DOCTORAL DISSERTATION SERIES

TITLE DURATION THRESHOLDS IN THE PERCEPTION OF UNPLEASANT WORDS

AUTHOR DAVID MORRIS STERNE

UNIVERSITY MICH. STATE COLL. DATE 1952

DEGREE Ph.D. PUBLICATION NO. 4034

UNIVERSITY MICROFILMS

ANN ARBOR MICHIGAN

DURATION THRESHOLDS IN THE PERCEPTION OF UNPLEASANT WORDS

By

David M. Sterne

A Thesis

Submitted to the School of Graduate Studies of Michigan State College of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

1952

ACKNOWLEDGMENT

The writer wishes to express his appreciation to Dr. Albert I. Rabin, his committee chairman, whose suggestions, encouragement, cooperation, and foresight contributed so significantly to the completion of this dissertation. Thanks are also due Dr. Milton Rokeach, Dr. M. Ray Denny, and Dr. C. A. Lawson, the other committee members, for their stimulating questions and advice which did much to help the writer clarify his own thinking in the formulation and treatment of the problem.

Dr. Gerald Hover, Chief Clinical Psychologist at the Saginaw Veterans Administration Hospital, was instrumental in providing the writer with time, materials, and "moral support" when these were most needed. To Mr. Richard Behan who was of valued assistance in the statistical treatment of the data, the writer feels particularly indebted. Further appreciation is also due the members of the Michigan State College Department of Psychology who permitted the writer to use members of their classes as experimental subjects, and to the staff of the Veterans Administration Hospitals at Ft. Custer, Michigan, and Saginaw, Michigan, who cooperated in the use of hospital patients as subjects. Above all, however, the writer wishes to make mention of his gratitude

to those patient individuals, the experimental subjects, without whose services this project and the many others like it would have been impossible.

TABLE OF CONTENTS

LIST OF TABLES

APPENDICES

INTRODUCTION

General Characteristics of Studies in the Problem Area

Since the time of Wundt, considerable attention has been focused upon the experimental exploration of perceptual processes. More recently, those aspects of perception which have seemed to demonstrate the effects of personal motivation have been examined with increasing concern. Within the last decade especially, the interest in experimentally investigating motivational determinants of perception has risen to the point of stimulating many ingenious studies, and these in turn have inspired a number of provocative hypotheses.

One avenue of investigation has employed stimulus material under a variety of circumstances in which numbers of relatively plausible interpretations could be given to the stimulus. The particular interpretations made by the experimental subjects have then been used as bases for making inferences as to the motivational background of the responses selected. The use of ambiguous stimulus material to provoke responses from which it is possible to infer antecedent motivational processes in the observer is hardly novel; it occupies a central position among the techniques familiar to the clinical psychologist today. A primary justification

for the use of such "projective" personality tests, as are the Rorschach and the Thematic Apperception Test, is that they frequently reflect motivational factors - often factors of which the test subject is unaware. In effect, such tests capitalize upon certain overt responses suggestive of other responses occurring below the limen of consciousness.

In an article in which he reviewed experimental studies of unconscious behavior in relation to perception, Miller (16) listed investigations carried on as early as 1863. These and somewhat later studies covered areas of subliminal perception, subliminal learning and conditioned response, reportable effects of subliminal neural stimulation, and related topics. In Miller's own study, in which he presented his subjects with geometric figures on a transparent mirror, the subjects were able to discriminate between the various figures at illumination intensities below the level at which the bare presence of the figures could be detected. His results have been rather generally accepted as evidence supporting the existence of discriminatory responses in which the subject was reacting to subliminal cues. This interpretation has been criticized by Lazarus and McCleary (8) on the grounds that: (a) the criterion of the subliminal perception was the correctness of the subject's verbal statement (and thus apparently was not as strictly "subliminal" as would have been an indication of response over which the

subject had no degree of voluntary control), and (b) the determination of each subject's threshold was a matter of statistical treatment combined with subjective judgment.

More recent studies (2, 6, 9, 10, 13, 24, 25, and 28) have made use of a wide variety of ambiguous stimuli, including patches of light to be equated with the size of discs bearing various purportedly affectively stimulating and neutral symbols, chromatic and achromatic figures, and dimly illuminated or indefinitely outlined pictures, drawings, and words. The tachistoscopic exposure of words or symbols for very short periods of time is a similar approach, designed like the others to force the observer to depend to a greater extent upon his own resources by rendering less obvious the perceptual cues to which he is ordinarily accustomed. The assumption is that under such ambiguous circumstances, whatever hypothetical interpretation of the stimulus is aroused will be influenced more forcefully by motivational factors and less effectively by the nature of the stimulus than would be the case with optimal methods of presentation.

Recent Studies

Explanatory Concepts for Motivational Factors in Perception

In the attempt to connect various affective states or motivational needs in the individual with response phenomena which have been observed in subjects in such states or with such needs, a number of explanatory principles have been invoked. Certain of these have been attacked on the grounds that they were derived as explanations of the results of experimental investigations, rather than having been hypothesized prior to the research and confirmed thereafter. The nature of their derivation, however, should not be considered as the important criterion of their worth. The "ad hoc" hypothesis is to be avoided, not because it arises through inductive thinking, but because it explains nothing besides the phenomenon which it was originally invoked to explain. Thus, if suggested explanatory concepts can be demonstrated to have predictive value, their contributions cannot legitimately be disdained.

In the studies to be summarized in the following pages, evidence for and against certain of these hypothesized intervening processes will be presented. Proponents of these concepts report results of experimentation which suggest that motivational determinants may play a significant role in the perception of visual stimuli. Critics present data

which they interpret as relegating personal motivation to a negligible or minimal role in perceptual processes. Certain investigators have proposed that a more profitable point of view would be to examine the phenomena in light of various antecedent conditions, attempting to define interactions of a number of variables which might be presumed to account for experimental results otherwise difficult to reconcile. This is the point of view basic to the treatment of the problem presented here.

Studies Supporting the Importance of Motivational Determinants in Perception

Personal values as perceptual determinants. In 1948, Postman, Bruner and McGinnies (19) presented the results of an experiment in which they administered the Allport-Vernon Scale of Values to a group of college students. This was followed by the tachistoscopic presentation of thirty-six words selected to represent the six value categories devised by Spranger and utilized by the authors of the scale. Each of the words was exposed three times at .01 seconds, repeated three times again at .02 seconds, and continued in this fashion with three exposures at each timer setting, the settings being increased in steps of .01 seconds until the word was correctly identified. This minimum length of exposure required for the correct identification of the stimulus word has been termed the "duration threshold".

Records were kept of the exposure duration at which each word was recognized and also of any "pre-recognition hypotheses" or guesses made prior to correct identification of the word.

Pre-recognition hypotheses were categorized into the following groups:

(1) Covaluant responses - those considered to represent clearly the same value as the stimulus words.

(2) Contravaluant responses - antonyms to the stimulus words or responses serving to derogate them.

(3) Structural responses - incorrect hypotheses prompted by structural similarities to the stimulus word.

(4) Nonsense responses.

(5) Unrelated responses - any responses not classifiable into one of the above categories.

It was found that there was a significant negative relationship between value orientation in terms of Allport-Vernon score rankings and the length of exposure needed for the recognition of the representative words for the values ranked; high value words were perceived more rapidly by the group of subjects as a whole than were low value words. Covaluant hypotheses were reported as occurring more frequently in response to high value words than to low, while both contravaluant and nonsense hypotheses appeared more often in response to low value words.

This experiment has been criticized (17) on the grounds that the statistical treatment of the data was inadequate, and differences in familiarity with certain classes of the words may adequately and more parsimoniously explain the results.

An investigation of the postulated process of "perceptual defense." Some time later, the report of a related experiment by McGinnies (14) was released. In this study, the investigator employed the following eleven ostensibly neutral and seven emotionally toned (underlined) words: apple, dance, raped, child, belly, glass, river, whore, sleep, kotex, broom, stove, penis, music, trade, filth, clear, and bitch. The emotionally toned words were apparently selected on the a priori assumption that they carried their affective connotation because of their association with subjects which were in some degree "socially taboo" and thus apt to be emotionally "loaded" for individuals educated in our culture.

McGinnies' experimental subjects were eight male and eight female college undergraduates in elementary psychology. The stimulus words were exposed tachistoscopically, once at .01 seconds, once at .02 seconds, and so on until the duration threshold had been determined. A galvanic skin reflex apparatus was used concurrently with the tachistoscopic presentation. McGinnies found a highly significant relationship between the GSR and word meaning during pre-recognition (and thus nominally subliminal level) exposures, with higher

GSR's for the emotionally toned words than for the neutral words. In addition, without exception, the observers displayed significantly higher duration thresholds for the emotionally toned words than for the neutral words, when mean thresholds for each class were compared.

McGinnies also attempted a content analysis, with a classification of the pre-recognition hypotheses into four groups:

(1) Structurally similar - resembling in structure the stimulus words.

(2) Structurally unlike - with no structural resemblance to the stimulus word.

(3) Nonsense - without dictionary meaning.

(4) Part - any letters not connected into a group.

He found that his subjects made proportionately more unlike and nonsense hypotheses to the critical words than they did to the neutral words, and fewer similar and part responses. The chi-square test of independence between type of hypothesis and type of stimulus words was significant below the .01 level of confidence.

McGinnies interpreted his experimental results as demonstrating the presence of "perceptual defense," a concept which had arisen earlier in connection with the experiment by Postman, Bruner and McGinnies previously cited (19). Perceptual defense refers to a raising of the duration threshold displayed generally by individuals when faced with

stimulus objects perceived as threatening. McGinnies has also called attention to occasional instances in his data of responses to taboo words at thresholds significantly below the mean thresholds for neutral words, and this phenomenon has been termed "vigilance." Both concepts are usually interpreted in terms of subliminal activity involving the autonomic response levels. In the case of the specific critical words used in his experiment, McGinnies attributed their threatening connotation to the effect of early emotional conditioning, with punishment by parents as the conditioning agent associated with the use of the socially taboo words.

As pointed out by McGinnies (14) this experiment has been criticized by Bruner on the grounds that the critical words used appear less frequently in print than do the others, and that therefore the increased threshold may be explained on the basis of a difference in familiarity with the words. In addition, the heightened GSR, it has been suggested, may have been an effect attributable primarily to the greater effort required to recognize the less familiar words. In a similar vein, Howes and Solomon (5) claimed the following:

(1) When the effects of differences in frequency of appearance are extracted from the data, the differences in threshold disappear.

(2) The experimental situation in which McGinnies used a female assistant would tend to promote inhibition

of overt report of the taboo words eliciting strong GSR's. This would constitute voluntary refusal to express verbally the nature of the immediate perceptual experience, rather than the effect of any genuine subliminal blocking processes.

A corollary possibility is that the heightened GSR's with the taboo words may have been due to conflict between the desire to cooperate with the experimenter's instructions and the urge to inhibit the verbal communication of the taboo word.

Further investigation of discrimination without awareness. An interim report on a study by McCleary and Lazarus (12) was provided by these investigators as further evidence that observers give discriminatory (GSR) responses when confronted with visual stimuli presented tachistoscopically at subthreshold durations. The implied perceptual process has been termed by them "subception." In their ingenious study, nonsense syllables were utilized as stimuli, with the ten syllables used with each subject divided into two groups of five each. The ten syllables thus paired were equated in terms of the number of times each word of each pair was used and the number of times it was recognized during a 100 response equation period of practice for each subject. One group of the experimental syllables used with each individual was conditioned to electric shock, using partial reinforcement,

to the point where consistent conditioned responses were obtained. During test administrations of the ten syllables, the subjects showed a reliable tendency to greater GSR at subthreshold durations in response to the syllables paired with shock. Of additional interest was the observation that when the accuracy of recognition for the two groups of syllables at each of the five exposure speeds used was equated for the frequency with which the syllables had been employed, the shock-conditioned syllables were recognized more accurately than the others, with four out of the five first subjects used. However, a later report (8) revealed that additional experimentation did not substantiate this trend - that in fact no significant difference was demonstrated between the accuracy of identification of shock as compared to non-shock paired syllables at various exposure speeds.

Generalization effects with perceptual defense. In a very recent experiment designed by the senior author as a follow-up of his study concerned with emotionality and perceptual defense (14), McGinnies and Sherman (15) have investigated further the hypothesis that perceptual defense represents a genuine interference or temporary repression of perceptual nature rather than merely suppression of verbal report. Eighteen ostensibly neutral words, each five letters in length and with approximately the same frequency rating

in terms of the Thorndike-Lorge list (27), and four of the socially taboo words used in the previous experiment were employed in this study. They are as follows: bitch, brand, cable, flush, frock, glide, grind, hound, legal, lucky, outer, penis, phone, quest, ranch, raped, rider, scent, spray, towel, weave, and whore. The subjects used were twenty naive male college undergraduates. The words were exposed tachistoscopically, with eight pairs of the words being shown to each subject. The first words presented in four of the pairs were neutral words, whereas in the remaining pairs, the initial word was one of the taboo words. In all of the pairs, the second word exposed was a neutral word. In each pair, the first word was shown for a period long enough to insure correct identification by all subjects. The second word of the pair was shown at much more rapid exposure settings.

The following procedure was observed. The first word was exposed at a duration of approximately two seconds, and immediately thereafter, the second word of the pair was exposed for .01 seconds. If the second word was not identified, the first word of the pair was again presented for the two-second period, and immediately followed by the second word of the pair, this time exposed for .02 seconds. This routine was continued, with the exposure period for the second word of the pair being increased in increments of .01 seconds to the point of recognition. Three pairs of the

neutral words were similarly presented before the test runs for practice purposes. The pairing was random, with a different order being used for each subject.

In general, the duration thresholds for the neutral words which followed the exposure of taboo words were greater than those obtained for neutral words preceded by other neutral words. The differences between the thresholds of these two groups of neutral words were significant at the .01 level of confidence. However, the proportion of hypotheses made (number of hypotheses made by the subject to each stimulus word, divided by the total number of exposures to the point of recognition) was approximately the same for both neutral and "critical" pairs. The results with respect to the duration thresholds were interpreted as supporting the assumption that the actual perceptual threshold is affected by embarrassment or anxiety producing stimuli, with the raised thresholds of the second words of each pair due to a generalization of perceptual defense effects rather than mere suppression of verbal report.

Differences of Opinion and Criticism

Not all of the experiments designed to test the postulated effect of motivational determinants on perception have lent as much support to the various hypotheses regarding their characteristics or effects as have those already mentioned, nor has there been an absence of pointed criticism.

On the basis of their own investigations, Klein, Schlesinger and Meister (6) believe that the discriminatory error found in their research which involved estimating and matching physical attributes of stimuli with value connotations, could not be unequivocally attributed to the influence of value aspects of the stimuli employed. They suggest the advisability of returning to individual variations, perceptual organization, and ego structures as areas of fundamental importance in the study of perception. Luchins (11) decries the emphasis on the hypothetical and deductive approaches to the problem, to the neglect of inductive observation and investigation. In addition, he stresses the importance of field conditions and the desirability of forsaking the tachistoscope in favor of techniques in closer conjunction to everyday life and the clinic.

Studies Casting Doubt on the Generality of the Results of the Foregoing Experiments

Word frequency vs. personal values in visual duration thresholds. Recently, Solomon and Howes (26) conducted what was, in effect, a replication of the earlier experiment done by Postman, Bruner, and McGinnies (19) using the Allport-Vernon Scale and tachistoscopic presentation of representative value symbols. They used words actually selected from the Allport-Vernon test, and compared with them the perception of synonyms or cognate words not differing greatly from them in

numbers of letters but considerably discrepant in terms of frequency - the matched cognates were all much less common words, in terms of Thorndike - Lorge (27) frequency ratings. The results bear attention.

Their data showed no indication of systematic variation of duration thresholds with value scores, but did reveal the definite influence of frequency. The difference between the mean thresholds for the frequent and infrequent words was appreciable, with the infrequent words exhibiting considerably greater variation in mean thresholds as a function of value ranks than did the frequent words. While the results do not bear out those obtained by Postman, Bruner, and McGinnies in significant degree, a trend seems in evidence pointing toward lower thresholds for words representing the high value ranks, but the authors point out that the evidence for the operation of emotional determinants in visual duration thresholds is inconclusive.

Perceptual selectivity with names of traits regarded as desirable or undesirable. In another study dealing with perceptual selectivity as influenced by value, Postman and Leytham (20) employed trait names as stimuli, presenting these tachistoscopically on a screen. Not only did their data provide no evidence for the faster recognition of traits rated by the subjects as desirable, but these names were in fact identified more slowly than others exposed. The

writers pointed out that the cases of heightened sensitivity to positively valued stimuli in psychological research are matched by reports of lowered thresholds for negative and threatening objects. In essence, however, they contended that it is the intensity rather than the nature of the consequences which is the important element in establishing the degree of motivational support for a perceptual hypothesis.

Task completion and its effect on perceptual sensitivity to stimuli symbolic of the task. Postman and Solomon (21) have endeavored to test the hypothesis that stimuli symbolic of completed and incompleted tasks will influence perception differently. They made no specific prediction as to the direction of the difference, assuming that this would be a function of the individual subject's attitude toward completion or incompletion of the task, in particular when his task performance represented success or failure to him. For their group of subjects as a whole, no significant differences were obtained between symbols of completed and of incompleted tasks, but the majority of individual subjects did deviate significantly in the direction of lowered thresholds for either success or failure words.

A relationship between personality dynamics and auditory perception. Lazarus, Eriksen, and Fonda (7) reported the results of a study in which they attempted to examine the

relationship between performance on a sentence completion task and the perception of sentences with sexual, aggressive, and neutral reference, partially masked by noise, on a wire recording. They found positive correlations of .46 to .74, significant beyond the .01 level when evaluated via Fisher's z, between the sentence completion test responses and the perceptual accuracy for sentences concerned with sex and hostility. Two basic types of reactions were observed:

(1) High accuracy in perception, accompanied by ready verbalization; and

(2) low accuracy with minimal verbalization and blocking.

Individuals were found to be consistent with both.

Summary

The recent intensification of interest in experimentally investigating the role and nature of motivational determinants of perception has stimulated ingenious experimentation and provocative hypotheses. Proponents of motivational determinants as important factors in perception have been led to the postulation of intervening processes such as perceptual defense and subception. These have been invoked in attempts to relate differences observed in the perceptual thresholds for high and low valued stimuli, socially taboo and neutral stimuli, shock-associated and non shock-associated stimuli, to differences in motivational connotations these

stimuli are assumed to have had for the subjects. Other investigators have found no significant differences in perceptual thresholds which they could relate to differences in value, degree of task completion, and similar motivational variables symbolized by the stimuli. Instead, their data have suggested variations in response availability on the basis of prior familiarity with the stimuli as being one significant factor in the differences of thresholds which have been observed. Uncontrolled variations in experimental field conditions have also been suggested as responsible for results obtained. Thus, the nature and role of motivational determinants of visual perception, and the existence and character of the intervening variables postulated to account for the observed phenomena remain controversial.

In short, the investigator in the field of the relationship of visual duration thresholds to motivational factors is confronted with a number of alternative conceptions:

(1) Duration thresholds of motivationally "loaded" stimuli are different from those of neutral stimuli.

(2) Duration thresholds of motivationally loaded stimuli are not different from those of neutral stimuli.

(3) When motivationally loaded stimuli have been observed to have different thresholds from those observed with neutral stimuli, the difference is

not a function of the motivational element, but of uncontrolled experimental field conditions which have biased the results in the predicted direction.

(4) Duration thresholds of motivationally loaded stimuli are different from those of neutral stimuli, but only when the motivational element acts in conjunction with certain other variables. It is the interaction of two or more variables, each accentuating the effects of the others, which is responsible for the differences, which do not occur prominently when the motivational element acts singly. The other variables thus limit the range under which the differential thresholds are exhibited and define more specifically the character of the motivational stimuli which produce this effect.

The last alternative provides a general frame of reference for the problem to be presented and for the mode of approach.

