	----------	----------	---------
mean:	18.6182	21.1636	28.2591	16.6864
SD:	15.11	13.62	16.08	12.31

ARC SINE

Subject	Target 1	Target 2	Context 1	Context 2
1	1.57	.7226	1.16	1.231
2	0	.9273	1.0472	0
3	0	1.1285	1.1285	1.3694
4	.9812	.9812	1.231	.9273
5	.9812	.9812	1.6811	0
6	.6435	.6792	.9812	1.231
7	0	.6435	.9273	.9273
8	.6435	.9812	1.6811	0
9	0	.9273	1.231	1.1593
10	.6435	0	1.4585	.9273
11	.7755	0	1.4282	.6792
12	.6435	0	1.231	.9273
13	1.231	1.3181	.9812	.6435
14	.7755	1.0472	.7755	.6792
15	.9812	1.231	.7226	.9812
16	1.1593	1.4585	1.7113	0
17	.9273	1.4585	1.0472	0
18	1.3694	.6435	.9812	.9273
19	.6435	.6435	0	.6435
20	.6792	1.1593	0	1.1593
21	1.57	.9812	.6792	1.0472
22	1.1593	1.3694	1.231	.6792
-------	----------	----------	----------	---------
mean:	.7899	.8765	1.0598	.7336
SD:	.4763	.4361	.4496	.4538

Data Analysis of Experiment 3
with the blanks included in the denominator:

Proportion Correct

Subject	Target 1	Target 2	Context 1	Context 2
1	40	10	30	30
2	0	20	20	0
3	0	20	20	40
4	20	20	30	20
5	20	20	50	0
6	10	10	20	30
7	0	10	20	20
8	10	20	50	0
9	0	20	30	30
10	10	0	40	20
11	10	0	30	10
12	10	0	30	20
13	30	30	20	10
14	10	20	10	10
15	20	30	10	20
16	30	40	40	0
17	20	40	20	0
18	40	10	20	20
19	10	10	0	0
20	10	30	0	30
21	50	20	10	20
22	30	40	30	10
mean:	16.82	19.09	24.09	15.91
SD:	13.93	12.31	13.68	11.82

Data Analysis of Experiment 3 with the blanks included in the denominator:

Arc Sine

Subject	Target 1	Target 2	Context 1	Context 2
1	1.37	.644	1.16	1.16
2	0	.927	.927	0
3	0	.927	.927	1.37
4	.927	.927	1.16	.927
5	.927	.927	1.57	0
6	.644	.644	.927	1.16
7	0	.644	.927	.927
8	.644	.927	1.57	0
9	0	.927	1.16	1.16
10	.644	0	1.37	.927
11	.644	0	1.16	.644
12	.644	0	1.16	.927
13	1.16	1.16	.927	.644
14	.644	.927	.644	.644
15	.927	1.16	.644	.927
16	1.16	1.37	1.37	0
17	.927	1.37	.927	0
18	1.37	.644	.927	.927
19	.644	.644	0	0
20	.644	1.16	0	1.16
21	1.57	.927	.644	.927
22	1.16	1.37	1.16	.644
mean:	.757	.828	.986	.714
SD:	.458	.409	.426	.443

Blanks Included in the Denominator:

Mean (SD)	Target 1 U	Target 1 L	Target 2 U	Target 2 L
% :	21.83 (23.02)	12.73 (15.79)	20.00 (17.46)	18.18 (17.36)
arc sine:	.785 (.693)	.522 (.567)	.771 (.584)	.709 (.591)

Mean (SD)	Context 1 U	Context 1 L	Context 2 U	Context 2 L
%:	19.09 (14.44)	29.09 (19.25)	17.27 (15.49)	14.55 (14.05)
arc sine:	.795 (.489)	1.035 (.573)	.691 (.562)	.608 (.538)

Blanks deleted from the Denominator:

Mean (SD)	Target 1 U	Target 1 L	Target 2 U	Target 2 L
% :	23.18 (25.66)	13.64 (16.34)	24.09 (20.85)	19.09 (18.17)
arc sine:	.839 (.806)	.544 (.586)	.838 (.637)	.729 (.608)

