1974

Frequency Judgments for the Wording and Meaning of Sentences

Christopher Gude
Loyola University Chicago

FREQUENCY JUDGMENTS FOR THE WORDING AND MEANING OF SENTENCES

by

Christopher Gude

A Thesis Submitted to the Faculty of the Graduate School

of Loyola University of Chicago in Partial Fulfillment

of the Requirements for the Degree of

Master of Arts

June

1974

ACKNOWLEDGMENTS

The author gratefully acknowledges the invaluable assistance and encouragement provided by Dr. Eugene B. Zechmeister at all stages of this project. Thanks are also extended to Dr. Frank Slaymaker and Dr. John Edwards for their helpful comments and suggestions during the preparation of this manuscript.

LIFE

Christopher Bernard Gude was born on June 26, 1949 in St. Louis, Missouri. He is the oldest of Mr. and Mrs. Henry B. Gude's nine children. He graduated from William Cullen McBride High, an Archdiocesan school conducted by the Brothers of Mary, in 1967. Christopher then attended the University of Missouri at St. Louis where he received a Bachelor of Arts degree in Psychology in June of 1971.

In September of 1971, Christopher entered the graduate program in Psychology at Loyola under the Experimental Division. During his first year, he served as a teaching assistant for Introductory Psychology courses under Drs. J. Johnson, T. Petzel, and R. Nicolay. In his second and third years, Christopher served as a research assistant under Dr. Eugene B. Zechmeister.

During his graduate study, Christopher has followed the course in Cognitive Psychology offered in the Experimental Division. His major areas of concentration have been memory and language processes.

Upon completion of his graduate work, Christopher will seek an academic position in which it will be possible to teach and conduct research projects.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure | Page

CONTENTS FOR APPENDICES

INTRODUCTION

The frequency with which a particular event has been experienced has long been considered as a variable which influences human perception and memory. Despite this pervasive treatment of frequency as an independent variable, the manner in which frequency is perceived and remembered has remained remarkably ambiguous. In this light, much research has treated frequency as the primary variable of investigation and has sought to explicate the representation of frequency in memory. Generally, frequency is thought of as some representation of the repetitiveness with which a particular event is noted and stored by an attending individual. Howell (1973), however, has pointed out that the nature of frequency representation for different types of stimulus events (i.e., verbal vs. non-verbal stimuli) has not at all been made clear.

If only verbal classes of stimuli are considered, the concept of frequency still has many different connotations. One traditional view of frequency has referred to the usage of commonality of a word in the natural language. The Thorndike and Lorge (1944) and Kucera and Francis (1967) word counts have provided some means of assessing frequency in these terms. These investigations literally counted the number of times that English words occurred in large samples of printed material (i.e., books, magazines, and newspapers). It is presumed that word counts of this type provide information regarding an individual's familiarity with a particular word. In studies of verbal processes, this dimension of commonality or familiarity has typically been titled

"background frequency." Numerous studies are available which demonstrate that "background frequency" influences behavioral tasks ranging from tachistoscopic word identification (Solomon and Howes, 1951) to the long-term recognition of words in a forced-choice situation (Gorman, 1961).

A second use of the term frequency is made in reference to the occurrence of a verbal event in a particular experimental context. This operational definition has been labelled "situational frequency." Underwood (1969) has maintained that frequency in these terms is closely correlated with the degree to which an individual learns that a verbal event occurred in a given experimental setting; the more times that an item occurs, the greater the probability that it will be recalled. Additionally, Underwood (1972) has pointed to "situational frequency" as a pre-eminent variable mediating the recognition and discrimination of words in verbal learning tasks. It is the nature and operation of "situational frequency" that will be given careful consideration. (Hereafter, the use of the term frequency will be limited to the "situational" definition.)

The Nature Of Frequency Representation

The manner in which frequency information is represented in memory is a topic which has generated considerable recent interest. (cf. Howell, 1973). The oldest and probably simplest position regarding the representation of frequency is the "trace-strength" hypothesis. This point-of-view holds that frequency information is mediated by some internal process that grows progressively stronger with repeated occurrences of a verbal event. As Howell (1973) points out, any behavioral manifestation of frequency such as discrimination or

estimation is simply "a matter of reading out current strength values (p.45)." The "trace-strength" hypothesis, while parsimonious, is plagued by several difficulties. For one, the operation of frequency is inextricably tied to the formation of an overall memory trace for an event which can undoubtedly be influenced by factors other than frequency. The meaningfulness and concreteness of verbal items, for example, have also been linked to the degree of learning and hence to the strength of a memory trace. Also, because the operation of such a mechanism is presumed to be cumulative in nature, only a general strength value would be available at a given time and little or no information would be present regarding how and when the repeated events had occurred. It is obvious, however, that individuals do possess information which allows fine discriminations to be made along temporal and modality dimensions. (Hintzman, 1969; Hintzman & Block, 1970, 1971; Hintzman and Waters, 1969, 1970)

The research of Hintzman and his associates has prompted the formulation of an alternative view regarding frequency representation which has been called the "multiple-trace" hypothesis. According to Hintzman and Block (1971), the "multiple-trace" hypothesis states that the effect of frequency is to increase the number of traces which exist in memory. With each repeated occurrence, an additional trace is laid down alongside those which are already present. The various traces are differentiated by some sort of "tag" which Hintzman and Block have assumed to be primarily of a temporal nature.

Support for the "multiple-trace" hypothesis has come from three experiments which demonstrate that individuals are sensitive to both temporal and frequency information and that frequency is dependent upon the temporal discriminability of repeated occurrences (Hintzman & Block, 1971). In the first experiment, evidence was presented indicating that individuals are fairly accurate in determining the

spatial-temporal position of a study list item. The subjects in this study were asked to judge in which 10th of a 50 item list a particular word had occurred.

A second experiment was more critical in formulating the "multiple-trace" hypothesis and sought to determine the extent to which two position judgments for a twice presented word could be made. The subjects were presented with a study list containing both once and twice occurring items. The twice presented words were of primary importance and each of these words occurred in two different sections of the list which were labelled A, B, C, and D. These "zones" corresponded to ordinal positions 3-8, 9-14, 15-20, and 43-48 in a 50 word study list and were chosen on the basis of the first experiment which indicated that these were the areas of maximum discriminability. Different combinations regarding the positions of the two occurrences resulted in words which were labelled as AC, AD, BC, and BD items.

Hintzman and Block expected that the judgment of the first occurrence of a word would be independent of the position of the second if independent memory traces with temporal "tags" are present in memory. Similar independence was, of course, anticipated for the second occurrence. The results of the experiment essentially confirmed the predictions. The first-position judgments for the AC and AD items did not differ significantly from one another and were consistently dissimilar to judgments made on the BC and BD items. Conversely, the second position judgments for AC and BC items were alike but differed from judgments made on the AD and BD items.

The third study demonstrated that individuals are capable of distinguishing recent from remote "situational" frequency. The subjects were presented with two lists of words separated by a five

minute interval. The two lists were constructed such that 36 words occurred in both of the lists. Four words were assigned to nine experimental conditions representing all combinations of three List 1 frequencies and three List 2 frequencies (0, 2, and 5 repetitions). As the "multiple-trace" hypothesis would predict, the subjects in this experiment were quite good in judging the number of times which a given item had occurred in each of the two lists.

The "trace-strength" and "multiple-trace" hypotheses do not exhaust the possibilities and other positions regarding frequency representation can be formulated by combining certain features of the "trace-strength" and "multiple-trace" views. Underwood's position (1969, 1972) regarding the nature of frequency can be cited as an example of a "multiple-process" hypothesis (cf. Howell, 1973). His position evolves from the conceptualization of memory for a verbal event as a collection of attributes or properties which can be encoded during the presentation of a given word. Both Underwood (1969) and Wickens (1970) have advocated that memory be viewed in these terms. In addition, Underwood has distinguished two major classes of attributes; one class serving to discriminate one memory from another, the second acting as retrieval mechanisms for accessing a target memory. The former class of attributes is represented by frequency, temporal, and modality information; all of which presumably provide dimensions along which differences among verbal events can be noted. These attributes, however, are not necessarily useful in the retrieval of the items. The latter class of attributes are associative in nature and do aid in the retrieval of verbal items from memory. Specific cues are established which allow otherwise unavailable items to be reproduced during a test of memory. For example, if a group of words all belong to a particular taxonomic category and

are encoded as such, memory for these items will be facilitated by invoking the category name as a retrieval cue at the time of recall.

The conceptualization of memory as a collection of attributes has given frequency a double role in Underwood's views of memory. On the one hand, Underwood has stated in his 1969 paper that "frequency is a major manipulable variable underlying learning; the greater the frequency the better the learning" and that further "frequency is normally associated with the strength of learning." (p. 563) This statement intuitively appears to mean that frequency (at least one) is necessary for any attribute to be encoded and that as frequency increases, more attributes are likely to be encoded thereby enhancing the memory for a given event. This is, of course, consistent with a "trace-strength" hypothesis.

Alternatively, however, Underwood (1969, 1972) has maintained that frequency, itself, is encoded as an attribute of memory in much the same way as acoustic orthographic, and associative properties of words. This assertion is based on research which has attempted to break the almost inevitable correlation between "strength" and frequency. For example, Underwood (1969b) has demonstrated that while words presented only once in a long study list are recalled much better if they occur at either the beginning or the end of the list than if they occur in the middle, judgments regarding the number of occurrences for these same items do not differ as a result of list position.

The studies of Hintzman and Block (1971) also make it necessary to consider frequency apart from the overall strength of a memory. Their results have shown that repeated occurrences of a verbal item can be identified fairly well along a temporal dimension and thus frequency and temporal information are, of necessity, dependent upon

one another to a large extent. Frequency, under the "multiple-trace" view, is derived from the retrieval of temporally-marked traces at the time of a memory test. Underwood, however, has preferred to consider frequency as an attribute distinct from temporal information and therefore his position is incompatible with the "multiple-trace" hypothesis. The reason is stated quite succinctly by Howell (1973). "If an event memory is defined as the sum total of stored attributes, and one attribute is frequency, then there is no vehicle to convey multiple traces" because "frequency must be represented in order to define an event, but multiple event representations define frequency." (p. 45)

Apart from the above considerations, there are a number of questions regarding the representation of frequency that are yet to be answered. It is probably safe to state that no one hypothesis can reconcile all of them and that some of these questions have not been reconciled by any of the existing hypotheses. For example, is the frequency encoded from a particular word with a number of different meanings (a homograph) specific to a given meaning or is it merely encoded via the generic representation of that word? Additionally, one wonders what constitutes a unit of frequency in different types of verbal materials if it is assumed that frequency serves as a general attribute of memory. Does frequency always accrue to an individual word even when it is presented in connected discourse or can frequency accrue to larger verbal units such as sentences?

The functional unit of frequency information is, of course, a relevant consideration for either the "multiple-trace" hypothesis or a position which considers frequency to be an encoded attribute of memory (cf. Underwood, 1969) If frequency is dependent upon the retrieval of

independent traces at the time of test, as the "multiple-trace" hypothesis implies, the question can be raised as to whether the frequency of sentence units as well as the frequency of the constituent words can be derived following experience with stimulus material comprised of meaningful sentences. Demonstration of such an ability would indicate that the derivation of frequency involves something more than the temporal-tagging of individual words in order to make them discriminable from one another. Other information regarding the properties of the sentences, themselves, (i.e. syntactic and semantic) would seem to be necessarily implicated. For example, an individual must be sensitive to changes in wording when various paraphrases of a sentence are presented for study, if he is to accurately assess the number of times that a particularly worded sentence has occurred. On the other hand, attention need not be given to wording if one is to judge the number of times that certain words have occurred regardless of the sentence in which they were included. Finally, an additional situation is confronted if one is asked to judge the frequency with which a particular idea or meaning has been expressed without regard to the manner in which it is stated. Again exact wording is not a relevant consideration but neither is the occurrence of individual words. In this case, the semantic content of the sentence whould be considered as the relevant unit of frequency. A demonstrated ability for judging the frequency of both sentences and words included in sentences would imply that if an explanation of frequency as an encoded attribute is to be accepted, frequency would have to be noted and maintained in an independent fashion for both types of verbal events (i.e. sentences and words).

It is suggested here that frequency is derived at the time of test as the "multiple-trace" hypothesis would contend, but that

frequency estimates are made on the basis of whatever information is relevant and necessary for the particular type of frequency judgment that is being requested. A temporal "tag," as Hintzman and Block have indicated, is, of course, one but not the only dimension that can function in judging the frequency of different types of verbal events. With these considerations in mind, attention will be turned to some of the experimental paradigms in which behavioral manifestations of frequency can be witnessed.