THE PROBLEM

The writer's main interest in investigating the phenomenon of raised duration thresholds with certain types of stimuli involved the conviction that the topic could be explored fruitfully in somewhat closer association to actual life situations than was the case with the experiments previously described. The question was raised, would not this phenomenon be expected to occur most forcefully and definitely with symbols representing particularly salient experiences for the individual, if it occurred at all? Further, if one may accept the conditioning of nonsense syllables with electric shock as a paradigm of associative learning with unpleasant or fear-provoking life situations, could not one also expect an individual to react unconsciously or subliminaly to verbal symbols associated with actual past experiences under situations of this nature? Would not one, in fact, expect perceptual defense phenomena to be at least as apparent with symbols particularly relevant to the individual in terms of his own learning experiences and personal associations as with those chosen because of the a priori assumption of their culturally stereotyped characters as symbols of the socially taboo?

To demonstrate the operation of the process of perceptual defense, one must show that the duration thresholds for the critical stimuli are higher than those observed for otherwise equated stimuli of neutral character. The nature of the motivation which provides the stimulus employed with its response arousing power must be specified, and its existence determined independently of the principal experimental procedure contemplated. Within this framework, a twofold problem was suggested for investigation:

(1) That of testing the primary hypothesis that perceptual defense would occur with words of idiosyncratically unpleasant significance for the individual involved. This involved (a) exploring the feasibility of using the tachistoscopic technique with idiosyncratically unpleasant and neutral words, equated in terms of frequency of appearance in general reading matter and in number of letters, to investigate perceptual processes; and (b) developing a procedure to determine words which could reasonably well be assumed to have definitely unpleasant significance for the individual in terms of whatever they represented to him personally as a function of their associations for him.

(2) That of investigating the secondary hypothesis that pre-recognition responses made to the idiosyncratically unpleasant words would differ from those made to neutral words equated with them in terms of frequency and number of letters, when the difference is examined in regard to the frequencies of structurally similar, structurally unlike, nonsense and part responses.

The primary hypothesis, of perceptual defense, was utilized as the focal point of the experimental design. Specifically it had been hypothesized that for words of idiosyncratically unpleasant connotation to the group members, the duration thresholds would be significantly higher than those observed for words of ostensibly neutral emotional reference, matched with the unpleasant words in terms of number of letters and Thorndike-Lorge frequency ratings. "Duration threshold", for reasons to be explained, is used in this study to apply to the number of exposures up to and including the point of word identification, rather than to the timer setting employed at the point of recognition.

The secondary hypothesis was that there would be significant differences between the frequencies of structurally similar, structurally unlike, nonsense, and part responses elicited by the unpleasant words and those elicited by the neutral words. Since it was of interest to deal with

interactions of a number of other variables which could conceivably act to promote or inhibit the phenomena observed by Bruner, Postman, McGinnies, and their fellow-workers (14, 15, 19 and 20), five subsidiary hypotheses were developed in order that treatment of the data might be rendered more systematic and effective. In these hypotheses, the differences predicted were in terms of the relationship between the unpleasant words and neutral or "control" words equated in number of letters and frequency.

The following subsidiary hypotheses were proposed:

(1) That words selected as strongly disliked, on the basis of composite weighted scores in terms of peer group ratings, would have greater duration thresholds than those selected on the basis of the same criteria as comparatively mildly disliked.

(2) That there would be significant differences between the duration thresholds for unpleasant words selected as strongly socially taboo on the basis of prediction ratings by clinical psychologists, and those for unpleasant words selected as mildly or not taboo, the thresholds being greater with the former.

(3) That unpleasant words rated as least-liked by individual subjects would have greater duration thresholds than would unpleasant words rated as

next least-liked.

(4) That there would be significant differences between the duration thresholds for uncommon unpleasant words and those for common unpleasant words, with the former being less rapidly perceived. On the basis of this hypothesis, the differences in threshold between uncommon unpleasant words and their matched control words were predicted to be greater than those between common unpleasant words and their controls.

(5) That there would be significant differences between the duration thresholds for long unpleasant words and those for short unpleasant words, the thresholds being higher with the former. On the basis of this hypothesis, it was predicted that the differences in thresholds between unpleasant and control words would be greater with the long than with the short words. The rationale for this and the preceding hypothesis was that with very short or common unpleasant words, there would be such rapid recognition that motivational determinants should be less effective than with longer or relatively uncommon words, assumed to be of somewhat more ambiguous appearance at very rapid exposures.

THE SELECTION OF STIMULUS WORDS

Methods Investigated

In order to investigate the role played by motivational determinants in the perception of tachistoscopically presented unpleasant words, the first step was to develop a method for finding suitable stimuli. As has been indicated, what were needed were words which could reasonably well be assumed to have definitely unpleasant significance for the individual subject, in terms of whatever they represented to him personally as a function of their associations for him. Further, it was necessary that these words be suitable for presentation within the framework of the experimental design - they had to be words for which it was possible to find a measure of frequency of appearance in the form of reading material (their recognition was to be a reading task), since it was essential that control words equated in terms of number of letters and frequency be selected for them. This consideration ruled out the use of any words considered profane or obscene which did not appear fairly frequently in print and had thus evaded the compilers of the Thorndike-Lorge word list (27), the criterion measure for frequency used throughout the study. The exclusion of such words did not appear

to constitute any appreciable handicap in finding suitable stimuli, since at no time was the assumption made that the words used were necessarily the most forceful or most disagreeable words within the individual subject's vocabulary. It was desirable, however, to approximate this extreme as closely as seemed feasible for each prospective subject.

In the search for an adequate technique for finding suitable stimuli, words selected by five different methods were considered and examined. In the first technique, words were abstracted from clinical folders of delusional neuropsychiatric patients as symbolizing stabilized delusional systems of unpleasant nature. In the second method, words were selected on the basis of the subject's responses to specific items in the Minnesota Multiphasic Personality Inventory[1] - responses suggesting fearful or avoidant tendencies toward environmental stimuli or stimulus situations, thus assumed to be unpleasant to him. Since neither of these two techniques was finally considered satisfactory for the purpose under consideration, they were discarded as far as the experiment was concerned, and no further consideration will be accorded them here.

---

[1]S. R. Hathaway and J. C. McKinley, Manual for the Minnesota Multiphasic Personality Inventory, 1951 revision, The Psychological Corporation, New York, 31 pp.

The three other techniques which were investigated were the following:

(1) Words were chosen by the writer on the basis of his inferences made to a subject's responses to a specially developed questionnaire (Appendix 1).

(2) Words were selected by the writer on the basis of his inferences made to a subject's responses to a sentence completion test (Appendix 2).

(3) Words were rated by the subject as those least liked of a list of words previously indicated as definitely unpleasant by a peer group (Appendeces 4 and 6).

Each of these techniques was subsequently employed and will be considered in greater detail as follows.

The questionnaire. The need for an adequate method of identifying words of idiosyncratically unpleasant significance suitable for use as tachistoscopic stimuli still remained, and a number of other techniques were tentatively considered. The possibility of developing a questionnaire specifically designed to explore situations and experiences which had been associated with personal discomfort or trauma of one kind or another was attractive, since it seemed to promise a more direct and economical solution to the problem than had either of the techniques already examined.

A questionnaire (Appendix I) of nineteen items was devised. The questionnaire was introduced with the following instructions mimeographed at the top:

> In requesting you to fill out this questionnaire as carefully and honestly as possible, we are asking your cooperation in gathering material to be used later in a psychological experiment. We wish to study certain things which most people like and others which they dislike, and we hope to learn more about how they react to these feelings. Take as much time as you feel is necessary. All answers will, of course, be held strictly confidential.

The questionnaire included items regarding objects and experiences of both pleasant and disagreeable nature, although only the latter were of real interest within the framework of the study. The questions covered such elements as lines of work liked and disliked, animals liked or found disagreeable or repulsive, and various personal experiences, actual or potential, which the subject might consider pleasing, irritating, embarrassing, frightening, etc.

The sentence completion test. Concurrently, the writer had been involved in another research project employing a battery of psychological tests, among which was the Adams Sentence Completion Test (Appendix 2). On the basis of the inspection of test responses of a number of patients, it was decided to try out this technique for the purposes of the study.

The sentence completion test was considered in this connection, because it seemed to have the possibility of covering a wide range of problems to which the patient was almost forced to respond in his own terms, if he responded at all. In this respect, it resembled the Rorschach and T. A. T., yet the responses obtained were often sufficiently limited in scope to suggest specific words as representative symbols for them. The particular test adopted for trial is one of sixty-one items, constructed with a simple vocabulary so that it is easily comprehended while providing a minimum of information for the subject to use as a frame of reference for his responses. The items themselves were considered of suitable character for the purpose, and the comparatively limited number of non-relevant items was an additional desirable feature.

The peer group word list. The need for some method to include more words likely to be idiosyncratically unpleasant to a major proportion of the experimental sample still remained, however. It was felt that the coverage of unpleasant objects, words, and situations could be profitably enlarged, and this provoked the development of the fifth technique examined - the "peer group word list."

Since it was obvious that any of the unpleasant words used to test the hypothesis of perceptual defense and related phenomena in the experiment would be limited to the confines

of the Thorndike-Lorge list (27), the scheme of using this list as the basic source of stimulus words seemed not impractical. Accordingly, the word list was examined carefully, and 955 words were selected by the writer as being in his opinion conceivably unpleasant to a sizable number of individuals. These words were roughly divided into four groups (240, 240, 239, and 236 words respectively) on a random basis, and each subgroup was then mimeographed on a separate sheet (Appendix 3). To each sheet was attached a small slip of paper bearing the following instructions:

> When people talk to one another, they do so with words which may mean pleasant or unpleasant things. When a person talks to a doctor or a psychologist, he talks about his problems and uses words which often mean unpleasant things. We are interested in finding more about how people react to the unpleasant words which they use and hear. In order to do so, we need to find words which a great many people feel are unpleasant. Please read through this list and put a check in front of any word which you find you definitely dislike. The word may be unpleasant because of the way it looks, the way it sounds, or what it represents. Do not mark any words which you do not know or fully understand. Your opinions will be held confidential and your name is not required.

The lists of words were then distributed to forty-nine patients of the Veterans Administration Hospital at Saginaw, Michigan, a general medical and surgical installation. Sixteen members of the hospital staff, nine of whom were attendants, were also employed as judges, with a few of these individuals making selections from all four of the sub-group lists, over a period of time. In all, eighty-four

of the lists were inspected and checked, with each subgroup being examined by at least twenty-one individuals. The frequency with which each of the 955 words had been checked was tabulated, and each word selected by five or more individuals (as definitely disliked) was used to compile a composite list of 96 words (Appendix 4). The composite list was mimeographed with the following instructions:

> Look over the following words and select the five words which you like least. Mark a 1 in front of the word you like least of all, a 2 in front of the word you like next least, a 3 in front of the word you like third least, and so on until you have indicated the five words you like least in the order of their unpleasantness for you. Make your choices from all the words listed below, those typed on the list as well as those mimeographed. Ignore any word which you do not recognize or which is unfamiliar to you.

Since no assumption was made that the words to be used in the tachistoscopic presentation were of any specific degree of unpleasantness (they were acceptable so long as they could be assumed to be "definitely unpleasant"), and since in any event, only five words were to be rated as least liked from a list of ninety-six words selected as disliked by members of a fairly uniform cultural group, the use of rigorous criteria for selection of the 65 judges did not seem warranted.

The peer group word list came to be the basic technique utilized in the study. The reference in the direction for rating typed words as well as those which were mimeographed

was developed in regard to those words which were added to the list on the basis of inferences made to responses to the questionnaire and the sentence completion test. These three techniques, the questionnaire, sentence completion test, and word list, were used with the first sixteen hospital subjects prepared for the tachistoscopic procedure. Inspection of the individual subjects' choices of the words rated numbers one and two on the augmented word list - the only unpleasant words used in the tachistoscope - revealed that out of the thirty-two choices, thirty had been drawn from the mimeographed section of the list, and only two had been elicited on the basis of questionnaire and test responses. Because of the rapid turn-over of hospitalized patients, it was often difficult to maintain contact with subjects long enough to complete the testing process, and the use of the questionnaire and sentence completion test necessitated an additional interview period to those required for the remainder of the procedure. Therefore, it was decided to use the questionnaire and sentence completion test only when time permitted, and to rely upon the ninety-six words of the peer group word list as the basic technique.

Summary

In the search for an adequate technique for finding stimulus words for the contemplated experiment, words selected by the following methods were considered and examined:

(1) Words abstracted from clinical folders of delusional neuropsychiatric patients as representative of stabilized delusional systems of unpleasant nature.

(2) Words selected on the basis of the subject's responses to specific items in the MMPI - responses suggesting fearful or avoidant tendencies toward environmental stimuli or stimulus situations, thus assumed to be unpleasant to him.

(3) Words chosen by the experimenter on the basis of his inferences made to a subject's responses to a specially developed questionnaire.

(4) Words selected by the experimenter on the basis of his inferences made to a subject's responses to a sentence completion test.

(5) Words rated by the subject as those least liked of a list of words previously indicated as definitely unpleasant by a peer group.

The last technique was adopted as the basic method employed in the study.

PREPARATIONS FOR THE EXPERIMENT

Special Problems

The primary purpose of the pre-experimental contacts with contemplated subjects was to determine the specific unpleasant words to be used as stimuli with them. The decision to use as the criterion of unpleasantness the subject's own verbal report raised certain difficulties which had to be circumvented in one way or another. A problem of major proportions, as has been the case with so many psychological experiments, was that concerned with the necessity of keeping the subject relatively naive as to the actual purpose of the experiment. Were he to realize that the words he selected as least liked would later be exposed to him tachistoscopically in an attempt to determine the effect of personal factors in his perception of the word, it appears highly likely that these words would gain particular salience for him, apart from that due to their unpleasant connotation. One would predict that under such circumstances, the mention of the tachistoscope alone would tend to reinstate associations with the word which would tend to lower the duration threshold for it significantly; the subject would, in effect, be looking for the word, and ordinarily

minimal cues could have pronounced and unambiguous meaning for him when the word was exposed.

A second problem involved no association with the tachistoscope itself. It was simply concerned with the point that under the limitations of time during which each subject might be available, especially within the hospital situation, learning effects generated in the rating of the words might carry over to the occasion of the tachistoscopic administration, despite the fact that the subject might have no specific knowledge of the experimental sequence nor any set for anything at all to follow the initial contact. Such learning would be especially prone to occur when emotionally loaded material was concerned, and particularly when references might be made to such material in not one but a number of psychometric techniques (another consideration against the use of the questionnaire and sentence completion test, in addition to the word list).

Even the fact that all contacts with the experimenter, for purposes of defining the stimuli as well as for the actual test runs, might take place in the same room, with the same administrator, was conducive to the reinstatement of various cues which could conceivably affect the experimental results. While it was impossible to control all of these factors perfectly, certain measures were adopted in the attempt to cope with the major difficulties which were foreseen.

The Hospital Group

Initial contacts with the subjects. The first group of subjects employed were patients of the Saginaw Veterans Administration Hospital. All of these subjects were seen in the writer's office, which was to serve as the experimental laboratory as well. The subjects could be classified into two groups on the basis of the nature of the writer's initial contact with them. Eight of the thirty-two had been referred to the writer for psychological evaluation; six of these had been referred as skin disorder cases in connection with a research project concerned with dermatological difficulties with possible psychosomatic implications. The remaining twenty-four subjects were simply requested by ward personnel to report to the writer's office at scheduled times. Thus, in the case of the first eight, psychological rapport had been already established, and the rating of the words, the use of the questionnaire, and the presentation of the sentence completion test were accomplished against a background of other psychological examinations - the Rorschach, the Bender-Gestalt, and other standard psychometric tools.

With the other twenty-four subjects, the entire procedure would have been meaningless and perhaps even somewhat threatening, had not a plausible rationale been provided. Since the cooperation of these subjects was important, even

after the actual testing was concluded, the maintenance of good relations with them was given constant attention; widespread release of the nature of the actual procedure would have rendered the results obtained with other, then non-naive, subjects suspect and probably unreliable. For these reasons, each non-referral subject was acquainted in a few brief words with the skin disorder research project concurrently in progress, and then asked whether he would volunteer as a control subject. It was explained to him that his contribution would involve taking a few short psychological tests, extending at the most over a week and a half and necessitating three or possibly four meetings with the psychologist. In no cases, was there refusal; in many cases, the subjects voiced the conviction that this would help pass the time, which apparently hung rather heavily on their hands.

Selecting the stimulus words. As soon as the cooperation of the subject had been enlisted, he was given a copy of the peer group word list to rate, or when it was considered that circumstances and time would permit, he was administered the Adams Sentence Completion Test and the specially devised and mimeographed questionnaire (this was accomplished with sixteen of the thirty-two subjects). In respect to the word rating, if the subject expressed his difficulty in establishing a frame of reference for

disliking the words, it was suggested to him that the words themselves were merely black marks on a white background, and that whatever the words signified to him was probably a function of his own acquaintance with them.

If the subject had filled out the sentence completion test and questionnaire during his first meeting with the psychologist, he was administered the word list, augmented by whatever additional words his responses to the first two instruments seemed to imply to the experimenter. The additional words, incidentally, were typed at the bottom of the mimeographed list, using the same typewriter and spacing employed in cutting the mimeograph stencil and thus rendering the difference in appearance between the mimeographed and typed words relatively small.

Selecting the control words. As soon after the words on the list had been rated as was feasible, those words rated number one and two, the least liked and next least liked, were recorded, and control words for each were selected from the Thorndike-Lorge list (27). In each case, the control words selected were of the same number of letters as the unpleasant word with which they were being matched. Each control word was one which appeared to have no particular possible emotional connotation of any consequence, in the writer's opinion. Because of the recognized fallibility of that criterion, and because of the possibility of such

words possessing rare but significant idiosyncratic affective loading for the subject, four control words were selected for each unpleasant word. Nearly all control words were matched exactly with the critical word in terms of its numerical Thorndike-Lorge rating. In the few cases when it appeared impossible to match four control words with the unpleasant word while observing all three criteria (affective neutrality and equality in letters and frequency rating), the last criterion was relaxed to the extent where a control word with the next greater or smaller numerical frequency rating was acceptible. In such case, however, this deviation was balanced by a deviation in the opposing direction in another of the control words. Words which met all of the criteria but appeared to have markedly esoteric meanings were avoided.

The writer is well aware of the fact that structural differences in the matched words were not being controlled. Unfortunately, as Potter (23) and others have pointed out, while differences distinguished between one word and another are easily described, little is understood about the processes of discrimination whereby the distinctions are drawn. One gesture in this direction was made, however. Whenever it was possible to do so, while observing the other three criteria for selection of control words, unpleasant words with two identical letters juxtaposed were matched with control words with the same arrangement. In this way, for the word, "nigger," the control word, "ballet," was employed as one of four.

Controls for possible learning effects. In an attempt to equalize for learning effects conceivably occurring in regard to the two unpleasant words when they were rated, the eight control words were shortly thereafter presented to the subject, on the day following the rating whenever this was feasible. The control words were also administered in the context of a rating procedure, in which the subject was required to indicate the eight words he liked least from a list of thirty-one (Appendix 5). Twenty-three of the words had been selected under the assumption that they would have a relatively pleasing connotation to most people - they included such words as: reward, considerate, vacation, attractive, delicious, amusing, etc. The other eight words of the list were, of course, the control words. More often than not, as had been the aim behind the design of the list, the control words were the majority of words checked as least liked.

Following the administration of this list, at least six days and in most cases more were allowed to elapse. At the end of this period, the subject was used to test out the main hypotheses of the study, with the tachistoscope.

The College Group

Use of a college group. In accordance with the experimental design originally set up for the study, the use of two similar groups of hospital subjects had been contemplated -

as cross validation rather than as experimental and control samples. It soon became apparent that possible influences on experimental results of learning and cue reinstatement arising from the use of the subject's verbal indication as the criterion of unpleasantness of prospective experimental stimuli, and from the procedures used to identify such stimuli, would be difficult to control with assurance. When the apparent superiority of the peer group word list as a technique for defining the idiosyncratically unpleasant words was demonstrated (thirty out of the first thirty-two words chosen were drawn from the standardized mimeographed word list rather than from the presumably personally relevant augmentations typed on it), the possibility of using a group technique in selecting these stimulus words received attention. While responses to the sentence completion test and the questionnaire often directly involved contemporaneous and highly personal material, such was not the case with the word list, in which the motivational determinants frequently seemed to function indirectly. Thus, while it was considered advisable to limit the use of the test and the questionnaire to the individual interview situation, no such restriction was felt necessitated by the rating procedure of the word list. This was the essential consideration which prompted the decision to use a college group as the second sample.

