

THE EXPERIMENTAL INDUCTION OF  
RESPONSE SET

by

WILLIAM BAHRET EDDY

B. S., Kansas State College  
of Agriculture and Applied Science, 1955

---

A THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE COLLEGE  
OF AGRICULTURE AND APPLIED SCIENCE

1957

LD  
2668  
T4  
1957  
E32  
C.2  
Document

TABLE OF CONTENTS

INTRODUCTION . . . . .	1
METHOD . . . . .	6
Experimental Materials. . . . .	6
Subjects. . . . .	10
Procedure . . . . .	10
Experimental Design . . . . .	11
RESULTS AND DISCUSSION . . . . .	13
SUMMARY . . . . .	20
ACKNOWLEDGMENT . . . . .	23
REFERENCES . . . . .	24
APPENDIX . . . . .	25

## INTRODUCTION

Those concerned with measuring human behavior with testing devices realize that the score which an individual receives on a test is not determined by his reaction to the content of the test items alone. Other factors which the individual brings with him to the test situation, or which are inherent in the test situation may influence his responses. The term response set has been used to define these phenomena which influence an individual's responses to items in a test. According to Cronbach (7) "response set is a tendency causing a person consistently to give different responses to test items than he would when the same content is presented in a different form". Content describes the actual substance of the test item itself. Form defines everything else present in the test situation. Most tests are designed to measure reaction to some particular content. If factors in the form of the test situation influence a subject's response, a response set is operating and the test score reflects something other than reaction to the content of the test items.

This investigation was carried out to determine the effect of a particular kind of form in influencing the responses of subjects to items in a rating scale test.

Specifically, it was designed to determine the effect of references to hypothetical norms supposedly obtained on subjects similar to those actually taking the test. Additionally, the study provided for a retesting of the experimental subjects in a context of altered form, i.e., references to norms different from those initially referred to in the first session. It was hypothesized that a phenomenon of social acquiescence would be operant in the test situation and would modify the "pure" response to content.

Investigators in the field of psychological measurement are aware of the effect of the form of the test situation in influencing the responses of subjects to items in various personality, interest, attitude, and ability tests. Studies (2, 9) have shown, for example, that response set is pronounced in several commonly used tests utilizing fixed response categories (Agree-Uncertain-Disagree, or Yes-?-No). It has been found that some individuals respond with one category more than with others, especially if the items are unclear or ambiguous. If a subject tends to use the "agree" category more than the other categories when he is actually uncertain as to how to respond, his test score will, of course, be distorted in that direction.

Many studies concerned with investigating various measuring devices have yielded information pertaining directly or indirectly to response set. Cronbach (7, 8) pioneered the organization of the data from these studies. Cronbach identifies the following kinds of response set: (1) tendency to gamble; caution versus incaution; (2) definition of judgment categories; (3) inclusiveness; (4) bias and acquiescence; (5) speed versus accuracy; and (6) response sets on essay tests (brevity, style, etc.).<sup>1</sup> These widely different response sets have been shown to be important factors in influencing subjects' responses on various kinds of tests. Cronbach emphasizes the importance of accounting for response set when evaluating data acquired through the use of testing devices.

In order to learn more about the nature of response set, other

---

<sup>1</sup> A complete discussion of these six categories of response set can be found in Cronbach (7).

investigators have developed scales designed to isolate it. Fricke (9) investigated the tendency of subjects to respond "yes", "true", or "agree" to test items. Using scores on his Opinions, Attitudes, and Interest Survey he developed a scale consisting of 69 statements in which a "true" answer was predictive of the criterion; and to which 40 to 60 per cent of the subjects tested marked true. Individuals receiving a high score on this "Set T" scale were described as having a strong tendency to answer true. Since response set is presumed to act to invalidate test results Fricke suggests that Set T be considered a suppressor variable by taking the subject's response set into account when evaluating the test results. The point is made that an imbalance of questions that require either a positive or negative answer to identify a particular characteristic should be avoided since the score of an individual with a tendency to use one of the response categories more than the others may be substantially affected.

Berg (5) devised two tests to measure tendencies to respond with the extreme position choice categories. By asking subjects to indicate the degree to which they liked or disliked abstract geometrical figures (the Perceptual Reaction Test) and meaningful words (the Word Reaction Test) he was able to measure tendency to respond on the extreme high or low ends of the choice category scale. The correlation between the two tests and between different administrations of the same test indicated that this particular response set was fairly stable within individuals. Significant sex, race, and anxiety level differences were found, indicating that extreme response set scores reflected certain personality and group characteristics.

