EFFECTS OF DEGREE OF CATEGORY SEPARATION AND WORD POLARIZATION ON SEMANTIC CONCEPT IDENTIFICATION

by

CURTIS L. TAYLOR

B. S., Kansas State University, 1966

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1967

Approved by:

Robert C. Haygood

Major Professor

ACKNOWLEDGMENT

This thesis is gratefully dedicated to . . .

Dr. Robert C. Haygood for his encouragement and guidance of the research project reported here.

Dr. Roy C. Langford, whose vivid introductory lectures made it necessary for me to change my plans and follow this path of knowledge.

My parents, Lloyd R. Taylor and Dorothy I. Taylor, for their unflagging enthusiasm and support even though at times they must have wondered what I was doing.

Wilma A. Taylor, my wife, whose efforts in the preparation of this manuscript were above and beyond the call of duty.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

The paradigm of concept-identification tasks is usually one in which S has to discover the correct method for sorting stimuli into different categories on the basis of common attributes.

The S is given a set of instructions to familiarize him with the nature of the task that he is to perform. The details of these instructions reflect the design of the problem. Briefly, the S is informed that he will be required to separate stimuli into two categories, those that are examples of the concept and those that are not examples. The S obtains information about the concept on a trial-by-trial basis until he reaches some criterion which reflects his grasp of the solution. Criterion may be a set number of errorless trials in which case the number of trials required to reach criterion provides an important dependent variable, along with number of errors. The number of trials may be predetermined, in which case the number of errors provides the major dependent variable. A trial consists of stimulus presentation, S's response, and some type of informative feedback. Any of these three events, their delays, or duration, may be experimentally studied.

The stimuli for concept-identification studies have, with few exceptions, consisted of abstract visual patterns, such as geometric designs (see, e.g., Bourne, 1965).

These patterns may be varied in shape, size, number, color, and physical orientation; these variations are advantageous in that they are universally known to and discriminable by *S*. Their main disadvantage is that they tend to be artificial and unrepresentative of the verbal concepts used in everyday situations.

One of the first experiments to employ verbal material in a concept-formation task was carried out by Reed (1946). The *S*'s task was to learn the nonsense syllable names of cards. Each card had four common words printed on it. Three of the words were unrelated; the remaining word gave the cue to the card's name. A KUN card was one which contained an animal name (e.g., horn leaf monkey debt), a DAX card contained a color name (e.g., answer highest airplane red); 42 cards with six concepts were used. On each trial *S* was shown the set of cards and asked to name each card. If he could not, he was told the name of the card. This procedure continued until *S* could name each card. One group of *S*s was instructed to learn the names of the cards and the other group was instructed to learn the concepts represented by the nonsense syllables. The *S*s set to learn the names also learned the concepts, but not so quickly as those *S*s set to learn the concepts. Reed described *S* as experiencing an initial period of disorientation, then searching for the appropriate mediating response, and finally evaluating its usefulness.

The facilitation of concept learning with instructions to learn concepts suggests that S alters the production of the stimulus responses or alters the scanning of the presented stimuli.

Using the method of free learning and recall, Bousfield (1953) studied the effect of grouping in verbal learning. The S was allowed to study a list and recall as many words as possible in a brief time interval. Sets of words in the original list were animal names or the names of professions. These items were frequently grouped together in the recall session, quite independently of their position in the original list. The clustering of certain words together when a category is available to S indicates the use of previously learned concepts; it also suggests that being able to group certain words facilitates their recall.

Underwood and Richardson (1956a) had college students give an association to the first sensory impression that came to mind for each of 213 nouns. The percentage of times that the same or a similar word is used to describe the sensory image for a given noun is defined as its level of dominance. Thus dominance is the relative strength of association between a noun and its sensory impression. The same response may be given to different nouns, with varying frequency. Using these materials to control dominance level, Underwood and Richardson (1956b) studied the effect of dominance on learning. Their lists contained six concepts with four examples of each concept. The S was told that there were six groups of four related words; his task was to guess the correct response for each noun as the list was presented serially. After each noun, the E informed S whether he was correct or not. The list was presented 20 times with the order of the nouns varied each time. Three levels of dominance were used in the selection of the nouns, 75%, 41%, and 16%. Performance was best at the higher dominance level; more concepts were learned and more correct responses were obtained. More erroneous responses occurred in the low and medium dominance groups, even after Ss had achieved knowledge of the appropriate class of responses. This would indicate that category labels with higher dominance levels serve as better mediators by reducing competing or interfering responses.

Coleman (1964), using the same materials, presented four nouns simultaneously and required S to give an adjective to describe all of them. Thirty-two tasks of this type were designed, half containing nouns of high dominance and the other half nouns of low dominance. The Ss attained the solution for high dominance significantly faster than the solution for low dominance.

Mayzner and Tresselt (1961) have used a judgement technique with 300 words, 213 of which were from the

Underwood and Richardson study (1956a). The Ss were given instructions to judge whether a word belonged in none, one, or more of the following six concepts: round, small, white, hard, smelly, and long. Rank-Order and Product-Moment correlations were computed between the per cent response frequencies obtained by Underwood's associational technique and the judgement score for words having values of 5% or above. All correlations were significant. Rank order correlations were .66 (round), .68 (small), .80 (white), .57 (hard), .75 (smelly), and .52 (long); while the corresponding r values were .60, .67, .80, .54, .72, and .49 respectively. The basic advantage of this system was that it allowed S to give more than one sense impression. The choice, however, was limited to the six concepts presented.

Using the materials developed in the previous study, Mayzner and Tresselt (1962) have explored verbal concept attainment as a function of the number and strength of positive instances. The response frequencies for words on three concepts, round, long, and hard, were used as a measure of response strength, and six levels of response strength were used in the experiment. For each concept, 11 words were chosen of a particular response strength. Each S had three lists of 11 words, representing each of the three concepts used. The number of instances required to get the right name was the dependent variable. A significant and systematic increase in the number of positive instances was required to discover the concept as response strength decreased.

Studies presented thus far have relied primarily upon either (a) associations between category labels and examples, or (b) connotative meaning. Associations are limited in their usefulness because the category members have no common meaning, other than that implied by high association with a label. This characteristic of association data makes it difficult to specify the concept that *S* is learning.

There are several difficulties in using connotative meaning in concept-formation studies. Connotative meaning usually produces dichotomies rather than continuous dimensions. As an example, usually a word is clearly either a food word or not a food word. Secondly, most meaningful words carry some affective or specific associational meaning for *S*s. Finally, the dimensionality of verbal materials is difficult to determine since it is possible for a particular word to belong to an extremely large number of categories which may or may not be available to *S*.

The Semantic Differential (SD), developed by Osgood and his associates (Osgood, Suci, & Tannenbaum, 1957), provides a methodology for verbal materials that appears to overcome these difficulties. Osgood et al. have shown that there are three major dimensions of meaningfulness;

they have labeled these Evaluation, Potency, and Activity. They are represented by the good-bad, hard-soft, and active-passive scales of the SD, respectively, which are essentially factorially-pure measures of the three major dimensions. These scales provide continuous dimensions of meaning of verbal material, and measure word positions on the dimensions with relatively high reliability. In addition, the dimensions so scaled appear to be rather general in the population, and they take into account the affective and associational aspects of the words (as opposed to the strict connotative meaning of the word).