Mean (SD)	Context 1 U	Context 1 L	Context 2 U	Context 2 L
%:	22.88 (18.25)	33.78 (21.37)	18.41 (17.14)	15.00 (14.31)
arc sine:	.881 (.562)	1.134 (.619)	.714 (.590)	.619 (.547)

Appendix G

Percentage Correct for 2AFC Frequency Task

Experiment 1:

S#:	t2-t1	c2-c1	t1-c1	t2-c2
1	35	50	45	60
2	45	45	35	65
3	50	50	75	70
4	35	50	70	55
5	30	30	55	55
6	80	60	90	75
7	75	60	65	90
8	50	55	60	45
9	60	55	75	60
10	65	55	80	60
11	45	45	60	35
12	80	45	50	50
13	60	40	65	60
14	70	45	55	95
15	53	35	65	60
16	65	35	65	90
17	65	65	50	50
18	80	60	65	55
19	75	60	85	55
20	65	65	55	40
21	60	40	35	85
22	80	75	50	50
23	45	55	65	50
24	65	50	60	55
25	80	65	65	70
26	95	45	60	85
27	60	60	65	70
28	80	85	70	60
29	40	45	55	65
30	75	70	50	55
31	90	85	50	70
32	75	45	85	90

Percentage Correct

Experiment 1:

S#:	t2-t1	c2-c1	t1-c1	t2-c2
33	50	50	70	85
34	75	65	45	60
35	80	65	55	55
36	70	60	45	45
37	65	60	50	70
38	65	85	40	50
39	60	60	65	50
40	85	75	60	65
41	75	60	47	70
42	80	65	60	60
m:	65.0714	56.4285	59.8095	62.8571
SD:	15.7486	13.2188	12.8238	14.4889

Percentage Correct for 2AFC Frequency Task

Experiment 2:

S#:	t2-t1	c2-c1	t1-c1	t2-c2	t2-c1	c2-t1
1	75	67	58	67	75	33
2	58	58	50	75	67	75
3	42	50	50	92	67	58
4	50	75	58	42	92	55
5	75	92	58	75	83	67
6	33	58	45	50	50	75
7	50	75	58	92	92	67
8	50	50	33	50	50	58
9	58	42	67	92	92	33
10	75	67	50	50	42	75
11	92	42	58	58	83	33
12	58	67	58	50	50	50
13	83	67	58	42	75	25
14	50	67	83	25	75	50
15	67	67	58	42	75	50
16	67	58	75	83	92	17
17	58	64	70	17	78	45
18	42	58	58	67	67	58
19	67	50	75	67	67	58

Percentage Correct

Experiment 2:

S#:	t2-t1	c2-c1	t1-c1	t2-c2	t2-c1	c2-t1
20	50	58	50	58	67	58
21	75	50	67	92	92	17
22	92	83	58	67	75	67
23	75	75	75	92	92	58
24	75	75	92	58	100	33
25	92	58	50	83	67	58
26	75	67	33	83	75	33
27	83	67	92	92	67	33
28	58	42	33	50	58	42
29	67	42	42	92	92	42
30	67	50	75	42	83	25
31	58	83	67	58	92	42
32	75	50	50	25	58	58
33	75	75	50	75	58	58
34	58	50	42	67	67	25
35	75	58	75	58	67	58
36	83	58	50	50	83	73
37	33	50	42	58	67	50
38	58	75	33	75	58	50
39	83	42	67	75	75	42
40	75	50	75	58	67	75
m:	65.8000	60.8000	58.4500	63.6000	73.3000	49.4750
SD:	15.3593	12.7766	15.2584	20.1466	14.2418	16.5002

Appendix H

Experiment 1 Normalized Total Confidence Ratings:

	Test Pair Categories			
Subject	t2-t1	c2-c1	t1-c1	t2-c2
1	.50	-1.50	.50	.50
2	.57	.95	-1.33	-1.90
3	.50	.-1.50	.50	.50
4	-.20	-.20	.20	.20
5	0	0	0	0
6	0	0	0	0
7	.50	-1.50	.50	.50
8	.26	1.31	-.78	-.78
9	.50	-1.50	.50	.50
10	1.41	-.71	0	-.71
11	.26	-.78	-.78	1.31
12	0	0	0	0
13	0	0	0	0
14	.50	.50	-1.50	.50
15	0	0	0	0
16	0	-1.22	0	1.22
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	0	0	0	0
21	1.02	.44	-1.32	-.15
22	1.30	-.78	-.78	.26
23	.87	-.87	-.87	.87
24	0	0	0	0
25	-.87	.87	-.87	.87
26	1.16	-.39	-1.55	.39
27	0	0	0	0
28	.87	.87	-.87	-.87
29	1.50	-.50	-.50	-.50
30	-.87	.87	-.87	.87
31	.50	.50	-1.50	.50
32	.50	-1.50	.50	.50
33	.24	-1.20	-.24	1.20