In a later paper Berg (6) further investigated the relationship between response set and personality, and developed a "deviation hypothesis". According to Berg, individuals with abnormal, i.e., deviant, personality characteristics tend to deviate from a "norm" established by normal subjects on a large number of testing devices. Several studies were cited which support this hypothesis. In one of them Barnes (2) reported that the tendency for subjects to give atypical answers to items on the Minnesota Multiphasic Personality Inventory was significantly correlated with the clinical scales in that testing device. Barnes (3) also found that response set as measured by Berg's Picture Reaction Test was related to psychiatrically diagnosed personality disorder.

Bass (4) developed a scale designed to measure response set to comply with social norms. He asked subjects to express an opinion of agreement, disagreement, or undecidedness concerning the validity of 56 well known but ambiguous proverbs. Those individuals who tended to accept or agree with most of the proverbs were judged to be more prone to accept or agree with the things they believed to be accepted by their society. The person who earned a high score on the Bass Scale was described as an "unquestioning conformer to social demands". Bass uses the term social acquiescence to describe this tendency to comply or conform to social norms. Respondents did, of course, vary in their degree of acquiescence as measured by the Bass Scale. This variation between degrees of acquiescence was related to level of education, sociability, social sensitiveness, and socio-economic status.

Sherif (11) in a non-test type of situation investigated the effect of group responses in their influence on the responses of a particular

subject. The situation was one in which the autokinetic effect could be observed.<sup>1</sup> Sherif's subjects, when placed alone in a room and asked to estimate the distance which the light seemed to move, tended to establish a frame of reference for themselves and estimate the movement of the light in terms of that frame of reference. Mean estimates by individual subjects varied from a fraction of an inch to several inches. Something different took place, however, when the subjects were placed in the dark room in groups. As each subject became aware of the estimates of the group, the estimates of all the members tended to become much more similar. From this study Sherif asserted that the group norm is important in influencing the reactions of group members. Later studies have supported Sherif's hypothesis.

Asch (1) investigated the behavior of individuals who found themselves in opposition to the opinion of the majority of their group. Groups of subjects were asked to match the length of a given line with one of three unequal lines. All but one of the subjects in each group were previously instructed to respond with incorrect - and unanimous - judgments. The one remaining member suddenly found himself a minority of one. There was a strong tendency for subjects to respond in agreement with the group, even though the majority opinion seemed obviously contrary to fact. Also, those subjects who did not revise their estimates toward the group norm experienced anxiety and discomfort about deviating.

The above studies by Sherif and Asch show that individuals tend to

---

<sup>1</sup> The autokinetic effect is one in which a stationary point of light in a darkened room is perceived to move.

conform to what they believe to be the group norm. This tendency to conform is similar to that described by Bass as social acquiescence. Bass's studies have pointed out that particular individuals in a test situation possess different degrees of acquiescence, while studies by Sherif and Asch emphasize acquiescence as a more general phenomenon capable of influencing the behavior of all, or at least most of the members of a group.

Acquiescence may be interpreted as a response set since it is a situational variable independent of the "real" stimulus. Asch's subjects did not acquiesce to the length of the lines, but to the group norm implied in the responses of other group members.

In the present study it is hypothesized that a response set to acquiesce may influence the responses of groups in a particular test situation if cues are present which indicate a group norm. Subjects who receive tests with the same item content but with instructions implying different group norms should differ in their responses. Furthermore, if acquiescence is operating, subjects receiving differing suggested group norms in two separate tests with identical content should give different responses — assuming, of course, that such identical content is not perceived as such.

#### METHOD

##### Experimental Materials

Cronbach (7, 8) suggested that the higher the degree of ambiguity, the more chance there is for response set to act. This is because, in a situation in which right or wrong, or good or bad are not clearly evident, the subject is forced to rely almost completely upon his own "best guess".

The less the amount of information the subject has about the material to which he is asked to react, the more he has to work from an internal frame of reference. Any cues from the outside, however small, may tend to have an important effect on the subject's responses. An attempt was made, therefore, to develop a measuring device in which the test items were highly ambiguous. Also, as much of the content as possible was eliminated from the measuring device. A characteristic of many of the tests previously used to measure response set is their relatively high degree of content. It is very difficult to separate response set from actual reaction to the content of the test items. Cronbach suggests that the use of nonsense syllables might be an effective way of satisfying the criteria of ambiguity and lack of content. This suggestion was utilized in the present study.