A typical example of an SD response sheet is shown in Figure 1. The scale positions have already been defined for the S in his instructions (extremely X, quite X, slightly X, neither X nor Y, slightly Y, quite Y, and extremely Y). Ideally, scales are chosen which tend to maximize only one dominant component and minimize all other factors. In practice, this has been almost impossible to accomplish. While it has been possible, in some cases, to select a pair of scales to represent a factor which met the criterion of independence from the other factors, these scales have not been highly correlated in their ratings on individual words. This covariation in their obtained ratings, indicates the inability of a scale to uniquely represent a factor.

Osgood postulates a semantic space, Euclidian in nature and of unknown dimensionality. All semantic scales are

Rating

	1	2	3	4	5	6	7	
(1) cruel	____ :	____ :	____ :	____ :	____ :	____ :	____	kind
(2) curved	____ :	____ :	____ :	____ :	____ :	____ :	____	straight
(3) masculine	____ :	____ :	____ :	____ :	____ :	____ :	____	feminine
(4) untimely	____ :	____ :	____ :	____ :	____ :	____ :	____	timely
(5) active	____ :	____ :	____ :	____ :	____ :	____ :	____	passive
(6) savory	____ :	____ :	____ :	____ :	____ :	____ :	____	tasteless
(7) unsuccessful	____ :	____ :	____ :	____ :	____ :	____ :	____	successful
(8) hard	____ :	____ :	____ :	____ :	____ :	____ :	____	soft
(9) wise	____ :	____ :	____ :	____ :	____ :	____ :	____	foolish
(10) new	____ :	____ :	____ :	____ :	____ :	____ :	____	old
(11) good	____ :	____ :	____ :	____ :	____ :	____ :	____	bad
(12) weak	____ :	____ :	____ :	____ :	____ :	____ :	____	strong
(13) important	____ :	____ :	____ :	____ :	____ :	____ :	____	unimportant
(14) angular	____ :	____ :	____ :	____ :	____ :	____ :	____	rounded
(15) calm	____ :	____ :	____ :	____ :	____ :	____ :	____	excitable
(16) false	____ :	____ :	____ :	____ :	____ :	____ :	____	true
(17) colorless	____ :	____ :	____ :	____ :	____ :	____ :	____	colorful
(18) usual	____ :	____ :	____ :	____ :	____ :	____ :	____	unusual
(19) beautiful	____ :	____ :	____ :	____ :	____ :	____ :	____	ugly
(20) slow	____ :	____ :	____ :	____ :	____ :	____ :	____	fast

Fig. 1. Form and order of the semantic differential scales as used by Jenkins, Russell, and Suci (1958) in compiling ratings for a semantic atlas.

assumed to be straight lines passing through the origin of the space. The definition of the space becomes better as the sample size is increased. To obtain maximum efficiency in defining the space, a minimum number of orthogonal dimensions or axes are needed, and in practice these are obtained by factor analysis of the SD ratings. The "meaning" of any concept in Osgood's system is determined by its location in semantic space, with emphasis on the three-dimensional space defined by the Evaluation, Activity, and Potency dimensions.

If SD dimensions represent real dimensions to S, then verbal materials that have been analyzed on the SD could be employed in learning experiments with results predictable from their ratings. Specifically, Ss should be able to identify these dimensions in a manner similar to that observed in the identification of geometric concepts. A pilot study conducted in this laboratory (Haygood, 1966) has established the feasibility of using SD dimensions in concept-learning experiments. Twenty-four Ss learned to sort words into two categories on the basis of conceptual dimensions (Evaluation and Potency) drawn from the SD. Another group of 24 Ss learned to categorize the same words under conditions in which the concept was irrelevant and only rote memory could be used. Performance of the groups for which the concept was available was significantly superior to that of the group to which it was not available.

The Haygood study did not attempt to determine the relationship between degree of category separation and performance, and used only words from the extreme ends of the SD scales. Thus it is not known whether categories less widely separated on the SD would be as easy to learn. The scale values of the words used by Haygood were normative, representing the average of ratings by many Ss (Jenkins, Russell & Suci, 1958). Thus individual differences in assessment of the words might lead to confusion between categories that are close together, and result in slower learning. This confusion should be particularly evident when both categories are immediately adjacent to the center of the scale (Scale Value 4 on a 1 to 7 scale), which represents a relatively neutral area. In addition, Archer (1962) has shown that increasing the scale separation of levels of the relevant dimension improves performance when geometric designs are used in a concept-identification task. Increasing the separation when the levels are already discriminably different seems to increase the obviousness of the relevant dimension, thus drawing S's attention to that dimension as a possible solution. If the same relationship holds for SD dimensions, performance should continue to improve as the categories become more widely separated.

The purpose of Exp. I was to determine the effects of variation of category separation on the identification of semantic concepts, throughout the range from zero separation (completely overlapping categories) to the most extreme separation possible.

CHAPTER II

GENERAL PROCEDURE AND APPARATUS

Task.--The task was essentially the same as that described by Haygood (1966). The Ss were required to learn to classify a list of 60 English words into two categories, X and NOT-X. For those groups to which the concept was available, the correct classification principle was one of the dimensions taken from the SD, Evaluative, Potency, or Activity (active-passive). The word lists were taken from the semantic atlas of Jenkins et al (1958). They were selected so as to be approximately equal on all dimensions not relevant to problem solution, and to be free of all obvious sources of bias such as differential word length.

Procedure.--The Ss were seated four abreast in a semicircle, with partitions that prevented them from seeing each other. They were given detailed instructions (Appendices I and II) at the outset, which explained the nature of the task, the method of responding, the meaning of the feedback signals, and the length of the experimental session. They were told that this was an experiment to determine how well they could learn to classify a list of words, with emphasis on discovery of a principle for correct classification. It was pointed out that the principle was one of meaning, and had nothing to do with beginning letter, length of word, parts of speech, or any other formal characteristics of the words. The instructions stressed accuracy rather than speed, although the nature of the group situation was such as provide some social pressure against inordinate delays in responding.

The words were projected one at a time on the wall in front of Ss. To each word, S was required to respond by pressing one of two buttons, labeled X and NOT-X, on a control panel located on the arm of his desk chair. After all Ss had responded, the apparatus automatically gave each S informative feedback by turning on a green signal lamp above the correct response button for 1.0 sec. After a postfeedback interval of 3.5 sec., during which the previous stimulus word was still visible to Ss, the next word was presented. At the end of the list, E reset the slide projector to the beginning of the list and so informed the group of Ss. The list was presented four times, each time through the list counting as a block of 60 trials.

Materials and apparatus.--In addition to Ss' control panels and the wall used for projection, the apparatus consisted of four major components: (a) a Kodak Carousel 35-mm slide projector used to present the stimuli, (b) an electronic timing unit used to control the delay and duration of feedback, and the length of the postfeedback interval, (c) a Western Union tape reader, used to control the feedback to Ss' signal lamps, and (d) an Esterline-Angus event recorder used to record Ss' responses. The stimulus words were photographed on 35-mm black and white slides in such a way that the projected words were black on a white background.

CHAPTER III

EXPERIMENT I

Subjects.--Eighty-eight summer-session students at Kansas State University served as Ss and were paid for their participation. The majority of Ss were run in groups of four which were assigned to treatment combinations in order of appearance to the laboratory. The occasional failure of one or more Ss to appear led to the use of some smaller groups; several Ss participated individually or in pairs to complete the 11 experimental groups of eight Ss each. One S was unable to complete the experiment because of apparent emotional disturbance and was replaced.