Experiment 1 Normalized Total Confidence Ratings:

	Test Pair Categories			
Subject	t2-t1	c2-c1	t1-c1	t2-c2
34	0	0	0	0
35	1.50	-.50	.50	.50
36	.50	-1.50	.50	.50
37	.55	-.55	-1.10	1.10
38	.13	.66	-1.45	.66
39	1.31	-.78	-.78	.26
40	.66	-1.45	.13	.66
41	.26	-.78	-.78	1.31
42	1.17	-.83	-.83	.50
mean:	.407	-.328	-.354	.268
SD	.565	.787	.650	.629

Experiment 1 Normalized Correct Confidence Ratings

	Test Pair Categories			
Subject	t2-t1	c2-c1	t1-c1	t2-c2
1	0	-1.22	1.22	0
2	-.78	.26	-.78	1.31
3	.50	-1.5	.50	.50
4	-1.31	-.26	.78	.78
5	0	0	0	0
6	0	0	0	0
7	.50	-1.5	.50	.50
8	.87	.87	-.87	-.87
9	1.22	0	-1.22	0
10	1.44	-.87	-.29	-.29
11	.26	-.78	-.78	1.31
12	0	0	0	0
13	0	0	0	0
14	.87	-.87	-.87	.87
15	0	0	0	0
16	-.83	-.83	.50	1.17
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0

Experiment 1 Normalized Correct Confidence Ratings

	Test Pair Categories			
Subject	t2-t1	c2-c1	t1-c1	t2-c2
20	0	0	0	0
21	.50	-.83	1.17	-.83
22	1.31	-.78	-.78	.26
23	.78	-1.31	-.26	.78
24	.50	.50	-1.5	.50
25	0	0	0	0
26	.83	-.50	-1.17	.83
27	0	0	0	0
28	.50	.50	-1.5	.50
29	-.50	-.50	1.50	-.50
30	.87	.87	-.87	-.87
31	.50	.50	-1.50	.50
32	.50	-1.50	.50	.50
33	.82	-1.22	-.41	.82
34	0	0	0	0
35	.50	-1.50	.50	.50
36	.50	-1.50	.50	.50
37	.85	-1.09	-61	.85
38	.78	.78	-1.31	-.26
39	.78	.78	-1.31	-.26
40	.5	-1.5	.5	.5
41	.26	-.78	-.78	1.31
42	1.45	-.66	-.66	-.13
mean:	.356	-.391	-.221	.257
SD:	.587	.736	.769	.559

Experiment 1 Normalized Wrong Confidence Ratings

	Test Pair Categories			
Subject	t2-t1	c2-c1	t1-c1	t2-c2
1	-.50	-.50	-.50	1.50
2	1.17	.50	-.83	-.83
3	.74	-1.37	-.11	.74
4	-.20	-.20	.20	.20

Experiment 1 Normalized Wrong Confidence Ratings

	Test Pair Categories			
Subject	t2-t1	c2-c1	t1-c1	t2-c2
5	0	0	0	0
6	.20	.99	.20	-1.39
7	1.16	-1.16	.84	-.39
8	-.78	1.31	.26	-.78
9	.50	-1.50	.50	.50
10	-.50	-.50	1.50	-.50
11	-.50	-.50	1.50	-.50
12	-.50	-.50	-.50	1.50
13	-.50	-.50	1.50	-.50
14	-1.38	.74	.74	-.11
15	0	0	0	0
16	1.02	-.15	-1.32	.44
17	0	0	0	0
18	-1.31	-.26	.78	.78
19	.50	-1.50	.50	.50
20	0	0	0	0
21	-.15	.44	-1.32	1.02
22	-.50	-.50	-.50	1.5
23	-.50	.50	-.50	1.5
24	0	0	0	0
25	-.87	.87	-.87	.87
26	.78	-1.31	-.26	.78
27	0	0	0	0
28	.20	-1.39	.99	.20
29	1.50	-.50	-.50	-.50
30	.12	-2.41	.12	1.09
31	0	.71	-1.41	.71
32	1.18	.45	-.99	-.64
33	.71	-1.41	0	.71
34	0	0	0	0
35	.87	-.87	.87	-.87
36	0	0	0	0
37	.20	.99	-1.39	.20
38	-2.72	.78	-1.83	.78
39	1.09	-.85	-.85	.61