The Test. A list of 60 nonsense syllables was obtained from Glazes's (10) lists of nonsense syllables with zero and seven per cent association value. The only criterion for selection was that an attempt was made to use syllables with a variety of beginning letters. The order in which the syllables were listed on the reaction sheets was determined with the aid of a table of random numbers. The syllables were placed on sheets in such fashion that the subjects could indicate with a check mark their reaction to each of them. (See Appendix for an example of the reaction sheets used.)

Subjects were asked to respond to each of the syllables according to the amount of "psychological value" each syllable had for them. The term psychological value was used in an effort to gain more ambiguity. Subjects were not told whether psychological value had to do with association value, emotional tone, sound, or any other possible means of evaluation. The term

"nonsense syllable" was never used; subjects were asked to react to a series of "letter groups".

Written Instructions. All subjects were asked to react to the same set of letter groups. In order to measure acquiescence the form was varied. Form, in this case was a reference in the written instructions to a social norm, i.e., the way in which other subjects had responded to the letter groups. Approximately one third of the subjects were given written instructions which contained the statement that other students had rated the letter groups as having high psychological value, one third received instructions containing the statement that other students had rated the letter groups as having low psychological value and one third of the students received written instructions which contained no information concerning how other students had rated the letter groups. All test forms were identical with the exception of one sentence near the middle of the instruction sheet. The instructions administered to one of the groups are given below.

#### INSTRUCTIONS

—Please Read Carefully—

On the following pages you will find a series of 60 three-letter groups. We are interested in determining the degree of personal psychological value of each of these groups for standardization purposes here at Kansas State.

For each item — that is, for each three-letter group — we would like you to indicate with a check mark in the appropriate column the actual psychological value that item has for you. Your rating on each item may be either Very Low or Low or Slightly Low or Average or Slightly High or High or Very High.

Look at each item carefully. Then rate that item accordingly, depending upon the degree of psychological value it has for you. We are interested in determining how you personally evaluate the items.

When you have read and understood the instructions, you may turn to the next page and look at the letter groups there. However, do not begin rating the items until you are instructed.

If you have any questions, hold up your hand.

The instructions shown are "neutral" instructions (designated Form N). In this form the subjects were given no information about how other students had rated the letter groups. Another form (designated Form H) contained the sentence, "This list of letter groups has been rated by other students and has been found to have overall high psychological value." This sentence was inserted in the third paragraph of the instructions between the sentence ending "...psychological value it has for you." and the sentence beginning, "We are interested in determining...". The third form (designated Form L) contained the sentence "This list of letter groups has been rated by other students and has been found to have overall low psychological value." This sentence was inserted at the same place as the sentence in Form H discussed above.

Written instructions for the second testing session were the same as for the first except for the statement that "this survey is similar to the one you were asked to participate in last week." The syllables were re-randomized for the second session and then arranged so that syllables which appeared first and last on the first test forms did not appear in the same position again. Subjects were not told that the syllables used were the same as in the first session.

Oral Instructions. The same oral instructions were used for both sessions and read as follows:

We are asking you to assist us with a research project. We are passing out to each of you a booklet. This is not a test of any kind. It has nothing to do with your grade in this course. What we want is your honest reaction to a series of letter groups.

Please read the directions carefully. Then wait for instructions to begin. There is no time limit; take your time. You will be able to finish in a few minutes. Please look only at your own paper; what we want is your own honest opinion.

When you finish, close the booklet and wait quietly for everyone to finish.

Are there any questions?

You may begin.

### Subjects

Subjects were students in four General Psychology classes at Kansas State College which met at different times during the same day. The first session consisted of testing each of the four classes on the same day. Two hundred and forty-six subjects were present for the first session. The second session consisted of testing the same four classes five days later. There were 229 subjects present for the second session. Forty-eight tests were discarded because subjects were present for the first session, but not for the second; 34 tests were discarded because subjects were present for the second session, but not for the first. The tests from three subjects were discarded because the subjects appeared to have toyed with the response sheets by making symmetrical designs with the check marks. One subject used a false name and two other subjects did not give their names. There were 116 men and 74 women in the final sample of 190.