Certain other advantages, however, also seemed to accrue from the use of the non-hospital population. It would

provide some indication as to the generality of any results observed in the hospital sample and would also provide a more suitable basis for comparison of findings with those of other investigations using college subjects. In addition it made possible the economy of the administrator's time and energy inherent in the use of the group technique as compared to that involving individual testing. Moreover the fact that a college population was a relatively stable one in terms of availability for testing and retesting over a sizeable period of time appeared to be a distinct asset. At the same time, from the point of view of mechanics of test administration, it seemed possible in the college situation to devise procedures whereby the element of cue similarity between the circumstances of stimulus determination and those of the tachistoscopic presentation would be minimal. Thus, consideration was given to developing an experimental design using college subjects, in which the rating of the word list would appear to have no connection with the tachistoscopic procedures, and in which there would be a sufficient time interval between the two to provide reasonable assurance against the probability of learning factors significantly influencing experimental results.

Developing a college peer group word list. Because of probable differences in intelligence, education, socio-economic level, and general cultural attributes which could be

presumed to exist between the members of the hospital group and those of the college sample, the use of the same peer group word list for both populations seemed inadvisable. Accordingly, the following procedures were instituted.

The 955 words abstracted from the Thorndike-Lorge list and mimeographed in four portions (Appendix 3) were administered to students in six elementary psychology classes at Michigan State College. The directions clipped to the lists of words were identical to those used with the Saginaw patients, requesting the students to check any words which they definitely disliked for any or a number of reasons. In this manner, 53 judgments were obtained in regard to the words on each of three of the four lists, while 54 judgments were obtained on the fourth. On the basis of the frequency with which the various words had been indicated as definitely disliked, a composite list of the 196 words checked by twelve or more judges was compiled and mimeographed (Appendix 6). It was considered that the intellectual level and reading skills of the college group could be assumed to be above that of the hospital patients used previously, and that therefore, a longer list of words to be rated, covering as wide a range of unpleasant words as feasible, could be used without the danger of producing excessive fatigue or boredom. The words themselves were of sufficiently stimulating nature, it was assumed, to tend to reduce these effects to a level

considerably lower than that which might have been observed with nonsense syllables, for example.

The new rating lists bore the following directions:

> Please read these instructions carefully: Look over the following words and select the five words you like least. Mark the number, 1, in front of the one word you like least of all the words. Next, mark a 2 in front of the word you like next least. Now mark a 3 in front of the word you like next least, and so on until you have indicated the five words you like least, in order of their unpleasantness for you. You may find them unpleasant because of their appearance, their sound, what they represent or remind you of, or for any number of these or other reasons. When you have finished, you should have five, and only five words marked, each with a different number. An explanation for this procedure will be supplied you later in the course.

Selecting the stimulus word. The word list developed for the college students was then administered to two classes in elementary psychology which had not participated in the compilation. In the case of one of these classes, the word lists were re-rated after a period of five weeks. The per cent of agreement in terms of all five of the words rated as least liked was 45%, on the basis of ratings by thirty-nine students. The agreement in respect to words rated numbers one and two only, the words selected as prospective stimulus words, was 34%. Since the ratings were made from 196 words, all of which had been selected by peer groups as definitely unpleasant to at least a quarter of the judges, the agreement cannot be regarded in the same terms applied to common measures of test-retest reliability.

It is felt that the percentages of agreement obtained represent a significant degree of stability in the choices of unpleasant words, particularly since many of the shifts from the first judgment to the second were to synonyms or cognate words, not taken into account in the percentage calculations.

It was from the second class to whom the word lists had been administered that the college subjects for the tachistoscopic procedure were drawn. The process of recording the words rated with numbers one and two, and the subsequent selection of four neutral control words equated with the unpleasant words in number of letters and frequency rating was identical to that followed with the hospital subjects.

After a period of at least two and a half weeks had elapsed since the administration of the word list for rating, appointments were scheduled for individual meetings with prospective subjects for the announced purpose of a "vision experiment." The experimenter was identified by name before the meetings took place, and the experiment was conducted in a small room with which none of the subjects had had previous contact. Neither the names of the apparatus nor of the experimenter had been mentioned in connection with the rating of the words, nor was the experimenter present at any time during the presentation or collection of the lists, so the possibility of his being connected with the ratings was remote.

After the tachistoscopic runs had been completed with all subjects, a brief mimeographed brochure explaining the background and purpose of the experiment was passed out to each student who had acted as a judge in selection of the words compiled into the word list or who had actually rated the completed word list. A short four item questionnaire was attached to the back of the brochures given to members of the class from which participants in the tachistoscopic tests had been drawn (Appendix 7). This questionnaire consisted of the following items:

(1) Did you know that vision experiment was connected with the ratings, before you saw any of the words in the tachistoscope? _____ If not,
(2) Did you have any idea that the two might be connected before you saw the words in the tachistoscope? _____ If not,
(3) Did you realize that the two were connected after seeing the first stimulus (unpleasant) word? _____ If not,
(4) Did you realize that the two were connected after seeing both stimulus words? _____

Comments:

Completed questionnaires were received from thirty of the thirty-three students used as subjects. Twenty-six of the subjects answered all of the questions in the negative. One student answered "yes" to question (3), while another stated that she felt some inclination toward the conclusion that the rating and tachistoscope run might have some connection, after identifying the first unpleasant word. Three students, including the individual who had had the

inclination, made the connection after seeing the second unpleasant word.

Summary

The decision to use the subject's own verbal report as the criterion of unpleasantness of the critical test stimuli raised the necessity of pre-experimental contacts in each case. With subjects from the hospital group, the following procedures were observed:

(1) Each patient was interviewed and asked whether he were willing to volunteer for a research project which would involve his taking a few short psychological tests and necessitating three or possibly four meetings with the psychologist.

(2) The subject was then asked to rate the peer group word list, or, when it was considered that circumstances and time would permit, he was administered the Adams Sentence Completion Test and the specially devised questionnaire.

(3) If the subject had filled out the sentence completion test and questionnaire during his first meeting with the psychologist, during the second meeting he was administered the word list augmented by additional words inferred by the writer on the basis of the subject's responses to the test and questionnaire given him previously.

(4) Four control words, matched in number of letters and frequency rating, were selected by the writer for each of the two unpleasant words rated respectively as least liked and next least liked from the word list by each subject. The control words were chosen from the Thorndike-Lorge list (27) and were all of ostensibly neutral affective connotation.

(5) The eight control words chosen for each subject were then administered to him in combination with twenty-three words assumed to be pleasing to most people. From this combined list of thirty-one words, the subject was asked to check the eight words he liked least. In this way he was exposed to the control words in a situation similar to that involved in his rating of the unpleasant words.

(6) Six days or more were allowed to elapse and the tachistoscopic presentation was employed.

With the college subjects, the following procedures were instituted:

(1) A peer group word list was developed for the college subjects in a manner similar to that employed with the patients.

(2) This word list was administered to two college elementary psychology classes, and after a period

of five weeks was re-administered to one of these groups for purposes of estimating the stability of the ratings.

(3) For each member of the class not re-administered the word list, the words rated least liked and next least liked were recorded, and four matched control words were selected for each of these rated words, meeting the same criteria used for control words employed with the hospital subjects. The writer was not present at any of the class meetings during which the word list was developed or rated, nor had his name been employed in connection with any of these procedures.

(4) At least two and one half weeks later, appointments were made for individual meetings with the students for whose rated words, the control words had been selected. The appointments were made for what was described as a "vision test."

(5) The tachistoscopic presentations were instituted at the time of the appointments.

(6) After all the subjects had been run through the tachistoscopic procedures, a four item questionnaire designed to provide an estimate of the extent to which the word ratings and the tachistoscopic presentation had been associated by the students was administered to each subject.

METHODOLOGY

Subjects. Two groups of subjects were selected for the study. Since no inter-group comparisons were contemplated and since each individual was to serve as his own control, no procedures of matching individual to individual or group to group were instituted.

The first group of subjects was made up of thirty-two male patients at the Veterans Administration hospital at Saginaw, Michigan. This hospital is a general medical and surgical institution and the patients selected had been hospitalized for a variety of reasons, with diagnoses covering such categories as hernias, rheumatoid arithritis, polio sequelae, fractures, burns, skin conditions, etc. The only criteria observed in the selection of these subjects were the following:

(1) That they were to be hospitalized for a period of at least a week and a half after the time of the initial contact with the experimenter. This was necessitated by the design of the study, so set up to minimize as much as possible memory effects carrying over from the initial selection of the test words to the actual test presentation.

(2) That they were ambulatory or at least wheelchair

cases. The apparatus to be used was unsuitable for bed patients nor was it practical to employ as subjects patients who required assistance from already heavily burdened hospital personnel.

(3) That they gave no evidence of being psychotic or subject to intellectual impairment of the type found in senile or brain-damaged individuals. It was, of course, desirable that reports of what was perceived be as reliable as possible.

(4) That they were not afflicted with an obviously disabling visual handicap. No attempt was made to control for differences in visual acuity and similar physiological factors, since each individual was to serve as his own control and since it was assumed that save for random variations, such inter-subject differences would remain fairly constant for all test stimuli.

(5) That they could be assumed to be reasonably familiar with the words used as test stimuli. This was determined at the time of the first exposure of the control words to each individual by simply asking him if any of the words were new or unfamiliar to him[1]. Each subject was specifically

[1]None of the words were indicated as being unfamiliar.

> instructed to ignore any words which were strange or unfamiliar to him in his unpleasant word ratings, and the control words were selected to meet the criterion of familiarity insofar as the experimenter could predict and choose from the words available which met the other criteria of length and frequency.

The ages of these patients ranged from twenty-one to sixty-four, with the majority of these below thirty-five. No assumptions were made in regard to differences in rate or character of perception as a function of age, intelligence, socio-economic status or similar variables, and no attempt to control for these was made.

The second group of subjects consisted of thirty-three college students from one elementary psychology class. Of these, seventeen were male and sixteen were female. The assumption was made that all of these subjects would be familiar with the comparatively common words used as test stimuli.[1] The selection of the college students as subjects, as has been pointed out, was undertaken in the attempt to avoid the influence of learning and memory factors in the hospital experimental situation which, because of practical

---

[1]With the possible exception of two words, this proved to be the case.

considerations, it seemed difficult to control.[2]

Equipment. The instrument employed for presentation of the stimulus words in this study was a Gerbrands Mirror Tachistoscope similar to that used by Postman and Bruner (18, 19) and by McGinnies (14, 15, 19) in their tachistoscopic experimentation. With this instrument, it was possible to expose the words at durations ranging from .01 seconds on upwards, with increments in exposure duration of even .01 seconds.

The words were typed on sheets of white bond paper, one word to a sheet, approximately centered on the page along the longer axis. Each word was typed on the same electric typewriter to minimize differences in legibility due to changes in pressure and stroke. The words were typed in capital letters, with one space intervening between adjacent letters.

Experimental Procedure. For the purpose of organizing the experimental design, the hypothesis of perceptual defense was utilized as a focal point. Specifically it had

---

[2]All of the hospital subjects were questioned immediately after the tachistoscopic presentation in regard to whether they had remembered any of the words which had been exposed, as coming from the word lists they had previously rated, and especially as to whether they felt that they had been able to anticipate words to any degree prior to their correct recognition. Three individuals answered the first question in the affirmative, and only one, the second.

been hypothesized that for a given group of experimental subjects, the duration thresholds for words of idiosyncratically unpleasant significance to the group members would be significantly greater than those for words of ostensibly neutral affective connotation, equated with the unpleasant words in terms of numbers of letters and Thorndike-Lorge frequency rating. It seemed desirable, however, to utilize three successive exposures of the stimuli at each duration in order to encourage maximal development of pre-recognition hypotheses as well as to minimize the effect of momentary vacillations in physiological efficiency and visual acuity, and of accommodation, fatigue phenomena, blinking, etc. With this in mind, the number of responses up to and including the correct identification of the stimulus word was adopted as the actual basis for comparison. To have used merely the timer setting of the duration threshold would have been to utilize a less sensitive measure, since with such a unit, it would not have been possible to differentiate between the recognition of words identified at the first exposure of a given duration, and those identified at the second or the third exposure with the same setting. To rephrase the hypothesis in terms of the number of exposures, then, it was postulated that for the group of subjects, the number of exposures required to identify unpleasant words would be significantly greater

than those needed for recognition of equated neutral words.

The experimental setting was held as uniform as possible for all the members of each group, and the actual procedure of the administration was essentially the same. All members of the hospital group were tested in the same office, and all the college students were seen in the same experimental room. Each individual was introduced to the tachistoscope with a few words of explanation in which the instrument was named and described as simply a machine with which pictures, designs, or words could be shown for very short periods of time. The subject was then informed that he would be shown seventeen words, and he was instructed to tell the experimenter whatever he saw or thought he saw. He was encouraged to guess whenever possible. If he was able to distinguish any word, he was asked to report this. In order to maximize any "projective" possibilities of the technique, if he could identify only certain letters but had any vague hunch or fancy that it might be an intelligible word, he was requested to follow this hunch and respond with a guess. He was urged to give the individual letters only when it was impossible for him to construct a word from what he saw. If he could identify nothing but a flicker, he was to say, "flash".

A slight variation was used with the college students who had been scheduled for a "vision experiment". They were introduced to the test situation as one which was concerned

with reading and vision, more specifically, with the identification of common words and the letters making them up, as a function of the shapes of the individual letters and their arrangements in words. Then, the same instructions for responding were used with them as with the other subjects.

The ten test words for both groups of subjects comprised two series of five words each. One word in each series was the word previously rated by the subject as least liked or next least liked of the prepared list of words selected as unpleasant by members of his peer group. In addition, with certain members of the hospital group, the lists had been augmented with words chosen on the basis of the subject's responses to the open-end questionnaire (Appendix 1) and the Adams Sentence Completion Test (Appendix 2). The remainder of each five-word series included four "neutral" or "control" words - words of no particular apparent emotional significance, equated with the unpleasant word of the series in terms of number of letters and frequency rating in the Thorndike-Lorge Word List (27).

Each series was presented with two of the control words first, followed by the unpleasant word and then the other two control words. The first series was preceded by the presentation of five "warm-up" words - common words of apparently neutral emotional connotation, made up of three,

four, five, six, and seven letters respectively, and shown in that order. These five words were used with all subjects: rug, loaf, short, summer, and opening. In addition, to establish a more similar spatial frame of reference for the test words to follow, as an aid in focusing, an additional warm-up word, a common neutral word which equaled in number of letters the words of the series to follow it, was shown immediately before each test series.

The following procedure was observed with each word, including those administered for warm-up purposes. The word was exposed three times at .01 seconds exposure duration, with as much time between showings as the subject required for his response. If the word had not been correctly identified, it was re-exposed three times at .02 seconds exposure setting, then at .03 seconds exposure, and so on until correct recognition occurred, with the exposure periods increasing in even units of .01 seconds. All responses were noted with the timer setting of the exposure which preceded each.

At the conclusion of the testing of the hospital patients, the general purpose of the experiment was explained to each subject. Whenever the scheduling of appointments permitted, each individual was asked for the reason for his choices of the two least-liked words. It was then requested that he not disclose any of the words used nor the specific nature

of the experiment.[1] When subsequent subjects were questioned as to whether they had any ideas of what the experiment was about, prior to their own initial introduction, none gave any evidence that the specific procedure or purpose of the study had "leaked out". Since the subjects were selected from four different hospital wards over a period of some months, and since most of them did not remain in the institution for more than two weeks, this confidence may have required less effort to maintain than might at first seem to be the case.

With the college students, no disclosure of the actual purpose of the experiment was considered desirable until the entire number of subjects had been run. Each subject was, however, requested not to divulge the specific words exposed to him, since, it was explained to him, prior knowledge would lead to fallacious findings or confusion on the part of the subject who came in expecting to see certain words. All subjects agreed to this request freely and apparently were most cooperative. In two cases, subjects stated their convictions that there was a connection between the rating lists administered previously and the tachistoscopic runs. When they faced the experimenter with

[1] In all cases the subjects willingly agreed to this request.

their conclusions, the connection was admitted and the reason for the deception explained. These subjects readily acquiesced to the request that they not reveal the connection until the experiment was over.

Summary

Two groups of subjects, thirty-two male patients of a Veterans Administration general hospital, and thirty-three college elementary psychology students, were used in the study. The instrument employed for presentation of the stimuli was a Gerbrands Mirror Tachistoscope with which it was possible to expose the typewritten stimulus words at durations ranging from .01 seconds on upwards, with increments in exposure duration of even .01 seconds. Each individual was shown seventeen words, of which seven were administered for practice purposes while ten were actual test words. Subjects were instructed to report whatever they saw or thought they saw, and they were encouraged to guess whenever possible. The words were shown three times at each exposure setting, beginning at .01 seconds, with the settings increased in steps of .01 seconds, until the words were correctly identified.

All responses were recorded, with the timer setting of the exposure which preceded each. The general purpose of the experiment was then explained to each of the hospital patients, after which he was requested not to divulge this

information nor specific particulars of the procedure. Whenever time permitted, each hospital subject was asked for the reasons for his choices of the two least liked words. He was also questioned as to whether he had picked up any significant information about the experiment prior to his own experiences.

With the college students, no routine disclosure of the purpose of the experiment was made, but each individual was asked not to divulge the specific words used with him.

TREATMENT OF THE DATA

Tests for Homogeneity and Transformation of the Data

While the design of the experiment was of a nature which would have permitted treatment by several statistical methods, the fact that the interaction of a number of variables was of particular interest suggested analysis of variance as an appropriate technique. This technique is commonly employed in testing the hypothesis that several independent samples of data have been drawn at random from a common normal population. In the case of this particular study, it was of importance to know whether the variation between the response of groups of subjects treated under diverse experimental conditions (in regard to characteristics of the stimulus material) was significantly greater than the uncontrolled variations in response between subjects treated alike. To rephrase the problem with an appropriate example, was the variation between the duration thresholds of subjects presented with unpleasant words and those of subjects presented with neutral words significantly more pronounced than the uncontrolled variations which would have been observed with the subjects when presented with words not differentiated in terms of the variable concerned? Could any significant differences in the means of the groups under the various

experimental conditions be legitimately attributed to the differences in the experimental conditions themselves?

Certain assumptions are involved in the use of analysis of variance:

(1) that the measures are independent;

(2) that they have been drawn at random;

(3) that they were taken from a common population, not differing significantly in distribution from normality.

There were no obvious reasons to suspect that the first two assumptions were not being met within each of the two samples used. It was, however, considered advisable to test the third assumption before attempting to estimate the normality of the measures when grouped according to the different experimental conditions.

Following the suggestion of Edwards (4), Bartlett's Test of Homogeneity of Variance was applied to determine whether the variances of the data from the college sample, from the hospital sample, and from the two combined and treated as a single sample were sufficiently homogeneous for the use of the contemplated method. Actually, there was no particular reason to presume that the two samples did come from the same population, since the experimental treatment was not strictly parallel nor was there any indication that the populations themselves were matched in respect to such variables as intelligence, education, socio-economic level,

and general cultural background, any one of which might logically have been predicted to effect certain of the variables under consideration. The chi-square obtained in the test of homogeneity of the duration threshold variance of the combined populations, grouped into the four subsamples of the thresholds for college unpleasant words, college control words, hospital unpleasant words, and hospital control words, was 233.2. This constituted statistical grounds at far beyond the .01 level of confidence that the combined samples were not homogeneous enough to be treated by analysis of variance, with the data combined in raw form. The size of the chi-square and the results of inspection of the data suggested that much of the variance might be a function of skewness in the distribution, with the means and variances tending to be correlated.

As has also been the case with much other data recorded in experiments involving timed measures, the distribution appeared to be positively skewed. That is, in comparison to the median thresholds for each series of five equated words (one unpleasant word with four matched controls), the more pronounced raw score deviations were predominantly in terms of raised thresholds. The timer settings of the tachistoscope ranged from .01 seconds to 1 second, and the correct identification of the words usually took place somewhere between exposure durations of .02 to .04 seconds. Yet

recognitions at durations beyond .07 seconds were not uncommon, and some were considerably greater, with a duration of .80 seconds in one extreme case. When the effect of physiological limitations are taken into account, as well as that resulting from the mechanical limitations of the timer, such a skewed distribution is readily understandable. Deviations, in terms of raw scores, treated either as normally or linearly distributed, would not accord sufficient significance to the deviations below the mean.