Verbal Discrimination Learning and Frequency Theory

Underwood (1969) has maintained that frequency is the predominant attribute of memory involved in discrimination and recognition processes. Basic support for this position has come from experimental work conducted within the verbal discrimination (VD) learning paradigm.

Generally, the VD paradigm involves the presentation of two items for study with one being arbitrarily denoted by the experimenter as correct. On a later test trial, the items are again presented to a subject and he is asked to indicate which of the two words had been previously labelled as correct. During the course of a VD experiment, study and test trials are typically alternated until some set criterion of learning is reached.

In 1966, Ekstrand, Wallace, and Underwood proposed the "Frequency Theory" of verbal discrimination to account for performance in this paradigm. As the name implies, frequency is considered as the primary determinant of learning and several mechanisms and rules are laid out to explicate its operation. The theory describes four basic ways in which frequency is accumulated for the presented words. First, a pair of words is presented; a unit of frequency accrues to each simply through the act

of perceiving the items. Ekstrand et al. have referred to this frequency input as a "representational response" (RR). Additional units of frequency, however, may be added to the correct words by having the subject pronounce the right items. The act of saying the correct item aloud (pronunciation response, PR) allows the correct items to accumulate a greater number of frequency units than the incorrect items during the course of study and test trials. Along with these overt mechanisms of accumulating frequency, the theory has also provided for two additional mechanisms which allow for covert accrual. It is assumed that during the course of VD learning, a subject will implicitly rehearse the correct and possibly the incorrect items. These "rehearsal responses" (RCR) also provide additional units of frequency. Finally, Ekstrand et al. (1966) also point out that the presentation of any word has the possibility of eliciting other words which one associates with the particular experimental item. These "implicit associative responses"(IARs) may be other correct or incorrect words which are present in the VD list and these implicit responses are assumed to increment the frequency of the actual list items (cf. Underwood, 1965).

Ekstrand et al. assume that performance in the VD task is mediated by a frequency differential which exists between members of a given pair. The manner in which this frequency information is stored is not clearly defined but the theory implies that some type of counter-mechanism tallies up the frequency units contributed through the different processes and produces a sum total for each word. The counter-mechanism seems to suggest that a strength-notion of storage is involved since both overt and covert presentations of a word contribute to the same frequency total and the theory provides no means by which the different types of responses can be distinguished from one another.

The frequency theory was originally developed in order to interpret the results of a verbal discrimination study conducted by Underwood, Jesse, and Ekstrand (1964), but was subsequently extended to other VD situations. (cf. Ekstrand et al., 1966) In the Underwood et al. experiment, the subjects were given an initial VD list and then transferred to a second list which was varied among the different groups. In one condition (R), the correct member of each pair in the first list was retained as the correct item in the second list while a new word was presented as the other member of the pair. In a second condition (W), the incorrect item from a pair in the first list was placed in the second list with a new word becoming the correct member. A control condition was also included in which the two lists were unrelated.

The performance on the second list showed essentially 100% transfer for subjects in condition (R), suggesting that frequency units can be transferred between lists. Transfer performance in condition (W), however, was quite different. Initially, the subjects performed quite well and were superior to the control group. As trials progressed, however, the subjects improved very little and eventually fell below the performance of the control group. Ekstrand et al. (1966) in their later interpretation of this finding, maintained that the new words which were correct in the second list rapidly gained frequency and soon achieved the frequency level held by the old words. As a result, the frequency discrimination broke down.

These results of the Underwood et al. (1964) experiment have been explained in frequency theory terms by the postulation of two rules which a subject may use in order to make a correct discrimination. Rule 1 states that the word with the highest level of frequency is chosen

as the correct item. Ekstrand et al. believed that this was the rule employed by the subjects in Condition R of the Underwood et al. (1964) study. Rule 2 states that the word with the lower frequency is chosen as the correct item. Presumably, this strategy was used by subjects in Condition W at the outset of the transfer list, but this rule became inappropriate as the frequency of the new items increased and thus the subjects were forced to switch to the Rule 1 strategy later in the transfer list. The breakdown of the frequency differential has been interpreted as the cause of the poor performance in the group with "old" items retained as incorrect. These results have been taken as strong support for frequency theory although a total breakdown to a chance level was not observed.

Ekstrand et al. (1966) demonstrated that this same rationale could be used to explain the manner in which a single VD list is learned. Their experiment involved repeating the presentation of correct or incorrect items in different pairs during the presentation of the same study list. It was found that presenting a correct item in two pairs facilitated overall VD performance. It is assumed that additional frequency units accrue to the repeated items thereby magnifying the frequency between the repeated items and their incorrect counterparts. Alternatively, presenting the same incorrect item in different pairs deterred VD performance. It is assumed here that the additional units added to the repeated incorrect items eliminated the advantage gained by the correct items through normal study-test procedures. A final condition was included in which the repeated item served as the correct word in one pair but as the incorrect item in another. Under this arrangement, VD learning proceeded with great difficulty since in one pair containing a repeated item, the correct word has a higher level of

frequency, while in another pair the incorrect item had a higher frequency level. This necessitated that Rule 1 (select the word with the highest frequency) and Rule 2 (select the word with the lower frequency) both be used in the same list; a task that subjects apparently could not do efficiently.

The Ekstrand et al. study also documented the operation of IARs in VD learning. This was accomplished by including a particular word and a high associate of that word in different pairs within the same study list. By manipulating the position of the given item and its high associate in conditions similar to those described above for repeated items, it was found that IARs do operate in the accrual of frequency units.

Additional supporting evidence for frequency theory has been provided by research that has familiarized subjects with words later included in a VD list. Underwood and Freund (1968), for example, built up the frequency of either correct or incorrect VD items through prior free-recall learning trials. Their results indicated that frequency transferred from free-recall trials to VD learning in basically the same manner that frequency transferred from one VD list to another (cf. Underwood et al., 1964)

Other research has indicated that perceived differences in situational frequency which mediate VD performance are dependent upon the initial or base level of frequency. Underwood and Freund (1970) found that verbal discrimination becomes quite inaccurate when the familiarization frequency of words later included in VD pairs is built up to a high level but the initial differences between the correct and incorrect items in the subsequent VD learning are slight. For example, an initial difference in frequency between correct and incorrect items was

much more discriminable if the familiarization inputs were respectively 3 vs. 1 as opposed to 7 vs. 5. Apparently, frequency operates under something akin to the psychophysical law of Weber which states that as the intensity of two stimuli is increased, the difference necessary to discriminate the two must also be of a greater magnitude.

A study conducted by King and Levin (1971) shows results which are somewhat anomolous regarding frequency theory predictions. In this experiment, the subjects were transferred from one VD list to another in much the same manner as the Underwood et al. (1964) investigation. One group of subjects received a second list which retained the wrong items from List 1 as the incorrect words (Condition W-W), while a second group was confronted with a second list in which the correct items from List 1 became incorrect (Condition R-W). In addition, the subjects were given varying numbers of trials (2, 4, and 8) during List 1 learning.

As would be expected from frequency theory, the initial trials on the second list showed superior performance for the two experimental groups when they were compared to a control group which learned unrelated lists. Also, the effect of varying List 1 trials was more pronounced in the R-W condition than in the W-W condition because of the greater frequency build-up for List 1 right items. However, the deterioration in performance over trials did not materialize as would be anticipated from frequency theory predictions. This was especially true in the case of eight List 1 learning trials.

The failure of King and Levin (1971) as well as the failure of Underwood et al. (1964) to obtain the predicted transfer effects for wrong items in a prior list can perhaps be explained by some suggestions made by Hintzman and Block (1971). As indicated earlier, they were

successful in demonstrating that subjects are capable of discriminating recent from remote frequencies. They have argued that deterioration to a chance level would not be expected because in later Test 2 trials a subject would be capable of ignoring frequency gained during List 1 and discriminating with only recent List 2 frequency.

A recent study by Pasko and Zechmeister (1973) has lent some experimental validity to this explanation. These investigators employed a VD transfer task in which the designation of correct and incorrect items learned in List 1 was reversed for the second list. It was found that a temporal separation between the two lists reduced the extent to which transfer performance deteriorated only when a relatively low degree of List 1 learning (4 trials) was given. Under conditions of high List 1 learning (8 trials), a temporal separation did not aid the subjects in distinguishing recent from remote frequencies and slight deterioration effects were found regardless of whether a temporal separation was provided.

Pasko and Zechmeister (1973) have argued that the differential results produced by varying the number of List 1 trials may be due to a confounding of the degree of List 1 learning and the temporal separation between the initial trials of the original and transfer lists. The frequency accumulated in early trials of original learning,is of necessity, more temporally removed when 8 as opposed to 4 List 1 trials are given simply because more presentation time is required. Perhaps this prolonged separation between the start of List 1 and List 2 learning in the case of 8 original learning trials allows the discrimination of recent and remote frequencies to be made without the addition of a temporal separation between the two lists.

Pasko and Zechmeister (1973) have also attempted to account for the

failure of King and Levin (1971) to obtain a deterioration in transfer performance following 8 List 1 learning trials. It was reasoned that increasing List 1 trials likewise increases an individual's information regarding List 1 context (i.e. the particular words included in the study list). If contextual information can aid in the discrimination of recent and remote frequencies, the failure of King and Levin (1971) to obtain a deterioration in transfer performance may have been due to the fact that the transferred items (either correct or incorrect) were paired with new items, thus substantially changing the context of the two lists. Remote from recent frequencies could then have been very efficiently discriminated on the basis of the contextual distinction. In the Pasko and Zechmeister (1973) study, however, the context of the two lists was not substantially changed because the same items were merely reversed in their designation of correct or incorrect. With no relevant contextual information for distinguishing List 1 and List 2 frequencies, the subjects were forced to rely on a temporal discrimination as Hintzman and Block have suggested and in accord with earlier studies (cf. Underwood et al.), a certain amount of deterioration in transfer performance was displayed.

Other research which has investigated the role of IARs in VD learning has also not wholly supported the expectations of frequency theory. A number of experiments have included transfer tasks that were designed in such a way that List 1 items were expected to elicit IARs which would be words subsequently included in the second list. (Kausler and Dean, 1967; Raskin, Boise, Rubel, and Clark, 1968; Cole and Kanak, 1972) In all of these studies, the predicted transfer effects were found only when the subjects were instructed in the relationships between the two lists. Although Ekstrand et al. (1966) have found that

IARs apparently do increment frequency in the learning of a single VD list, the incrementation of frequency through the production of IARs does not appear to readily transfer across lists unless individuals are given a specific set to do so.

Recognition Memory and Frequency Information

In many respects, a recognition memory task can be likened to verbal discrimination learning. In a typical forced-choice recognition task, a list of study words is presented to a subject who is later asked to identify these "old" words when they are intermingled among distractor or "new" items. According to the frequency theory, the study list presentation allows at least one unit of situational frequency to accrue to the study words. Therefore, the theory assumes, a recognition decision can be made on the basis of a 1 vs. 0 frequency differential at time of test. This situation is, of course, not unlike the strategies which are purportedly used by subjects in a VD task.

If both recognition and discrimination are governed by frequency information, subjects who are required to make either frequency judgments or recognition decisions should show similar performance when the input conditions are identical. Underwood (1972) has shown that such does appear to be the case. In this experiment, two groups of subjects viewed identical study lists containing once and twice presented words, and were later tested with a multiple-choice recognition task. The groups, however, differed in terms of the experimental instructions. One group was told before presentation of the study list that they would later have to choose the most frequently presented word when the words appeared in a multiple-choice set. The second group was instructed that they were to learn the study list words in order to correctly recognize them in the multiple-choice test. The results did, in fact,

indicate that errors resulting from frequency judgments and recognition decisions for both once and twice presented words were of the same order of magnitude.

Frequency theory predictions regarding recognition memory have also been investigated by building up the frequency of distractor items during the testing phase of an experiment. (Underwood & Freund, 1970a) In a series of experiments, new (or wrong) items were repeatedly paired with several different right items during the recognition test. It was expected that the recognition of the old items would become progressively more difficult as the number of previous experiences with the distractor items during the recognition test became larger. In general, these studies confirmed the view that as the frequency of the wrong item increases, recognition decisions should become progressively more difficult.

The role of IARs in recognition memory has also been of considerable importance in extending frequency theory to the recognition paradigm. The theory stated by Ekstrand et al. (1966) maintains that the situational frequency of a word can be incremented through the word's occurrence as an IAR. This assumption should lead to differential results in recognition memory depending upon the associative relationships among old and new items.