### Procedure

In the first session the three test forms (Forms H, N, L) were stacked alternately (H-N-L-H-N-L) and passed down rows in the four classes;

every third person receiving the same test form. In the second session the forms were stacked in alternate groups of six of the same form (H-H-H-H-H-N-N-N-N-N-L-L-L-L-L-H-H-H-H-H) and again passed down rows; the forms were thus distributed among the subjects in blocks of six of the same form. In each session one third of the subjects should have received Form H. For each group of subjects which received a particular form during either session, one third should have received the same form during the other session, one third should have received another form and another third should have received still another. For instance, one third of the subjects which received Form L during the second session should have received Form L during the first session, one third should have received Form N, and one third should have received Form H. This did not hold exactly true because some subjects were absent during one of the sessions, some subjects shifted seat positions, and because it was not always possible to adhere precisely to the prescribed method of distribution. However, the sampling procedure worked sufficiently well to yield a sample of subjects for each of the nine possible session-to-session combinations of test forms.

#### Experimental Design

Table 1 shows the total number of subjects in each of the groups in the first and second sessions. Groups in each session are designated by the test form which they received. That is, a group which received Form L is called Group L. Group L ("low" instructions) in the first session may be seen to contain 67 subjects. Group H ("high" instructions) in the second session contains 59 subjects.

Subjects who received a particular set of instructions during the

first session may have received any one of the three sets of instructions during the second session. The "subgroup" of subjects which received Form L ("low" instructions) during the first session and Form N ("neutral" instructions) during the second session was designated Subgroup L-N. The subgroup which received Form H ("high" instructions) during the first session and Form L during the second session was designated Subgroup H-L. The first letter in the designation refers to the test form received during the first session and the second letter refers to the test form received during the second session. This system was used to assign a designation to each of the nine subgroups. The subgroups and number of subjects in each are shown in Table 1. Subgroup L-L may be seen to contain 16 subjects; subgroup L-N contains 30 subjects.

Table 1. Groups which received each test form and subgroups which received each combination of test forms

		<u>Second Session</u>		
		Group L N = 64	Group N N = 67	Group H N = 59
First Session	Group L N = 67	Subgroup L-L N = 16	Subgroup L-N N = 30	Subgroup L-H N = 21
	Group N N = 62	Subgroup N-L N = 27	Subgroup N-N N = 21	Subgroup N-H N = 14
	Group H N = 61	Subgroup H-L N = 21	Subgroup H-N N = 16	Subgroup H-H N = 24

The design of this experiment made it possible to make the following comparisons: (1) difference between groups which received different test forms (instructions) during the first session; (2) differences between

groups which received different test forms during the second session;  
 (3) differences in change in response between sessions for subgroups which received different order-form combinations of test, i.e., differences in change in response between sessions for the nine subgroups shown in Table 1.

#### RESULTS AND DISCUSSION

As described earlier, subjects indicated the degree of psychological value each letter group had for them. Ratings were made on a seven point scale with scale values assigned to the response categories as shown in Table 2.

Table 2. Response categories and assigned scale value

Response Category	Scale Value
Very Low . . . . .	1
Low . . . . .	2
Slightly Low . . . . .	3
Average . . . . .	4
Slightly High . . . . .	5
High . . . . .	6
Very High . . . . .	7

The "basic" or "raw" score obtained for each experimental category was the mean response value per stimulus (letter group). For example in the group of 67 subjects that received "low" instructions during the first session, the mean response per item was obtained by adding the assigned scale values for all responses made by all subjects in this group and dividing by 4020 (67 subjects multiplied by 60 responses per subject).

Table 3 shows the mean response per item for each of the groups which received different instructions during the first session. This table shows that the group receiving high instructions rated the items as having the highest psychological value, while the group receiving low instructions rated the items lowest.

Table 3. Mean response per item for groups during the first session

Group	N	Mean Response Per Item
L	67	2.22
N	62	2.81
H	61	3.01

A between-within analysis of variance was carried out on the three groups of subjects, differentiated as described, in the first session. The results of this analysis are summarized in Table 4. The resulting F-ratio of 14.84 which is significant at the .01 level of confidence indicates that the difference in mean responses between the three groups is significantly greater than could be expected by chance, i.e., there are real differences between the L, N, and H groups in terms of mean response to the same content. The attempt to induce a response set was obviously successful. The significant differences in average response per item between the three groups in the first session indicates the importance of the form of the test situation in influencing subjects' responses.