40	1.24	-1.02	-.56	.61
41	-.71	-.71	0	1.41
42	.26	-.78	-.78	1.31
mean:	.043	-.288	-.108	.296
SD:	.824	.829	.807	.741

Experiment 2 Normalized Total Confidence Ratings:

			Test Pair Categories			
Subject	t2-t1	c2-c1	t2-c1	c2-t1	t2-c2	t1-c1
1	-.913	-.913	-.913	.913	-.913	.913
2	-.598	-.598	1.793	-.598	-.598	-.598
3	2.043	-.407	-.407	-.407	-.407	-.407
4	0	0	0	0	0	
5	0	0	0	0	0	0
6	.848	-1.186	.848	-.169	.848	-1.186
7	-.639	-.639	1.297	-.639	1.297	-.639
8	-1.186	-.169	-.169	-.169	-.169	1.865
9	.913	-.913	.913	-.913	.913	-.913
10	.816	-.409	-1.634	.816	.816	-.409
11	-1.118	-1.118	0	1.118	1.118	0
12	0	0	0	0	0	0
13	1.187	.169	1.187	-.848	-.848	-.848
14	.647	.647	-1.29	.647	.647	-1.29
15	1.348	-1.101	.55	-.275	.55	-1.101
16	.409	-2.041	.409	.409	.409	.409
17	0	0	0	0	0	0
18	-.598	-.598	1.793	.598	-.598	-.598
19	0	-1.118	1.118	-1.118	1.118	0
20	.645	-1.291	.645	.645	.645	-1.291
21	.627	-1.630	.627	.627	.627	-.877
22	-1.088	-.363	1.088	-1.088	1.088	.363
23	.627	-1.630	.627	-.877	.627	.627
24	.645	-1.292	.645	-1.292	.645	.645
25	.407	.407	-2.043	.407	.407	.407

Experiment 2 Normalized Total Confidence Ratings:

Subject	t2-t1	c2-c1	t2-c1	c2-t1	t2-c2	t1-c1
			Test Pair Categories			
26	1.581	0	0	0	0	-1.581
27	.598	-1.793	.598	.598	.598	-.598
28	.169	.169	.169	-1.865	1.186	.169
29	.981	-1.215	.981	.249	.249	-1.215
30	.816	-.409	-.409	.816	.816	-1.634
31	0	0	0	0	0	0
32	-.487	-.487	.244	1.708	.244	-1.219
33	-.407	-.407	-.407	-.407	-.407	2.043
34	.886	-1.77	.222	-.442	.886	.222
35	-.142	-.142	-.997	.713	1.569	-.997
36	-1.118	0	0	0	0	0
37	1.107	-.886	.442	-.886	1.107	-.886
38	.713	-.142	-.142	.713	.713	-1.853
39	.477	-1.430	1.43	-.477	-.477	.477
40	0	0	0	0	0	0
mean:	.281	-.618	.230	-.037	.368	-.393
SD:	.758	.677	.848	.751	.625	.837

Experiment 2 Normalized Correct Confidence Ratings:

Subject	t2-t1	c2-c1	t2-c1	c2-t1	t2-c2	t1-c1
			Test Pair Categories			
1	-.125	-.125	-.125	1.379	.627	1.379
2	.639	.639	.639	.639	-1.297	-1.297
3	-.407	-.407	2.043	-.407	-.407	-.407
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6.	.817	-1.02	1.429	-.408	.205	-1.02
7	.281	-.545	1.107	-1.371	1.107	-.545
8	-.647	-.647	1.289	-.647	-.647	1.289
9	913	-.913	.913	-.913	.913	-.913
10	.412	-2.043	.412	.412	.412	.412
11	-.816	-.816	.409	.409	1.634	-.816
12	0	0	0	0	0	0
13	.408	.408	.408	-2.042	.408	.408
14	1.187	.169	.169	.169	.169	-1.865
15	1.292	-.645	1.292	-.645	-.645	-.645