Design.--The experimental design was an incomplete 4 X 3 X 4 repeated measures factorial, with four levels of separation from category midpoint (0, 1, 3, and 5 scale units), three different relevant dimensions (Evaluative, Potency, and Activity), and four successive presentations of the list of words (four 60-trial blocks). The design was necessarily incomplete because the lack of extreme words on the Activity dimension.

The lists differed in the degree of scale separation between the category midpoints. There were four degrees of scale separation. The largest was 5 scale units, for which the X category contained words rated between 1 and 2 on the relevant scale, and the NOT-X category contained words rated between 6 and 7. Intermediate scale separations of 3 and 1 scale units were also used. For the separation of 1 scale unit, the X words were rated between 3 and 4, and the NOT-X words were between 4 and 5. The smallest separation was 0 scale units. For this condition, the same words were used as in the one-scale-unit lists, but the words were scrambled so that each category contained words rated between 3 and 5.

The separation of 5 scale units could not be used with the Activity dimension because of the lack of extreme words for that dimension. A slight shortage of extremely bad and extremely soft words also resulted in an imbalance of four and three words, respectively, for the Evaluative and Potency 5-unit separations. The stimulus sequences were arranged so that the imbalance occurred late in the list.

Results and discussion.--Because all Ss were run for the same number of trials, the principal dependent variable was the number of correct responses in each 60-trial block. The incomplete nature of the experimental design necessitated two separate statistical analyses. In the first, the Activity dimension was dropped to create a 4 X 2 X 4 design; in the second, the 5-unit separation was dropped to create a 3 X 3 X 4 design.

Figure 2 shows the mean number of correct responses in each 60-trial block for category separations of 0, 1, 3, and 5 scale units (Evaluative and Potency combined). Increasing scale separation clearly improved performance, $F\ (3,56) = 37.20$, $p < .01$. The Evaluative dimension was

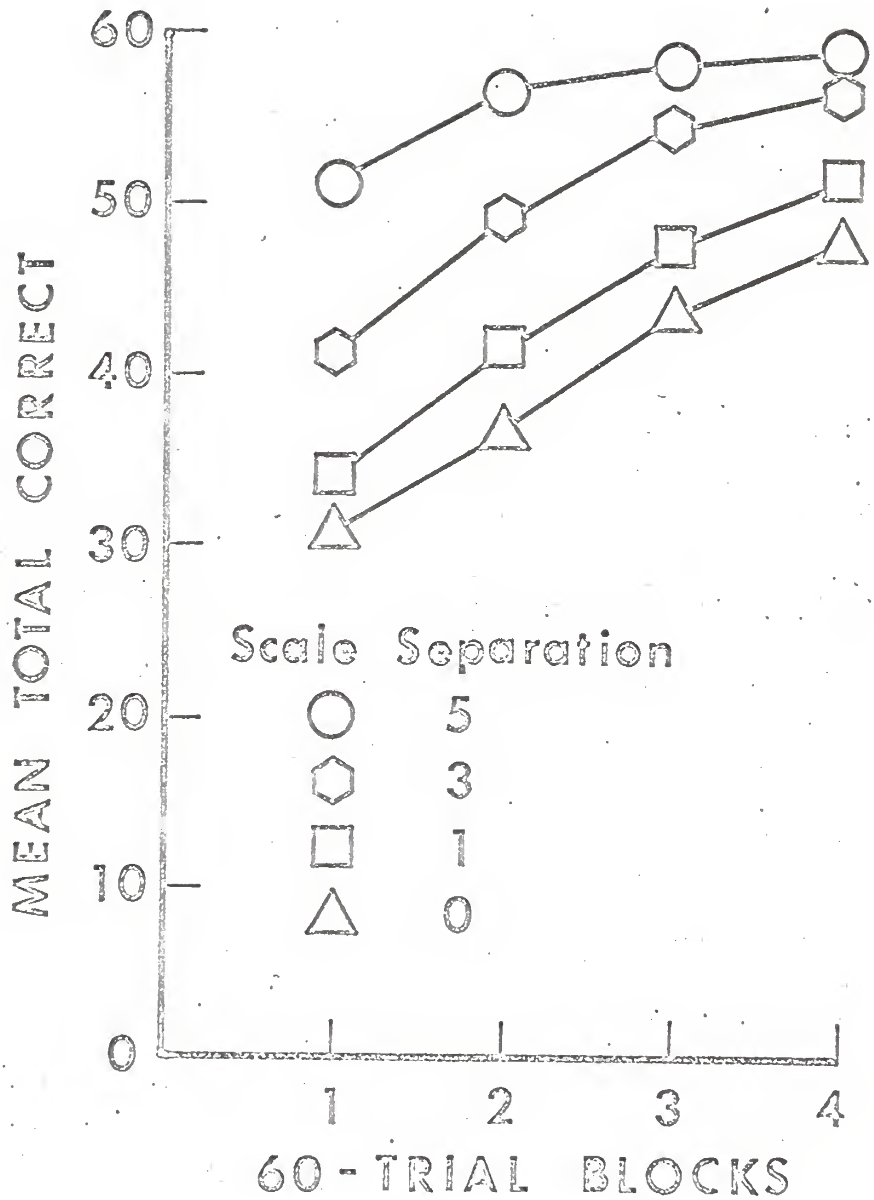

Fig. 2. Mean number of correct responses in each 60-trial block for four levels of category separation (0,1,3, and 5 scale units), Evaluative and Potency dimensions combined.

slightly easier than the Potency dimension, F (1,56) = 5.60, $p < .05$, but there was no interaction between concept and scale separation. Overall improvement over successive blocks was significant, F (3,168) = 165.69, $p < .01$. The interaction of separation and blocks was significant, F (9,168) = 5.24, $p < .05$, but none of the remaining interactions involving blocks was significant. The results of this analysis are shown in Table 1. The difference between the 0 and 1 conditions was significant both by Duncan's Multiple Range Test and by a separate analysis of variance for those conditions, F (1,28) = 9.71, $p < .01$. However, it should be noted that instructions to "find the concept" may have interfered with performance of the 0-separation groups.

In the second analysis, shown in Table 2, increasing separation also improved performance, F (2,63) = 26.13, $p <$ .01, and dimensions again differed significantly, F (2,63) = 4.27, $p < .05$, with Evaluative the easiest and Potency the hardest. The interaction of separation and dimension was also significant, F (4,63) = 4.04, $p < .01$. This resulted primarily from the fact that the original scrambling of the 3-5 Activity words to create the 0-separation list resulted in a bias which made that list somewhat easier to learn than the corresponding 1-separation list. After discovery of the bias, the list was re-scrambled for the remainder of the group; however, the data from all *S*s were retained in the statistical analysis.

Table 1

Summary of Analysis of Variance of Errors, Exp. I.[a]

Source of Varistion	df	SS	MS	F
Between Ss	63			
Separation (S)	3	9,855.71	3,285.23	37.20**
Concept (C)	1	495.06	495.06	5.60*
S X C	3	689.66	229.88	2.60
SS: S X C	56	4,945.07	88.30	
Within Ss	192			
Blocks (B)	3	7,267.71	2,422.50	165.69**
B X S	9	689.96	76.66	5.24*
B X C	3	31.73	10.57	.72
B X S X C	9	282.00	31.33	2.14**
Residual	168	2,457.71	14.62	
Total	255	26,759.00		

**$p < .01$

*$p < .05$

(a) This analysis includes the dimensions of Evaluation and Potency with 0, 1, 3, and 5 degrees of separation.