Underwood (1965) demonstrated that in a continuous recognition paradigm, the probability of falsely identifying a word as "old" was significantly higher if it was preceded in the list by a word which was likely to elicit it as an IAR. In a similar manner, Underwood and Freund (1968b) have found that recognition memory suffers when the distractor items are especially designed as possible IARs of the study list words. Both of these experiments have been interpreted as strong support for

the pre-dominance of the frequency attribute in recognition memory.

The production of IARs has been somewhat indirectly demonstrated as a variable affecting forced-choice recognition in a study conducted by Underwood and Freund (1970b). Previous research (Gorman, 1961) had indicated that background frequency as determined by the Thorndike and Lorge count influences recognition memory. In particular, the recognition of words is poorer when high-frequency words are used as study and distractor items than when all low-frequency words are used. It was believed that these findings could at least be partially explained by frequency theory predictions regarding the role of IARs. Underwood and Fruend maintained that during the presentation of a study list, IARs will naturally be produced and that a greater number will be elicited when the study words have a high background frequency. These IARs can produce two different effects. First, they can increment the situational frequency of another study word which would facilitate recognition performance, but alternatively, they could increment the frequency of a distractor item which would produce a deleterious effect.

Word association norms show that low-frequency words typically elicit associations that are higher frequency words. Therefore, the role of IARs should be negligible under conditions where all low-frequency words are used as study and distractor items. Recognition decisions could then proceed, as frequency theory would predict, on the basis of the situational frequency accrued to the study items during the study list presentations.

The case in which all high-frequency words are used is considerably different. At first glance, it would seem that the opposing effects of IARs would cancel one another and recognition could be carried out in the predicted manner. Underwood and Freund (1970),

however, made a crucial assumption in order to explain the poorer performance with high-frequency words. In their own words, "adding an additional unit of frequency to an old word produces a relatively small increase in discriminability...therefore, the negative effect should be greater than the positive effect when high-frequency words are presented for study and new words are also high frequency words." (p. 345)

In order to provide a theoretical accounting of the role of IARs in the recognition of words of varying background frequency, an experiment was conducted in which high and low frequency words were used in all possible combinations of study and distractor words. It was predicted that the best recognition performance would be found in the condition where high-frequency words were used as study items with low-frequency words serving as distractors (Condition HF-LF). Under this arrangement, only a facilitating effect due to IARs would be present. Likewise, it was expected that condition HF-HF would show the poorest performance while the other two conditions (LF-HF and LF-LF) would be intermediate.

The results of the experiment clearly bore out the predicted effects. Underwood and Freund were also careful to point out that the superior performance in condition HF-LF could not have resulted from the subjects use of background frequency information (i.e. picking the high-frequency word in a multiple-choice pair). If such had been the case, conditions HF-LF and LF-HF should not have differed significantly, for in condition LF-HF, the subjects simply could have reversed this strategy.

While considerable evidence seems to weigh favorably towards the view that frequency is the predominant attribute in recognition, it should be noted that a frequency discrimination is, of course, not the

only manner in which a recognition decision can be made. Two recent investigations (Hall and Pierce, 1972; Zechmeister and Gude, 1973), in particular, have shown that the production of IARs may operate in a fashion somewhat different from what frequency theory would predict. In these experiments, the subjects were asked to study a long list of high-frequency words for a subsequent recognition test. In some conditions, the subjects were encouraged to produce IARs while in others, the subjects were simply told to repeat the word over and over to themselves.

Considering the nature of the study lists, it would seem that the IAR-producing group would be handicapped in recognition performance for the distractor items were also high-frequency words. (cf. Underwood and Freund, 1970) Both studies, however, found that the IAR group was superior to repetition and control subjects. It was suggested by Zechmeister and Gude (1973) that the act of making an association, itself, may serve as a discriminative cue with which a recognition decision could be made. By this, it is meant that the subject actually remembers that he produced an IAR to a particular study word and this retention bolsters a discrimination.

Frequency Judgments and Temporal Information

Research which has been identified with the multiple-trace hypothesis has linked the frequency attribute quite closely to temporal information. To restate the hypothesis, frequency serves to increase the number of independent traces for a particular event which exist in memory. These traces are temporally marked and by retrieving the various traces, an individual is capable of giving some accounting of the number of times that an event has occurred.

Hintzman and Waters (1969) provided evidence that the occurrence

of a word in a particular list of items is tagged with temporal information. In this study, the subjects were presented with two successive lists which were either separated in time by a 15 minute interval or were presented with no appreciable interval between the two. Upon completion of the study list presentations, the subjects, at varying intervals, were again exposed to study items and asked to place them in one of the two lists. Basically, it was discovered that as the interval between study and test increased, this task became progressively more difficult in both conditions. However, when the discrimination judgments took place during the same experimental session, Hintzman and Waters found that the subjects were more adept at determining the list to which a particular item belonged in the situation where a 15 minute interval had been placed between the two lists.

In another investigation, Zimmerman and Underwood (1968) rather strikingly demonstrated the amount of temporal information which a subject possesses following the presentation of study items. During the course of this experiment, the subjects were consecutively presented with 12 short lists containing 8 to 12 items. Following each list, a recall period ensued in which the subjects were asked to reproduce only the items in the preceding set. Two additional tasks were administered after all of the lists had been presented. First, the subjects were given 12 cards, each with one of the study lists printed on it, and were told to arrange the cards in the sequence which the 12 study lists had followed. In a second task, the subjects were given two items from each list and asked to judge the ordinal relationship of these items in their respective lists.

The critical manipulation in this experiment was embodied in instructions which were given to the subjects prior to the presentation

of any of the study lists. These instructions differed in the extent to which they informed the subjects to attend to temporal or ordinal information. One group was given normal free-recall instructions with no mention being made that temporal information would later be tested. Two other groups were respectively told that word position or word position and list position would be tested.

Zimmerman and Underwood (1968) found that there were essentially no differences among the three instruction groups regarding performance in either the word or list position tasks although accuracy in all groups was substantial on both measures. It seems clear from this study that spatial-temporal information is naturally acquired without specific instructions to do so and that there is no obvious expense to an individual's ability to recall.

As has already been mentioned, Hintzman and Block (1971) carefully laid out the multiple-trace hypothesis and in so doing made a crucial connection between a frequency attribute and temporal information such as that described in the Hintzman and Waters (1969) and Zimmerman and Underwood (1968) studies. Frequency under the multiple-trace hypothesis is not perceived or encoded directly but is rather inferred from the number of temporally discriminable traces which can be retrieved for a particular repeated occurrence.

If frequency judgments are, in fact, determined by temporally distinguished traces, increasing the temporal discriminability of repeated occurrences should facilitate a subject's ability to make accurate judgments of frequency. One variable which should influence temporal discrimination is the spacing of item repetitions. Melton (1970) has pointed out that spacing the repetitions of items in a free-recall task enhances the overall number of recalled words as compared to a situation

where the repetitions occur in a successive manner. Further, Melton (1967) and Madigan (1969) have found that as the number of items intervening between repetitions increases so too does the performance in free-recall increase.

If the spacing variable influences the retrieval and recall of words in free-recall, it seems reasonable to expect that spacing would similarily facilitate the retrieval of the time-tagged traces which Hintzman and Block (1971) imply as the basis of frequency information. Two studies indicate that this does appear to be the case. Hintzman (1969) found that as the spacing of repetitions was increased, the subjects' estimated frequency more closely approximated the actual frequency with which the words in a study list had occurred. A later study by Hintzman and Block (1970) essentially confirmed these results. Apparent or perceived frequency does increase with the spacing of repetitions thus bolstering the position that frequency is dependent upon temporal information.

Summary and Some Remaining Questions

The frequency-counter notion of Underwood receives some support from verbal discrimination and recognition performance. However, certain failures of frequency theory, particularly, in regards to the transfer of frequency units across lists force its proponents to rely on discrepencies between apparent and actual frequency as explanations for the anomolous data. The multiple-trace hypothesis, on the other hand, is amenable to the findings of most research involved with the nature of frequency information. But according to Hintzman and Block (1971), frequency does not appear to be frequency at all but rather an extension of temporal discriminations.

In any case, a number of questions raised at the outset of this

paper have not been adequately assessed. Namely, what exactly constitutes a unit of frequency and if frequency is derived at the time of test, what type of information is used to infer the frequency with which different types of verbal events have occurred? Although these questions are closely related, an attempt will be made to separately give each a more detailed consideration.

What Is a Frequency Unit?

In nearly all of the studies which have been considered, the stimulus material has been individual words. In these experiments, therefore, an isolated word is assumed to be the unit of investigation. It should be remembered, however, that individuals do not typically encounter isolated words much less attempt to remember them. Words serve as constituent parts of larger language units and therefore are undoubtedly influenced by the context in which they occur.

Regarding the encoding and representation of frequency, the question has already been raised as to whether a word is generically encoded (i.e. in terms of its orthographic properties) or is considered in terms of its specific semantic properties. In studies where only isolated words are presented for study, this consideration becomes relatively unimportant because there is no reason to assume that different semantic encodings will occur each time that an item is repeated. However, when a word can be biased to a number of different meanings depending upon its context, the question is one of utmost importance if the manner in which words are encoded is to be understood.

Several investigations have recently provided some understanding of the effects of context on memory for words. Bobrow (1970) investigated the recall of words contained in sentences. It was hypothesized that if word meanings are remembered then repeating nouns in paraphrased

sentences which preserved the original meaning of the nouns would result in higher recall than repetitions of the nouns in sentences which changed the meanings of the words. To test this hypothesis, Bobrow used homographs as the critical nouns and included same meaning (SM), different meaning (DM), and exact repetition (E) conditions. At the time of test, the subject-noun was presented to a subject and recall of the object-noun was requested. The proportion of object-nouns recalled in the SM and E conditions was not significantly different but both were evidently superior to the condition in which different meaning repetitions were presented.

Thios (1972) has expanded on these initial findings of Bobrow. In addition to varying the context in which critical nouns were repeated (SM,DM, and E conditions), the repetitions were also varied according to the number of intervening sentences. This manipulation of the spacing of repetitions warrants some explanation. Madigan (1969) has suggested that the effect of spacing is to enhance the retrievability of repeated items by providing additional retrieval cues associated with the different lists contexts in which the repetitions occur. By presenting items in widely disparate portions of a study list, Madigan maintained that variable encodings of the word would result because the items would be experienced in the presence of a different set of list members and would likely enter into rehearsal strings made up of different study items. In addition, Madigan (1969) found that the spacing effect was eliminated if another cue-word presented along with the to-be-remembered item was changed at each repetition. Apparently, providing the subjects with these additional retrieval cues at recall attenuated the facilitative effect of spacing. Even with these single-word cues, the context of encoding was already suitably different

according to Madigan's interpretation. Thios (1972) anticipated that semantic context of a sentence would serve similarily if not more efficiently than a single word as a retrieval cue for the nouns contained in sentences.

In this experiment, the spacing effect was eliminated when the retrieval cues were changed (i.e. DM repetitions). However, the recall of nouns from the DM sentences showed the poorest overall performance. Although the DM condition showed a slight advantage over the E condition at very short lags, the E condition became progressively better as the number of intervening sentences increased while the DM condition showed no significant changes. The SM condition also displayed a spacing effect and was superior to the other two conditions across all spacing lags. Apparently, maintaining a certain amount of similarity in context is crucial for the recall of words when they are presented in the body of a sentence. If such were not the case, the DM condition should have shown performance more equivalent to the SM condition in, at least, the short lag situations.

Other experiments in which the recall of homographs has been investigated under varying conditions of spacing have produced somewhat different results. In these studies, however, the meaning of the homographs has been biased by simply presenting a single adjective relevant to one of the various meanings. Gartman and Johnson (1972), for example, found that the spacings effect was attenuated by changing cues but that overall recall was higher when the context cues were varied than when they were the same. (i.e. the homographs were biased to different meanings) This finding is, of course, contrary to the results reported by Thios (1972).

It is possible, however, that including words in sentences as

Thios did, causes subjects to pay more attention to the meaning of the sentence as a whole rather than keying in on the occurrences of particular words as might be the case in the Gartman and Johnson experiment. Attending to the meaning of a whole sentence may effectively mask-out the second occurrence of critical words especially when they are embodied in a sentence expressing a totally new idea. This would not be the case where only two words are presented and the subjects may make special note that the same stimulus word was present but that the cue word was changed.