Experiment 2 Normalized Correct Confidence Ratings:

			Test Pair Categories			
Subject	t2-t1	c2-c1	t2-c1	c2-t1	t2-c2	t1-c1
16	.408	.408	.408	-.204	.408	.408
17	0	0	0	0	0	0
18	-1.55	-.222	1.107	-.222	1.107	-.222
19	.728	.104	.728	-1.769	.728	-.520
20	.105	.105	-.598	.105	.105	-1.793
21	.408	-2.04	.408	.408	.408	.408
22	-.848	-.848	1.186	.169	1.186	-.848
23	1.55	-1.107	.222	-1.107	.222	.222
24	.645	-1.292	.645	-1.292	.645	.645
25	.409	-2.041	.409	.409	.409	.409
26	1.55	.222	.222	-1.107	.222	-1.107
27	.550	-1.927	.550	-.276	.550	.550
28	1.379	-.125	-.125	-1.629	-.125	.627
29	1.118	-1.118	1.118	0	0	-1.118
30	.329	-.989	-.329	1.648	.329	-.989
31	0	0	0	0	0	0
32	-.169	-.169	-.169	1.865	-.169	-1.186
33	-.409	-.409	2.041	-.409	-.409	-.409
34	.627	-1.629	.627	-.877	.627	.627
35	-.645	-.645	-.645	1.292	1.292	-.645
36	-1.118	-1.118	0	1.118	1.118	0
37	1.633	-.817	.408	-.817	.408	-.817
38	.409	.409	.409	.409	.409	-2.041
39	.571	.082	1.061	-1.877	.082	.082
40	0	0	0	0	0	0
mean:	.291	-.528	.492	-.236	.301	-.293
SD:	.745	.734	.632	.980	.579	.811

Experiment 2 Normalized Wrong Confidence Ratings:

			Test Pair Categories			
Subject	t2-t1	c2-c1	t2-c1	c2-t1	t2-c2	t1-c1
1	-.598	-.598	-.598	.598	-.598	1.793
2	1.641	-.808	.416	-.808	-.808	.416
3	1.266	.253	.253	.253	-1.770	-.253

Experiment 2 Normalized Wrong Confidence Ratings:

			Test Pair Categories			
Subject	t2-t1	c2-c1	t2-c1	c2-t1	t2-c2	t1-c1
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	.848	-1.186	.848	-.169	.848	-1.186
7	.105	.729	-1.768	.729	-.519	.729
8	-.852	.852	-.852	.284	-.852	1.420
9	-1.865	-.169	-.169	-.169	-.169	-1.186
10	1.088	.363	-1.088	-.363	1.088	-1.088
11	-2.042	.408	.408	.408	.408	.408
12	-1.278	.658	-1.278	.658	.658	.658
13	.645	-1.151	.645	.645	.645	-1.151
14	0	-.791	-1.581	.791	.791	.791
15	1.388	-1.061	-1.061	-.081	.898	-.081
16	-.325	-1.496	.845	.845	-.715	.845
17	0	0	0	0	0	0
18	-.848	-.848	1.187	1.187	-.848	.169
19	-.645	-.645	-.645	-.645	1.292	1.292
20	.550	-1.101	1.376	-.275	.550	-1.101
21	.889	.534	-1.245	.889	.178	-1.245
22	-1.413	.061	.798	-1.044	.798	.798
23	1.187	-.848	1.187	.169	-.848	-.848
24	1.107	-.222	-1.550	-.222	-.222	1.107
25	-1.107	.222	1.107	.222	.222	1.550
26	-.877	-.877	.627	.627	1.380	-.877
27	-.090	.999	.455	.999	-1.181	-1.181
28	.329	.329	-.989	-.989	1.688	-.329
29	-.995	-.995	.332	.332	.332	-.995
30	.647	.647	-1.296	.647	.647	-1.296
31	.550	-.276	.550	.550	.550	-.276
32	-.408	-.408	-.408	1.225	1.225	-1.225
33	.081	.081	.081	1.061	-1.878	.571
34	1.181	-.999	-.454	.091	1.181	-.999
35	1.254	-.358	-.358	-.895	1.254	-.895
36	-.989	.329	-.329	.329	-.989	1.648
37	.598	-.598	-.598	-.598	1.793	-.598
38	-1.118	0	-1.118	1.118	1.1181	0
39	-.325	-.714	-.714	.844	-.714	1.624