Table 2

Summary of Analysis of Variance of Errors, Exp. I.[a]

Source of Variance	df	SS	MS	F
Between Ss	71			
Separation (S)	2	3,536.47	1,768.23	26.13**
Concept (C)	2	578.86	289.43	4.27*
S X C	4	1,094.39	273.59	4.04**
SS: S X C	63	4,263.23	67.67	
Within Ss	216			
Blocks (B)	3	12,392.38	4,130.79	281.01**
B X S	6	50.19	8.36	
B X C	6	123.25	20.54	1.39
B X S X C	12	322.99	26.91	1.83
Residual	189	2,778.90	14.70	
Total	287	25,140.66		

**$p < .01$

*$p < .05$

(a) This analysis includes the dimensions of Evaluation, Potency, and Activity with 0, 1, and 3 scale units of separation.

The second analysis also showed overall improvement across blocks to be significant, $\underline{F}$ (3,189) = 281.01, $\underline{p} < .01$. As in the first analysis, none of the interactions involving blocks was significant. Despite the difficulty with the Activity dimension zero list, the difference between the 0- and 1-unit separations was significant, $\underline{F}$ (1,42) = 4.34, $\underline{p} < .01$.

To examine the course of learning of the concepts, performance within the first block of 60 trials was examined. Figure 3 shows performance in 10-trial segments during the first 60-trial block for each of the four scale separations (Evaluative and Potency combined). The rapid improvement at the most extreme separation provides a striking contrast to the essentially flat curve for 0 separation. The results of the analysis of variance are summarized in Table 3. Improvement over blocks was significant, $\underline{F}$ (5,280) = 2.51, $\underline{p} < .05$. The interaction was statistically significant, $\underline{F}$ (15,280) = 3.25, $\underline{p} < .01$.

A word-by-word analysis did not disclose any words which were consistently missed by most or all Ss. This suggests that the population of Ss used in this experiment tends to agree, in general, with the pattern of word ratings found by Jenkins et al. (1958). Although the same order of words was used in each block, there was no evidence of the traditional bowed serial position curve in any block.

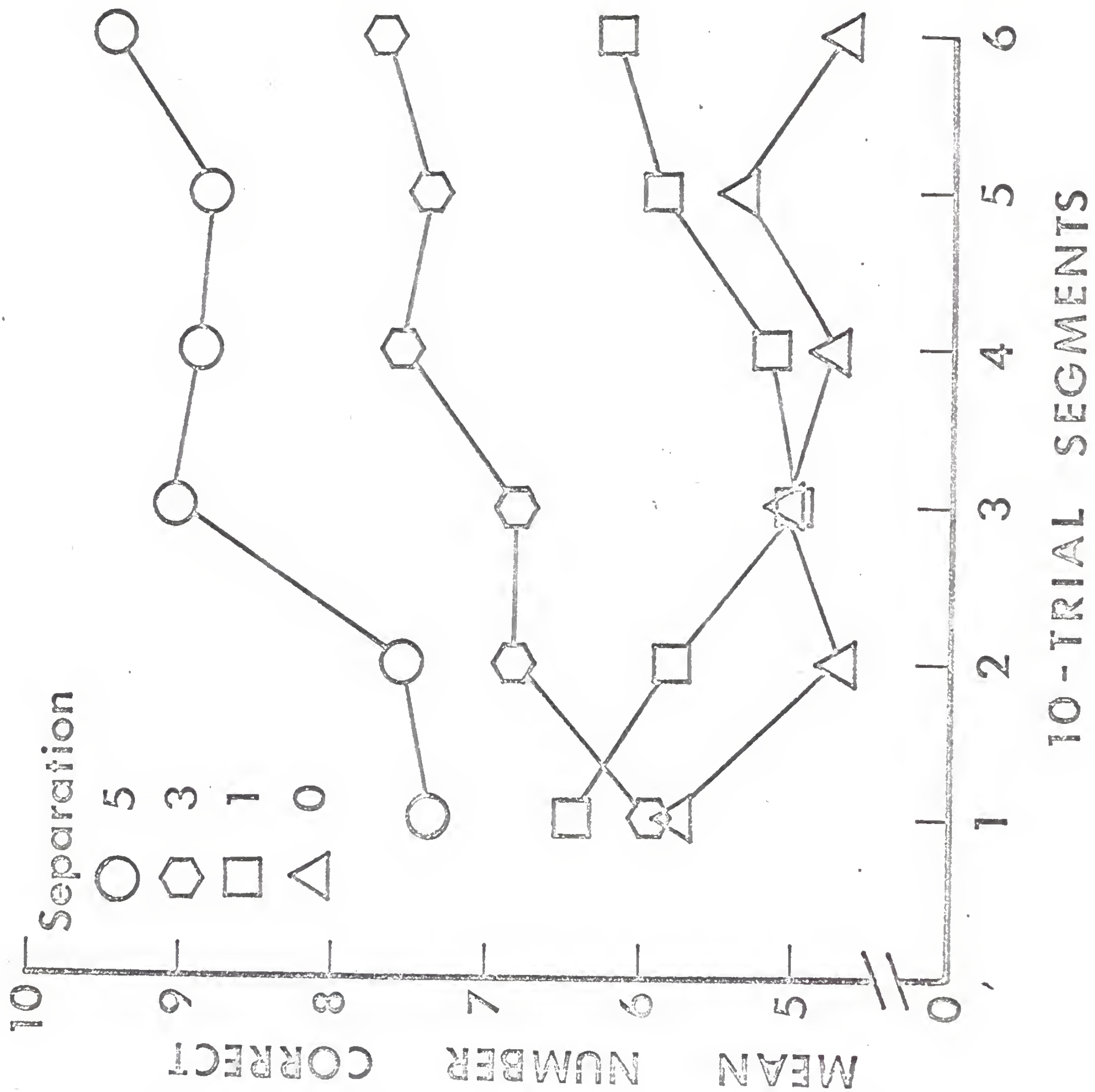

Fig. 3. Mean number of correct responses in 10-trial segments of the first 60-trial block for four levels of category separation, Evaluative and Potency dimensions combined.

Table 3

Summary of Analysis of Variance of Errors, Block 1 of Exp. I.[a]

Source of Variance	df	SS	MS	F
Between Ss:	63			
Separation (S)	3	670.01	233.33	23.66**
Dimension (D)	1	45.37	45.37	4.81*
S X D	3	121.65	40.55	4.30**
SS: S X D	56	528.63	9.44	
Within Ss:	320			
Blocks (B)	5	27.49	5.49	2.51*
S X B	15	106.49	7.10	3.25**
D X B	5	21.32	4.26	1.95
S X D X B	15	58.91	3.92	1.80
Ss X B	280	611.12	2.18	.20
Total		2190.99		

(a) This analysis includes the dimensions of Evaluation and Potency with 0, 1, 3, and 5 scale units of separation.