Since inspection of the data suggested that the principal source of variation lay in the thresholds observed for the college group, the test of homogeneity was applied to this group and the hospital group separately. In each case, the thresholds were classified into the two subsamples of responses to unpleasant and to control words which were then compared for homogeneity of variance. The chi-square obtained for the college group alone was 129.95, a very significant difference again, and one which would not permit the acceptance of the hypothesis of homogeneous variance. With the hospital group treated separately, however, the obtained chi-square of .11 was not significant, suggesting that the analysis could be applied to this data in its present form.

The distribution of the raw data was of a nature intimating that the means and standard deviations would tend

to be roughly proportional, thus violating a fundamental condition of the analysis of variance which depends upon the independence of the measures of inter-group variation from those of inter-individual variability. This suggested that a transformation be employed in the attempt to convert the original scale of measurement used to another with which the analysis of variance would more likely be valid, the variances being more homogeneous. The transformation which appeared appropriate was one of logarithmic character.[1] Accordingly, the logarithm for the number of exposures to the point of recognition for each word was recorded, and this figure, multiplied by 100 and rounded off to the nearest whole number, was the one adopted as the score unit to be used. This, it was felt, would stabilize the variance and in addition would decrease the skewness of the raw data distribution, adequately reflecting the comparative significance of both positive and negative deviations from the mean while providing a convenient unit for computation.

The logarithmic transformation was first applied to the data combined for the two groups, since it was still of interest to determine whether or not they could be considered as coming from similar populations in respect to the variables to be investigated. The chi-square, computed with the transformed data for the combined groups, divided into

---

[1] cf. Edwards (4) p. 199.

four subsamples as before (college unpleasant, college control, hospital unpleasant, and hospital control words), was 13.9, significant at between the .05 and .01 levels of confidence. While this was interpreted as still admitting the possibility of adequate homogeneity, under the more rigorous criterion it was considered advisable to treat the college and hospital groups as separate and different samples. Consequently, the test was then applied to these two groups after transformation. Chi-squares for the college and hospital samples under these circumstances were 1.63 and .04 respectively, statistically insignificant in both cases and interpreted as indicating adequate homogeneity for the analysis. The results of the tests of homogeneity are summarized in tabular form as follows.

TABLE I.

HOMOGENEITY OF VARIANCE
OF DURATION THRESHOLDS

Sample	Untransformed data		Logarithmic data	
	X^2	p	X^2	p
Combined hospital and college	233.2	.01	13.9	.05-.01
College alone	129.95	.01	1.63	N. Sig.
Hospital alone	.11	N. Sig.	.04	N. Sig.

Since after conversion to logarithms the variances were appreciably stabilized with both samples, it was decided to use the transformation with all the data despite the fact that none of the basic assumptions for the analysis would appear to have been violated, had the raw, untreated data been utilized for the hospital group alone. In addition, with all data similarly treated, cross comparisons would be more readily grasped, should they be desired. The same transformation was applied to the measures of frequency of the pre-recognition hypotheses.

Variables Investigated

With the data transformed logarithmically, it was now possible to set up designs for statistical treatment. It may be recalled that the primary hypothesis being investigated was that which dealt with perceptual defense - the raising of duration thresholds in the perception of unpleasant words. The secondary hypothesis was concerned with differences in the frequencies of structurally similar, structurally unlike, nonsense, and part responses with the pre-recognition hypotheses to unpleasant as compared to neutral words. Five subsidiary hypotheses had also been advanced as a framework for studying the interaction of the main variables with others considered of possible consequence in explaining the conflicting experimental results obtained

by various investigators in the field. The five subsidiary hypotheses concerned the following variables:

(1) The intensity with which the unpleasant word was disliked by members of a peer group.

(2) The degree to which an unpleasant word was socially taboo.

(3) The degree to which the unpleasant word was considered unpleasant by the individual subject.

(4) The relative familiarity with the word, control words as well as unpleasant.

(5) The relative length of the word, control words as well as unpleasant.

An additional variable was also considered to be of some possible interest in relation to the college sample - that of sex (the hospital group was entirely male).

It was in its interaction with each of the subsidiary variables that the primary variable, the unpleasantness of the critical word, was of maximum interest in respect to the experiment; these were the relationships therefore explored by the analysis of variance used. With the pre-recognition hypotheses, the interaction of the primary variable, word unpleasantness, was explored in its relation to the frequencies of the four categories of pre-recognition responses which had been recorded, and in relation to the intensity with which the unpleasant word was disliked by members of a

peer group. The analysis of variance of the pre-recognition hypotheses was designed after that of the duration thresholds had been computed, and was developed to investigate the effects of word unpleasantness for the individual and for the peer group.

Sub-categorizing the Variables

The primary and the secondary variables. In order to treat the variables in their interaction, it was necessary to sub-categorize each into two or more values or intensities by employing arbitrary cutting points. The division of the primary variable of word unpleasantness in terms of unpleasant and control or neutral words was explicit in the formulation of the basic hypothesis of the study, and the cutting point here was determined essentially by the subject's ratings of the unpleasant words and by the writer's selection of the controls. In regard to the secondary hypothesis concerning the frequencies of the various types of pre-recognition responses, the mere adoption of McGinnies' (14) fourfold system of classification with his published definitions of the criteria for each category was not found sufficiently discriminating in all cases. Consequently it was considered advisable to re-define the criteria for three of the groupings.

As was the case with McGinnies' original use of this categorization, the criterion of a nonsense response was simply that it could not be found in a dictionary (29). Under part responses were included any disconnected groups of letters or any responses which the subject clearly identified as incomplete or fractional representations of the stimulus as he perceived it. In order to differentiate between structurally similar and structurally unlike responses, it was considered expedient to establish a uniform criterion which could be applied to all the responses which did not fall into part or nonsense groupings.

A criterion was needed which would take into account the differential cue strength of letters as a function of their position in the word as well as of their structural characteristics. It was not at all rare to find a subject making a whole series of pre-recognition responses which had little in common with the stimulus word or with each other save for the first letter in each, yet in such a case the evidence for an element of structural similarity was too consistent to deny. On the other hand, simply to make the assumption that because two words have one letter in common they are genuinely similar in structure, especially when both words are long words, seems to leave more up to chance than would be warranted or desired in the classification contemplated. On these grounds, the following alternative arbitrary standards for characterizing words as structurally

similar to the stimulus word were adopted:

(1) The first letter or the last letter of the hypothesis and the stimulus word were common to both.

(2) In the case of hypotheses of three or fewer letters, one or more of the letters were common both to the hypothesis and the stimulus.

(3) With hypotheses of four or more letters, two or more consecutive letters of the hypothesis were found in the same order in the stimulus.

(4) With hypotheses of four or more letters, at least half of the letters of the hypothesis were found, in any order, in the stimulus.

Words found unpleasant by peer groups. The first of the subsidiary hypotheses was concerned with the intensity with which words were disliked by members of a peer group. Essentially the same procedure was followed with both the college and hospital samples in securing a rough estimate of the relative degree of unpleasantness with which each of the unpleasant words used as stimuli were regarded by the group members. The method of securing this estimate was as follows.

With the college group, the five words rated as most unpleasant by each of the 93 students administered the peer group word list were recorded in the numerical order of the

ratings (a rating of number 1 represented the word designated as most unpleasant or least liked, while a number 5 indicated the fifth least liked). With the hospital group, the same process was used with the ratings of the 48 individuals prepared for testing. All choices of the college group came from the peer group word list, and all but 22 out of 240 selections of the hospital subjects came from their word list (the 22 were chosen from words added to the list on the basis of responses to the questionnaire and the sentence completion test).

Each word was then given an arbitrary weighted score from 1 to 5 in inverse order to its numerical rating - words rated number 1 or least liked were given a weight of 5, while words rated number 5 were given a weight of 1. Total weighted scores for each of the words were then computed, and in each sample the unpleasant words used as stimuli were divided into two groups with the cutting score being approximately the median of the weighted score values in each distribution. Thus, in the college group, the unpleasant words having a composite weighted score of 18 or more were characterized as "strongly disliked," while those with a score of 17 or less were termed "mildly disliked," the assumption being that the frequency with which the words were chosen as the five least liked and the order in which they were selected for this dubious distinction were functions of their unpleasantness

for sample members as a group. With the smaller number of raters in the hospital group, scores of 13 and above were considered indications of relatively strong dislike, while scores of 12 and below were considered representative of a milder degree of disliking. The resulting frequencies in these and the related groupings which follow are summarized in Table II, page 81.

Words considered socially taboo. In two of the more important experiments in the area of this study, significant differences in response were attributed to the effect of motivational determinants. These experiments were concerned with critical words described as "socially taboo" (14, 15). Consequently it was considered important to explore the possible effects of this variable within the framework of the analysis; this consideration was the basis for the second subsidiary hypothesis. It thus seemed desirable to differentiate among the various unpleasant words used as stimuli, those which might arbitrarily be termed "strongly taboo" and those which could be compared as "mildly taboo." Once such a distinction was made, it appeared feasible to examine the responses to the two categories of stimuli for evidence of significant differences.

Rather than depend upon the judgment of a single individual in the proposed classification, the writer felt that judgments of a number of clinically trained individuals

should be pooled to provide a more reliable basis for the grouping. A rating scale for all of the words used as unpleasant stimuli was developed (Appendix 8), which bore the following introduction:

> It has been suggested that most individuals learn early in life that certain words carry with them a "socially taboo" connotation. When these words are used by a child, this use generally results in chastisement by parents, with the consequent establishment of a conditioned reaction to the verbal symbol. This pattern of conditioned emotional response may be considered to be one of anxiety or fear aroused by symbols having sexual, excretory, or otherwise unpleasant or "immoral" connotations. (McGinnies, Elliott, "Emotionality and Perceptual Defense," Psychological Review, 56, 1949, p. 244.)
>
> Please rate the following words in terms of the extent to which each word seems "socially taboo" by putting a check in the appropriate space after it.

The words were then listed in a vertical column at the left-hand edge of the page, and after each word, four spaces were provided for the rater's checked indication of his opinion in respect to the word. The categories of rating were: highly taboo, moderately taboo, slightly taboo, and not taboo.

This rating scale was administered to sixteen college faculty members and graduate students in clinical psychology, each of whom had had a minimum of a year of clinical experience. The raters were asked to predict how people in general would have reacted to the words, in terms of the definition of socially taboo used on the scale. Each rating of highly taboo was given an arbitrary weighted score of 3, while moderately taboo, slightly taboo and not taboo were

accorded scores of 2, 1, and 0 respectively. Composite weighted scores for each word were computed, and the approximate median score value of 13.5 was used as the cutting point, with all words scoring above it termed "strongly taboo" and those below termed "mildly taboo."

The subject's dislike for his own selections. Since the selection of the unpleasant stimulus words used with each subject was at least in part a function of his own choosing, it seemed desirable to investigate whether differences in the rating of these words by the subject would be reflected in a differential response to them as compared to their control words. That is, would differences, if any, in the strength of the effective response to a least liked word as compared with that to a next least liked word be reflected in the duration thresholds of the two words in relation to the thresholds of their controls? This was the basis for the third subsidiary hypothesis.

In order to provide some information in regard to this question, the duration thresholds for the unpleasant and control words were analyzed in interaction with a classification of series of words presented first and series presented second. Since in the college experimental situation there seemed little reason to anticipate significant learning effects carrying over from the rating task to that of the actual test situation, for matters of convenience the first

series shown to each individual included the unpleasant word selected as least liked. In the hospital group, where the probability of learning effects and cue reinstatement, especially after the recognition of the first unpleasant word, seemed greater, 15 of 32 subjects were shown the least liked word in the first series administered, while 15 other subjects saw the least liked word in the second series (2 of the subjects had not been able to discriminate in their ratings of the least liked and next least liked words, and thus their thresholds could not be employed in this particular context). With this split treatment used in the hospital sample, it was thus possible to avoid confounding the effect of the variable under consideration with the next one to be regarded.

Order of presentation. Although not specified in terms of a subsidiary hypothesis, the order of presentation of the stimuli was felt to be of sufficient interest to be concomitantly investigated with the others. Since six practice words had been presented before the first word in the initial test series, and since the results of inspection of the data divulged no consistent changes suggestive of improvement through learning, practice effects were not felt to be of great consequence in the test task. It was of more importance, however, to secure some indication as to whether or not the recognition of the first unpleasant word seemed to have

markedly influenced the identification of the second.[1] Since such effects, if consistently present, could markedly distort the data in half of the cases, some appraisal of possible differences attributable to this factor was thought prudent, especially in the hospital sample. With this in mind, the data was again classified in two subgroupings - words presented in the first series, and words presented in the second.

Familiarity with the word. The effect of stimulus familiarity has long been recognized in studies concerned with learning and discrimination, and it was hardly surprising that criticism of certain of the experiments previously cited here was focused on the lack or inadequacy of controls over this variable. The attempt having been made to exert as rigorous control of frequency in the selection of the stimulus material as seemed feasible, efforts were then turned to measuring effects attributable to controlled statistical manipulation of this factor, the consideration basic to the fourth subsidiary hypothesis.

---

[1]It was anticipated that the words most likely to be fixed and thus later to be recalled were the unpleasant words rather than the controls, because of the subject's own participation in their selection and possibly because of their stronger emotional connotation. This proved to be the case.

In order to do this conveniently within the framework of the experimental design, all of the words, unpleasant and control, were divided into two groups on the basis of the criterion of frequency used throughout - the Thorndike-Lorge ratings (27). These ratings are in numerical form and, for the first and more common approximately 20,000 words, they range from 49 to 1. Each number represents the number of times words with such ratings have been observed to occur per million words in a variety of printed matter. In addition to the numerical ratings, the symbols "A" and "AA" are also employed, representing frequencies of observation of from 50 to 99, and 100 or more respectively, per million words.

The Thorndike-Lorge frequency ratings for all of the unpleasant words (and consequently for their matched control words) were recorded and tallied for each sample separately. In this way it was possible by inspection to find convenient cutting scores (frequency ratings) between rough groupings which had thus occurred as a function of the words selected by the subjects. For the college subjects, words occurring six or more times per million, as reflected in their frequency ratings, were designated as "common" words, while those appearing with a frequency of five or fewer per million were termed "uncommon." For the hospital subjects, a frequency of seventeen times per million was adopted as the cutting score,

with words above that frequency considered as common, and words below as uncommon.

Word length. The fifth subsidiary hypothesis presented involved the variable of word length, including the assumption that duration thresholds for long words would be greater than for short words. While such an assumption required no astute prophetic capacities and in fact appeared quite logical, it nevertheless seemed appropriate to include the variable concerned as one of those to be examined alone and in interaction with the variable of primary interest.

In order to examine the effect of word length most conveniently and investigate whether it might contribute a differential tendency in response to motivational factors, it was desirable once again to split the samples into two sub-groups - in this case, long words and short words. Accordingly, the number of letters in each of the unpleasant words (and thus also the control words equated in this respect) was recorded and tallied with others of equal size. Each sample was arbitrarily divided into groups of approximately equal size - "long" words and "short" words. The same cutting line was used with both samples; words of 6 or more letters were designated as long, while those of 5 or fewer letters were termed short.

Sex differences. No specific hypotheses had been made in regard to the effect of sex differences on the verbally reported responses to the unpleasant words within the experimental situation. Many of the words used carried colloquial connotations of sexual character in addition to their dictionary definitions. Therefore, since approximately half of the college subjects were female, in light of the criticism made by Howes and Solomon (5) of McGinnies' procedure (14) using sexually loaded stimulus words with male subjects in the presence of a female assistant, it was of some interest to investigate the possibility that experimental results might have been influenced by the reactions of female subjects in the presence of a male experimenter. With this in mind, comparisons were made of the variances in the duration thresholds of the male and female subjects of the college group, in interaction with the primary variable of the study.

Grouping the pre-recognition hypotheses. With the pre-recognition hypotheses, the effects of the following variables were examined: categorization of the hypotheses, unpleasantness of the stimulus words prompting the hypotheses, and the intensity with which the unpleasant words were disliked by members of peer groups. In each case the pre-recognition responses were grouped on the basis of the stimulus words, and the criteria for these groupings were similar to

those employed in the treatment of the duration threshold data.

TABLE II FREQUENCIES IN THE ARBITRARY GROUPINGS OF UNPLEASANT STIMULUS WORDS

Categories	Hospital	College
Strongly disliked	33	45
Mildly disliked	27	21
Strongly taboo	22	26
Mildly taboo	38	40
Series presented first	30	33
Series presented second	30	33
Least liked words	30	33
Next least liked words	30	33
Common words	23	13
Uncommon words	41	53
Long words	32	29
Short words	32	37
Male		32
Female		34

The Analyses of Variance

With each classification of the various samples into sub-groupings now achieved, it was possible to begin the analyses proper. The following procedure was observed. Four separate designs were developed for each sample, with the operations essentially parallel for the two samples with the following exceptions:

(1) No comparison on the basis of sex differences could be made for the all-male hospital group.

(2) The variables of relative degree of dislike by the individual subject for the two unpleasant stimuli and that of order of presentation were treated together in the college sample but separately in the hospital group. (Only in the hospital sample did there appear reason to suspect that order of presentation might be a significant influence on the duration thresholds.)

(3) For the college group, the number of threshold measures employed throughout the analysis remained constant (an N of 33 subjects, permitting 330 threshold determinations of which 66 were for unpleasant words and the remainder for controls. With the hospital group, the thresholds of all of the 32 subjects were employed in three of the four designs, but it was necessary to omit those of two

of the subjects from that design involving the individual designations of least-liked and next least-liked unpleasant words. Since inspection of the data of the two individuals thus omitted suggested that neither deviated markedly from their peer groups in respect to the variables of chief interest, and especially in virtue of the relatively homogeneous nature of the hospital sample implied by the results of the tests for homogeneity, this omission seemed to provide little cause for concern.

The designs having been developed, the following formulae were applied:

Total sum of squares: $\Sigma X^2 - \frac{(\Sigma X)^2}{N}$

Sum of squares between groups:

$$\left[\frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \ldots\ldots \frac{(\Sigma X_n)^2}{n_r} \right] - \frac{(\Sigma X)^2}{N}$$

with n equivalent to the number of observations in each subgroup represented by the subscript numbers. Interaction sums of squares were obtained by subtracting the sums of squares already calculated for the variables concerned from the total between groups sum of squares of the groups involved in the interaction:

$$\left[\frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \ldots\ldots \frac{(\Sigma X_n)^2}{n_r} \right] - (A + B + \ldots\ldots r)$$

with the sums of squares of variables represented by the alphabetical designations.

The sum of squares within groups was obtained by consolidating all of the data from those designs involving the same numbers of observations from the same samples, subtracting the sums of squares between groups so obtained from the total sums of squares. F was computed by dividing the mean squares between groups by the mean squares within groups.

Sums of squares were computed for the following groupings studied in the duration thresholds data:

(1) Unpleasant vs. control words.

(2) The series in which the unpleasant words were categorized as strongly disliked by the peer group vs. the series with the unpleasant words designated as mildly disliked.

(3) The series with the unpleasant words classified as strongly taboo vs. those series with the unpleasant words termed mildly taboo.

(4) The series of words presented first vs. the series of words presented second.

(5) The series in which the unpleasant words had been rated as least liked by the subject vs. those in which the unpleasant word was rated next least liked.

(6) Common vs. uncommon words.

(7) Short vs. long words.

(8) Responses of male vs. those of female subjects.

Interactions were obtained for the first grouping with each of the others, since these were the relationships of primary interest.

The following groupings were used for computing sums of squares with pre-recognition hypotheses:

(1) The four categories of pre-recognition hypotheses: structurally similar, structurally unlike, nonsense and part responses.

(2) Hypotheses made to unpleasant vs. those made to control words.

(3) Hypotheses made to those series of words in which the unpleasant word was designated as strongly disliked vs. those in which the unpleasant word was classified as mildly disliked.

Interactions were obtained between the unpleasant vs. control word grouping and each of the other two.

RESULTS

Presentation of the results. The results to be presented here are also summarized in Tables III, IV, V, VI, and VII, on pages 98 to 103. In addition, for purposes of comparison, the results of the analyses for both the college and hospital samples are given in abbreviated form in Table VIII, page 104. Those results considered to be of particular significance for the study are included in greater detail in Tables IX, X, and XI, found on pages 106, 107 and 108. The means presented below and in the tables are means of the duration thresholds and frequencies of pre-recognition hypotheses after the application of the logarithmic transformation described earlier.