Each of these studies in which recall was investigated, nevertheless, support the conclusion that words are not encoded generically. A recent study in which frequency judgments and recognition of homographs was investigated is more pertinent to our primary concern with frequency information. Rowe (1973) presented homographic words at frequency levels ranging from 1 to 5 while orthogonally varying the nature of contextual information. The repetitions of the homographs were carried out in four different ways. In one condition, the homographs were repeated in phrases that were intended to bias the same semantic encoding (SM), while in a second condition, the repeated phrases were intended to evoke different semantic encodings (DM). Two control conditions were included in which the homographs were either repeated in an isolated fashion (RW) or were repeated in identical phrases (RP). All subjects were subsequently asked to judge the frequency with which a particular homograph had occurred.

Rowe clearly found evidence that frequency input is influenced by semantic context. As would be expected, the various groups did not differ in their judgments when the presented frequency was only one and all of the conditions showed an increase in perceived frequency as the actual presented frequency became progressively larger. The increase in

perceived frequency, however, was not identical across the treatment conditions. The best estimates of frequency were made in condition RW where the homographs were repeated in an isolated fashion. In the conditions where the homographs were embedded in the context of a phrase, frequency judgments suffered as the context of the homographs became less similar during the successive repetitions of the phrases. The exact phrase repetitions were most conducive to the judgment of frequency while the SM repetitions showed a significant advantage over the DM repetitions. Rowe's findings undoubtedly allow the conclusion that frequency input is specific to the semantic encoding of a word.

While it appears evident that the perceived frequency of words is influenced by semantic context, all of these studies varying the semantic encoding of homographs have still implicitly assumed that an individual word is the basic unit of frequency. Semantic context does produce differential results in recall depending upon whether a homograph is presented with a single word cue or embedded in the context of a meaningful sentence. Likewise, frequency judgments are differentially affected depending upon whether the judged words are presented alone or in a phrase context. In recall, it has been suggested that the discrepent findings are the result of different word processing strategies in the two experimental situations. In particular, sentences may be viewed as intact units expressing a unitary idea rather than as a mere composite of individual words that are noted and stored independently. For frequency judgments, it is reasonable to suggest that under conditions where intact sentences are presented for study, it is the sentence, itself, and not the individual word which is the unit of frequency. Information about specific words may be merely incidental in such a task. There appears to be only one study in which frequency judgments on

sentences have been attempted.

Jacoby (1972) presented his subjects with a series of sentences and later required frequency judgments on either intact sentences or the subject-nouns found in the sentences. During study, the sentences were repeated under varying conditions; some of the sentences were repeated with the subject-noun intact but with a different synonomous adjective and verb in each repetition, others were repeated with intact subject-nouns but completely different modifiers (i.e. adjective and verb), and a final condition contained exact repetitions of the original sentences. In addition, the number of items intervening between the repetitions (0, 3, and 11) and the frequency of presentation were orthogonally combined with the types of repetition.

One general finding of Jacoby's is particularly relevant to the suggestion made above that a sentence as a whole may be the relevant unit of frequency. In the case of exact repetitions, the frequency judgments of sentences were more accurate than those of the subject-nouns. This result implies that the sentence is, indeed, the relevant unit for frequency estimation when meaningful sentences are presented for study. If such were not the case, there would be no reason to expect the frequency of the whole sentences and the subject-nouns to differ in the exact repetition condition for their actual frequencies of presentation were identical.

Looking specifically at the sentence judgments, Jacoby found that the subjects were quite efficient in detecting slight modifications in the wording of the sentences and thus avoiding confusions with sentences containing synonomous modifiers. The fact that frequency judgments were little affected by the similarily worded versions is indicative of this basic result. In the same vein, the sentences which

were followed by similar repetitions were not judged differently from the sentences with completely different repetitions. Apparently, the slightest change in wording made it possible to discriminate the original sentences from their modified repetitions.

These results are somewhat surprising in light of a number of investigations which have assessed memory for sentences in both recall and recognition paradigms. Sachs (1967) and Begg (1971) have respectively shown that primarily semantic rather than syntactic information is retained in long-term recognition and recall. These studies point out that it is the general meaning as opposed to the specific wording which is remembered. However, another study by Bregman and Strasberg (1971) has indicated that this generalization does not necessarily hold in all situations. These latter authors have demonstrated that subjects, when pressed to do so, are capable of remembering specific wording information. The subjects in Jacoby's experiment were, of course, pressed for this type of information and the simple constant structure of the experimental sentences may have made wording information relatively easy to remember. It will be of interest to discover whether or not frequency information regarding the occurrence of similar sentences expressing the same semantic content is also present in memory.

<u>How Is Frequency Derived?</u>

Jacoby (1972) has maintained that frequency of presentation is derived at the time of test rather than encoded as an attribute of memory during study. This position is, of course, consistent with the multiple-trace hypothesis. Hintzman and Block (1971) have considered a temporal "tag" discriminating the repeated occurrences of an item as the principal means by which frequency information is retrieved. Other researchers who have adopted the multiple-trace hypothesis have presented

a somewhat different view regarding the identification of the differentiating tags. Anderson and Bower (1972), for example, have postulated the existence of "list markers" which link an event's occurrence to a set of contextual elements. These list markers are presumed to bring into play information of a temporal sort, as well as information pertinent to other events which have immediately preceded and followed the event in question, and to subjective feelings of an individual (i.e. boredom) which might have been experienced at the time which a particular event occurred. Rowe (1973) has also supported the derived view of frequency and has suggested that his research with homographs marks the semantic encoding of an item as a possible dimension along which frequency information can be derived.

Jacoby (1972) has, in some respects, shown that when the stimulus materials are identical, frequency judgments of different types (in his experiment, sentence vs. noun judgments) will be made and influenced along those dimensions of the stimulus material which are relevant to the type of judgment being requested. Considering the judgments of sentence frequency, the spacing of repetitions (i.e. a temporal attribute) was a facilitating variable in estimating frequency of occurrence for only those sentences which had identical repetitions. (Jacoby, 1972) In this situation where repetitions are exact, a temporal tag such as that proposed by Hintzman and Block (1971) appears to be an efficient manner in which to derive frequency. With the similar and different modifier repetitions, however, it has already been pointed out that any change in wording was sufficient to discriminate the repeated occurrences of the sentences and thus the spacing variable is not operative.

On the other hand, the frequency judgments for nouns were

influenced by spacing regardless of the type of modifiers. While the subjects were not completely accurate in their frequency judgments, they were, nevertheless, sensitive to the repetitions of nouns in differently worded sentences but at the same time were capable of detecting the dissimilarity of the sentences. As Jacoby has pointed out, to consider frequency as an encoded attribute would necessitate that an index be maintained for the occurrence of particularly worded sentences as well as an index for a particular word regardless of its sentence context. The preferred interpretation was, of course, that the ability to judge the frequency of both sentences and words viewed in the same stimulus material is largely a matter of deriving the information at the time of test on the basis of task instructions and the retrieval of cues which allow the judgments to be carried out in the most efficient manner.

THE PRESENT EXPERIMENT

Recent research has provided a number of insights concerning the representation of frequency in memory. For example, Howe (1973) has shown that the perceived frequency of a homograph is influenced by the manner in which semantic context biases encoding of meaning. Jacoby (1972) has also demonstrated that the frequency judgments of nouns are influenced by the type of sentences in which nouns are included. Additionally, Jacoby (1972) has found that the frequency of entire sentences can be somewhat adequately assessed. Perhaps, it is more reasonable to ask the question "what can be a frequency unit" rather than the question "what is a frequency unit."

While Jacoby (1972) has presented evidence indicating that the repetitiveness of a sentence can be accurately judged, it is not possible to conclude that these judgments were made on the basis of the semantic content of the sentence as a whole. The fact that sentences followed by synonomous repetitions were judged no higher than sentences followed by completely different repetitions (i.e. the noun was the same but the modifiers had different meanings) indicates that sentence judgments in Jacoby's experiment were made on the basis of the actual physical presence or absence of the modifying words. It is obvious, however, that people do not remember exactly what they hear or see. (Sachs, 1967; Begg, 1971)

Therefore, the present experiment addresses itself to two basic issues. First, is it possible to adequately assess the frequency with which the semantic content of a sentence has been expressed without

reference to the actual words in which that idea has been stated? Second, if meaning or semantic content can be judged, will the spacings of meanings facilitate frequency judgments of semantic content in a manner similar to the facilitation found by the spacing of words or intact sentences? In short, is it reasonable to speak of the frequency of sentence meaning?

Specifically, the subjects in the present experiment were presented with a series of simple active declarative sentences in which a critical set were repeated in either an exact or a paraphrased form. Paraphrases were constructed by substituting synonyms for all content words. For example, "the huge policeman halted the expensive automobile" would be rephrased as "the large cop stopped the high-priced car." Consensus that a sentence and its paraphrase did express the same content was obtained by asking a group of subjects to rate the degree of semantic similarity existing between a pair of sentences. Only those sentences which were judged to be highly similar were subsequently included in the experimental lists.

All subjects viewed the series of sentences under general memory instructions. Following study, however, groups of subjects were given different instructions regarding the type of frequency judgments which they were to make. In one condition, the subjects were asked to judge the number of times that a sentence had occurred during study in the exact form (i.e. same wording) shown on the test sheet. A second group of subjects was instructed to judge the frequency with which the particular meaning expressed by a test sentence had occurred regardless of the exact wording. In order to perform this task, the subjects must ignore the particular words comprising a sentence and focus on the semantic content.

Performance in this second condition is of importance in two specific respects. First, if the frequency judgments of meanings can be reasonably accomplished, this will indicate that frequency can be reported on yet another dimension of the stimulus material. Jacoby (1972) has already demonstrated that the frequency of specifically-worded sentences and the occurrences of nouns in different sentences can be derived from the same study presentations. Additionally, such a finding would be consistent with research indicating that semantic content is the primary component of long-term memory for sentences.

Secondly, the role of the spacing of repetitions in the meaning judgment condition is also of importance in supporting a notion that frequency is derived on the basis of whatever dimensions are relevant for the type of judgments being requested. In Jacoby's (1972) experiment, it was found that spacing was only effective for exact repetitions of experimental sentences. Any change in wording was sufficient to discriminate a similarly worded repetition from the original sentence and thus frequency judgments were not influenced by the similar repetitions regardless of the level of spacing. In the present experiment, however, a difference in wording is not a relevant dimension for discriminating paraphrased sentences when meaning judgments are requested. It is expected that spacing will facilitate the judgments of frequency in the meaning judgment regardless of whether a sentence is repeated in an exact or paraphrased form. Spacing under these circumstances will always provide a relevant temporal cue for the determination of frequency. This finding would provide considerable support for a position maintaining that frequency is derived on the basis of whatever dimensions are pertinent for the task at hand.

METHOD

Materials. Sixty sentences were employed during the actual experiment, each having a constant syntactical frame. Thirty-six of these sentences were essential to the experimental manipulations (critical sentences); 14 sentences served as filler items in the study lists (filler sentences); and 9 sentences were not seen until the time of test (new sentences). For each critical sentence, a paraphrased version was constructed by substituting synonyms for all content words (i.e. adjectives, nouns, and verbs) while maintaining the same syntactic structure. Roget's Thesarus of the English Language and the Funk and Wagnall's Standard Handbook of Synonyms, Antonyms, and Prepositions were used in selecting synonyms. The 36 critical sentences and their paraphrases were selected from a larger set of synonomous pairs which had been rated by a group of 20 Ss for their degree of semantic similarity. A five-point scale was used for these ratings: 5-exactly the same in meaning; 4-very similar in meaning; 3-similar in meaning; 2-different in meaning; 1-very different in meaning. A number of sentence pairs expressing divergent content were included in the to-be-rated set to ensure that the entire range of the scale would be employed by the Ss. Those 36 pairs receiving the highest ratings of semantic similarity were subsequently included in the experimental manipulations. No two sentences (i.e. critical sentences, their paraphrases, filler sentences, or new sentences) had any content words in common. In order to control for the imagery value of the sentences, the 36 critical sentences (only the original version) were submitted to four independent judges for rating on

concreteness. Again, a five-point scale was used: 5-very easy to imagine; 4-easy to imagine; 3-fairly difficult to imagine; 2-difficult to imagine; 1-very difficult to imagine. The 36 sentences were then grouped into 3 sets of 12 sentences for control purposes in the study lists; representing high, medium and low levels of concreteness.

Design and Procedure. Visual presentation of the sentences was used. Each subject received a deck of 3X5 inch cards with one sentence typed per card. Prior to study, all Ss were given the following instructions:

> This experiment is concerned with your memory for sentences. During the series of sentences which you will be asked to study, some of the sentences will be repeated and some of the sentences will have very similar meanings.