Table 4. Summary of analysis of variance for mean response scores in first session

Source of Variation ;	df	Mean Square	F-Ratio
Between H,N,L Groups	2	39775.09	14.84**
Within H, N, L Groups	187	2690.08	
Total	189		

\*\*p < .01

Before discussing separately the results of the second half of this study, the following correlations are reported to indicate the degree of consistency between subjects' average responses on the first test and their average responses on the second test. Correlations for each cell in the design were computed separately based on the respective  $N_s$  for the different cells.

The magnitude of the relations between scores in the first and second sessions indicates significant tendencies for group shifts with different suggested norms and significant tendencies for group stability with the same suggested norms.

Table 5. Product moment correlation coefficients between scores in the first and second sessions

		<u>Second Session</u>		
		L	N	H
First Session	L	.84	.88	.64
	N	.80	.85	.52
	H	.72	.88	.79

Table 6 shows the mean response per item for each subgroup for both sessions and the change in mean response per item between sessions.

Table 6. Mean response per item.

		<u>Second Session</u>						Weighted Means	
		Group L :		Group N :		Group H :			
First Session	Group L	S <sub>1</sub> <sup>*</sup>	2.10	S <sub>1</sub>	2.14	S <sub>1</sub>	2.42	S <sub>1</sub>	2.22
		S <sub>2</sub>	2.15	S <sub>2</sub>	2.03	S <sub>2</sub>	2.64	S <sub>2</sub>	2.25
		D	.05	D	-.11	D	.22		
	Group N	S <sub>1</sub>	3.06	S <sub>1</sub>	2.77	S <sub>1</sub>	2.44	S <sub>1</sub>	2.81
		S <sub>2</sub>	2.69	S <sub>2</sub>	2.74	S <sub>2</sub>	2.74	S <sub>2</sub>	2.72
		D	-.37	D	-.03	D	.30		
	Group H	S <sub>1</sub>	3.05	S <sub>1</sub>	3.00	S <sub>1</sub>	2.99	S <sub>1</sub>	3.01
		S <sub>2</sub>	2.76	S <sub>2</sub>	2.85	S <sub>2</sub>	3.03	S <sub>2</sub>	2.87
		D	-.29	D	-.15	D	.04		
	Weighted Means	S <sub>1</sub>	2.82	S <sub>1</sub>	2.54	S <sub>1</sub>	2.66		
		S <sub>2</sub>	2.58	S <sub>2</sub>	2.45	S <sub>2</sub>	2.82		

\*S<sub>1</sub> refers to the first session, S<sub>2</sub> refers to the second session, and D refers to the difference or change between sessions which was obtained by subtracting S<sub>1</sub> from S<sub>2</sub>.

In Table 6, subgroups which changed least between sessions were Subgroups L-L (.05), N-N (-.03) and H-H (.04). This is not surprising since each of these subgroups received the same instructions in both sessions. Subgroups which changed most in mean response per item were (1) the Subgroup H-L which received "high" instructions during the first session and "low" instructions during the second session (-.29); (2) the Subgroup L-H which received "low" instructions during the first session and "high" instructions during the second session (.22); ( 3 and 4 ) the

Subgroups N-H and N-L which received neutral instructions during the first session and either "high" or "low" instructions during the second session (.30 and -.37). Subgroups which received "high" or "low" instructions during the first session and neutral instructions during the second session (Subgroups H-N and L-N) changed to a small degree. In general, instructions in the second session which induce a set different from that induced by the instructions in the first session seem to have the greatest effect in changing response.

It may be further pointed out that all subgroups which received "high" instructions during the second test session (Subgroups L-H, N-H, H-H) responded with a higher mean response per item. All subgroups which received "low" or "neutral" instructions in the second session responded with a lower mean response per item (except for Subgroup L-L). These trends seem to re-emphasize the importance of the second set of instructions. As previously noted, the subgroups which received neutral instructions during the second session decreased in mean response per item. This might lead to the hypothesis that the neutral instructions were not really neutral in that when received during the second session they influenced the mean response per item downward. The inherent lack of content in the test items may have in itself constituted a kind of content. Subjects who received no information about how the letter groups had been rated by other subjects may have responded to the lack of content by rating the items low. An examination of the mean score per test item as given in Table 6 shows that the range of responses for all subgroups is between 2.03 and 3.05. Since the possible range is between 1 and 7, it can be seen that the mean response per item for all subgroups in both sessions fall toward the low end

of the scale. This indicates that all subgroups, not just those receiving "neutral" instructions, were affected to some degree by the content or lack of content inherent in the test items. The increase in mean response per item for subgroups which received high instructions during the second session may not fully indicate the degree of the influence of the induced "high" response set since these instructions had to overcome not only the previously induced set but the low content of the test items. This point might be tested by asking subjects to react to nonsense syllables with a known degree of high association value.