The Hospital Sample

Duration thresholds in terms of number of exposures to the point of word recognition. With one degree of freedom and with 300 threshold values (thresholds for each of ten test words exposed to 30 subjects) in one treatment of the data, shown in Table III, and with 320 threshold values in the other, presented in Table IV, unpleasant and control

words are indicated as not significantly different.[1] With means of 83.2 and 86.8 for the unpleasant and control words in the first treatment, and 83.6 and 88.0 in the second, F's of less than 1 were found in both instances.

The following data are summarized in Table III and were gathered from 300 threshold determinations. No significant differences were demonstrated between the series of words (one unpleasant and four control words matched for length and frequency) in which the unpleasant word had been classified as strongly disliked by the peer group and the series in which the unpleasant word had been considered mildly disliked. With means of 86.3 and 85.8 respectively, and with one degree of freedom, the F found was less than 1. The groupings here did not provide controls for the effects of possible differences in word length or frequency; this was however accomplished later through the calculation of the interaction with the unpleasant-control dichotomy.

In comparing the series with the unpleasant word designated as strongly taboo with the series with the unpleasant word designated mildly taboo, the thresholds for

---

[1]The two separate treatments indicated here were necessitated because two of the thirty-two hospital patients were unable to specify as least liked and next least liked the two words used as unpleasant stimuli. Consequently, it was necessary to omit their responses from part of the analysis. This resulted in one analysis with an N of 320 (thresholds for ten words with each of 32 individuals), and another with an N of 300 (ten words with each of 30 subjects).

the former were found longer. The means here were 92.2 and 82.5. With one degree of freedom, the F of 5.381 was significant at between the .05 and .01 levels of confidence. In this form again, the comparison does not take into account differences of word length or frequency.

This is also the case with the differences found between the series of words presented first compared with the series of words presented second, where the means were 90.7 and 81.4. In this instance, with the same number of degrees of freedom as the previous analysis, the obtained F of 5.312 is significant at between the .05 and .01 levels of confidence.

Again with one degree of freedom, the comparison between series in which the unpleasant word was selected as least liked by the subject and series in which the unpleasant word was selected as next least liked yields no indication of significant differences. With means of 84.6 and 87.5 respectively, the F was less than 1. Here, as well, since the unpleasant and control groups were not separated and compared, no controls over word length or frequency were being exerted.

In Table IV are summarized the data used in the comparison of common and uncommon words, and short and long words, drawn from 320 threshold values and making use of one degree of freedom in each case. The means for the common and uncommon words were 77.8 and 92.3 respectively, while the means

for short and long words were 78.5 and 95.7. Very significant differences were obtained for both groupings, with F's of 13.543 and 20.670 respectively. Unpleasant and control words were combined for both of these analyses.

Interactions. The following results are summarized in Table III, with the N of 300. Investigation of the interaction between word unpleasantness and the degree to which the unpleasant word of the series was indicated as disliked by the peer group revealed that this effect was not significant. With one degree of freedom, the F computed was 1.772. The interaction between word unpleasantness and the degree to which the unpleasant word of the series had been designated as socially taboo, with a similar number of degrees of freedom, was also not significant, with an F of less than 1. With one degree of freedom, no significant interactions were found between word unpleasantness and the order of presentation of the series, and between word unpleasantness and the degree to which the unpleasant word of the series had been rated as disliked by the individual subject. F's in both instances were less than 1.

Table IV reveals that with an N of 320 and one degree of freedom, interactions between word unpleasantness and frequency ("commoness"), and between word unpleasantness and length, were not significant. F's here were 2.101 and less than 1 respectively.

Pre-recognition hypotheses with the hospital patients. For the results of this analysis, 451 responses were recorded; the data are summarized in Table V. With three degrees of freedom, the F of 5.708 obtained between the frequencies in four categories of pre-recognition hypotheses was very significant. The means were as follows: structurally similar, 34.0; structurally unlike, 17.4; nonsense, 33.7; and part, 36.4. This grouping does not differentiate between the hypotheses made to unpleasant and control words. When these latter two groupings of pre-recognition responses are compared, with one degree of freedom, the obtained F is less than 1 and thus not significant. The means obtained were 31.4 and 33.5 respectively. The comparison of hypotheses to words of series in which the unpleasant word had been categorized as strongly disliked on the basis of peer choices, with those to words of series in which the unpleasant word had been classified as mildly disliked showed no significant differences. With means here of 31.2 and 35.6 respectively, the F, with one degree of freedom, was 2.576.

Interactions. In the investigation of interaction between word unpleasantness and the categorization of pre-recognition hypotheses into the four types of responses, very significant differences were revealed with controls for length and frequency differences here in effect. With three

degrees of freedom, the obtained F was 56.828. As reference to Table IX discloses, the significant differences between the unpleasant and the control group may be attributed to the structurally unlike group, in which more responses were recorded to the unpleasant words. The basis for the significance of the interaction itself appears to reside in the reversal of trend in the structurally unlike and nonsense groups as compared with that exhibited by the structurally similar and part groups. In this regard, while the structurally similar and part hypotheses tend to appear in greater number in response to control words than to unpleasant, structurally unlike and nonsense hypotheses tend to appear more profusely in response to unpleasant than to control words.

Returning to Table V, the interaction between stimulus word unpleasantness and the degree to which the unpleasant stimulus word of the series was indicated as disliked by a peer group was insignificant in respect to pre-recognition hypothesis frequencies. With one degree of freedom, the obtained F was 1.694.

The College Sample

Duration thresholds. The following results, presented in Table VI, were obtained from an N of 330 duration threshold values, with one degree of freedom applicable to the

interpretation of each F value obtained. When the duration thresholds for the unpleasant words were compared to those recorded for their matched control words, an F value of less than 1 and thus insignificant differences were disclosed. The means were 98.5 and 96.4 respectively.

Those series of five words (one unpleasant and four controls) in which the unpleasant words had been designated as strongly disliked on the basis of peer selection were not perceived differently in respect to threshold than were those series in which the unpleasant words had been designated as mildly disliked. The means obtained here were 95.0 and 100.7 respectively. Since this preliminary grouping did not distinguish between thresholds obtained for unpleasant and for control words, the results here, an F of 3.000 which was not significant, were of little particular consequence per se.

An F of 2.366, also not significant, was provided by the analysis of the series of words in which the unpleasant words had been classified as strongly taboo, as opposed to the series in which the unpleasant words had been classified as mildly taboo. The means obtained here were 99.7 and 94.9 respectively. This too was a computation preliminary to the interaction with the primary variable and was of minor pertinence by itself.

In another preliminary calculation, the series of words in which the unpleasant words had been selected as least liked

of those with which the individual subject had been confronted, were identified at thresholds not different from those observed with series of words in which the unpleasant words had been selected as next least liked. With means of 98.0 and 95.6, the F here was less than 1 and therefore not significant. The series with the least liked words were also the series which had been presented first to the students.

When the thresholds for common words were compared with those for uncommon words, unpleasant and controls combined, a very significant difference with an F of 7.201 was disclosed, with means of 88.5 and 98.8. This grouping did not control for possible effects of length differences. The uncommon words, as had been anticipated, were correctly identified at higher thresholds.

When the thresholds for short and long words were compared, a non-significant F of 3.294 was obtained. This grouping did not control for possible effects of frequency differences. The means were 94.3 and 99.9.

Very significant differences were found between thresholds for words, unpleasant and control combined, administered to male students and those administered to female students. This grouping did not control for differences in both word length or frequency. The means were 92.9 and 100.9 respectively. The F obtained was 6.816.

Interactions. The interaction of the primary variable, word unpleasantness, with that of peer group dislike of the unpleasant word of the series was significant at between the .05 and .01 levels of confidence with an F of 5.189. The significance of the interaction is attributed to the two groupings of unpleasant words (strongly disliked vs. mildly disliked) which were perceived at appreciably differing thresholds not manifested by their respective controls. Words classified as strongly disliked were perceived correctly by the students as a group at lower thresholds than were those words designated mildly disliked. This is demonstrated in Table X. While the mildly disliked words appeared to be perceived significantly more slowly than their controls as well, the differences between the strongly disliked words and their controls did not seem significant.

Returning to Table VI, one may observe that the interaction between word unpleasantness and the degree to which the unpleasant word of the series had been rated socially taboo was not significant, with an F of 2.003. The interaction of word unpleasantness and individual choice of the unpleasant word of the series as least liked as opposed to the next least liked of the word list was also not significant, with an F of less than 1. As was indicated before, the least liked word series were also those administered first.

Interaction of word unpleasantness with word frequency or "commonness" was significant at between the .05 and .01 levels of confidence, with an F of 5.018. As indicated in Table XI, while the students did not react differently to the control common and uncommon words respectively, they did respond differently to the unpleasant common as compared to the unpleasant uncommon words. The common unpleasant words were perceived at lower thresholds than were the uncommon unpleasant words. Further, as may be seen by reference to the table, it was not the uncommon words which were the most prominently deviant; the significant differences between the unpleasant and the control words lie in the common words only.

Again returning to Table VI, one may discern that no significant degree of interaction was disclosed between the variables of word unpleasantness and word length, nor between word unpleasantness and the sex of the subject, with F's of 2.666 and less than 1 respectively.

Pre-recognition hypotheses with the college students. The following data are summarized in Table VII and were drawn from an N of 630 responses. When the frequencies of pre-recognition hypotheses falling into the four categories were subjected to analysis, very significant differences between them were indicated. The means were as follows: structurally similar, 45.1; structurally unlike, 22.4; nonsense,

38.8; and part, 31.3. With three degrees of freedom here, the obtained F of 9.845 was very significant. As with the hospital subjects, the grouping involved here in no way discriminates between unpleasant and control words, and consequently it is in itself of little meaning to the study.

When the number of hypotheses in response to unpleasant words is compared to that of those prompted by the control words, with one degree of freedom, the F of less than 1 indicates no significant differences. The means were 38.1 and 35.8 respectively.

When the number of hypotheses generated in response to series of words with the unpleasant word strongly disliked by the peer group is compared to the number of responses prompted by series in which the unpleasant word was mildly disliked, the differences are significant at between the .05 and .01 levels of confidence. The means were 34.0 and 41.0 respectively. With one degree of freedom, the F was 5.138. Once again, since this comparison does not involve any control over frequency or word length, its importance as such to the study is minor.

Interactions. When the interaction of word unpleasantness with the categorization of the pre-recognition hypotheses is examined, with an F of 1.801 and three degrees of freedom, the results are not significant. The interaction of word unpleasantness and the degree of peer group dislike for the

unpleasant word of the series with relation to the resulting frequencies of pre-recognition responses also indicates no significant effects of one variable upon the other. With one degree of freedom, the F obtained was 2.430.

TABLE III SUMMARY OF ANALYSIS OF VARIANCE OF THE NUMBER OF EXPOSURES TO THE POINT OF WORD RECOGNITION BY THIRTY HOSPITAL PATIENTS

Source of Variance	d. f.	Mean Square	F	p
Total	299			
Unpleasant vs. control words	1	606	< 1	N. sig.
Series with unpl. word strongly disliked by peer group vs. series with unpl. word mildly disliked	1	26	< 1	N. sig.
Series with unpl. word strongly taboo vs. those with unpl. word mildly taboo	1	6,532	5.381	.05-.01
Order of presentation (first series presented vs. second series)	1	6,449	5.312	.05-.01
Series with unpl. word least liked vs. series with unpl. word next least liked	1	613	< 1	N. sig.
Interactions				
Unpleasant vs. control words x strongly disliked vs. mildly disliked words	1	2,151	1.772	N. sig.
Unpleasant vs. control words x strongly taboo vs. mildly taboo	1	608	< 1	N. sig.
Unpleasant vs. control words x order of presentation	1	129	< 1	N. sig.
Unpleasant vs. control words x least liked vs. next least liked words	1	529	< 1	N. sig.
Error	290	1,214		

TABLE IV SUMMARY OF ANALYSIS OF VARIANCE OF THE NUMBER OF EXPOSURES TO THE POINT OF WORD RECOGNITION BY THIRTY-TWO HOSPITAL PATIENTS

Source of Variance	d. f.	Mean Square	F	p
Total	319			
Unpleasant vs. control words	1	992	< 1	N. sig.
Common vs. uncommon words	1	15,439	13.543	V. sig.
Short vs. long words	1	23,564	20.670	V. sig.
Interactions				
Unpleasant vs. control words x common vs. uncommon words	1	2,395	2.101	N. sig.
Unpleasant vs. control words x short vs. long words	1	273	< 1	N. sig.
Error	314	1,140		

TABLE V SUMMARY OF ANALYSIS OF VARIANCE OF THE NUMBER OF PRE-RECOGNITION HYPOTHESES MADE BY HOSPITAL PATIENTS

Source of Variance	d. f.	Mean Square	F	p
Total	450			
Categories of hypotheses	3	4,846	5.708	V. sig.
Unpleasant words vs. control words	1	293	< 1	N. sig.
Series with unpl. words strongly disliked by peer group vs. series with unpl. word mildly disliked	1	2,187	2.576	N. sig.
Interactions				
Unpleasant vs. control words x categories of hypotheses	3	48,247	56.828	V. sig.
Unpleasant vs. control words x strongly disliked vs. mildly disliked words	1	1,438	1.694	N. sig.
Error	441	849		

TABLE VI SUMMARY OF THE ANALYSIS OF VARIANCE OF THE NUMBER OF EXPOSURES TO THE POINT OF WORD RECOGNITION BY COLLEGE STUDENTS

Source of Variance	d. f.	Mean Square	F	p
Total	329			
Unpleasant vs. control words	1	240	< 1	N. sig.
Series with unpl. word strongly disliked by peer group vs. series with unpl. word mildly disliked	1	2,298	3.000	N. sig.
Series with unpl. word strongly taboo vs. those with unpl. word mildly taboo	1	1,812	2.366	N. sig.
Series with unpl. word least liked vs. series with unpl. word next least liked (confounded with order of presentation)	1	470	< 1	N. sig.
Common vs. uncommon words	1	5,516	7.201	V. sig.
Short vs. long words	1	2,523	3.294	N. sig.
Male vs. female	1	5,221	6.816	V. sig.
Interactions				
Unpl. vs. control x strongly disliked vs. mildly disliked	1	3,975	5.189	.05-.01
Unpl. vs. control x strongly taboo vs. mildly taboo	1	1,534	2.003	N. sig.
Unpl. vs. control x least liked vs next least liked (confounded with order of presentation)	1	572	< 1	N. sig.

TABLE VI continued

Source of Variance	d. f.	Mean Square	F	p
Unpl. vs. control x common vs. uncommon words	1	3,844	5.018	.05-.01
Unpl. vs. control x short vs. long words	1	2,042	2.666	N. sig.
Unpl. vs. control words x male vs. female	1	31	< 1	N. sig.
Error	316	766		

TABLE VII SUMMARY OF THE ANALYSIS OF VARIANCE OF THE NUMBER OF PRE-RECOGNITION HYPOTHESES OF COLLEGE STUDENTS

Source of Variance	d. f.	Mean Square	F	p
Total	629			
Categories of hypotheses	3	13,064	9.845	V. sig.
Unpleasant vs. control words	1	522	< 1	N. sig.
Series with unpl. word strongly disliked by peer group vs. series with unpl. word mildly disliked	1	6,818	5.138	.05-.01
Interactions				
Unpl. vs. control x categories of hypotheses	3	2,390	1.801	N. sig.
Unpl. vs. control x strongly disliked vs. mildly disliked	1	3,225	2.430	N. sig.
Error	620	1,327		

TABLE VIII A COMPARISON SUMMARY OF ANALYSES OF VARIANCE WITH THE COLLEGE AND HOSPITAL SAMPLES

Source of Variance	Hospital F	Hospital p	College F	College p
Number of exposures to the point of recognition:				
Unpleasant vs. control words	< 1	N. sig.	< 1	N. sig.
Series with unpleasant word strongly disliked by peer group vs. series with unpl. word mildly disliked	< 1	N. sig.	3.000	N. sig.
Series with unpleasant word strongly taboo vs. series with unpl. word mildly taboo	5.381	.05-.01	2.366	N. sig.
Order of presentation (first series vs. second series)	5.312	.05-.01	< 1	N. sig.
Series with unpl. word least liked vs. series with unpl. word next least liked	< 1	N. sig.	< 1	N. sig.
Common vs. uncommon words	13.543	V. sig.	7.201	V. sig.
Short vs. long words	20.670	V. sig.	3.294	N. sig.
Male vs. female			6.816	V. sig.
Interactions:				
Unpleasant vs. control x strongly disliked vs. mildly disliked	1.772	N. sig.	5.189	.05-.01
Unpleasant vs. control x strongly taboo vs. mildly taboo	< 1	N. sig.	2.003	N. sig.
Unpleasant vs. control x order of presentation	< 1	N. sig.	< 1	N. sig.

TABLE VIII continued

Source of Variance	Hospital F	Hospital p	College F	College p
Unpleasant vs. control x least liked vs. next least liked	< 1	N. sig.	< 1	N. sig.
Unpleasant vs. control x common vs. uncommon	2.101	N. sig.	5.018	.05-.01
Unpleasant vs. control x short vs. long	< 1	N. sig.	2.666	N. sig.
Unpleasant vs. control x male vs. female			< 1	N. sig.
Number of pre-recognition hypotheses:				
Categories of hypotheses	5.708	V. sig.	9.845	V. sig.
Unpleasant vs. control words	< 1	N. sig.	< 1	N. sig.
Series with unpl. word strongly disliked by peer group vs. series with unpl. word mildly disliked	2.576	N. sig.	5.138	.05-.01
Interactions:				
Unpl. vs. control x categories of pre-recog. hypotheses	56.828	V. sig.	1.801	N. sig.
Unpl. vs. control x strongly disliked vs. mildly disliked	1.694	N. sig.	2.430	N. sig.

TABLE IX INTERACTION OF UNPLEASANT VS. CONTROL x CATEGORIES OF PATIENTS' PRE-RECOGNITION HYPOTHESES

Mean Log. of Number of Pre-recognition Hypotheses

Categories	Unpleasant	Control	Total
Structurally similar	30.9	34.7	34.0
Structurally unlike	23.1	16.1	17.4
Nonsense	37.4	32.6	33.7
Part	31.4	37.4	36.4
Totals	31.4	33.48	

The differences here between the unpleasant and control words lie with respect to the structurally unlike group. No simple relationship between SS, SU, N, and P categories with respect to the unpleasant and control words is suggested. The reversal of the tendency in the structurally unlike and nonsense groups as compared with that exhibited by the structurally similar and part groups is the basis for the interaction significance.

TABLE X INTERACTION OF UNPLEASANT VS. CONTROL x STRONGLY DISLIKED VS. MILDLY DISLIKED WITH STUDENTS

Mean Log. of Number of Exposures to Point of Recognition			
Categories	Strongly Disliked	Mildly Disliked	Total
Unpleasant	92.0	112.5	98.5
Control	95.5	97.7	96.4
Total	95.0	100.7	

The significance of this interaction can be traced to the unpleasant words. In relationship to the control words, the unpleasant words demonstrate opposing tendencies in respect to their classification as strongly disliked by a peer group or as mildly disliked. This difference is not demonstrated by the controls.

TABLE XI INTERACTION OF UNPLEASANT VS. CONTROL x COMMON VS. UNCOMMON, WITH STUDENTS

Mean Log. of Number of Exposures to Point of Recognition			
Categories	Unpleasant	Control	Total
Common	76.5	91.6	88.5
Uncommon	103.9	97.5	98.8
Total	98.5	96.4	

This interaction is significant because while the students did not react to the control common and uncommon words differently, they did react differently to the unpleasant common and uncommon words respectively, as shown in the table. Further, it may be seen that it is not the uncommon words which are perceived in different fashion. The significant differences in duration threshold between unpleasant and control words lie in the common words only.

DISCUSSION OF THE RESULTS

The Hypotheses

Since the results of the study can perhaps be most meaningfully examined in relation to the hypotheses originally advanced, a reconsideration of these seems appropriate.

The primary hypothesis: perceptual defense. This hypothesis was used as the focal point of the experimental design. It predicted that the duration thresholds (the number of exposures required for word recognition) for words of idiosyncratically unpleasant connotation would be significantly higher than those for words of neutral connotation, matched with the unpleasant words in number of letters and frequency.