The Ss were then paced through their deck of cards at a rate of 7 seconds per card. Within a study deck, the frequency with which the critical sentences were presented was varied. There were 18-once-presented (1P) and 18 twice-presented (2P) sentences. Regarding the 2P sentences, two additional factors were manipulated: the type of sentence repetition (identical or paraphrased) and the number of other sentences intervening between the repetitions of a 2P sentence (spacing of 0, 4 and 8 sentences). Combination of these factors resulted in 6 types of 2P sentences. There were three instances of each type within a study deck. Additionally, a number of sentences were repeated three times in order to increase the Ss' range of possible responses.

Following study, the Ss were presented with test sheets of sentences. Testing instructions were varied between subjects. One-half of the Ss received instructions which stressed that frequency judgments of the sentences were to be made only on the basis of the specific wording of the test sentences (Condition JOW).

The instructions were as follows:

> The sheets which have been handed-out to you contain a list of sentences. Some of these sentences were included in the deck of cards which you have just studied; others you have not seen. You are asked to judge the number of times which you saw each of the sentences during study. Your judgments may range from 0 to 3. Be sure to judge only the number of times that you have seen the sentence exactly as it is worded on the test sheets given to you.

The other one-half of the Ss received a second set of instructions which emphasized that the frequency judgments be made on the basis of the content or meaning of the sentences. (Condition JOM) These instructions were stated in the following manner.

> The sheets which have been handed-out to you contain a list of sentences. Some of these sentences were included in the deck of cards which you have just studied; others you have not seen. You are asked to judge the number of times that the idea or meaning expressed in each sentence occurred in the study deck regardless of the exact manner in which it was worded. Your judgments may range from 0 to 3. Remember that your judgments are to be made on the idea or meaning of the sentence and not on the exact wording.

Examples of appropriate frequency judgment procedures for the two tasks were given to the respective instructional groups.

Two control groups were included in which the study decks were unmixed with respect to the repetitions of the 2P sentences. In one control group, the repetitions of the 2P sentences were always identical in form and JOW instructions were given. In the second control group, the repetitions of the 2P sentences were always paraphrased and JOM instructions were given. The study and test instructions given to the control groups were slightly modified (i.e. the first control group was not told to expect different sentences with similar meanings etc.) and the spacing of repetitions

was varied in both control groups.

Study Decks. A study deck consisted of 79 card presentations. A primacy and recency buffer, respectively, occupied ordinal positions 1 to 8 and 72 to 79. Within each buffer, there were three 1P sentences, one 2P sentence, and one 3P sentence (the 2P and 3P sentences were repeated in identical form). The buffers were constant across all study decks. The central portion of a study deck was divided into three blocks of 21 sentences. Each block was composed of six 1P sentences, six 2P sentences (representing all six of the 2P sentence types), and a 3P filler sentence which was repeated in an identical form. Within a block, two of the 1P sentences represented each of the three levels of concreteness. There was one sentence from each level of concreteness within the three instances of a given 2P sentence type. Additionally, the six 2P sentences in a given block were equally divided among the three concreteness levels. Each block followed a completely different order of sentence-type presentations.

The 36 critical sentences functioned as both 1P sentences and as the six 2P sentence types during the course of the experiment. This necessitated that 12 different study decks be formed. Eighteen critical sentences first served as 1P sentences while the other 18 served as 2P sentences. Groups of three 2P sentences were then systematically rotated through the six 2P sentence types producing six different decks. The designation of 1P and 2P sentences was then reversed and the same procedure was followed with six more decks resulting. While the position of the sentence types within a block remained constant throughout the experiment, the blocks, themselves, were rearranged in three different manners (i.e. B-1, B-2, B-3; B-3, B-1, B-2, and B-2, B-3, B-1). The position of sentence types in the study decks was partially

controlled by presenting four of the different study decks under these three block orders.

The study decks for the control conditions were similar in that the buffers, the filler items, and the division of the central portion of the deck into blocks were maintained. In the case of 2P sentences, there were two instances of each spacing level within a given block of sentences. Two critical sentences from each level of concreteness were included among the six 2P sentences in a block. Six different study decks were needed for each critical sentence to serve in all of the sentence types (1P sentences and three spacings of 2P sentences). These decks were formed by rotating groups of six critical sentences through the different sentence types via the procedures described above. Again, the position of the sentence types was partially controlled by rearranging the block orders in three different manners. Two study decks were studied under each of the three different orders.

<u>Test Booklets</u>. The booklets were composed of three sheets of paper with 16 sentences typed per sheet. Each page included six 1P sentences, six 2P sentences, one 3P filler sentence taken from the central portion of the study decks, and three new sentences. Two sentences from each level of concreteness were included among the six 1P sentences on a page. In addition, all six types of 2P sentences were found on a given page. In the control conditions, two instances of each 2P type were found per page. Order of test items was randomized on each sheet of the test booklet. Paraphrased 2P sentences were always tested by presenting the arbitrarily designated original critical sentence during test.

<u>Subjects</u>. A total of 128 introductory psychology students at Loyola University served as Ss. Participation partially fulfilled a

course requirement. Twenty Ss served in the initial rating procedure. The remaining 108 Ss were randomly assigned upon arrival at the experimental room to one of the between Ss' treatment or control groups. In the experimental conditions, the treatment groups were formed by combining the two types of test instructions (JOW and JOM) with the 12 different study decks. In the control conditions, the groups were formed by combining the two test instructions with the six study decks. Ss were tested in groups of three under the same treatment conditions. Thus, 36 Ss served in each experimental instruction condition and 18 Ss served in each control instruction condition.

RESULTS

Mixed List Conditions

An initial three-way analysis of variance (Instructions X Type of Item X Sentence Concreteness) indicated that the concreteness level of the sentences produced no main effect and did not enter into any interactions with other experimental variables. Therefore, the mean judged frequency for once-presented (1P), twice-presented (2P-S), and paraphrased items (2P-P) in the two instructional conditions (JOW and JOM) have been collapsed over the levels of concreteness and are shown in Table 1. The means indicate that Ss displayed differential performance depending upon the type of test instructions which they received. When frequency judgments were requested for the exact wording of the sentences, the 2P-P items were judged in a manner similar to the judgments made on the 1P items. Alternatively, the 2P-P items were judged like 2P-S items when Ss were asked to make frequency estimates on the basis of the general meaning expressed by the sentences.

This description of performance was statistically analyzed by a series of orthogonal planned comparisons performed on the simple effects of Item Type (1P, 2P-S, 2P-P) in the two instructional groups. For Ss making judgments of sentence wording, the judged frequency of 1P and 2P-P items did not differ significantly, but the weighted comparison between 1P and 2P-P items vs. 2P-S items was highly significant, $F(1, 70) = 194.48$, $p < .001$. Two planned comparisons were also performed for condition JOM. The first comparison indicated that mean judged frequency for 2P-S and 2P-P items did not differ significantly while the second comparison found that the judged frequency of 1P items was significantly

TABLE 1

Mean Frequency Judgments in Mixed List Conditions

TEST	TYPE OF ITEM 1P	2P–P	2P–S
JOW	.844	.924	1.481
JOM	1.000	1.728	1.836

different from frequency judgment performance for the 2P-S and 2P-P items, $F(2, 70) = 226.43$, $p < .001$. In sum, the Ss were quite capable of performing two types of frequency judgment tasks following study under identical conditions.

One additional point, however, needs to be made regarding the means shown in Table 1. The Ss tested under JOM instructions tended to produce generally higher estimates of frequency than Ss tested under JOW instructions. This observation is true for 1P and 2P-S items where the two test instructions under optimal performance should both lead to the same frequency estimates. Comparison of the means in the first and third columns of Table 1 makes this tendency clear. Simple effects analyses for Instructions statistically confirm the higher frequency judgments in the JOM condition for both the 1P items, $F(1, 70) = 31.16$, $p < .001$ and the 2P-S items, $F(1, 70) = 4.27$, $p < .01$.

The tendency for higher frequency estimates in Condition JOM is also reflected in Ss' performance on sentences presented only at the time of test (not-presented or NP items). The false-alarm rates (i.e. the proportion of total responses to NP sentences that were 1 or greater) were 12.6% for Condition JOM and 5.8% for Condition JOW. The corresponding mean judged frequencies were .19 in Condition JOM and .06 in Condition JOW. However, t-test performed on mean frequency judgments for NP items failed to reach statistical significance.

In order to assess the role of the spacing of item repetitions, frequency estimates for the 2P-S and 2P-P sentences were further analyzed in a 2 X 2 X 3 X 3 analysis of variance (Instruction X Type of Repetition X Sentence Concreteness X Spacing of Repetitions). This analysis produced significant main effects for Instructions (JOW or JOM), $F(1, 70 = 63.70$, for the Type of Repetition

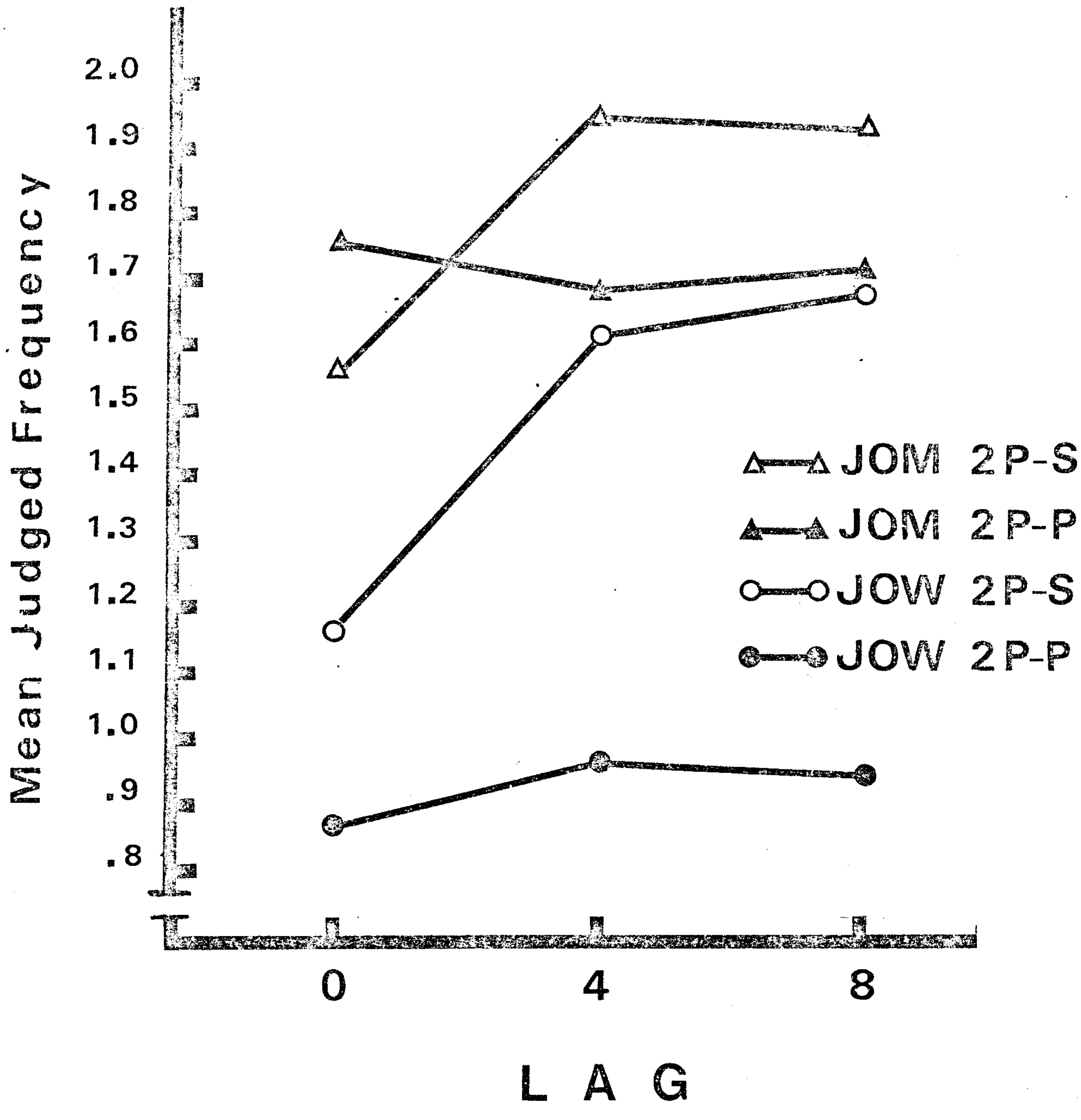

FIGURE 1. Frequency Judgment Performance For 2P-S and 2P-P Sentences in the Mixed List Conditions.