40	0	0	0	0	0	0
mean:	.827	-.242	-.175	.231	.141	-.025
SD:	.951	.652	.863	.622	.939	.975

REFERENCES

Anderson, J.R. (1990). Cognitive Psychology and Its Implications. 3rd Ed. New York: W.H. Freeman & Co.

Begg, I., Maxwell, D., Mitterer, J.O., & Harris, G. (1986). Estimates of frequency: attribute or attribution. Journal of Experimental Psychology: Learning, Memory & Cognition, 12 (4), 496-508.

Blaxton, T.A. (1989). Investigating dissociations among memory measures: support for a transfer-appropriate processing framework. Journal of Experimental Psychology: Learning, Memory & Cognition, 15 (4), 657-668.

Bransford, J.D., Franks, J.J., Morris, C.D., Stein, B.S. (1978). An analysis of memory theories from the perspective of problems of learning. In L.S. Cermak & F.I.M. Craik (Eds.), Levels of processing and human memory, Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, N.R. (1995). Estimation strategies and the judgment of event frequency. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,(6), 1539-1553.

Brown, N.R. (1997) Context memory and the selection of frequency estimation strategies. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23, (4), 898-914.

Bruning, J. L. & Kintz, B.L. (1968). Computational Handbook of Statistics. Illinois: Scott, Foresman and Company.

Carrol, J.B., Davies, P., & Richmond, B. (1971). The American Heritage Word Frequency Book. New York: American Heritage Publishing Co./Houghton Mifflin.

Craik, F.I.M. & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. Journal of Experimental Psychology: General, 11, 671-684.

Craik, F.I.M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 104, 268-294.

Fisk, A.D. & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. Journal of Experimental Psychology: Learning, Memory & Cognition, 10 ,. 181-197.

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

Graf, P. & Ryan, L. (1990). Transfer-appropriate processing for implicit and explicit memory. Journal of Experimental Psychology: Learning, Memory & Cognition, (6), 978-992.

Graf, P., Shimamura, A.P., & Squire, L.R. (1985). Priming across modalities & priming across category extending the domain of preserved function in amnesia. Journal of Experimental Psychology: Learning, Memory & Cognition, 11, (2), 386-396.

Graf, P., Squire, L.R. & Mandler, G. (1984). The information that amnesic patients do not forget. Journal of Experimental Psychology: Learning, Memory and Cognition, 10, 164-178.

Greene, R. (1986). Effects of intention and strategy on memory for frequency. Journal of Experimental Psychology: Learning, Memory & Cognition, 12,(4), 489 – 495.

Greene, R. (1989). Spacing effects in memory: evidence for a two-process account. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,(3), 371-377.

Greene, R. (1990). Memory for pair frequency. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,(1), 110-116.

Greene, R. and Thapar, A. (1994). Mirror effect in frequency discrimination. Journal of Experimental Psychology: Learning, Memory & Cognition, 20, (4), 946-952.

Hasher, L. & Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? Journal of Verbal Learning and Verbal Behavior, 16, 173 –184.

Hasher, L., & Zacks, R.T. (1984). Automatic processing of fundamental information: The case of frequency of occurence. American Psychologist, 39 ,(12),1372 – 1388.

Hintzman, D.L. (1976). Repetition and Memory. In G.H. Bower (Ed.), The Psychology of Learning and Motivation (Vol. 11, pp. 47-91). New York: Academic Press.

Hintzman, D.L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. Psychological Review, 95, (4); 528-551.

Howell, W. C. (1973). Storage of events and event frequencies: a comparison of two paradigms in memory. Journal of Experimental Psychology: Learning, Memory & Cognition, 98 (2), 260-263.