**$p < .01$

*$p < .05$

CHAPTER IV

EXPERIMENT II

A large part of the benefit of increased category separation obviously results from quicker discovery of the concept. Presumably S's discovery of the correct concept (from among the other possible concepts) is facilitated by the presence of more extreme, and hence more noticeable, examples of the concept. Once the correct concept is learned, however, it is not clear that S will have any difficulty classifying any of the stimuli, except possibly those very close to the center of the relevant SD scale. On the other hand, it seems possible that words farther from the center of the scale (more highly polarized words) may be more easily recognized as exemplifying a known concept. If this is correct, problems in which broad categories are used (e.g. 1-4 vs. 4-7 on the relevant SD scale) should show fewer errors for individual words that are more highly polarized, with the less extreme words "catching up with the extreme words." Exp. II was designed to test this interpretation.

A secondary purpose of Exp. II was to examine the possibility that the benefits of the concept are achieved largely during the first time through the list, and that thereafter, the process is primarily one of rote memorization. This possibility was suggested by the fact that performance curves over successive blocks in Exp. I were essentially parallel for the four degrees of separation; the same result was found by Haygood (1966). A rote memorization interpretation implies that words which are missed during a given block should have the same probability of being classified correctly during the next block regardless of polarization. Specifically, the contingent probability of an error on a word in Block _n_, given that an error was made on the same word in Block _n_ - 1, should be the same for all levels of polarization. Exp. II was designed to compare these contingent probabilities for different levels of polarization. Because errors in Block 1 can also represent words missed because the concept has not yet been discovered, three degrees of pretraining were used. This makes possible a comparison of groups beginning with a high degree of knowledge of the concept with those beginning relatively naive.

Method

Materials, apparatus, task and procedure.--The materials and apparatus were the same as in Exp. I. The task and procedure were also the same as those of Exp. I, with the following exceptions.

First, category width was 3 scale units, with words rated between 1 and 4 contained in the X category and words between 4 and 7 in the NOT-X category. Polarization levels were defined by the distance from the midpoint of the scale. For example, words rated between 3 and 4, and between 4 and

5, were defined as having polarization of 0-1. At the extremes, words between 1 and 2, and between 6 and 7, were defined to have polarization 2-3. The words selected for each category were spread evenly across the scale, so that one third fell in each polarization level, and the sequence of stimuli was constructed so that each 6-trial block contained one word from each of the three levels of polarization in the two categories.

Second, all Ss were given pretraining by being shown a separate list of 20 words (except for the familiarization group to be discussed) with feedback. An Extreme Separation (ES) pretraining group was shown 20 words rated either between 1 and 2 (X) or between 6 and 7 (NOT-X) with the correct category indicated by E. A Narrow Separation (NS) pretraining group was shown 20 words rated either between 3 and 4 (X) or 4 and 5 (NOT-X), also with feedback. A Familiarization (F) group was shown the 60 words of the list that they were to sort without feedback.

Subjects and design.--The Ss were 120 students from introductory psychology classes at Kansas State University, who received class credit for participation. They served in groups of one to four Ss each and were assigned to treatment combinations in stratified random manner.

The experimental design was a 3 X 2 X 3 X 3 repeated measures factorial, with three kinds of pretraining (ES, NS, and F), two relevant dimensions (Evaluative and Potency), three successive 60-trial blocks and three levels of polarization within each category in each list (0-1, 1-2, 2-3).

Results and discussion

The results of the analysis are summarized in Table 4. Pretraining with ES was superior to that with either NS or F, F (2,114) = 6.50, $p < .01$, but the latter two did not differ significantly. The Evaluative dimension was clearly easier than Potency, F (1,114) = 27.63, $p < .01$. The interaction of pretraining and relevant dimension was not significant.

Turning to the within-S variables, polarization was a significant and striking determiner of the ease of classifying words, F (2,912) = 223.80, $p < .01$. Figure 4 shows mean number correct as a function of polarization for the three pretraining groups. The interaction of pretraining and polarization was significant, F (4,912) = 8.08, $p < .01$, as was the polarization X dimension interaction, F (2,912) = 5.55, $p < .01$. The effect of polarization diminished across blocks, F (4,912) = 22.29, $p < .01$, representing a ceiling effect for the more extreme words.

Performance improved across blocks, F (2,912) = 491.54, $p < .01$ and several significant interactions (dimensions X blocks, pretraining X blocks and dimensions X pretraining X blocks) all reflected the "catching up" of the more difficult conditions as the easier conditions approached a ceiling on performance.

Table 4

Summary of Analysis of Variance of Errors, Exp. II.[a]

Source of Variance	df	SS	MS	F
Between Ss	119			
Dimensions (D)	1	487.35	487.35	27.63**
Pretraining (P)	2	271.74	135.87	6.50**
D X P	2	9.96	4.98	
SS: D X P	144	2540.16	17.64	
Within Ss	960			
Separation (S)	2	1204.24	602.12	111.71**
S X D	2	29.86	14.93	2.76
S X P	4	86.98	21.74	4.03**
S X D X P	4	4.14	1.04	
Ss X S	228	1228.72	5.39	
Blocks (B)	2	2644.84	1322.42	267.16**
B X D	2	114.20	57.10	11.54**
B X P	4	102.90	25.72	5.20**
B X D X P	4	75.68	18.92	3.82**
Ss X B	228	1129.12	4.95	
B X S	4	239.86	59.96	260.70**
B X S X D	4	26.42	6.60	28.70**
B X S X P	8	47.58	5.94	25.83**
B X S X D X P	8	20.36	2.54	11.04**
Ss X B X S	456	104.06	.23	
Total	1079			

**p .01
*p .05

(a) This analysis includes the dimensions of Evaluation and Potency with Extreme Separation, Narrow Separation, and Familiarization pretraining.

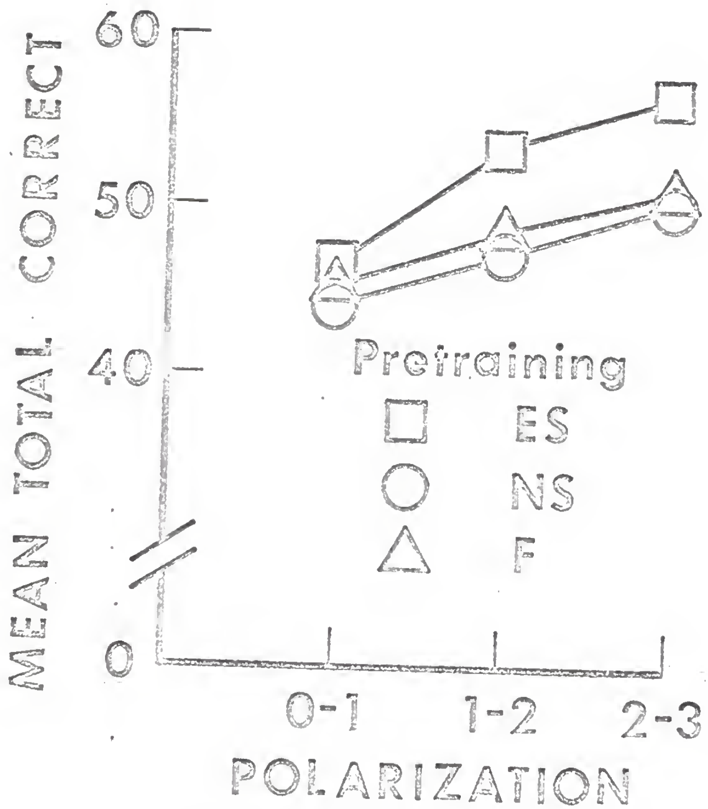

Fig. 4. Mean number of correct responses as a function of word polarization for three different types of pretraining. The maximum number correct at any level of polarization is 60.