Unmixed List Conditions

In the unmixed list conditions, the Concreteness of the study sentences again entered into no significant sources of variation in any of the analyses performed. Therefore, no further mention of this variable will be made. Frequency judgment performance in the JOW and JOM condition was analyzed separately. In both conditions, Frequency of Presentation (1P and 2P) produced a significant main effect; for Condition JOW, $\underline{F}(1, 17) = 76.84$ and for Condition JOM, $\underline{F}(1, 17) = 147.75$ (both $\underline{p}$'s $< .001$).

Performance on the NP sentences was quite comparable in the two instructional groups. The false-alarm rates were 5.5% for Condition JOW and 4% for Condition JOM. A $\underline{t}$-test performed on the mean judged frequencies for NP items proved nonsignificant, $\underline{t}(70) = 1.74$, $\underline{p} < .10$.

Figure 2 illustrates the mean frequency judgments made by the instructional groups for 1P and 2P items at each of three levels at which sentences repetitions were spaced (0, 4 and 8). It should be remembered that twice-presented items in Condition JOW were repeated in the same form (2P-S) while the twice-presented items in Condition JOM consisted of an original sentence followed by a paraphrased version (2P-P). The two unmixed list conditions were comparable, however, in terms of the number of 1P and 2P (either 2P-S or 2P-P) items in the study lists. Inspection of Figure 2 shows that frequency judgments in Condition JOW increased from 0 spacing to 4 spacing and then remained relatively unchanged at spacing level 8. In Condition JOM, however, judged frequency first increased and then decreased across the different spacing levels. The amount of increase from 0 spacing to 4 spacing was considerably larger in Condition JOW than in Condition JOM (1.35 to 1.94 for Condition JOW vs. 1.72 to 1.85 for Condition JOM).

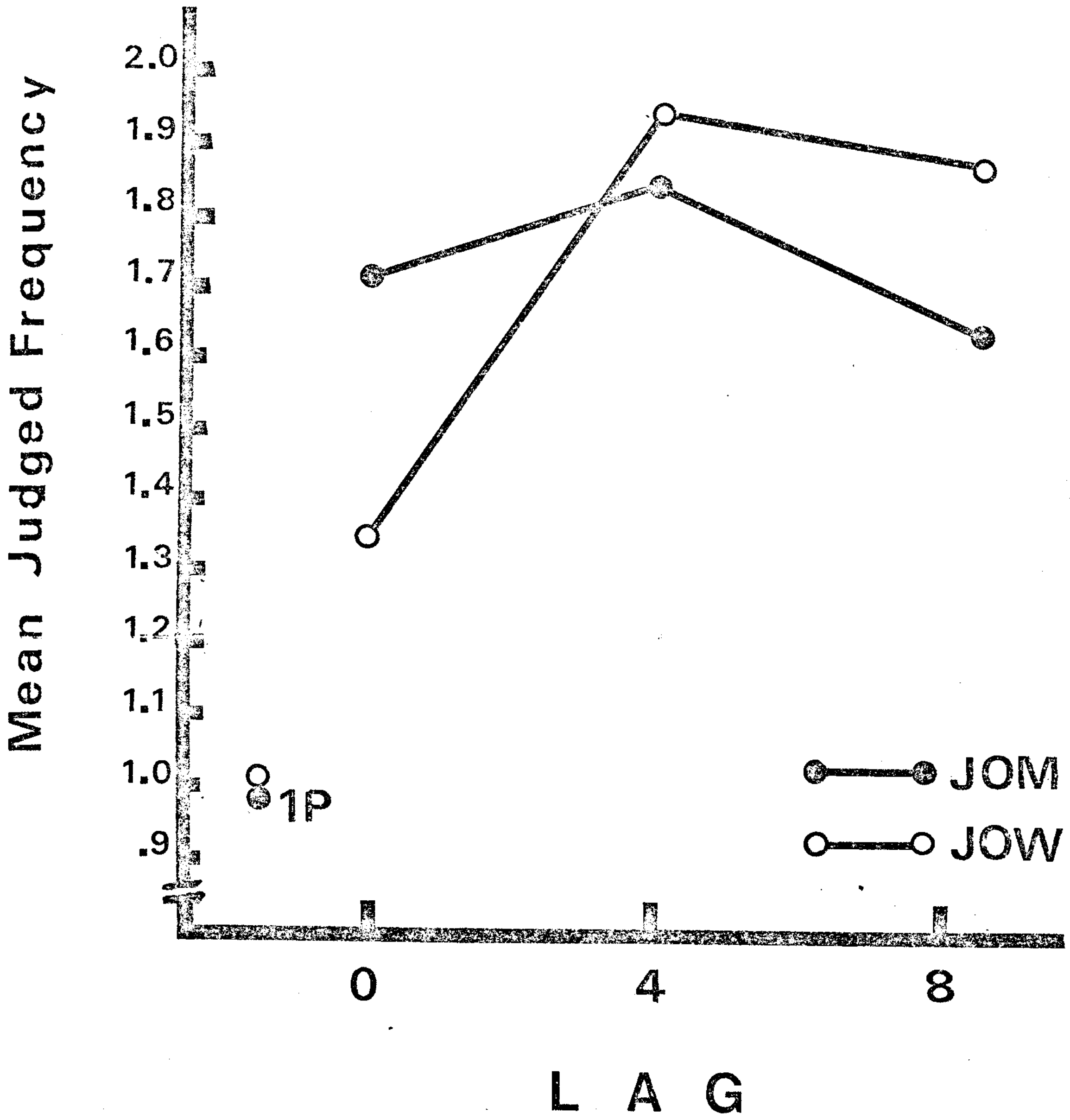

FIGURE 2. Frequency Judgment Performance in the Unmixed Lists.

(Same or Paraphrased), $F(1, 70) = 70.28$, and for the Spacing of Repetitions (0, 4, or 8), $F(2, 140) = 17.86$ (all p's $< .001$). Additionally, a significant Instruction by Repetition Type interaction, $F(1, 70) = 31.38$ and a significant Repetition Type by Spacing interaction, $F(2,140) = 46.58$, were found (both p's $< .001$).

The Instruction by Repetition Type interaction can be attributed to the differential performance displayed in the two instructional groups for 2P-P sentences. The nature of the interaction can be seen in Table 1. The 2P-P sentences were treated as 1P items in Condition JOW and as 2P-S items in Condition JOM.

Figure 1 shows the mean judged frequency for the two repetition types (2P-S and 2P-P) at the different levels of spacing in both Condition JOW and JOM. The Repetition Type by Spacing interaction is reflected in the fact that 2P-S sentences resulted in higher frequency judgments at the longer lags in both instructional conditions, while the 2P-P sentences showed no effect due to spacing in either Condition JOW or JOM. This observation is true although the 2P-P sentences were judged as 1P sentences in Condition JOW and as 2P-S sentences in Condition JOM. The three-way interaction of Instructions X Repetition Type X Spacing was not significant.

Simple effects of Spacing for the 2P-S items collapsed over instruction groups indicates significant variation due to the level of spacing, $F(2, 140) = 39.57$, $p < .001$. Newman-Keuls tests performed on the means for each spacing level indicate that levels 4 and 8 are significantly higher than the 0 level but do not differ between themselves (significant p's $< .05$). The simple effects analysis for spacing showed no significant effect on the performance for 2P-P sentences where the second occurrence of an item was a synonomous version of the first occurrence.

Statistical confirmation of this description of Figure 2 is provided by two 3 X 3 (Spacing X Concreteness) within-subject analyses of variance. The analysis performed for Condition JOW produced a significant main effect due to the level of spacing, $F(2, 34) = 13.50$, $p < .01$. Individual comparisons among the means at each spacing level were conducted using the Newman-Keuls procedure. It was found that the judged frequency at spacing levels 4 and 8 did not differ significantly but was substantially higher than the judged frequency at the 0 spacing level (significant p's $< .05$). For Condition JOM, no significant variation in the frequency judgments could be attributed to the influence of spacing.

DISCUSSION

The pattern of results obtained in the present experiment provide two significant findings regarding frequency judgment performance with sentence material. First, it has been shown that reliable frequency information can be reported along the dimensions of either the exact wordings or the basic "gists" of study sentences. Second, frequency estimates for paraphrased sentences in Condition JOM demonstrates a situation in which the facilitative effects of distributed repetitions fails to occur although the two synonomous sentences are evidently viewed in conjunction by those Ss making judgments of sentence meaning.

In the present study, the Ss were given sentences for study but were not told until the time of test what the nature of the impending test would be. Performance by Ss in the JOM condition clearly indicates that the frequency of some abstract representation of sentence's meaning can be assessed even when the manner of phrasing the meaning varies markedly at each occurrence. In this experiment, two synonomous versions of the same basic meaning were properly seen as a twice-occurring event when Ss were told to make their frequency judgments on the basis of the study sentences underlying meaning without regard to the specific wording. On the other hand, Ss in Condition JOW were quite capable of discriminating twice-presented (2P-S) and paraphrased (2P-P) sentences when the test instructions placed emphasis on the judgment of a sentence's exact wording. Under these conditions, the paraphrased items were treated in a manner similar to once-presented sentences.

These results would necessitate that an index for a sentence's

exact wording and an index for some abstract representation of the sentence's meaning would have to be maintained if frequency information is directly encoded into memory as is asserted by Underwood (1969). In short, the Ss must necessarily perform a summing operation along two dimensions of the study sentences (i.e. the wordings and the meanings). That such operations might be carried out does not seem unreasonable for the Ss were told prior to study that some sentences would be repeated in identical form while in other cases two different sentences would be similar in meaning. The two salient dimensions of the sentence material, therefore, were made relatively clear.

The spacing of sentence repetitions has resulted in consistent findings in the mixed and unmixed list conditions of the present study. When the study sentences were repeated in identical form, frequency judgments much more closely approximated the actual presented frequency when the two occurrences of the sentences were separated by either 4 or 8 intervening items than if the two repetitions came in immediate succession. This facilitative effect of spacing for 2P-S sentences was observed in Conditions JOW and JOM for both mixed and unmixed study lists. It should be noted that a normal "spacing effect" implies that frequency judgments should become progressively higher as the lag between the two repetitions becomes larger. In the present study, frequency judgments for 2P-S items did not increase beyond the 4 level of spacing. However, it may be suggested that a "ceiling" level was reached at spacing level 4 due to the restricted range of possible responses available in this experiment.

For the paraphrased sentences, no spacing effect was found; even in Condition JOM where the 2P-P items were treated as twice-occurring events. In the unmixed list conditions, the Ss making

judgments of sentence meaning did appear to be influenced to some extent by the lag between the two synonomous sentences. It may be noted that frequency judgments did increase from the 0 to 4 levels of spacing. However, the frequency judgments for paraphrased sentences at spacing level 8 dropped below the performance shown at the 0 level and none of the differences among the lags were found to be significant. In mixed list conditions, judged frequency for paraphrased sentences was virtually identical at the three levels of spacing in Condition JOM. These results indicate that the lag between paraphrased sentences played little or no role in the frequency judgments made on this type of study item.

One additional aspect of the present results warrants some interpretation. While the findings of this experiment clearly indicate that both the wording and the meaning of study sentences can be adequately assessed, the question may still be raised as to whether or not the judgment of meaning and the judgment of wording tasks are of equal difficulty. Studies such as those conducted by Sachs (1967) and Begg (1970) would suggest that general meaning as opposed to the exact wording of sentences would more likely be remembered. In terms of the present experiment, such findings would imply that the judgments of sentence meaning should more closely approximate the actual presented frequency of meanings than the judgments of sentence wording should approximate the actual occurrence of specifically worded sentences.

In the mixed list conditions of this study, the JOM instructions did, in fact, tend to produce generally higher frequency judgments than Condition JOW. This tendency is reflected in the mean frequency judgments for NP, 1P, and 2P-S sentences where the two instructions under optimal performance would result in the same judgments. In terms

of the "summed" or encoded view of frequency, this finding would imply that the Ss employed different criteria in reporting their stored frequency information depending upon the test requirements. These observations, however, must be interpreted with caution because the overall means for once-presented (1.09) and twice-presented (1.73) sentences were identical for the JOM and JOW conditions in the unmixed lists.