Since the subjects were tested twice in order to study the effect of a change in induced set, an analysis of variance was carried out to test the significance of the differences between subgroups in amount of change in mean response per item between sessions. These changes in mean response per item are shown in Table 6 and were discussed above. Results of the analysis of variance are summarized in Table 7.

Table 7. Summary of analysis of variance for difference in mean response

Source of Variation	df	Mean Square	F-Ratio
Between subgroups which received the same test form in the first session	2	1465.32	1.12
Between subgroups which received the same test form in the second session	2	9352.01	7.11**
Interaction	4	1563.84	1.19
Replication	181	1314.65	
Total	189		

\*\*p < .01

The  $F$ -ratio for the amount of change in mean response per item between groups which received the same set of instructions during the first session is 1.12 which is not statistically significant at the .05 level of confidence. This lack of significance indicates that the difference in change in response between groups which received the same test form during the first session is no more than could be expected by chance.

The analysis of variance for the data of the first session showed that there were significant differences between the groups (or sets of 3 subgroups) which received different test forms in the first session, and that these differences were the result of the differences in instructions. It can be seen that subgroups which received the same set of instructions during the first session differed significantly in mean response per item, but that the difference in the amount which they changed in mean response per item from first session to second session was not significant.

The  $F$ -ratio for the amount of change in mean response per item between the three sets of subgroups which received the same instructions during the second session is 7.11 which is statistically significant at the .01 level of confidence. These results indicate that the difference in change in mean response per item within sets of three subgroups which received the same instructions during the second session is more than can be accounted for by chance. The instructions given during the second session must have accounted for the difference between the first and second sessions.

The  $F$ -ratio for interaction is 1.19 which is not significant at the .05 level of confidence.

The results of the above analyses indicate clearly that the change in mean response per item between sessions for the subgroups was the result

of the instructions in the test form. The "set" induced by the instructions in the first session did not persevere enough to significantly effect the changes in response between sessions. It has been shown that it is possible to change response set by changing the form of the test situation. The "tendency to conform" discussed by Sherif seems to be an important factor in influencing responses of groups of subjects to ambiguous test items.

#### SUMMARY

This study was carried out to determine the effect of suggested hypothetical "social norms" on the responses of subjects to ambiguous items in a rating scale test. The experiment was designed to answer the following questions. (1) Will subjects receiving test forms containing the same item content but instructions suggesting differing evaluations of the items in terms of social norms differ in their responses to the items? (2) Will subjects who evaluate the items in a rating scale test in terms of a suggested social norm change their evaluation of the items in a later test session if a different norm is suggested?

Three test forms were developed to measure the effect of suggested social norms on the responses of groups of subjects. The three test forms contained exactly the same content, 60 nonsense syllables with low association value. Subjects were asked to respond to the nonsense syllables by indicating the degree of "psychological value" each syllable contained for them on a seven point scale from very low to very high. The written instructions were varied so that those contained in one test form stated that other students had rated the syllables as having low psychological

value; those contained in another form suggested that students had rated the items as having high psychological value; and those contained in the third form contained no statement concerning how other students had rated the items.

Subjects were tested twice with a five day interval between sessions. The sampling procedure made it possible to measure the change in response between sessions for each particular subgroup which received one of the three test forms during the first session followed by one of the three test forms during the second session.

The three groups of subjects which received different test forms i.e., instructions pertaining to social norms, during the first session differed significantly in their evaluations of the test items.

There were marked positive correlations between subjects' scores in the two test sessions. Those correlations indicated a consistency in subjects' responses to the test items in the test sessions.

Analysis of variance techniques showed that groups which received different sets of instructions in the first session did not differ significantly in the way they changed their responses in the second session. Groups which received different instructions during the second session did differ significantly in their change in response to the test items. These results indicate that the instructions given during the first session had no statistically significant effect on the degree to which subgroups changed their responses between test sessions; and that the instructions given during the second session were very important in determining the degree to which subgroups changed their responses from the first to the second session.