The contingent error probabilities for Block 1-Block 2 and Block 2-Block 3 are shown in Table 5. The probabilities for the combined conditions are not radically altered by using unweighted means instead of the weighted combinations shown. Using the conservative assumption that the probabilities are uncorrelated, none of the differences is significant, though the largest (.33 vs. .38) approaches significance, *t* (1407) = 1.84, $p < .10$. Such analyses are not entirely suitable, however, because each *S* makes a different contribution to the probabilities reported. A better analysis would be one in which each *S* was scored separately of the six probabilities, and an analysis of variance computed on these scores. Unfortunately, that analysis would have many missing values, since the better *S*s did not make errors at all levels of polarization. Furthermore, the missing values are heavily biased since they represent the easier levels of polarization, pretraining, and dimension. For groups in which a reasonable number of *S*s did make errors at all levels, the analysis has been computed. The only analysis which produced a significant effect was a Block 1-Block 2 analysis for F and NS groups, Potency dimension only, covering 38 of the 40 *S*s. In this analysis, increased polarization did decrease the contingent error probability, *F* (2,72) = 4.82, $p < .05$.

Table 5

Contingent Probability Analysis[a]

Experiment II

Type of Pretraining	Blocks 1 and 2 Polarization			Blocks 2 and 3 Polarization		
	0-1	1-2	2-3	0-1	1-2	2-3
Familiarization						
Evaluative	.40	.31	.40	.40	.40	.31
Potency	.37	.33	.31	.44	.39	.46
Narrow Separation						
Evaluative	.36	.35	.32	.31	.23	.22
Potency	.44	.39	.39	.35	.48	.37
Extreme Separation						
Evaluative	.33	.44	.37	.49	.32	.14
Potency	.39	.33	.21	.41	.32	.35
Weighted mean of all conditions	.38	.35	.33	.39	.37	.35
Number of errors in earlier block	955	663	454	542	359	273

(a) The figures in this table represent the probability of and error on a word in Block n, given that it was missed in Block n - 1, summed across all words of a given level of polarization for all Ss in a group.

CHAPTER V

DISCUSSION

The results of Experiment I are consistent with those of Haygood (1966) and reaffirm the utility of SD dimensions for the study of concept learning. It is clear from the results that rather small separations will allow Ss to discover and utilize relevant semantic dimensions, and that performance improves continuously with increases in category separation. Thus further confirmation is provided for Archer's (1962) finding that increased separation between levels of the relevant dimension leads to improved performance in a concept-identification task.

The results of Experiment II verified the expectation that the farther a word is from the neutral point of the scale, the more easily it is learned, or recognized as belonging to its proper category. The size of the effect varies somewhat, depending on the relevant dimension, type of pretraining, and degree of learning--an embarrassment of riches in significant interaction effects contributed in large part by the extreme sensitivity of the design and methodology employed.

The contingent probability analysis of Exp. II strongly suggests that the concept is not providing any striking benefit to S after the first time through the list. The results thus indicate that the concept may aid S on an occasional word during the second and third blocks, but that, in general, once S learns the concept, he either recognizes a word as belonging to a category or he does not. He must then memorize the categorization of those words which are not recognized. This finding points up one important difference between the use of visual forms and meaningful words in concept identification. Following solution of a problem involving geometric designs, S makes no further errors except through carelessness. In contrast, with verbal materials errors continue at a minimal rate because there is always the likelihood that S will fail to recognize or agree with the popular semantic meaning of a word.

The above results provide ample evidence that some semantic concepts are easier to learn than others. This is also true of abstract visual materials, even when the same formal logical solution is used. The difficulty of obtaining a solution in a concept formation task may be affected by the negative or positive transfer from previous experience with the material being learned. While there are many possible mediating processes, their mediating function is the same; to provide an indirect association between perception and behavior. As each word is presented to S, it evokes mediators whose individual strength is determined by their distance from the origin. As a mediator is repeatedly evoked, the S comes to recognize it as a solution. Recognition of the correct mediator is affected by the number of words that have been presented and by the degree of word polarization. The amount of mediational information conveyed by a highly polarized word can be equated by using a number of less polarized words. Rote memorization occurs, but apparently only as an adjunct to concept learning, and only to the degree that concept learning fails to provide correct responses.

Any measure that could provide an index of the degree to which a concept instance is governed by a particular dimension or attribute could serve as an organizational scheme in concept-learning studies. That rating of words by the SD technique provides some of these indices is evidenced by the results of this study. The main orthogonal dimensions that are postulated in the model of the semantic space, are Evaluation, Potency, and Activity. Osgood (1957) has found that the Evaluation factor represents approximately twice the variance of either potency or activity and each of these two accounts for about twice the variance of any other factors. There is nothing in the present results to contradict these approximations.

According to Deese (1965), the SD ratings are limited by the selection of scales upon which any concept is judged, the applicability of the properties implicit in any psychological scale with particular anchors, and the bipolar nature of the scale (comparison to something less than 180 degrees). While free associations are not constrained by these features, it would require sampling an entire linguistic universe before dimensionality could be determined by factor analysis or other techniques.

Despite its limitations, the verbal dimensions of the SD can be effectively used in studies of concept learning, thus opening the way for a variety of more meaningful studies of concept learning.

REFERENCES

REFERENCES

Archer, E. J. Concept identification as a function of obviousness of relevant and irrelevant information. *J. exp. Psychol.*, 1962, 63, 616-620

Bourne, L. E. *Human conceptual behavior*. Boston: Allyn and Bacon, Inc. 1966.

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates, *J. gen. Psychol.*, 1953, 49, 229-240.

Coleman, E. B. Verbal concept learning as a function of instructions and dominance level. *J. exp. Psychol.*, 1964, 68, 213-214.

Deese, J. *The structure of associations in language and thought*. Baltimore: John Hopkins Press, 1965.

Haygood, R. C. The use of Semantic Differential dimensions in concept learning. *Psychon. Sci.*, 1966, 5 (8), 305-306.

Jenkins, J. J., Russell, W. A., & Suci, G. J. An atlas of semantic profiles for 360 words. *Amer. J. Psychol.*, 1958, 71, 688-699.

Mayzner, M. S., Tresselt, M. E. A comparison of judgement and associational techniques in developing verbal concept formation materials. *J. Psychol.*, 1961, 51, 331-342.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.

Reed, H. B. Factors influencing the learning and retention of concepts. *J. exp. Psychol.*, 1946, 36, 71-87.

Underwood, B.J., Richardson, J. Some verbal materials for the study of concept formation. *Psychol. Bull.*, 1956a, 53, 84-95.

Underwood, B. J., Richardson, J. Verbal concept learning as a function of instructions and dominance level. *J. exp. Psychol.*, 1956b, 51, 229-238.

A P P E N D I C E S

APPENDIX I

Instructions and Word Lists for Experiment I

This is an experiment to see how well you can learn to classify a list of words. I'm going to show you a list of 60 words one at a time. Each of these words is either a type X or a not X. Your task is to discover what it is that makes a word an X or not X. This will be related entirely to the affectual meaning of the word, feelings towards the word or aroused by it. Sorting will have nothing to do with dictionary definitions, length, first letter, parts of speech, or any unusual characteristics that the slide might have.