The higher frequency judgments for the JOM condition in the mixed lists might be more parsimoniously explained in terms of differential response biases in the two instructional conditions which were built into the experimental design. In the JOW condition, the task demands required that three-fourths of the experimental sentences be called once-occurring while the demands for appropriate performance in the JOM condition specified that only one-half of the experimental sentences be judged as once-presented. These differences in response probabilities were not present in the unmixed lists.

In order to test this possible explanation for the higher frequency judgments in Condition JOM, a testing effects analysis was conducted for the frequency judgments made in the mixed list conditions. This analysis considered the frequency judgments for the different types of items (1P, 2P-S, and 2P-P) in the two instruction groups according to whether they occurred in the first, second or third portion of the test list. Thus, the position in the test list at which the sentences were presented was taken into account. If the response bias described above were actually operating, it would be expected that the judgments in Condition JOM should become higher in each third of the list and the judgments in Condition JOW lower in each portion of the test list as the differential response probabilities became apparent

to the Ss.

The testing effects analysis, however, found no main effect for position nor any interaction of position with either the instruction groups or the type of item. Therefore, the results of this experiment are in agreement with the finding of Sachs (1967) and Begg (1971) that the meaning rather than the wording of sentences is the more likely display of memory for sentence material. The failure to find an overall difference in frequency judgments between Conditions JOW and JOM in the unmixed lists may be attributed to the fact that in these conditions, the Ss were presented with only one type of repetition in the study lists. As compared to the mixed list conditions where both types of repetitions were presented, performance with the unmixed lists most probably constituted an easier task in which frequency information would only be summed or encoded along one dimension (i.e. wording or meaning) of the study sentences.

While the results of the present experiment have been considered in light of an encoded view of frequency, Jacoby (1972) has suggested an alternative position regarding frequency information. In his study, sentences were also presented for study under neutral instructions. At the time of test, Jacoby's Ss were asked to judge either the frequency of the intact sentences or the frequency of subject-nouns contained in the sentences. In some cases, the subject-nouns were repeated in identical sentence frames, while in other cases, subject-nouns were repeated in different sentence frames. Like the present experiment, Jacoby (1972) discovered that Ss were fairly accurate in reporting frequency information on both dimensions (i.e. sentences and nouns). In interpreting this finding, Jacoby maintained that frequency information is not directly entered into memory but is rather derived

at the time of test on the basis of contextual information. Context for Jacoby can be classified as of two basic types; temporal context referring to the point in the study list at which a given event occurred and semantic context referring to either the sentences immediately preceding and following a given sentence or to the sentence frame in which a given subject-noun occurred.

Jacoby (1972) maintained that the occurrences of both sentences and nouns result in independent traces as the Hintzman and Block (1971) multiple-trace hypothesis would suggest and that further, the different occurrences of an item result in traces which are in some way marked for both the temporal and semantic context in which they occurred. At the time of test, a S would then retrieve the different traces requested by task instructions. The extent to which the independent traces of a given event overlapped in contextual features would then dictate a S's success in estimating frequency along different dimensions of the study material (either nouns or sentences in the Jacoby study). Howe (1973) has provided a more detailed analysis of the derived view of frequency based on contextual feature tagging.

While context undoubtedly may influence frequency judgment performance such as that found in the Jacoby study, it does not seem reasonable to deny that a S will ever sum events of different types during study. Jacoby (1972) has argued that the encoded view of frequency would necessitate that frequency information in his study must be stored for each sentence and additionally, for each word found in the sentences regardless of the particular sentence frames in which they occurred. His conclusion was that such an analysis of frequency information would "soon become unwieldy." Nonetheless, it is possible that Ss may sum frequency along a number of particularly salient

dimensions of the study material as has been suggested for the findings of the present experiment. Subject-nouns along with the intact sentence units appear to be likely candidates for the salient dimensions in the Jacoby (1972) study. It is unlikely that the Ss in the Jacoby experiment would have been so successful in reporting the frequency of less essential words found in the sentences such as adjectives or articles. The position taken here is that frequency may be summed (and thus encoded) along a number of salient dimensions found in study material and necessarily derived along any other dimensions which a S does not view as salient during study. The number of salient dimensions along which frequencies can be summed at any one time remains a question for future investigation.

Jacoby (1972) has also considered the role of spacing in frequency judgment processes. Basically, Jacoby discovered that spacing effects were markedly apparent when the to-be-judged events were repeated in identical fashion (i.e. the intact sentences and subject-nouns repeated in the same sentence frames) and considerably reduced when the context of the to-be-judged events varied at each repetition (i.e. subject-nouns repeated in different sentence frames). These results are, of course, not unlike the influence of spacing found in the present research; clear spacing effects were only present for those sentences repeated in identical fashion. The role of spacing must then be considered in light of the encoded and derived views of frequency information.

Under the derived point-of-view, frequency estimates depend upon a S's ability to retrieve the independent traces associated with the the repetitions of a given item, and the various traces are more readily retrievable if they are readily discriminable.

According to Jacoby (1972), discriminability depends upon the number of contextual features which the different traces for a repeated event have in common. Therefore, the derived view of frequency would explain the "spacing effect" by asserting that the distribution of repetitions in a study list would provide the different occurrences of an item with considerably more distinct contextual environments than would be the case if the repetitions occurred in a successive fashion. Lower frequency judgments for items at a 0 spacing level can then be interpreted as a failure to discriminate the independent traces of a repeated event on the basis of their contextual features. The derived point-of-view could also explain the lack of spacing effects for paraphrased sentences in the present study if the exact wordings of study sentences could be viewed as the semantic context in which the abstract representations of sentence meaning occurred. Context would then make the two occurrences of a paraphrased sentence readily discriminable no matter how widely the two synonomous sentences were dispersed in a study list. Again, Howe (1973) provides a more detailed analysis of such processes.

The "encoded" view of frequency suggested in the discussion of present findings, alternatively, maintains that frequency information is directly entered into memory as a <u>S</u> sums stimulus events along different dimensions of the study material. The encoded position does not rely on the discriminability and retrievability of contextual information at the time of test and therefore, different explanations must be sought for the presence of spacing effects in judgments of twice-presented (2P-S) sentences and the absence of spacing effects in judgments of paraphrased (2P-P) sentences in the present study.

Such explanations may be found if it can be assumed that lower

frequency judgments under massed conditions of presentation are due to a failure to adequately encode the second presentation of an item rather than to a failure to distinguish between two adequately encoded traces. Hintzman (1974) has recently reviewed several theories of the spacing effect; two of which interpret the relatively poor performance with massed items as a failure to properly encode the second presentation. These theories have been labelled the "attention hypothesis" and the "habituation hypothesis."

The attention hypothesis, on the one hand, assumes that the inefficient storage of the second occurrence of a massed item is due to a voluntary process on the part of the subject. Very basically, the hypothesis states that during the second presentation of a massed item, a subject simply ceases processing of the item and rests or redirects his attention to the processing of other items in the study list. The reason that a subject would undertake either of these two activities is presumably because he believes that he has sufficiently encoded the item during the first occurrence.

The habituation hypothesis, alternatively, views the inefficient encoding of the second occurrence of an item as due to some underlying process over which the subject has no control. Hintzman (1974) has likened the notion of habituation to a psychological refractory period during the threshold "for the response of storing a particular kind of memory trace" (p. 21) is raised. Habituation, under this view, will occur regardless of the subject's level of attention. Either hypothesis potentially provides the encoded view of frequency with an explanation as to why massed repetitions of items would be less efficiently summed than distributed repetitions of the items.

While the present study was not designed to specifically evaluate

these two proposals, the findings can provide some insight into which of the two hypotheses would be more amenable to an explanation of spacing effects found in the judgment of sentence frequency. If the assumption can be made that an abstract representation of a sentence's meaning is stored in memory and also, that the synonomous versions of paraphrased items result in the same abstract representation, it appears that a situation has resulted in which a subject will not attenuate his attention for the second occurrence of a paraphrased sentence. In the JOM condition, the subject's task is to judge the underlying meanings of the sentences but because the underlying meaning of a sentence will not become apparent until the sentence has been decoded and comprehended, the subject will not alter his attention to paraphrased sentences even when the two versions are presented in immediate succession. The altering of attention would, of course, be expected for sentences repeated in identical form. The failure to find a spacing effect for paraphrased items in Condition JOM provides some support for the attention hypothesis.

If the habituation notion of spacing is to be accepted, the present findings suggest that it is the response of storing a rather true copy of the sentence which habituates in the processing of sentence material. The failure to find a spacing effect for paraphrased sentences in the JOM condition seems to indicate that the storage of the abstract representation of a sentence's meaning is not susceptible to the habituation process.

<u>Summary</u>

Subjects in the present experiment succeeded in judging the frequency of either the exact wording or the underlying meaning of study sentences. Frequency estimates for paraphrased sentences clearly

indicated that subjects could focus on either surface properties of the two synonomous sentences or on the underlying "gist" of the two sentences depending upon the task demands. Because these task demands were not specified until after the sentences had been presented for study, it has been argued that frequency information can be summed simultaneously on a number of salient dimensions found in the sentence material. The occurrence and nonoccurrence of spacing effects in different conditions of the present study have been considered in light of an attention and a habituation hypothesis. The findings of this experiment seem more amenable to an attention explanation. If a habituation process is occurring, it appears that the response of storing an exact representation of a sentence is decremented rather than the response of storing an abstract trace of the sentence's meaning.

REFERENCES

Anderson, J.R., & Bower, G.H. Recognition and retrieval processes in free recall. Psychological Review, 1972, 79, 97-123.

Begg, I. Recognition memory for sentence meaning and wording. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 176-181.

Bobrow, S.A. Memory for words in sentences. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 363-372.

Bregman, A.S. & Strasberg, R. Memory for the syntactic form of sentences. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 396-403.

Cole, L. & Kank, N.J. Transfer of implicit associative responses between free-recall and verbal discrimination tasks. Journal of Verbal Learning and Verbal Behavior, 1972, 95, 110-115.

Ekstrand, B.R., Wallace, W.P. & Underwood, B.J. A frequency theory of verbal discrimination learning. Psychological Review, 1966, 73, 566-578.

Gartman, L. & Johnson, N.F. Massed vs. Distributed practice of homographs: A test of the differential encoding hypothesis. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 808-12.

Gorman, A.M. Recognition memory for nouns as a function of abstractness and frequency. Journal of Experimental Psychology, 1961, 61, 23-29.

Hall, J.W. & Pierce, J. Recognition and recall as a function of variations in memory-encoding instructions. Paper presented at the Thirteenth Annual Meeting of the Psychonomic Society, St. Louis, November, 1972.

Hintzman, D.L. Apparent frequency as a function of frequency and spacing of repetitions. Journal of Experimental Psychology, 80, 139-145.

Hintzman, D. L. Theoretical Implications of the Spacing Effect. In R. L. Solso (Ed) Theories of Cognitive Psychology: The Loyola Symposium Washington D.C.: Earlbaum Press, 1974.

Hintzman, D.L. & Block, R.A. Memory judgments and the effects of spacing. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 561-566.

Hintzman, D. L. & Block, R.A. Repetition and memory: Evidence for a multiple-trace hypothesis. Journal of Experimental Psychology, 1971, 88, 297-306.

Hintzman, D. L. & Waters, R.M. Interlist and retention intervals in list discriminations. Psychonomic Science, 17, 384-387.

Hintzman, D.L. & Waters, R.M. Recency and frequency judgments as factors in list discriminations, Journal of Verbal Learning and Verbal Behavior, 1970, 9, 218-221.

Howell, W.C. The Representation of frequency in memory. Psychological Bulletin, 1973, 80, 44-53.

Jacoby, L.L. Context effects on frequency judgments of words and sentences. Journal of Experimental Psychology, 1972, 94, 255-260.

Kausler, D. H & Dean, J. Verbal Discrimination and implicit associative responses. Psychonomic Science, 1967, 7, 285-288.

King, M.E. & Levin, I.P. Tests of frequency theory in verbal discrimination transfer with varying paradigms and List 1 trials Psychonomic Science, 1971, 25, 226-228.

Kucera, H. & Francis, W.N. Computation analysis of present-day American English, Providence: Brown University Press, 1967.

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

Melton, A.W. The situation with respect to spacing of repetitions in memory. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 595-606.

Pasko, S.J. & Zechmeister, E.B. Temporal separation in verbal discrimination learning. Journal of Experimental Psychology 1973, in press.

Raskin, D.C., Boise, C., Rubel, E.W., & Clark, D. Transfer tests of the frequency theory of verbal discrimination. Journal of Experimental Psychology, 1968, 76, 521-529.