Several trends concerning the direction and degree to which subgroups changed their responses between test sessions were pointed out. Subgroups which received the same instructions during both sessions changed very little in their responses to the test items. Subgroups which received opposite instructions (high followed by low, or low followed by high) changed considerably as did subgroups which received neutral instructions in the first session and either high or low instructions in the second session. Subgroups which received neutral instructions during the second session changed very little from their responses in the first session.

The results of this study show conclusively that a response set can be induced in groups of subjects, and that this response set can influence the responses of the subjects to test items. Further, this response set and its effect can be changed by changing the form of the test situation, i.e., the way in which the test items are presented.

The tendency of individuals to conform to what they believe to be the group norm has been shown by other investigators to be a personality characteristic common to most individuals. The present study has shown that this tendency, which may be called social acquiescence, can affect the responses of groups of subjects to test items. It is apparent that those concerned with measuring human behavior with testing devices must be concerned not only with the effect of various response sets in influencing the responses of particular individuals in particular situations, but also they must be concerned with the effect of social acquiescence response set in influencing the responses of groups of subjects in any situation which may contain cues concerning group norms relevant to the test items.

## ACKNOWLEDGMENT

The author wishes to express his sincere gratitude to Dr. Lowell W. Schipper whose counsel and encouragement made possible the development of the present study from its conception to the completed thesis. Appreciation is also expressed to Dr. Arthur H. Brayfield who posed the questions which lead to the study. The author is also indebted to the other members of staff of the Department of Psychology for advice given on many points, and for allowing students in their classes to be used as subjects. Further gratitude is extended to the undergraduate and graduate students who helped with the administration of the test form and other details connected with the study.

## REFERENCES

- (1) Asch, S. E.  
Effects of group pressure upon the modification and distortion of judgments, Groups, Leadership, and Men, Pittsburgh: Carnage Press, 1951, Edited by H. Guetzkow
- (2) Barnes, E. H.  
Response bias and the MMPI, Journal of Consulting Psychology, 1955, 40, 1956
- (3) Barnes, E. H .  
The relationship of biased test responses to psychopathology, Journal of Abnormal and Social Psychology, 1955, 51, 286-290.
- (4) Bass, B. M.  
Development and evaluation of a scale for measuring social acquiescence, Journal of Abnormal and Social Psychology, 1956, 53, 296-299.
- (5) Berg, I. A.  
The reliability of extreme position response sets in two tests, Journal of Psychology, 1953, 36, 3-9.
- (6) Berg, I. A.  
Response bias and personality; the deviation hypothesis, Journal of Psychology, 1955, 40, 61-72.
- (7) Cronbach, L. J.  
Response sets and test validity, Educational and Psychological Measurement, 1946, 6, 475-495.
- (8) Cronbach, L. J.  
Further evidence on response sets and test designs, Educational and Psychological Measurement, 1950, 10, 3-31.
- (9) Fricke, B. G.  
Response sets as a suppressor variable in the OAS and MMPI, Journal of Consulting Psychology, 1948, 12, 20-31.
- (10) Glaze, J. R.  
The association value of nonsense syllables, Journal of Genetic Psychology, 1928, 35, 255-267.
- (11) Sherif, M.  
A study of some social factors in perception, Archives of Psychology, 1935, 36, 3-9.

APPENDIX

	Very Low	Low	Slightly Low	Average	Slightly High	High	Very High
1. <u>GAH</u>							
2. <u>MEF</u>							
3. <u>VAF</u>							
4. <u>VUK</u>							
5. <u>XAD</u>							
6. <u>QIF</u>							
7. <u>JID</u>							
8. <u>KEZ</u>							
9. <u>BOF</u>							
10. <u>JUH</u>							
11. <u>QAD</u>							
12. <u>TOV</u>							
13. <u>ZIK</u>							
14. <u>GIX</u>							
15. <u>VUF</u>							
16. <u>FEH</u>							
17. <u>JEQ</u>							
18. <u>ZUB</u>							
19. <u>XEH</u>							
20. <u>GUK</u>							
	Very Low	Low	Slightly Low	Average	Slightly High	High	Very High

	Very Low	Low	Slightly Low	Average	Slightly High	High	Very High
21. <u>QIH</u>							
22. <u>JEX</u>							
23. <u>WUH</u>							
24. <u>GAX</u>							
25. <u>GID</u>							
26. <u>WUB</u>							
27. <u>KOJ</u>							
28. <u>VUH</u>							
29. <u>ZOT</u>							
30. <u>CEF</u>							
31. <u>ZOF</u>							
32. <u>GEX</u>							
33. <u>GIW</u>							
34. <u>GOO</u>							
35. <u>YOF</u>							
36. <u>DAX</u>							
37. <u>YIL</u>							
38. <u>GEC</u>							
39. <u>WEZ</u>							
40. <u>QIM</u>							
	Very Low	Low	Slightly Low	Average	Slightly High	High	Very High