Here is how the experiment works. Each time I show you a word, I want you to decide if it is an X or not X and then push the appropriate red button on the control panel. When everyone has done this, either of the two green lamps will come on. You can tell by the way the buttons and the lights line up whether you are right or wrong. In either event, the green light that comes on is the correct answer. The principle may be a little hard to catch on to, and at first you will be guessing. With practice you will be able to predict where some of the words belong without having seen them previously.

We will go through the list four times. Of course you can memorize some of the words, but it is more efficient if you can discover the correct principle. We are more interested in how accurate you can be, rather than how fast you can push buttons. If, on the other hand, you don't know an answer--make a guess. Questions?

GOOD-BAD

3-4/4-5

MALLET*	INTERMITTENT	ARMY*	FLASHY
ABRUPT	BOX*	BOULDER	CONTINUOUS*
BOY*	STIFF	LIQUOR	HIGH*
DEEP*	SNAIL*	RAPID*	LOW
KNIFE	NARROW	HARD	GOJEY
WAGON*	PLAIN*	BOTTOM	FAST*
TRUNK*	BLOCK*	SUDDON*	EDGED
FIRE	RUGGED*	FAR*	FIERY
QUOTA*	HASTY	PUNGENT	DIM
OBSCURE	DREAM*	BODKIN	OBVIOUS*
CURLED*	COILED	BASE*	COARSE
JUMP*	DIRT	SEVERE	POLITICIAN*
CITY*	DARK	HEAVY	WINTER
JOY*	RED*	FROSTY*	ARCHED
SLACK	MILLIONAIRE*	ARGON*	RIGID

2-3/5-6

LIFT*	HUNGRY	ART*	FLEA
AFRAID	BOY*	BLEAK	COP*
CONSTANT*	ROUGH	INFERIOR	GLEAMING*
DELIBERATE*	TALL*	ME*	LAGGING*
INDIFFERENT	LATE	GLOOMY	GLARING
SHINY*	LUMINOUS*	BITTER	FIRM*
RIVER*	BOAT*	RICH*	DELAYED
FEAR	OCEAN*	ELEVATED*	FAT
LUSCIOUS*	GRIEF	MAD*	DANGER
LAZY	DEBATE*	BEGGAR	MILLIONAIRE*
DOOR*	COLD	BODICE*	CLUMSY
HAND*	DEFORMED	PAIN	MAN*
COAL*	CROOKED	HIT	SOUR
HAMMER*	NAIL*	EVEN*	ANGER
PYTHON	MONEY*	ADORNED*	OVERCASTE

1-2/6-7

PIANO*	DAWN*	HEALTH*	RANCID
COURAGE*	THIEF	THINK*	MOSQUITO
ENGINE*	SICKNESS	DEVIL	JUSTICE*
ABORTION	MIND*	PROGRESS*	LAMP*
STAGNANT	FRAUD	TROUBLE	NASTY
WASH*	COMPLETE*	CRIMINAL	RELIGION*
PUTRID	BATH*	BARN*	EAT*
WISE*	HURT	SUCCESS*	INCOME*
TORNADO	RIGHT*	BAD	AMERICA*
WAR	HATE	SIN	DIVORCE
STUDY*	CAR*	TRUTH*	BIBLE*
EFFORT*	TREES*	HOSPITAL*	DOCTOR*
PATRIOT*	DISCOMFORT	RAGE	STOVE*
SLIME	STARVING	BRAVE*	BEAUTIFUL*
HEARTLESS	SCORCHING	STEAL	CHURCH*

* Words that were in the X category.
Note - The word sleep was erroneously substituted for the word slime.

HARD-SOFT

3-4/4-5

INTERMITTENT*	GREEN	AGILE	EVEN
ADORNED	BRILLIANT*	BIRTH	CONTINUOUS*
COMPLETE*	ME	HEAL	HEALTH*
DEEP*	NEGRO*	PATRIOT*	HOT
HAND	INFERIOR	FAMILY	FAITH
SPICY*	MAN*	BEGGAR	FATHER*
ROOT*	BIBLE*	RICH*	DOCTOR
ELEGANT	PLAIN*	FARM*	EATING
MOSQUITO*	FLEA	KITCHEN	DELAYED
JOY	ELEVATED*	ART	LOFTY*
DARK*	CHAIR	BARN*	CANDY
HOSPITAL*	DIM	LEG	MILD*
COURAGE*	CHURCH	GOD	MEMORY
HIGH*	PIANO*	FOREIGNER*	AMERICA
LEMUR	LAMP*	AFRAID*	LAGGING

2-3/5-6

PROGRESS*	LAZY	RAPID*	HAPPY
BATH	CONTROVERSY*	CHARMING	DELIBERATE*
COP*	SEX	MUSIC	LIFT*
FAST*	COMPLETE*	SUCCESS*	NICE
MOTHER	NURSE	HOLY	HARMINOUS
BOAT*	STOP*	CALM	JUSTICE*
TALL*	BOY*	TREES*	GARMENT
GRACEFUL	TRUNK*	INCOME*	GLOW
STUDY*	LADY	PIG	EASY
PEACE	GLEAMING*	BED	SHINY*
EFFORT*	DAWN	BRAVE*	CHILD
OCEAN*	EAT	RIPE	STOVE*
DEBATE*	DUSKY	LAKE	SISTER
MONEY*	SUPERIOR*	LEADERSHIP*	BEAUTIFUL
ROUND	MILLIONAIRE*	WINDOW*	PLIABLE

1-2/6-7

TENSE*	FIRM*	KNIFE*	KITTENS
ENGINE*	ROSE	HIT*	FRAGRANT
HARD*	LOVEABLE	DOUGH	STATUE*
BABY	TABLE*	STIFF*	STREET*
PRETTY	FEATHER	SLEEP	GIRL
BASE*	DOOR*	CUSHION	BOULDER*
GLOVE	BITTER*	BOX*	NAIL*
ANGER*	FLOWERS	SEVERE*	GLARING*
SILK	RIGID*	SOFT	BRISTLY*
SNOW	FLEECY	LOVELY	BODICE
RUGGED*	COAL*	ABRUPT*	BLOCK*
HAMMER*	LIQUOR*	MALLET*	MOUNTAIN*
WAGON*	DOWNY	JELLY	FLOWER
MILD	PUPPIES	CAR	ROUGH*
FAT	LENIENT	RELAXED	CITY*

* Words that were in the X category.