The secondary hypothesis. This hypothesis predicted that there would be significant differences between the frequencies of structurally similar, structurally unlike, nonsense, and part responses elicited by the unpleasant words and those elicited by the neutral words.

Subsidiary hypotheses. Five subsidiary hypotheses were as follows, with the differences predicted in terms of the

relationships between the unpleasant words and their control words matched in word length and frequency:

(1) Words selected as strongly disliked on the basis of composite weighted scores in terms of peer group ratings would have greater duration thresholds than those selected as mildly disliked on the basis of the same criteria.

(2) There would be significant differences between the duration thresholds for unpleasant words selected as strongly socially taboo on the basis of prediction ratings by clinical psychologists, and those for unpleasant words selected as mildly or not taboo, the thresholds being greater with the former.

(3) Unpleasant words rated as least liked by individual subjects would have greater duration thresholds for them than would unpleasant words rated as next least liked.

(4) There would be significant differences between the duration thresholds for uncommon unpleasant words and those for common unpleasant words, with the former being higher. On the basis of this hypothesis, it was predicted that the differences in threshold between unpleasant and control words would be greater with the uncommon than with the

common words.

(5) There would be significant differences between the duration thresholds for long unpleasant words and those for short unpleasant words, with the former being higher. The prediction made here was that the differences in threshold between unpleasant and control words would be greater with long than with short words.

Interpretations and Conclusions

Perceptual defense. The primary hypothesis involving the concept of perceptual defense was not supported by the experimental results from either the hospital or college sample. In each case, the differences obtained between thresholds to unpleasant and to control words were statistically insignificant. This is interpreted as indicating that neither sample of subjects responded differently to the total group of idiosyncratically unpleasant words than they did to the ostensibly neutral words, in terms of the number of exposures required for correct word identification. These findings taken by themselves lend support to the contention of Howes and Solomon (5) that the influence of affective determinants on duration thresholds is minimal.

Categories of pre-recognition hypotheses. With respect to the secondary hypothesis involving differences in the

frequencies of the four categories of pre-recognition responses, the hypothesis was supported by the responses of the hospital group, but was not supported by those of the college students. This may be concluded through reference to the interaction obtained between word unpleasantness and categories of responses (Table VIII, Results). The frequencies of responses falling into the four categories were significantly different for both groups, but only in the hospital sample did differentiating these categories in terms of unpleasant and neutral words produce any significant changes in the frequencies. In this case, as may be seen in Table IX, there were significantly more structurally unlike hypotheses in response to the unpleasant words than to the neutral words. In addition, the reversal of trends responsible for the significant interaction term can be attributed to the greater number of structurally unlike and nonsense responses produced after exposure of the unpleasant words, and the greater number of structurally similar and part responses produced after the neutral words had been exposed.

The fact that the hypothesis was upheld for one sample but not for the other suggests that attributes peculiar to the samples or to the words chosen for each were involved. This demands that care must be used in generalizing from the results obtained. However, the fact remains that within

the hospital group, in spite of insignificant deviations in terms of duration thresholds, there apparently were differences in the way in which unpleasant words were perceived prior to the actual recognition from the way in which the neutral words were perceived. This in turn intimates that the individuals manifesting such behavior were particularly sensitive to the connotations of the words selected by them as especially disliked, as if set to avoid verbal corroboration of subliminal pre-recognition hypotheses carrying such unpleasant meanings. One may thus suggest as did McGinnies in explanation of similar results (14), that the relative profusion of unlike and nonsense responses to the critical words compared to those elicited by neutral words may represent a type of defensive reaction or blocking.

A further comment in this connection is stimulated by the fact that the differences in category frequencies, when these are examined in interaction with unpleasant as opposed to control words, were significant in the hospital group, while there were no significant differences in the total numbers of pre-recognition hypotheses to unpleasant as compared to control words. This implies that in few cases was there any total blocking of responses in reaction to the unpleasant words, but rather, a substitution of nonsense or structurally unlike responses for a portion of the part and structurally similar responses made to the neutral words.

This in turn suggests the possibility that had McGinnies and Sherman (15) categorized their pre-recognition responses as did McGinnies in his earlier experiment (14), they might have found differences in respect to the hypotheses generated by taboo as opposed to neutral words, differences not evident from the mere comparison of proportions of hypotheses made to the two classes of stimulus words.

Peer group dislike. The first subsidiary hypothesis, that words designated as strongly disliked on the basis of peer group ratings would have higher thresholds than words classified as mildly disliked, appears not only not to be substantiated by the results secured with the college students, but to be contradicted by these data. The interaction term computed between word unpleasantness and the degree of peer group disliking (Table VI) shows differences significant at between the .05 and .01 levels of confidence, in the direction opposite to that predicted, when the unpleasant words are considered apart from their controls.

The more detailed presentation of the results in Table X reveals that the strongly disliked words were correctly perceived sooner than those selected as mildly disliked. This may in part have been a function of the frequency of the words so designated. Nevertheless, the thresholds of the strongly disliked words and their matched controls combined, compared with those of the mildly disliked words with their

controls, were not significantly different. The thresholds for the unpleasant and controls words were likewise not significantly different. The fact that significant reversals of trend were demonstrated by the strongly disliked words, in respect to thresholds for unpleasant and control words, as compared to the mildly disliked words, seems to indicate that the strongly disliked words were really perceived differently from those mildly disliked. However, it is not the former which appear to be the center of the deviation, but it is the mildly disliked words. These deviated prominently, not only from the strongly disliked words, but also from their own controls and those of the strongly disliked words as well.

This may involve idiosyncratic connotations of the mildly disliked words which were not involved in the ratings by the peer group, or the latter may simply not have been a reliable indication of actual intensity of disliking. At any rate, the results suggest that the concept of perceptual defense may have application to responses to certain unpleasant stimuli made by certain individuals. It is difficult to explain away the fact that, in spite of careful controls for variations which might otherwise have been attributed to discrepancies in length and frequency between the critical and control words, there remained differences statistically significant between the mildly disliked words

and their matched controls.

Furthermore, the differences obtained cannot be explained by reference to the accentuation of an effect which in itself was sufficiently potent to provoke significant threshold differences. In the case of the mildly disliked words, it was apparently only in interaction that the two variables involved were sufficiently powerful to produce significant effects with a portion of the stimuli, and only with those administered to one of the two samples. This in turn, however, suggests that the effect is by no means such a general one as that of word frequency; with both samples, threshold differences presumably due to the effects of frequency were very significant.

McGinnies' earlier experiment using socially taboo words was criticized by Howes and Solomon (5) on the grounds that no controls for differences in word frequency were employed, and that the significance of the results might be more parsimoniously attributed to these frequency differences than to any effects of tabooness per se. The later experiment by McGinnies and Sherman (15) did employ controls for frequency, yet results significant at the .01 level were reported, differentiating between thresholds for words following taboo words and those following neutral words. In view of this and of the careful equation of frequencies of the unpleasant and control words used in the present experiment, it seems

somewhat premature to dismiss the concept of perceptual defense entirely, yet there certainly appear to be grounds for doubting the generality of its application. There seems little basis, however, for according to it the importance credited to word frequency as a determinant of duration thresholds.

The socially taboo. The second subsidiary hypothesis involved the prediction that words of strong socially taboo character would stimulate higher duration thresholds than would those of relatively mild taboo nature. This hypothesis was not upheld by the data secured with either of the samples, with insignificant interaction terms obtained between word unpleasantness and degree of taboo in each case.

It is of some interest in this connection to compare the socially taboo words employed by McGinnies in his experiment (14) and in that conducted by him and Sherman (15), with some of those designated as strongly taboo in this study. In the experiment by McGinnies alone, the socially taboo words were as follows: raped, belly, whore, kotex, penis, filth, and bitch. In the second experiment cited above, the following of the taboo words were employed: bitch, penis, raped, and whore.

The words designated as strongly taboo in the present study, it may be recalled, were classified on the basis of clinical psychologists' predictions of how people in general

would react to them. The ten words with the highest weighted composite ratings of "tabooness" are presented here in descending order of their numerical scores: whore, bitch, ass, bastard, ejaculate, nigger, prostitute, syphilis, rape, and vomit. The overlapping of the three selections of taboo words is evident. On the basis of the data collected from both the college and hospital samples, no significant differences were noted in the thresholds of unpleasant as compared to control words, as a function of tabooness. It is suggested that while socially taboo words may serve as adequate critical stimuli in experiments such as those of McGinnies', their effectiveness may stem from other qualities than their character as socially taboo words per se.

The difficulties experienced by this and other investigators in securing significant evidence for perceptual defense in respect to duration thresholds suggests that McGinnies' results may be due to the effects of a number of unidentified variables consistently confounded with the critical variable in his experimental procedure. Among these may be uncontrolled factors in the experimental field affecting subjects' "set" toward the test procedure or toward subliminal pre-recognition responses to critical words, especially those after the initial subliminal reaction to the first critical word. It may be recalled that a definite set was instilled in members of both experimental

samples used in this study. The hospital subjects were either referred directly for psychological evaluation or were introduced to the experiment as "control subjects for a research project in connection with skin disorders." To the college subjects, the tachistoscopic presentation was introduced as a "vision test." In both instances, it seems plausible that the words used as stimuli were perceived within the framework of the introduction used, but in neither case were there evidences for perceptual defense for the unpleasant words as a whole or for those designated as strongly taboo.

Under the conditions of testing employed by McGinnies, no definite set appears to have been given the college subjects. If subliminal responses to McGinnies' stimuli did occur as he suggests, then such responses would involve early discrimination of the critical socially taboo words as such, and even the initial subliminal reaction could have instilled a set to perceive the critical word being re-exposed and the critical words to follow, in such a manner as to avoid corroborating the subliminal discrimination at a verbal level. Equally probable is the possibility that a specific set was unwittingly given to his subjects.

Individual dislike. The third subsidiary hypothesis was concerned with the question of whether or not differences

in the unpleasantness of words rated numerically as least liked and next least liked by the individual subjects would be sufficiently pronounced and stable in this attribute to influence significant threshold differences. The prediction, that the least liked words would be perceived less rapidly, was not upheld with either group. The test of this hypothesis was confounded in the college group with the effect of possible differences due to order of presentation, since all of the least liked words were exposed in the first test series shown to these subjects. However, since no differences were found in the college group, as anticipated, it may be assumed that there were no differences attributable to either of the variables confounded.

Word frequency. The fourth subsidiary hypothesis was that there would be a significant difference in duration thresholds for uncommon unpleasant words as compared with common unpleasant words, with the former recognized less rapidly. Differences between uncommon and common words, unpleasant and controls together, were significant in both samples, thus apparently justifying the importance accorded to word frequency differences and their control in the experimental design. These findings were similar to those emphasized by Solomon and Howes (5, 26).

It was thus conceivable that these data might possibly have reflected only the effects of differences in word

familiarity, and in such case, interaction terms with word unpleasantness would have been insignificant. However, when the unpleasant and control words were treated separately by means of the interaction computation, significant differences in threshold to common and uncommon words were also revealed between the unpleasant and control categories with the students but not with the patients. Reference to Table XI demonstrates that the interaction was significant because the students did react differently to the unpleasant words, common and uncommon, but not to the controls. Moreover, it was the common words which departed most significantly from the controls, being recognized more rapidly than the uncommon words. These same common unpleasant words were also recognized more rapidly than the common control words. Thus, the hypothesis itself seems to be supported by the data from the college sample but not from the patients. The prediction based on the hypothesis, anticipating greater differences in threshold between unpleasant and control words for the uncommon rather than the common words, was, of course, not verified.

Once again, the difference between the samples with respect to the significance of the results obtained here suggests reservations in the application of the conclusions to other groups. It is possible that one reason for the observed difference between the samples may be attributed to probable discrepancies in intelligence, vocabulary, and

reading skills of the two groups. One might rationalize the discrepancy by postulating that with the patients, all of the unpleasant words were relatively equally unfamiliar in terms of reading material, while this was not the case with college students accustomed to study and to reading as a recreation.

The fact that the discrepancy between the common unpleasant words and their controls was significant suggests a number of possible alternative explanations. The first is that there were genuine tendencies for certain common unpleasant words to influence perceptual processes subliminally, altering the thresholds normally manifested to neutral words. That is, that certain of the subjects were hypersensitive to the connotations of the critical words chosen by them as least liked and next least liked, and these individuals reacted to such words more rapidly or "vigilantly," either because of consistent idiosyncratic response tendencies toward such words or because of specific salience of the particular words.

A second possibility is that either the common unpleasant words were often encountered in sources not included in the Thorndike-Lorge list; e.g., used in casual conversation between members of the same sex or read from restroom walls, so that the matching of these words with their controls was based on false premises; or, more likely, that

generalization in terms of familiarity with the words occurs through transfer from spoken vocabulary skills to those employed in reading.

A third possibility is that the common unpleasant words appear so often in print and over the air that, partly because of their repetition and partly because of their emotional salience, certain subjects develop a set to perceive these words when they are presented through any sense modality. Consequently these words tend to be more often employed in relatively accurate pre-recognition hypotheses, verbalized or not, to relatively ambiguous stimuli. Were this the case, pre-recognition hypotheses reported might be expected to show a larger number of structurally similar and part responses and a smaller number of nonsense and structurally unlike responses, than those observed with less familiar or salient unpleasant words.

This rationalization receives some degree of indirect support from the fact that with the patients, where the common unpleasant words were not perceived at significantly different thresholds from the common controls (although common and uncommon words as a whole were differentially perceived), there were more structurally unlike and nonsense responses and fewer similar and part responses to the unpleasant words. This was not the case with the students whose pre-recognition hypotheses were not different for

unpleasant as compared to control words, but who perceived the unpleasant common words significantly more rapidly than the control common words. The available data permit only speculation as to the actual explanation.

Due consideration should be given to the very significant differences in both groups of subjects with respect to thresholds for common as compared to uncommon words. The differences were beyond the .01 level in both instances, whereas differences attributable to affective factors, here word unpleasantness, were observed only with the students, and the significance was between the .05 and .01 levels. This suggests that the variable of frequency is a more potent and generally effective agent in duration thresholds than is word unpleasantness, and perhaps than other affective determinants as well. What seems to have occurred in the present study was that the effect of word frequency, powerful in itself, was particularly accentuated with idiosyncratically unpleasant words, to a significant extent with a significant proportion of the college subjects. In these instances and apparently only in these instances, were the unpleasant words more rapidly perceived than their matched neutral controls. The fact that this effect occurred only in conjunction with the common words renders the concept of vigilance somewhat beclouded by the possible connection with aspects of frequency and set involved in the three alternative explanations previously presented by the writer.

Word length. The fifth subsidiary hypothesis predicted that duration thresholds for long unpleasant words would be significantly higher than those for short unpleasant words, in relation to their respective matched controls. With the effects of differences in word length and frequency thus controlled, this hypothesis received no support from the data recorded for either of the two samples; apparently differences in word length do not have significant effect upon the extent to which word unpleasantness influences duration thresholds.

With the patients, very significant differences were observed in the thresholds for short words, both unpleasant and neutral, as compared to long words of both affective groupings, with the short words perceived more rapidly. However, when unpleasant and neutral words were examined separately in respect to length and their thresholds compared in interaction, the resulting non-significant term provides no support for the contention that word unpleasantness has a differential effect for short as compared to longer words. That is, the relationship between long and short words was not significantly different with unpleasant words than with control words.

Not only were there no significant differences obtained for long unpleasant words as compared to short unpleasant words with the college sample, but none were obtained from comparisons of all of the long words with the short words

used, or the long control words with the short control words. Thus, with both samples, differences in the length of those words used as stimuli, varying from three to ten letters in the college group, seemed to be a relatively negligible factor in the effect of word unpleasantness on duration thresholds. Conversely, the effect of differences in word unpleasantness, unpleasant words opposed to the neutral control words, seemed to be of no significance in the effect of word length on duration thresholds.

The likely explanation for the differences observed in respect to the effects of word length on the thresholds for all the words combined, when exposed to the hospital subjects as compared to the college group, is perhaps appropriate here. The mean logarithmically transformed thresholds of the hospital subjects for the short and long words were 78.5 and 95.7 respectively, while those for the college students were 94.3 and 99.9. It is the differences between the short and long words observed in each sample and the significances of these differences which are important here, and not the differences between the means of the two samples per se. In respect to the latter, however, while the differences between the samples may appear of consequence, they can largely be explained with reference to the frequency of the words chosen in the two samples as the least liked and next least liked words, since this also determined the frequencies of the matched control words. It may be remembered

that the threshold differences between common and uncommon words were highly significant for both samples. When the mean Thorndike-Lorge frequency ratings are computed for short and long words in both groups,[1] and these are compared with the means of the thresholds, they appear as shown in Table XII, presented on the following page.

Note that within each sample there is an inverse relationship between magnitude of threshold and that of frequency rating. As expected, the common words, those with the higher mean frequency ratings were seen at lower thresholds than the uncommon words with the lower mean frequency ratings.

The writer attributes the relatively slight and in all likelihood insignificant difference between the thresholds of the long words for the hospital sample and the short words of the college group (which are in direct rather than inverse relationship to the frequency differences) to the effect of uncontrolled discrepancies between the two samples. Since the experiment was not designed to include inter-sample comparisons except as incidental aspects of possible

---

[1]Thorndike-Lorge non-numerical ratings of A and AA were given the minimum numerical values of the ranges they represent. Thus, A, representing words appearing between 99 and 50 times in 1,000,000, was assigned a value of 50, while AA, representing words occurring 100 or more times per million, was assigned the value of 100.

TABLE XII COMPARISON OF MEAN DURATION THRESHOLDS AND MEAN THORNDIKE-LORGE FREQUENCY RATINGS FOR SHORT AND LONG WORDS

Sample	Hospital		College	
Category	Short	Long	Short	Long
Mean log. threshold	78.52	95.68	94.34	99.92
Mean Thorndike-Lorge frequency	36.53	12.00	6.73	4.79

interest, no attempt was made to match for a number of variables which could account for these discrepancies. It was felt that the most effective method of controlling such complex variables as those of physiological efficiency, visual acuity, and allied factors, was to use each subject as his own "matched control," assuming that the variations affecting the experimental results would be relatively constant for all stimuli administered to each subject. The most likely factor accounting for the threshold differences not attributable to frequency here, would appear to be that of age difference with its concomitant retardation of physiological response processes - the patients were in general somewhat older than the students, and it seems quite likely that their reactions were relatively slow.

With reference to the differences between thresholds for long and short words, differences significant for the patients but not for the students, it is suggested then that the higher thresholds observed for longer words with the patients reflected differences in frequency essentially, rather than differences in length. Reference again to Table XII indicates that the short words perceived at a mean duration threshold of 78.52 by the patients had a mean frequency rating of 36.53, whereas the longer words perceived at a significantly higher mean threshold, that of 95.68, had a mean frequency rating of only 12.00. Thus,

in the process of selection of the least and next least liked words, and/or in the compiling of the peer group word list, the patients appear to have chosen short words which were very common, but long words which were considerably less frequently encountered. With the students, on the other hand, there were comparatively small differences in the frequency ratings of the long and short words which they selected as least and next least liked words, both being somewhat less common than the patients' long words, and markedly less common than the patients' short words. Here with the students, where no prominent differences existed between the mean frequencies of long and short words, no significant threshold differences were found. Thus, it would seem that when the effects of differences in frequency are taken into account, the effects of differences in length upon duration thresholds for control and unpleasant words combined are minimal with both groups of subjects. It is suggested that with the tachistoscopic procedure employed in the study, once the threshold has been reached, long words are recognized approximately as rapidly as are short words. This may be a function of the distance between the eye and the tachistoscope exposure back, as a result of which presumably both long and short words of the length employed may be perceived with a single fixation, within the limits of the subject's eye span.

Additional Conclusions and Comments

Order of presentation. It was of particular interest with the hospital group to investigate whether or not the identification of the first unpleasant word had an appreciable effect on the duration threshold of the second unpleasant word. This comparison was made, and while differences significant at between the .05 and .01 levels of confidence were obtained when the series of five words presented first were compared with the series presented second, this seems to have been largely a function of differences in frequency, uncontrolled in this grouping. When the duration thresholds of each series were differentiated in terms of unpleasant and control words, thus involving controls over the factors of frequency and word length, no significant differences were revealed. This seemed to indicate that there was no appreciable carry-over from the identification of the first unpleasant word to the recognition of the second, with respect to the duration threshold of the latter. As already indicated, no significant differences had been obtained with the college group, in which this variable had been confounded with that of the degree of the subject's dislike for his choices of unpleasant stimulus words.