Hyde, T.S. & Jenkins, J.J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 12, 471-480.

Jacoby, L.L. (1983). Remembering the data: analyzing interactive processes in reading. Journal of Verbal Learning and Verbal Behavior, 22, 485-508.

Jones, C.M. and Heit, E. (1993). An evaluation of the total similarity principle: effects of similarity on frequency judgments. Journal of Experimental Psychology: Learning, Memory, & Cognition, 19, (4), 799-812.

Jonides, J. & Jones, C. M. (1992). Direct coding for frequency of occurrence. Journal of Experimental Psychology: Learning, Memory & Cognition, 18, (2), 368-378.

Jonides, J. & Naveh-Benjamin, M. (1987). Estimating frequency of occurence. Journal of Experimental Psychology: Learning, Memory & Cognition, 13, (2), 230-240.

Lachman, R., Lachman, J.L., & Butterfield, E.C. (1979). Cognitive Psychology and Information Processing: An Introduction. New Jersey: Lawrence Erlbaum Associates.

Macmillan, N.A. & Kaplan, H.L. (1985). Detection theory analysis of group data: estimating sensitivity from average hit & false alarm rates. Psychological Bulletin, 98, (1), 185-199.

Macmillan, N.A. & Creelman, C.D. (1990). Detection Theory: A User's Guide. Cambridge: Cambridge University Press.

Madigan, S.A. (1969). Intraserial repetition and coding processes in free recall. Journal of Verbal Learning and Verbal Behavior, 8, 828-835.

Merikle, P.M. & Reingold, E.M. (1991). Comparing direct (explicit) and indirect (implicit) measures to study unconscious memory. Journal of Experimental Psychology: Learning, Memory, & Cognition, 17,(2), 224-233.

Morris, C.D., Bransford, J.D., & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Behavior, 16, 519-533.

Naveh–Benjamin, M. & Jonides, J. (1986). On the automaticity of frequency encoding: effects of competing task load, encoding strategy, and intention. Journal of Experimental Psychology: Learning, Memory & Cognition, 12,(3), 378 – 386.

Neill, W.T., Beck, J.L., Bottalico, K.S., & Molloy, R.D. (1990). Effects of intentional versus incidental learning on explicit and implicit tests of memory. Journal of Experimental Psychology: Learning, Memory & Cognition, 16,(3), 457-463.

Perfetti, C.A., Lindsay, R., & Garson, B. (1971). Association & Uncertainty: Norms of Association to Ambiguous Words. University of Pittsburgh: Learning Research & Development Center.

Postman, L. (1964). Short term memory and incidental learning. In A. W. Melton (Ed.), Categories of human learning. New York: Academic Press.

Reber, A. S. (1995). The Penguin Dictionary of Psychology, New York: Penguin Books., Ltd.

Reber, A. S., Allen, R., Reber, P. J. (1999). Implicit and explicit learning. In R. J. Sternberg (Ed.), The Nature of Cognition. Boston, MIT Press.

Richardson-Klavehn, A. & Bjork, R.A. (1988). Measures of memory. Annual Review of Psychology, 36, 475-543.

Rosenthal, R. & Rosnow, R.L. (1984). Essentials of Behavioral Research: Methods and Data Analysis. (1st Ed.). New York: McGraw-Hill.

Schacter, D.L. (1992). Understanding implicit memory: a cognitive neuroscience approach. American Psychologist, 47,(4), 559-569.

Schacter, D.L. (1987). Implicit memory: history and current status. Journal of Experimental Psychology: Learning, Memory & Cognition, 13, 501-518.

Snodgrass, J.G. (1977). The Numbers Game: Statistics in Psychology. Baltimore: Waverly Press, Inc.