Rowe, E.J. Frequency judgments and recognition of homonyms. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 440-447.

Sachs, J.S. Recognition memory for syntactic and semantic aspects of connected discourse. Perception & Psychophysics, 1967, 2, 437-442.

Solomon, R.L. & Howes, D.H. Word probability, personal values, and recognition thresholds for words. Journal of Experimental Psychology, 1951, 43, 195-201.

Thios, S.J. Memory for words in repeated sentences. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 789-793.

Thorndike, E.L. & Lorge, I. The Teacher's Word Book of 30,000 Words. Teachers College, Columbia University, Bureau of Publication, 1944.

Underwood, B.J. False recognition produced by implicit verbal responses. Journal of Experimental Psychology, 1965, 70, 122-129.

Underwood, B.J. Attributes of memory. Psychological Review, 1969, 76, 559-573. (a)

Underwood, B.J. Some correlates of item repetition in free-recall. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 83-94 (b).

Underwood, B.J. Are we overloading memory. In A.W. Melton & E. Martin (Eds) Coding Processes in Human Memory. Washington: V.H. Winston & Sons, 1972, 1-23 (a).

Underwood, B.J. Word recognition memory and frequency information. Journal of Experimental Psychology, 1972, 94, 276-283. (b)

Underwood, B.J. & Freund, J.S. Two tests of a theory of verbal discrimination. Canadian Journal of Psychology, 1968, 22, 96-104.

Underwood, B.J. & Freund, J.S. Testing effects in the recognition of words. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 117-125 (a).

Underwood, B.J. & Freund, J.S. Relative frequency judgments and verbal discrimination learning. Journal of Experimental Psychology, 1970, 83, 279-285 (b).

Underwood, B.J. & Freund, J.S. Word frequency and short-term recognition memory. American Journal of Psychology, 1970, 83, 343-351.

Wickens, D.D. Encoding categories of words: An emperical approach to meaning. Psychological Review, 1970, 77, 1-15.

Zechmeister, E.B. & Gude, C. B. Instruction effects in recognition memory. Bulletin of the Psychonomic Society, 1973, in press.

Zimmerman, J. & Underwood, B.J. Ordinal position knowledge within and across lists as a function of instruction in free-recall learning. Journal of General Psychology, 1968, 79, 301-307.

APPENDIX A

APPENDIX A

Sentence Material Used in the Mixed and Unmixed Lists

Experimental sentences presented during study:
the original sentences and their paraphrases.
Mean Concreteness ratings are indicated in parentheses.

High Concreteness Sentences

The impenetrable barricade obstructed the primary road.
The impassable barrier blocked the main highway. (4.25)

The absent-minded attorney mishandled the critical case.
The forgetful lawyer bungled the important trial. (4.50)

The pushy journalist questioned the dishonest politician.
The aggressive reporter quizzed the corrupt official. (4.25)

The frugal tourists rested in a spotless hotel.
The thrifty travelers relaxed in an immaculate inn. (4.75)

The novice skydiver came down on the solid earth.
The amateur parachutist landed on the hard ground. (4.25)

The murky fluid spoiled the clear stream.
The dark liquid ruined the unpolluted creek. (4.25)

The curious researcher investigated an appealing subject.
The inquisitive experimenter studied an interesting topic. (4.50)

A permanent stain tarnished the plush carpet.
An indelible blot discolored the thick rug. (4.75)

The famished flock drank from the quiet pond.
The starving herd watered at the calm pool. (4.25)

The large policeman stopped the expensive automobile.
The huge cop halted the high-priced car. (4.25)

The brave horsemen carried brilliant banners.
The courageous cavalry bore brightly-colored flags. (4.50)

The intoxicated tramp staggered into the grimy tavern.
The drunken bum stumbled into the dingy bar. (4.75)

Medium Concreteness Sentences

The angry soldier fired the deadly pistol.
The irritated G.I. shot the dangerous revolver. (3.75)

The community citizens arranged the yearly feast.
The local residents organized the annual banquet. (3.75)

The muscular boxer registered a brutal punch.
The burly fighter dealt a crushing blow. (3.50)

The nosey foreman observed the employee's actions.
The snoopy supervisor watched the worker's behavior. (3.50)

The cunning robbers stole the precious timepiece.
The clever bandits swiped the invaluable clock. (3.75)

The feared sickness swept the little village.
The small town was ravaged by the dreaded disease. (3.75)

The tired army defeated the hated enemy.
The weary troops conquered the despised foes. (3.50)

The cranky spinster slipped on the slick pavement.
The crabby old-maid fell on the slippery sidewalk. (4.00)

The repulsive creatures crawled through the spooky cemetery.
The ugly monsters crept through the eerie graveyard. (3.50)

The popular combo played a unique tune.
The well-liked band performed an unusual number. (3.50)

The bold explorers traversed the desolate prairie.
The adventurous pioneers crossed the barren plain. (3.50)

An unexpected disaster ended the exhausting journey.
A sudden accident concluded the fatiguing trip. (3.50)

Low Concreteness Sentences

The controversial proposal caused a heated argument.
The debatable suggestion produced a violent dispute. (2.00)

The upper class ruled over the common folk.
The elite caste reigned over the plain people. (3.00)

The unsightly rubbish filled the empty street.
The unattractive trash covered the vacant avenue. (3.25)

The scanty salary enraged the straightforward assistants.
The meager pay enfuriated the outspoken assistants. (2.25)

The wretched junkie longed for the prohibited dope.
The depraved addict craved the illegal drugs. (2.00)

The winning team appeared to be an unbelievable long-shot.
The triumphant club seemed to be an incredible dark-horse. (3.25)

A pungent odor filled the intire home.
An unpleasant smell permeated the whole house. (2.75)

The good-looking lady was deceived by the handsome lad.
The pretty woman was tricked by the attractive boy. (3.00)

The famed singer released a wonderful album.
The noted vocalist put out a fine record. (3.00)

The growing business showed a sizable profit.
The prospering company displayed a considerable gain. (3.00)

The apprehensive pupil wanted a good schooling.
The anxious student desired a sound education. (2.00)

The loyal typist defended her distinguished boss.
The faithful stenographer protected her respected employer. (3.00)

Sentences Presented in the Primacy and Recency Buffers of the Study Lists.

Once-presented Buffer Sentences

The wholesome cafe served good food.

The displeased executive pardoned his aloof colleague.

The slimy frogs lived in the dreary swamps.

The imaginative artist stared at the shapely model.

An unshakeable faith in God fortifies religious people.

The decrepit docks could only accomodate a feq boats.

Twice-presented Buffer Sentences.

A rock edifice marked the remote boundary.
A stone monument signified the outmost border.

The shrill sound upset the contented chickens.
The piercing noise alarmed the satisfied hens.

Buffer Sentences Presented Three Times.

The cheerful child picked the lovely blossoms.
The happy youngster plucked the beautiful petals.
The gleeful tot gathered the gorgeous flowers.

The skilled physician bandaged the painful wound.
The proficient surgeon covered the throbbing injury.
The expert doctor wrapped the aching sore.

Filler Sentences Presented Three Times in the Study and Test Lists.

The wild tyrant persecuted the radical sect.
The cruel ruler suppressed the dissenting faction.
The hot-blooded dictator oppressed the revolutionary party.

A strong breeze heartened the sturdy sailors.
A vigorous gale perked up the rugged boatmen.
A brisk wind gladened the hardy seafarers.

An unexpected disaster ended the exhausting journey.
A sudden accident concluded the fatiguing trip.
An unforeseen tradegy interupted the tiring voyage.

Sentences Presented Only at the Time of Test.

The devout minister counseled the parishoners.

The penniless serfs reaped the autumn harvest.

The tan thoroughbred galloped the circular course.

The disorderly mob left the bustling station.

The modern slang astounded the scholarly instructors.

The vain bookkeeper purchased a fashionable wig.

The crude laborer gulped the icy beverage.

The insane moron hurt the tiny dog.

The secret agent received a coded message.

APPENDIX B

APPENDIX B

Analysis of Variance Summary Tables

Analysis of Variance for 2P-S and 2P-P Sentences in the Mixed Lists (Instructions X Subjects X Type of Repetition X Lag X Concreteness Level)

Source	Sums of Squares	df	Mean Squares	F
Between Ss				
Instructions (I)	99.445	1	99.445	63.703**
Subjects w I	109.276	70	1.561	
Within Ss				
Type of Rep (T)	38.371	1	38.371	70.280**
Lag (L)	19.473	2	9.737	17.864**
Concreteness (C)	1.061	2	.531	1.124
IT	17.130	1	17.130	31.376**
IL	.659	2	.329	
TL	14.826	2	7.413	17.871**
IC	2.247	2	1.123	2.378
TC	1.284	2	.642	1.238
LC	1.901	4	.475	
ST w I	38.218	70	.546	
SL w I	76.309	140	.545	
SC w I	66.133	140	.472	
ITL	.863	2	.431	1.040
ITC	.914	2	.457	
ILC	1.142	4	.285	
TLC	.290	4	.072	
STL w I	58.070	140	.415	
STC w I	72.569	140	.518	
SLC w I	159.825	280	.570	
ITLC	.142	4	.035	
STLC w I	149.625	280	.534	
Total	1030.032	1295		

** denotes significance at the .001 level

Testing Effects Analysis for 1P, 2P-S, and 2P-P Sentences in the Mixed Lists (Instructions X Subjects X Type of Sentence X Position in Test List)

Source	Sums of Squares	df	Mean Squares	F
Between Ss				
Instructions (I)	22.842	1	22.842	25.834**
Subjects w I	61.892	70	.884	
Within Ss				
Type of Sen (T)	58.352	2	29.176	107.897**
Position (P)	.799	2	.399	1.579
IT	8.594	2	4.297	15.890
IP	.924	2	.462	1.828
TP	.671	4	.168	
ST w I	37.857	140	.270	
SP w I	35.394	140	.253	
ITP	1.153	4	.288	1.254
STP w I	64.351	280	.229	
Total	292.809	647		

** denotes significance at the .001 level

Analysis of Variance for 1P and 2P Sentences in Condition JOW for Unmixed Lists (Type of Sentence X Concreteness Level X Subjects)

Source	Sums of Squares	df	Mean Squares	F
Between Ss	7.180	17		
Within Ss				
Type of Sen (T)	14.208	1	14.208	76.840**
Concreteness (C)	.141	2	.070	
TC	.146	2	.073	
T by Ss	3.143	17	.185	
C by Ss	4.528	34	.133	
TC by Ss	2.780	34	.082	
Total	32.126	107		

**denotes significance at the .001 level

Analysis of Variance for 2P Sentences in Condition JOW of the Unmixed Lists (Lag X Concreteness Level X Subjects)

Source	Sums of Squares	df	Mean Squares	F
Between Ss	14.327	17		
Within Ss				
Lag (L)	10.892	2	5.446	13.497**
Concreteness (C)	.114	2	.057	
LC	3.265	4	.816	1.950
L by Ss	13.719	34	.404	
C by Ss	13.164	34	.387	
LC by Ss	28.456	68	.419	
Total	83.937	161		

** denotes significance at the .001 level

Analysis of Variance for 1P and 2P Sentences in Condition JOM for Unmixed Lists (Type of Sentence X Concreteness Level X Subjects)

Source	Sums of Squares	df	Mean Squares	F
Between Ss	1.965	17		
Within Ss				
Type of Sen (T)	14.199	1	14.199	147.753**
Concreteness (C)	.097	2	.049	
TC	.084	2	.042	
T by Ss	1.633	17	.096	
C by Ss	2.755	34	.081	
TC by Ss	1.176	34	.035	
Total	21.909	107		

**denotes significance at the .001 level

Analysis of Variance for 2P Sentences in Condition JOM of the Unmixed Lists (Lag X Concreteness Level X Subjects)

Source	Sums of Squares	df	Mean Squares	F
Between Ss	6.304	17		
Within Ss				
Lag (L)	1.948	2	.974	3.140
Concreteness (C)	.522	2	.261	1.130
LC	.599	4	.150	
L by Ss	10.552	34	.310	
C by Ss	7.812	34	.230	
LC by Ss	19.985	68	.294	
Total	47.722	161		

APPROVAL SHEET

The thesis submitted by Christopher Bernard Gude has been read and approved by the following Committee:

Dr. Eugene B. Zechmeister, Director
Associate Professor, Psychology, Loyola

Dr. Frank Slaymaker
Assistant Professor, Psychology, Loyola

Dr. John Edwards
Assistant Professor, Psychology, Loyola

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the Committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

5/17/74
Date

E B Zechmeister
Director's Signature