Sex differences. McGinnies (14) was criticized by Howes and Solomon (5) for using a female assistant in his

experiment which made use of sexually loaded words with subjects of both sexes. This was construed as possibly affecting the interpretation of his observations, tending to confuse the effect of an inhibition of verbal report with what had been considered as actual perceptual blocking. Since the present study also included words frequently interpreted with sexual connotations, and since approximately half of the members of the college sample were women and thus might conceivably feel some constraint in reporting their perceptions to a male experimenter, it was of interest to explore the possibility that this might have affected the experimental results.

While very significant differences were obtained for the duration thresholds of the ten words recorded for the female as compared to the male subjects in the college group, the point of importance is that when these results were examined in regard to differences observed between the unpleasant and control words, no significant interaction was revealed. Thus, the apparent differences noted between the men and women in respect to the thresholds seems to have been a function of the variables uncontrolled in comparing the duration thresholds of various individuals for different words but controlled in the interaction, the variables of frequency and word length, especially the former.

Comparison of Results with those of Related Studies

Some implications of the results obtained in this study and their connections to conclusions drawn from certain other experiments have already been mentioned. It may, however, be relevant to extend this discussion to studies less intimately related, pointing out significant similarities and differences. Before doing so, however, certain peculiarities in the technique used in this study should be made explicit.

Comparisons with this study. The propriety of making direct comparisons of results and conclusions drawn from this study to others which are related is open to considerable question. In the first place, the critical stimuli used were somewhat different in nature; they consisted of idiosyncratically unpleasant words rather than words assumed to be socially taboo, words symbolic of personal values, conditioned nonsense syllables, or auditory stimuli of sentence length. In the second place, the measure of duration threshold was re-defined for use in the procedure employed. Third, the fact that the subjects were strongly urged to make pre-recognition hypotheses may have produced or accentuated differences which would not have been obvious, had this not been done. The experiment by Blake and Vanderplas (1) would suggest that this might be the case. Fourth, the use of a non-college group makes certain comparisons with the typical sample of sophomore psychology students used as

subjects in many of the related studies somewhat inappropriate.

Related studies. The results reported by Lazarus, Eriksen, and Fonda (7), in common with the findings of the present study, seem to point toward the definite possibility of obtaining either raised or lowered perceptual thresholds as results of the emotional connotations of stimuli of specified character. With college students in the present study, raised thresholds were observed for words classified as mildly disliked by peer groups, while lowered thresholds were noted for common unpleasant words. Lazarus, Eriksen and Fonda, dealing with auditory perception, found two basic and individually consistent reactions to threatening stimuli: (1) high perceptual accuracy and ready verbalization; and (2) low perceptual accuracy and minimized verbalization and blocking. McGinnies (14) and McGinnies and Sherman (15) found predominantly raised thresholds to their critical words, while Postman and Solomon (21) and Postman and Leytham (20) seem to have concluded from their results that lowered thresholds may occur with stimuli of either rewarding or punishing connotation.

The significant threshold differences found in our college sample with the common unpleasant words as compared to the uncommon unpleasant words is somewhat difficult to reconcile with conclusions drawn from certain other investigations. It may be that the use of nonsense syllables

by Lazarus and McCleary (8, 12) was one factor in their failure to find significant differences in the accuracy of perception of shock-associated nonsense syllables as compared to those not so conditioned. As indicated, the significant threshold differences in the present experiment occurred with the common unpleasant words, and it seems quite possible that the nonsense syllables were never learned to the point achieved in the use of common words. In addition, the exact nature of the effect of the conditioning procedure used by them as an analogue to reality situations is not fully understood, and thus the analogue itself, ingenious as it was, is open to some question.

Also, in apparent contradiction to the writer's findings, Solomon and Howes (26) in their experiment using Allport-Vernon value ranks noted that it was the infrequent words, not the common words, which exhibited the greater effect of value ranks, the criteria of the intensity of motivational connotation with their stimuli. Postman and Schneider (22), too, observed the same tendency in another study using the Allport-Vernon values. It may be that the key to the discrepancies lies in differences in the types of ostensibly motivational stimuli used.

Summary

In summary, then, the following findings were noted in respect to this experiment which employed a sample of general hospital patients and a sample of college students:

(1) Unpleasant words which had been designated as mildly disliked on the basis of peer group ratings were perceived at significantly higher duration thresholds by college students than were neutral "control" words matched with them in terms of number of letters and frequency. No significant differences were obtained with the hospital patients with respect to this variable. No differences were observed with either sample in regard to the thresholds compared for unpleasant words designated as strongly disliked and their matched control words.

(2) Common words, unpleasant and control, were perceived at very significantly lower duration thresholds than were uncommon words, in both samples.

(3) Common unpleasant words were perceived by college students at significantly lower thresholds than were matched common control words. Uncommon unpleasant words were not perceived differently from their controls. With the patients, common unpleasant words were perceived no differently in relation to their matched common control words

than were uncommon unpleasant words in relation to their matched uncommon control words; in neither case were there significant differences between unpleasant and control words.

(4) Short words were perceived at very significantly lower duration thresholds by the patients than were long words. With the students, the differences were in the same direction but were not significant.

(5) There were significant differences in the frequencies of pre-recognition hypotheses within the categories defined as structurally similar, structurally unlike, nonsense, and part responses made by hospital patients to unpleasant as compared to control words. No such differences were found to be significant with the pre-recognition responses of the college students.

No significant differences in duration thresholds were noted in either sample with respect to the following:

(1) word unpleasantness;

(2) the degree to which unpleasant words were rated socially taboo;

(3) the order in which the unpleasant word was presented (in the first five-word series or the second);

(4) The degree to which the unpleasant word had been rated as least liked by the subject (least liked vs. next least liked);

(5) the influence of word length upon the effect of word unpleasantness; and

(6) sex of the subject (college group alone).

In regard to the frequencies of pre-recognition hypotheses falling into the four categories previously cited, no significant differences were noted for either sample when responses to unpleasant and neutral matched words were compared for series of words in which the unpleasant words had been indicated as strongly disliked by a peer group as opposed to series in which the unpleasant words had been indicated as mildly disliked.

General Conclusions

It is concluded on the basis of the results obtained in this experiment, that the concept of perceptual defense is not supported by the response to unpleasant words as such. No evidence for this concept was found when the duration thresholds for unpleasant words were compared with those observed for words of ostensibly neutral affective connotation matched with the unpleasant words in number of letters and frequency. Raised duration thresholds were observed with college students, but only with certain unpleasant words, here classified as mildly disliked in terms of peer group ratings. On the other hand, lowered thresholds were observed with college students for common

unpleasant words. On the basis of these results, the conclusion is drawn that perceptual defense and perceptual vigilance may perhaps operate for certain individuals or groups of individuals, in response to certain types of unpleasant words, but the evidence seems at best inconclusive for postulating these two concepts in reference to general modes of reaction to unpleasant visual stimuli.

Lower perceptual thresholds were observed for common words than for uncommon words with both samples, and with the patients, lower thresholds were also recorded for short words than for long words. Differences in respect to these variables, word frequency and length, appeared highly significant, but computation of the mean frequency ratings of the long and short words indicated that in all likelihood the differences attributed to word length were largely a function of frequency. The actual effect of word length upon duration thresholds for the words employed in the study appeared minimal.

It is also concluded that with certain individuals or groups of individuals, differences in pre-recognition hypotheses to unpleasant as compared to pleasant words do occur, with the most prominent differences attributable to the greater number of structurally unlike hypotheses occurring to unpleasant words. In view of the fact that such

differences were observed only with one of the two samples employed, the hospital patients, the evidence for assuming this to be a general mode of reaction to unpleasant visual stimuli is considered inconclusive.

BIBLIOGRAPHY

1. Blake, R. R., and J. M. Vanderplas. The effect of pre-recognition hypotheses on veridical recognition thresholds in auditory perception. J. Personality, 1950, 19, 95-115.

2. Bruner, J. S., and C. C. Goodman. Symbolic value as an organizing factor in perception. J. Soc. Psychol., 1948, 27, 203-208.

3. Bruner, J. S., and D. Krech, Editors. Perception and Personality. Durham, North Carolina, Duke University Press, 1949, 266 pp.

4. Edwards, A. L. Experimental Design in Psychological Research. New York, Rinehart & Company, 1950. 446 pp.

5. Howes, D. H., and R. L. Solomon. A Note on McGinnies' "Emotionality and perceptual defense." Psychol. Rev., 1950, 57, 229-234.

6. Klein, G. S., H. J. Schlesinger, and D. E. Meister. The effect of personal values on perception: an experimental critique. Psychol. Rev., 1951, 58, 96-112.

7. Lazarus, R. S., C. W. Eriksen, and C. P. Fonda. Personality dynamics and auditory perceptual recognition. J. Personality, 1951, 19, 471-482.

8. Lazarus, R. S., and R. A. McCleary. Autonomic discrimination without awareness: a study of subception. Psychol. Rev., 1951, 58, 2, 113-122.

9. Levine, R., I. Chein, and G. Murphy. The relation of the intensity of a need to the amount of perceptual distortion. J. Psychol., 1942, 13, 283-293.

10. Luchins, A. S. Social influences on perception of complex drawings. J. Soc. Psychol., 1945, 20, 257-273.

11. Luchins, A. S. An evaluation of some current criticisms of Gestalt psychological work on perception. Psychol. Rev., 1951, 58, 69-95.

12. McCleary, R. A., and R. S. Lazarus. Autonomic Discrimination Without Awareness: An Interim Report. Pp. 171-179. Bruner, J. S., and D. Krech, Editors. See reference No. 8.

13. McClelland, D. C., E. A. Clark, T. D. Roby, and J. W. Atkinson. Projective expression of needs; the effect of need for achievement on thematic apperception. J. Exper. Psychol., 1949, 39, 242-255.

14. McGinnies, E. Emotionality and perceptual defense. Psychol. Rev., 1949, 56, 244-251.

15. McGinnies, E., and H. Sherman. Generalization of perceptual defense. J. Abnorm. Soc. Psychol., 1952, 47, 81-85.

16. Miller, J. G. Discrimination without awareness. Amer. J. Psychol., 1939, 52, 562-578.

17. Pastore, N. Need as a determinant of perception. J. Psychol., 1949, 28, 457-476.

18. Postman, L., and J. S. Bruner. Multiplicity of set as a determinant of perceptual behavior. J. Exper. Psychol., 1949, 39, 369-377.

19. Postman, L., J. S. Bruner, and E. McGinnies. Personal values as selective factors in perception. J. Abnorm. Soc. Psychol., 1948, 43, 142-154.

20. Postman, L., and G. Leytham. Perceptual selectivity and ambivalence of stimuli. J. Personality, 1951, 19, 390-405.

21. Postman, L., and R. L. Solomon. Perceptual sensitivity to completed and incompleted tasks. J. Personality, 1950, 18, 347-356.

22. Postman, L., and B. Schneider. Personal values, visual recognition, and recall. Psychol. Rev., 1951, 58, 271-284.

23. Potter, M. C. Perception of Symbol Orientation and Early Reading Success. New York, Teachers College, Columbia University, 1949. 69 pp.

24. Proshansky, H. and G. Murphy. The effects of reward and punishment on perception. J. Psychol., 1942, 13, 295-305.

25. Shafer, R., and G. Murphy. Role of autism in visual figure-ground relationships. J. Exper. Psychol., 1943, 32, 335-343.

26. Solomon, R. L., and D. H. Howes. Word frequency, personal values, and visual duration thresholds. Psychol. Rev., 1951, 58, 256-270.

27. Thorndike, E. L., and I. Lorge. The Teacher's Wordbook of 30,000 words. New York, Teachers College, Columbia University, 1944. 274 pp.

28. Verville, E. The effect of emotional and motivational sets on the perception of incomplete pictures. J. Gen. Psychol., 1945, 69, 133-145.

29. Webster's Collegiate Dictionary. Fifth Edition. Springfield, Mass., G. C. Merriam Company, 1948. 1275 pp.

In requesting you to fill out this questionnaire as carefully and honestly ssible, we are asking your cooperation in gathering material tc be used in a psychological experiment. We wish to study certain things which most e like and others which they dislike, and we hope to learn more about how react to these feelings. Take as much time as you feel is necessary. All rs will, of course, be held strictly confidential.

_____ AGE_____ WARD _____

re there any lines of work which you admire greatly? _____

hat occupation would you admire most? _____

re there any kinds of work which you would dislike? _____

hat line of work would you dislike most of all? _____

re you fond of animals? _____ which animal would you like

most as a pet? _____

re there any animals which you dislike or which bother you? _____

_____ which animal bothers you most? _____

Do you find any insects disagreeable or repulsive? _____

What insect do you find most disagreeable? _____

are there any words in the English language which you find particularly

pleasant? _____ Name three: _____, _____

_____.

What three words are the most unpleasant ones you can think of? _____

_____, _____, _____.

If you were being chased with something that could be used as a weapon against

you, what weapon would frighten you most? _____

What way of dying would seem most terrifying to you? _____

_____

Are there any kinds of people whom you dislike? _____

What kind of person do you dislike most? _____

What kind of person do you like best? _____

Have any experiences in your life been particularly pleasing? _____

What experience was the most pleasing? _____

_____

Have you had any experiences which were particularly irritating to you? _____ What was it that was the most irritating?

_____

Have you ever gone through any particularly frightening experiences? _____ _____ What was it that frightened you most of all? _____

_____

Do you remember ever having been extremely embarrassed? _____

What was it that caused your embarrassment? _____

_____

Have you ever been severely ashamed of yourself? _____

What was it that was responsible for this? _____

Has the possibility of having any illness or disease ever been disturbing to you? _____ Of what illness or disease have you been most frightened? _____

Does any illness or disease condition appear particularly disagreeable or repulsive to you? _____ Which illness or disease condition appears most disagreeable? _____

When you think of all the things in the world that could possibly happen to you:

(a) What do you think would make you the most angry? _____

_____

(b) What would make you the happiest? _____

_____

(c) What would make you the most frightened? _____

_____

(d) What would make you the most unhappy? _____

_____

SENTENCE COMPLETION TEST (Adams)

TIENT	NUMBER	DATE

L

OOD

OUBLE

NT TO KNOW

EDTIME

ANNOYS ME

EL ASHAMED WHEN

OME

EGRET

BEST

ER PEOPLE USUALLY

MY MOTHER

AT PUZZLES ME

I HAD MY WAY

ST BOSSES

FEEL HURT WHEN

NERVES

HEN I WAS A CHILD

THER PEOPLE

HE MOST DANGEROUS

Y GREATEST FEAR

WAS MARRIED

ED TO

WHEN

I HAVE WORKED AROUND

28. A WIFE

29. THE HAPPIEST TIME

30. THE ONLY TROUBLE

31. MY GREATEST HOPE

32. I HATE

33. I AM VERY

34. MOST DOCTORS

35. THE WAR

36. A MOTHER-IN-LAW

37. THE FUTURE

38. THE STRANGEST THING

39. MY MIND

40. I FAILED

41. MY EDUCATION

42. THE BEST JOB

43. I'M AFRAID

44. I CANNOT UNDERSTAND WHAT MAKES ME

45. THE WORST THING

46. MOST GIRLS

47. MY FAMILY NEVER

48. MY MOST IMPORTANT DECISION WAS

49. MY GREATEST WORRY IS

50. I OBJECT

51. A FATHER-IN-LAW

52. TODAY, I

53. MY SCHOOL WORK

54. I BECOME EMBARASSED

55. I GET ANGRY WHEN

56. MY JOB

people talk to one another, they do so with words which may mean pleasant or unpleasant things. When a person talks to a doctor or a psychologist, he talks about his problems and uses words which often mean unpleasant things. We are interested in finding more about how people react to the unpleasant words which they hear and hear. In order to do so, we need to find words which a great many people agree are unpleasant. Please read through this list and put a check in front of each word which you find you definitely dislike. The word may be unpleasant because of the way it looks, the way it sounds, or what it represents. Do not mark words which you do not know or fully understand. Your opinions will be held confidential and your name is not required.

PAGAN
TERROR
REBUKE
NIGHTMARE
KILL
MALIGNANT
OBNOXIOUS
NEGLECT
MURDER
GANGRENE
THROTTLE
PARAMOUNT
SIEGE
NUPTIAL
REPULSE
TOILET
PRISONER
SIN
MELANCHOLY
RETREAT
OATH
SCORCH
LOSS
TORTURE
HATE
SKELETON
PROFANE
SKEWER
MISER
TORTURE
PATRONIZE
SCORPION
WRETCHED
REVENUE
PROSTRATE
SCREAM
PENITENTIARY
TRAMP
REPRESSION
SCUTTLE
RHEUMATISM
PAY
SLAP
RICH
QUACK

PAIN
TEST
RECKLESS
NOOSE
SHRIMP
MANGLE
OCTOPUS
REPEL
MAR
THRASH
PANIC
SICK
NUDE
REPROACH
TIMID
PREJUDICE
SILLY
MAUL
RESENT
OBSCENE
SCAVENGER
MENACE
TOMB
PARASITE
SINISTER
MILDEW
PROFITEER
SKIMP
MISERY
TOTTER
PAUNCH
SCOUNDREL
WRECK
REVOLTING
PROTRUDE
SCREECH
PERJURY
TRAMPLE
PEST
SEDUCE
RIBALD
PSYCHIATRIST
SLASH
RIOT
PUTRID

THEFT
RECOIL
NORMAL
SHRIVEL
MANIAC
OBSTINATE
REPRIMAND
MASH
THREATEN
PANSY
SICKNESS
NUISANCE
REPTILE
TITTER
PREY
SIMPER
MEAGER
RESTRICT
OBNOXIOUS
SCOFF
MENTAL
TORMENT
PASSION
SINUS
MILITARY
PROMISCUOUS
SKULK
MANGE
TRAGIC
PAUPER
SCOURGE
WORM
REVOLUTION
PROVOKE
SCREW
PERSECUTE
TRAP
PESTILENCE
SEIZE
RIDICULE
PSYCHOLOGY
SLAT
RISK
PUSSY

THIEF
REMORSE
NOTORIOUS
SHROUD
MANIPULATE
OBSESSION
REPRISAL
MASSACRE
THROB
PARALYZE
SIDLE
NUMB
REPUGNANT
TOAD
PRICK
SIMPLE
MEDAL
RETALIATE
OBESITY
SCOLD
MERCILESS
TORNADO
PASSIVE
SISSY
MIRE
PROPAGANDA
SKULL
TORRENT
PATHETIC
SCORN
WRITHE
REVENGE
PROSTITUTE
SCOWL
PAWN
TRAITOR
PERVERT
SCUM
REVOLVER
PROWL
SLANDER
RIGID
QUAKE
PURGE