ACTIVE-PASSIVE

3-4/4-5

GREEN*	HOUSE	ARGON*	FLOWERS
ADORNED	BARN*	CANDY	COAL*
BASE*	RIGID	LEPER	FRAGRANT*
DIRT*	PUTRID*	LEMON*	LOW
LATE	MOLD	GLOVE	GARMENT
PIGMENT*	HIGH*	BOTTOM	ELEGANT*
OBVIOUS*	BABY*	NICE*	EVEN
FLEECY	NAIL*	DREAM*	FEATHER
LADY*	GRADUAL	NARROW	DOWNY
MOUNTAIN	DOOR*	BLOCK	HEARTLESS*
COLOR*	COMFORT	ART*	CLUMSY
GLOW*	EASY	PEACE	INTERMITTENT*
CHARMING*	DOUGH	HEAVY	STIFF
FRAUD*	LIZARD*	FOOD*	BED
PIG	HARD*	AFRAID*	OVERCASTE

2-3/5-6

GIRL*	MOON	AGILE*	INFERIOR
BEGGAR	BATH*	CHAIR	COARSE*
BIBLE*	STAGNANT	PLAIN	FRIGHTFUL*
DEVIL*	QUOTA*	JEW*	RELAXED
OBSCURE	SILK	LEISURELY	LAGGING
NASTY*	HATE*	CALM	EDGED*
LIGHT*	BAD*	LEG*	FRIGID
INDIFFERENT	KNIFE*	DIVORCE*	GLOOMY
INCOME*	LINGERING	SLEEP	DREARY
SLACK	DISCORDANT*	BOTTOM	HAND*
CRIMINAL*	DEFORMED	ARCHED*	CUSHION
GLARING*	FAT	TABLE	HEAL*
CHURCH*	DIM	MILD	SOMBER
FIRM*	KITTENS*	FEAR*	BLEAK
STATUE	GOD*	ABORTION*	TRUNK

* Words that were in the X category.

APPENDIX II

Instructions and Word Lists for Experiment II

This is an experiment to see how well you can learn to classify a list of words. I'm going to show you a list of 60 words one at a time. Each of these words is either a type X or a not X. Your task is to discover what it is that makes a word an X or not X. This will be related entirely to the affectual meaning of the word, feelings towards the word or aroused by it. Sorting will have nothing to do with dictionary definitions, length, first letter, parts of speech, or any unusual characteristics that the slide might have.

Familiarization

I am now going to show you all of the words in the list. This will allow you to get an idea of what the words are like.

Pretraining

I am now going to show you some example words and indicate which category they belong in, either X or NOT-X. You will not see these words in the list that you sort, but it will give you an idea of what the list words are like.

Here is how the rest of the experiment works. Each time I show you a word, I want you to decide if it is an X or not X and then push the appropriate red button on the control panel. When everyone has done this, either of the two green lamps will come on. You can tell by the way the buttons and the lights line up whether you are right or wrong. In either event, the green light that comes on is the correct answer. The principle may be a little hard to catch on to, and at first you will be guessing. With practice you will be able to predict where some of the words belong without having seen them previously.

We will go through the list three times. Of course you can memorize some of the words, but it is more efficient if you can discover the correct principle. We are more interested in how accurate you can be, rather than how fast you can push buttons. If, on the other hand, you don't know an answer--make a guess. Questions?

PRETRAINING GOOD-BAD

3-4/4-5		1-2/6-7	
JUMP*	ARGON*	TREES*	LAMP*
RIGID	INTERMITTENT	HATE	THIEF
ARMY*	BLOCK*	STOVE*	DAWN*
PUNGENT	BODKIN	WAR	DEVIL
OBVIOUS*	CURLED*	AMERICA*	INCOME*
NARROW	GOJEY	ABORTION	BAD
RAPID*	BASE*	SUCCESS*	BIBLE*
COILED	FLASHY	HEARTLESS	SCORCHING
CITY*	JOY*	BEAUTIFUL*	COMPLETE*
LOW	DIM	RANCID	HURT

GOOD-BAD MIXED LIST

Polarization

3 MILLIONAIRE*	4 DARK	2 COAL*	4 FIRE
4 HEAVY	3 RED*	3 MALLET*	3 FROSTY*
5 FEAR	1 HEALTH*	4 QUOTA	6 SIN
6 STEAL	4 DIRT	1 BARN*	1 STUDY*
2 BOY*	3 POLITICIAN*	5 CROOKED	2 DOOR*
1 CHURCH*	5 DEFORMED	6 STARVING	5 BEGGAR
5 FLEA	2 BOAT*	4 COARSE	6 SICKNESS
2 ADORNED*	1 MIND*	5 CLUMSY	4 BOTTOM
1 DOCTOR*	6 TROUBLE	3 HIGH*	3 DREAM*
4 BOULDER	5 DELAYED	6 FRAUD	2 EVEN*
3 ARGON*	3 PLAIN*	2 CONSTANT*	6 STAGNANT
6 MOSQUITO	6 DIVORCE	1 BRAVE*	1 JUSTICE*
6 TORNADO	2 BODICE*	5 BLEAK	3 DEEP*
5 DANGER	4 KNIFE	1 BATH*	4 ARCHED
2 ART*	1 ENGINE*	2 DELIBERATE*	5 AFRAID

* Words that were in the X category.

PRETRAINING HARD-SOFT

3-4/4-5		1-2/6-7	
INTERMITTENT*	HIGH*	TENSE*	RUGGED*
ADORNED	GREEN	BABY	MILD
COMPLETE*	COURAGE*	ENGINE*	WAGON*
HAND	ME	PRETTY	FAT
DEEP*	BRILLIANT*	HARD*	FIRM*
ELEGANT	INFERIOR	GLOVE	ROSE
SPICY*	MAN*	BASE*	HAMMER*
JOY	FLEA	SILK	LOVEABLE
ROOT*	MOSQUITO*	ANGER*	TABLE*
LEMUR	LAGGING	SNOW	FEATHER

HARD-SOFT MIXED LIST

Polarization

3 DARK*	4 CHURCH	2 TALL*	4 KITCHEN
4 CHAIR	3 NEGRO*	3 PLAIN*	3 PIANO*
5 BATH	1 RIGID*	4 FAMILY	6 CUSHION
6 FLOWERS	4 BIRTH	1 LIQUOR*	1 BOX*
2 PROGRESS*	3 AGILE*	5 LAZY	2 DEBATE*
1 DOOR*	5 PEACE	6 DOUGH	5 LADY
5 MOTHER	2 EFFORT*	4 BEGGAR	6 SOFT
2 FAST*	1 COAL*	5 SEX	4 ART
1 BITTER*	6 PUPPIES	3 ELEVATED*	3 LAMP*
4 DIM	5 ROUND	6 SLEEP	2 MONEY*
3 HOSPITAL*	3 BIBLE*	2 STUDY*	6 LOVELY
6 FLEECY	6 LENIENT	1 HIT*	1 SEVERE*
6 DOWNY	2 BOAT*	5 NURSE	3 PATRIOT*
5 GRACEFUL	4 HEAL	1 STIFF*	4 LEG
2 COP*	1 KNIFE*	2 OCEAN*	5 DAWN

* Words that were in the X category.

EFFECTS OF DEGREE OF CATEGORY SEPARATION AND WORD POLARIZATION ON SEMANTIC CONCEPT IDENTIFICATION

by

CURTIS L. TAYLOR

B. S., Kansas State University, 1966

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1967

The first of two experiments demonstrated that the ease of identification of semantic concepts improves continuously as category separations increase from 0 separation (completely overlapping categories) to 5 units on the relevant Semantic Differential scale. The second experiment used categories consisting of entire halves of the relevant dimension, and demonstrated that increased distance of an individual word from the center of the semantic scale leads to improved performance on that word. Further analysis suggested that most of the benefit of the concept is achieved during the first and second presentations of the list of words, and that improvement thereafter is primarily the result of rote memorization. The use and limitations of Semantic Differential ratings of verbal materials was discussed. The Semantic Differential provides a potent technique for assessing dimensions in verbal concepts and their degree of saliency.