Squire, L. (1987). Memory and Brain. New York: Oxford University Press

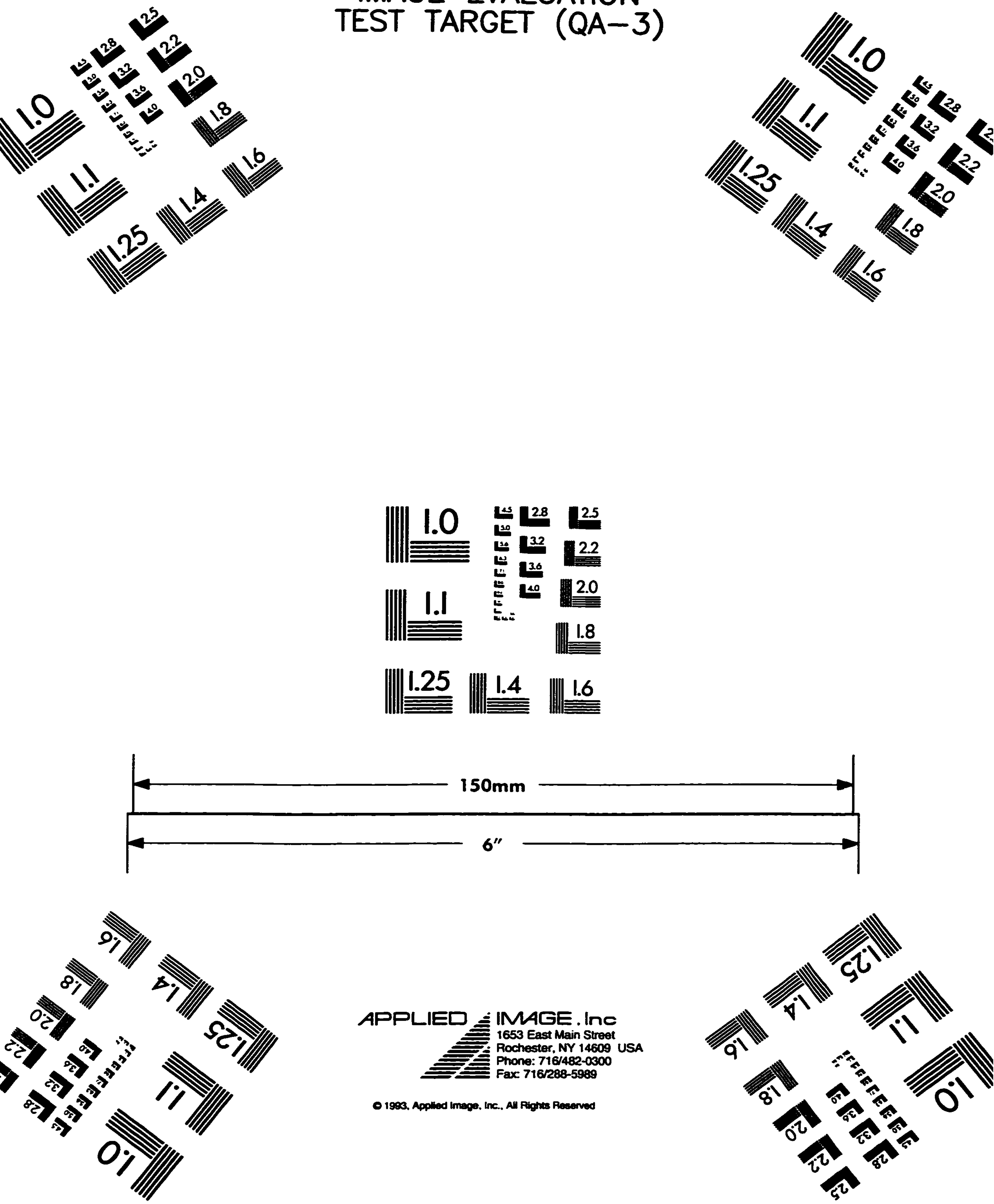
Squire, L. (1991). Talk given at the Cognitive Science Conference, Univ. of Chicago.

Voss, J.F., Vereb, C. & Bisanz, G. (1975). Stimulus frequency judgments and latency of stimulus frequency judgments as a function of constant and variable conditions. Journal of Experimental Psychology: Learning, Memory & Cognition, Human Learning and Memory, 104 (3), 337-350.

Zacks, R. T., Hasher, L., & Sanft, H. (1982). Automatic encoding of event frequency: further findings. Journal of Experimental Psychology: Learning, Memory & Cognition, 8,(2), 106 – 116.

Zola-Morgan, S. & Squire, L.R. (1990). The neuropsychology of memory: parallel findings in humans & nonhuman primates. Annals of the New York Academy of Sciences, 608, 434-456.

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989