OON	SANITARIUM	SANITARY	MUSS
CH	MURDER	MURKY	POUT
TLY	POUCH	POUNCE	SCALD
AGE	SCAB	SCAFFOLD	PREGNANT
ERTY	PRATTLE	PREACH	NAUSEA
ILATE	NAKED	NASTY	RAUCOUS
H	RAT	RATTLE	TEDIOUS
NT	TAVERN	TAWDRY	WELT
PON	WEASEL	WEEP	OFFAL
E	OGLE	OFFENSIVE	REBEL
AGE	RAVISH	RAW	NIGGER
VOUS	NEUROTIC	NIGGARD	TERRIFY
PTATION	TENTACLE	TERRIBLE	SCARE
ALP	SCALY	SCANDAL	MALICE
ADJUSTMENT	MALADY	MALARIA	PALSY
AN	PAIN	PALLOR	THIEF
ERCH	TEST	THEFT	REMORSE
UKE	RECKLESS	RECOIL	NOTORIOUS
ETRADE	NOOSE	NORMAL	SHROUD
ILL	SHRIMP	SHRIVEL	MANIPULATE
LIGNANT	MANGLE	MANIAC	OBSESSION
OUS	OCTOPUS	OBSTINATE	REPRISAL
EGADE	REPEL	REPRIMAND	MASSACRE
URE	MAR	MASH	THROB
NGRENE	THRASH	THREATEN	PARALYZE
ROTTLE	PANIC	PANSY	SIDLE
AMOUR	SICK	SICKNESS	NUMB
EGE	NUDE	NUISANCE	REPUGNANT
PTIAL	REPROACH	REPTILE	TOAD
PULSE	TIMID	TITTER	PRICK
ILET	PREJUDICE	PREY	SIMPLE
ISONER	SILLY	SIMPER	MEDAL
IN	MAUL	MEAGER	RETALIATE
LANCHOLY	RESENT	RESTRICT	OBESITY
ETREAT	OBSCENE	OBNOXIOUS	SCOLD
ATH	SCAVENGER	SCOFF	MERCILESS
OURGH	MENACE	MENTAL	TORNADO
SE	TOMB	TORMENT	PASSIVE
RIEDO	PARASITE	PASSION	SISSY
ATE	SINISTER	SINUS	MINE
KELETON	MILDEW	MILITARY	PROPAGANDA
ROFANE	PROFITEER	PROMISCUOUS	SKULL
KEWER	SKIMP	SKULK	TORRENT
ISER	MISERY	MANGE	PATHETIC
ORTURE	TOTTER	TRAGIC	SCORN
ATRONIZE	PAUNCH	PAUPER	WRITHE
SCORPION	SCOUNDREL	SCOURGE	REVENGE
RETCHED	WRECK	WORM	PROSTITUTE
EVENUE	REVOLTING	REVOLUTION	SCOWL
ROSTRATE	PROTRUDE	PROVOKE	PAWN
SCREAM	SCREECH	SCREW	TRAITOR
PENITENTIARY	PERJURY	PERSECUTE	PERVERT
TRAMP	TRAMPLE	TRAP	SCUM
PESSIMISM	PEST	PESTILENCE	REVOLVER
SCUTTLE	SEDUCE	SEIZE	PROWL
RHEUMATISM	RIBALD	RIDICULE	SLANDER
FIX	PSYCHIATRIST	PSYCHOLOGY	RIGID
SLAP	SLASH	SLAT	QUAKE
RIGOR	RIOT	RISK	PURGE
QUACK	PUTRID	PUSSY	

F	DEARIE	DEATH	DEBACLE
ICIT	IMBECILE	IMPAIR	IMPOTENT
AMITY	CALLOUS	CANNIBAL	CAT
	LIFELESS	LILY	LIMB
BRIS	DEBT	DESTRUCTIVE	DETECTIVE
DEST	INDECENT	INFECTION	INFERIOR
ACASS	CASKET	CATHARTIC	CHEAT
AFT	CRATE	GREEDY	GRIEF
TERIORATION	DETEST	DEVIL	DIAGNOSIS
MP	LIQUIDATE	LISP	LIVER
IPE	GROSS	GRUDGE	GRUESOME
ALECT	DIAPER	DIAPHRAGM	DICTATE
FINAL	INJURY	INJUSTICE	INSECT
CKIE	DIE	DINGY	DIPHTHERIA
SANE	LIZARD	FATAL	FATTY
NATIC	FAMINE	DIRTY	DISAGREEABLE
SAPPOINTMENT	DISTASTE	LOATHE	LOCUST
DITER	LONELY	DISEASE	DISCIPLINE
ISGRACE	DISGUST	FALSE	FAKE
TRY	FAILURE	MISTRESS	MOAN
B	MOCK	QUIT	RABBI
ABBLE	RACIAL	RACKET	DISHONEST
ISMAL	DISOBEY	DISORDER	SEXUAL
RABBY	SHACK	SHADOW	OINTMENT
MINOUS	OOZE	OPERATION	EXTRAVAGANT
XECUTION	EXCRETION	EXCOMMUNICATE	LOON
OOT	LOUD	LOUSY	HOOK
OOSTER	ROTTEN	ROUGH	EXAGGERATE
VOLUTION	EVIL	EVACUATE	PIRATE
ISTOL	PITH	PITY	DISTORT
ISTRESS	DISTRUST	DIVORCE	MONGREL
ONOPOLY	MONOTONOUS	MONSTER	SHAKY
RAL	SHAMELESS	SHAME	DOG
ON INATE	DOSE	DRAGON	OPPRESS
RDER	ORGAN	ORGANIC	ROUTINE
OWDY	RUBBISH	RUDE	ESPIONAGE
PILEPSY	ENVY	EMOTIONAL	LOUT
UCHE	LUMP	LUNATIC	SHANTY
SHAPELESS	SHARK	SHATTER	DRAMA
DREADFUL	DREAM	DREARY	MORASS
ORBID	MORGUE	MORTAR	RADICAL
RAGE	RAGGED	RAID	DROWN
DRUNK	DULL	DUMB	PLAGUE
PLUG	PLUMP	PLUNDER	EMBARRASS
EJACULATE	EGOTISM	EFFEMINATE	ORGY
OUST	OUTCAST	OUTLAW	SHERIFF
SHIFTLESS	SHIRK	SHOCK	DUNG
DWARF	DYNAMITE	DYSENTERY	LURCH
LUSH	LUSTY	LYMPH	RUIN
RUMP	RUPTURE	RUT	MORTGAGE
MOULDY	MOURN	MOUSE	RAKE
RANKLE	RAPE	RASCAL	LYNCH
MAD	MADAM	MADNESS	MAGGOT
POISON	POLICE	POLLUTE	POMPOUS
SHOOT	SHRAPNEL	SHREW	SHRIEK
QUID	QUEER	QUAVER	QUARREL
RAIM	MUCK	MUCOUS	MUD
RUTHLESS	SACRIFICE	SAD	SAG
SALIVA	OUTRAGE	OVARY	OYSTER

AGONY
COBWEB
DECOMPOSE
FERTILIZE
GUILT
INSULT
ASSASSINATE
DISCREDIT
FIEND
GUT
INTERFERE
IRRITABLE
ATROCITY
CONCEIT
FOOL
JAIL
DEFORMITY
NAUGHTY
BANDIT
CONSPIRACY
FOREIGN
JUNK
DEGRADE
HEARTLESS
BEAST
DEMAND
JUNK
KNIFE
CORPSE
FRANTIC
BITTER
DEPRAVE
MIND
CRAB
FRIGID
BLAME
HORROR
CRIPPLE
LAME
DERISION
FUSS
GASH
CROWD
HURT
BOMB
LAW
CUCKOO
HYSTERICAL
DESPISE
BRIBE
ILLEGITIMATE
DAMAGE
LAWYER
GREEDY
BULGE
LEPROSY
DARK
GLUTTON
BUTTOCKS

CHOKE
DECAY
FAULT
GUILLOTINE
GURGLE
INTERCOURSE
COLLISION
DEFEAT
FIGHT
HAG
INTESTINE
APOLOGIZE
COMMAND
FLABBY
ITCH
DEFIANCE
HARLOT
AWKWARD
BARBARIC
CONTAMINATE
FORGERY
JILT
DELAY
HELL
BEGGAR
DEMON
KIDNAP
CONTEMPTIBLE
FORMAL
BELLY
DENOUNCE
DEPRESSED
HISS
CRASH
FROZEN
SPITE
HOSTILE
CRITICISM
LARCENY
DESERTER
GANGSTER
CROAK
HUMILIATE
BLIND
LARD
CRUCIFY
HUSSY
DESOLATE
BOWEL
IDIOT
CURSE
DAMN
LAZY
GERM
BULL
LEWD
DEAD
GOODY
BUZZARD

CLINIC
DECEASE
FEAR
GUILTY
INSIST
ARROGANT
COLON
DEFECT
FILTHY
HAIRY
INTIMIDATE
ASYLUM
COMMUNIST
FLATTER
JACKASS
DEFICIENCY
HARSH
BABOON
CONDEMN
FORCE
JEALOUS
DEFRAUD
DELINQUENT
HERETIC
BELCH
DENIZEN
KIDNEY
CONTROL
FOUL
BIGOT
DENY
HICK
CORRUPT
FRAUD
BLACKMAIL
HOG
CRAZY
LABOR
DEPRIVE
FUNERAL
GAPE
CROOK
HUMP
BLOOD
LASCIVIOUS
CRUDE
HYENA
DESPAIR
BRAG
IGNORANT
CYNICAL
DANGER
LEECH
GHOST
BURN
LIAR
GIZZARD
GOSSIP
CACKLE

CLINK
DECEIVE
FEEBLE
GULCH
INSOLENT
ASS
COMBAT
DEFENSELESS
FIRE
HANG
IRK
ATHEIST
COMPULSORY
FLESH
JAGGED
DEFILE
HATE
BALD
CONFESS
FORECLOSE
JEER
DEGENERATE
HEARSE
BASTARD
DELUSION
JOWL
KILL
CONVULSION
FRAME
BILIOUS
DEPLORABLE
HIDEOUS
COWARD
FRIGHT
BLADDER
HONK
CRIMINAL
LAGGARD
DERELICT
FUNGUS
GARBAGE
CROTCH
HUNGER
BOAST
LASH
CRUEL
HYPOCRISY
DESPERATE
BRAT
ILLEGAL
DAGGER
LAWLESS
GASP
BRUTAL
LEER
DANK
GLOOMY
BUTT

ELFISH	SENSUAL	SERGEANT	SERVENT
RITUAL	ROACH	ROB	RODENT
REW	PULP	PUCKER	PEWTER
BRUM	SEVERE	SEWER	SLAUGHTER
WASH	TREACHEROUS	TREASON	TREMBLE
PHANTOM	PHYSIC	PIAZZA	PIG
SLAVERY	SLAY	SLEET	SLICE
PILE	PILL	PIMPLE	PINCH
SLIME	SLINK	SLIPPERY	SLOPPY
LAP	WOMB	WOLF	WITCH
WIPE	TREMOR	TRENCH	TRESPASS
THICKERY	SPLINTER	SPOIL	SPURN
SPURT	WINCE	WILT	WILDLY
WICKED	TRIPE	TROUBLE	TRUSS
TUBERCULOSIS	STRICKEN	STRICT	STRIKE
STRINGY	UDDER	UGLY	ULCER
UNBALANCED	SLOUCH	SLOVENLY	SLUG
SLUGGISH	WHORE	WHIPPING	WHINE
WHIMPER	TUMOR	TWIST	TWITCH
TYPHOID	SPUTTER	SQUABBLE	SQUALOR
SQUANDER	UNBEARABLE	UNDERTAKER	UNEASY
UNFRIENDLY	TYRANT	TYPHUS	STRUGGLE
STRUT	STUD	STUFFY	WHELP
WHEEZE	WENCH	WEAK	STUN
STUPID	SQUASH	SLUM	UNGRATEFUL
UNKIND	UNJUST	UNSAVORY	SLUMP
SLUSH	SQUAT	STUPOR	UNSCRUPULOUS
UNSTABLE	UNTIDY	URINE	SLY
SMACK	SQUAWK	STUTTER	URN
USURP	VEIN	VENGEFUL	SMASH
SMEAR	SQUEAK	SUBDUE	VENOM
VERMIN	VICIOUS	VICTIM	SMELL
SQUEAL	SUBMARINE	SMIRK	VICTUAL
VILLAIN	VIOLATE	VIOLENCE	SQUEAMISH
SUBMIT	SMOTHER	SQUIRM	VIPER
VIRGIN	VIRTUOUS	VOID	SUCKER
SUCCOR	SMUDGE	SQUIRT	SUCCULENT
SMUT	STAB	SUCCOMB	VOMIT
VULGAR	VULTURE	SNAIL	STAGGER
SUCK	SNAKE	STAGNANT	SUFFER
SNARL	STAIN	SUFFOCATE	WADDLE
WALLOW	WANTON	WARNING	SNEER
STALE	SUICIDE	SNICKER	STAMMER
SULK	SNOB	STARK	SULLEN
SNOOP	STARVE	WARP	WASP
WASTE	SNUB	STEAL	SUPERIOR
SOB	SOCIALIST	STEALTHY	SUPPRESS
SURGERY	STENCH	SOCKET	SODDEN
STERILIZE	SURRENDER	SOGGY	STICKY
SUSPICION	SORDID	STIFLE	SWAB
SWADDLE	STING	SORE	SORROW
STINGY	SWAGGER	SWEAR	STINK
SOUR	SOUSE	STOLEN	SWEAT
SWELTER	STOOL	SPASM	SPATTER
STOOP	SWINDLE	SWINE	STRADDLE
STORM	SPERM	SPIDER	SPINE
STRAGGLE	SWOLLEN	SYPHILIS	STRAIN
SPINSTER	SPIT	STRANGE	STRANGLE
TAINT	TANTRUM	TARNISH	STRETCHER

Look over the following words and select the five words which you .ke least. Mark a 1 in front of the word you like least of all, 2 in front of the word you like next least, a 3 in front of the ›rd you like third least, and so on until you have indicated ıe five words you like least in the order of their unpleasantness ›r you. Make your choices from all the words listed below, those ⁄ped on the list as well as those mimeographed. Ignore any word ıich you do not recognize or which is unfamiliar to you.

)VERTY	OGRE	RAVAGE	MALIGNANT
ANGRENE	MELANCHOLY	TORTURE	SICK
IGGARD	NASTY	PUTRID	MANGE
ROSTITUTE	OBESITY	NIGGER	AGONY
IEND	GUT	CORPSE	LEPROSY
ASTARD	BRAT	HAG	FLABBY
ILL	BRUTAL	BELLY	CROAK
ROTCH	ILLICIT	HUSSY	IDIOT
YPOCRISY	LICE	LAZY	ARROGANT
ARCASS	GRIPE	CRAZY	ASS
INSANE	DISGRACE	MOB	EXECUTION
EPILEPSY	DRUNK	POISON	INDECENT
MORGUE	RAPE	SHOOT	MOCK
HEARSE	DEATH	ROTTEN	CASKET
ANNIBAL	DESTRUCTIVE	INJURY	DIE
REEDY	DEVIL	DISASTER	EVIL
GRUDGE	FATAL	DIRTY	DIVORCE
LOUSY	LUNATIC	POLLUTE	CHEAT
DISAGREEABLE	DISHONEST	ORGY	DUNG
LYNCH	SHRIEK	SELFISH	TRASH
SLAVERY	YAP	SLIME	TRICKERY
USURP	TAINT	SPITE	WHORE
SNAKE	SPIT	TREASON	SLOVENLY
SLUM	SYPHILIS	VOMIT	STINK

Appendix 5

(Reproduction of a Checked List)

Indicate the eight words you like least from this list by putting a check in front of each one you select:

REWARD	✓ BIDDER	VACATION	ATTRACTIVE
✓ CHUBBY	HELPFUL	ADMIRE	SPORTS
FOOTBALL	✓ FELLER	INTERESTING	PRIZE
CONSIDERATE	GAME	✓ GOGGLE	AMUSING
SKILLFUL	HUMOROUS	SINCERE	✓ WEIGHTY
GENEROUS	✓ TARNISH	APPROVAL	ENTERTAINING
APPLAUD	PEACE	✓ CORDIAL	DELICIOUS
✓ SUNSPOT	[illegible]	ROADCOAT	

(The underlinings have been employed here to indicate which of the words were the control words).

Last name Doe, First Jane

Sex (M or F) _____

Please read these instructions carefully:
Look over the following words and select the five words you like least. Mark the number; 1, in front of the one word you like least of all the words. Next, mark a 2 in front of the one word you like next least. Now mark a 3 in front of the word you like third least, and so on until you have indicated the five words you like least, in order of their unpleasantness for you. You may find them unpleasant because of their appearance, their sound, what they represent or remind you of, or for any number of these or other reasons. When you have finished, you should have five, and only five words marked, each with a different number. An explanation for this procedure will be supplied you later in the course.

AGONY	FIEND	GUT	ATROCITY
GASH	LEPROSY	ABORTION	FLABBY
BELLY	HUSSY	BOWEL	LEWD
FILTHY	[redacted]	BELCH	LEECH
ADULTERY	ASS	CROTCH	BRAT
POVERTY	MUTILATE	OGRE	TERROR
MALIGNANT	MANURE	GANGRENE	TORTURE
SCORPION	MURDER	SCAB	SCALY
MANGLE	MAUL	PUTRID	NIGGARD
SHRIVEL	MANIAC	OBNOXIOUS	PUSSY
NAUSEA	NIGGER	MASSACRE	SISSY
WRITHE	PERVERT	SCUM	LICE
INCEST	CARCASS	DETERIORATION	GRIPE
INSANE	FANATIC	DISGRACE	FAIRY
RABBLE	DISMAL	SHABBY	ROWDY
EPILEPSY	MORBID	RUMP	MOULDY
MAIM	DEARIE	CASKET	MOCK
OOZE	EXCRETION	ROTTEN	MORGUE
RUPTURE	RAPE	QUEER	DEATH
CANNIBAL	INFECTION	CATHARTIC	GREEDY
DIRTY	LOATHE	DISEASE	LOUSY
SHAME	LUNATIC	DREARY	EFFEMINATE
DYSENTERY	POLLUTE	MUCOUS	CHEAT
GRUESOME	FATTY	LOUT	ORGY
DUNG	MAGGOT	PURGATORY	SELFISH
TRASH	SLAVERY	SLIME	WICKED
TUBERCULOSIS	STRINGY	SLUGGISH	WHIMPER
TYPHOID	WHEEZE	STUPID	USURP
VERMIN	SMUT	VULGAR	SUCK
SNOOP	ROACH	SLINK	WHORE
WENCH	SUICIDE	STENCH	SOUSE
SPIT	PUNY	SEWER	SLOVENLY
TYPHUS	SLUM	URINE	SMIRK
SYPHILIS	PUNK	RODENT	PIG
SLOPPY	WITCH	WHINE	WHELP
VENOM	SUCKER	VOMIT	PEW
WALLOW	SNOB	SQUAT	SLOUCH
TRIPE	UGLY	STAGNANT	ULCER
GAUDY	BUTTOCKS	CROAK	HELL
HAG	BILIOUS	SALOON	ODOROUS
TOILET	MESS	WRETCHED	QUACK
PANIC	NASTY	DEBRIS	DISTRESS
EJACULATE	RUTHLESS	SALIVA	FAILURE

Last Name _____, First Name_____

Please do not fill this slip out unless you took part in the tachistoscopic "vision experiment" in which you identified 17 words for Mr. Sterne in Room 18 of the Psychology Building on South Campus, and also had previously rated the list of unpleasant words, with the numbers from 1 to 5. If you did take part in both parts of the experiment, please answer yes or no to the following questions as honestly and carefully as possible.

1. Did you know that the vision experiment was connected with the ratings, before you saw any of the words in the tachistoscope? _____ if not,

2. Did you have any idea that the two might be connected before you saw the words in the tachistoscope? _____ If not,

3. Did you realize that the two were connected after seeing the first stimulus (unpleasant) word? _____ If not,

4. Did you realize that the two were connected after seeing both stimulus words? _____

Comments:

REMOVE THIS PAGE AND HAND BACK TO YOUR INSTRUCTOR, IF YOU HAVE ANSWERED THE QUESTIONS.

It has been suggested that most individuals learn early in life that certain words carry with them a "socially taboo" connotation. When these words are used by a child, this use generally results in chastisement by parents, with the consequent establishment of a conditioned reaction to the verbal symbol. This pattern of conditioned emotional response may be considered to be one of anxiety or fear aroused by symbols having sexual, excretory, or otherwise unpleasant or "immoral" connotations. (McGinnies, Elliott, "Emotionality and Perceptual Defense", Psychological Review, 56, 1949, p.244.)

Please rate the following words in terms of the extent to which each word seems "socially taboo" by putting a check in the appropriate space after it.

Word	Highly taboo	Moderately taboo	Slightly taboo	Not Taboo
Bastard				
Fatty				
Fairy				
Queer				
Nigger				
Dearie				
Few				
Dung				
Ooze				
Effeminate				
Suck				
Smut				
Ass				
belly				
Syphilis				
Slime				
Rape				
Belch				
Gut				
Stench				
Nausea				
Scab				
Tripe				
Slavery				
Leprosy				
Suicide				
Putrid				
Rabble				
Usurp				
Maggot				
Failure				
Murder				
Gangrene				

Word	Highly taboo	Moderately taboo	Slightly taboo	Not Taboo
Mangle				
Ejaculate				
Stink				
Vomit				
Bilious				
Lynch				
Divorce				
Bitch				
Snob				
Tuberculosis				
Snake				
Maim				
Whore				
Prostitute				
War				
Evil				
Hate				
Shrapnel				
Kill				
Disaster				
Lice				
Slum				
Death				
Morgue				
Rotten				
Disgrace				
Torture				
Die				
Cannibal				
Fiend				
Dishonest				
Brutal				
Poverty				
Execution				
Poison				
Disagreeable				