	Very Low	Low	Slightly Low	Average	Slightly High	High	Very High
41. <u>MEC</u>							
42. <u>DAQ</u>							
43. <u>MEQ</u>							
44. <u>WOJ</u>							
45. <u>CIJ</u>							
46. <u>VEC</u>							
47. <u>MIB</u>							
48. <u>XUC</u>							
49. <u>WOG</u>							
50. <u>ZIP</u>							
51. <u>XAT</u>							
52. <u>NUX</u>							
53. <u>QIJ</u>							
54. <u>QUG</u>							
55. <u>CEJ</u>							
56. <u>QOB</u>							
57. <u>QOW</u>							
58. <u>KEB</u>							
59. <u>JIC</u>							
60. <u>JUQ</u>							
	Very Low	Low	Slightly Low	Average	Slightly High	High	Very High

THE EXPERIMENTAL INDUCTION OF  
RESPONSE SET

by

WILLIAM BAHRET EDDY

B. S., Kansas State College  
of Agriculture and Applied Science, 1955

---

AN ABSTRACT OF A THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE COLLEGE  
OF AGRICULTURE AND APPLIED SCIENCE

1957

A study was carried out to determine the effect of suggested hypothetical "social norms" on the responses of subjects to ambiguous items in a rating scale test. The experiment was designed to answer the following questions. (1) Will subjects receiving test forms containing the same item content but instructions suggesting differing evaluations of the item in terms of social norms differ in their responses to the items? (2) Will subjects who evaluate the items in a rating scale test in terms of a suggested social norm change their evaluation of the items in a later test session if a different norm is suggested?

Three test forms were developed to measure the effect of suggested social norms on the responses of groups of subjects. The three test forms contained exactly the same content, 60 nonsense syllables with low association value. Subjects were asked to respond to the nonsense syllables by indicating the degree of "psychological value" each syllable contained for them on a seven point scale from very low to very high. The written instructions were varied so that those contained in one test form stated that other students had rated the syllables as having low psychological value; those contained in another form suggested that students had rated the items as having high psychological value; and those contained in the third form contained no statement concerning how other students had rated the items.

Subjects were 190 students in four General Psychology classes. They were tested twice with a five day interval between sessions. A sampling procedure was used which made it possible to measure the change in response between sessions. The three different test forms were given during the first session, and the same three forms were given during the second session. By this method samples or "subgroups" were obtained for each of the nine

possible session-to-session combinations of test forms.

The three groups of subjects which received different test forms, i.e., instructions pertaining to social norms, during the first session differed significantly in their evaluations of the test items in the first session.

There were marked positive correlations between subjects' scores in the two test sessions. These correlations indicated a consistency in subjects responses to the test items in the test sessions.

Analysis of variance techniques showed that the groups which received different sets of instructions in the first session did not differ significantly in the way they changed their responses in the second session. Groups which received different instructions during the second session did differ significantly in the degree to which they changed their responses to the test items from the first session to the second session. These results indicate that the instructions given during the first session had no statistically significant effect on the degree to which subgroups changed their responses between test sessions; and that the instructions given during the second session were very important in determining the degree to which subgroups changed their responses from the first to the second session.

Several trends concerning the direction and degree to which subgroups changed their responses between sessions were pointed out. Subgroups which received the same instructions during both sessions changed very little in their responses to the test items. Subgroups which received opposite instructions ("high" followed by "low", or "low" followed by "high") changed considerably as did subgroups which received neutral instructions in the

first session and either "high" or "low" instructions in the second session. Subgroups which received neutral instructions during the second session changed very little from their responses in the first session.

The results of this study show conclusively that a response set can be induced in groups of subjects, and that this response set can influence the responses of the subjects to test items. Further, this response set and its effects can be changed by changing the form of the test situation, i.e., the way in which the test items are presented.

The tendency of individuals to conform to what they believe to be the group norm has been shown by other investigators to be a personality characteristic common to most individuals. The present study has shown that this tendency, which may be called social acquiescence, may affect the responses of groups of subjects to test items. It is apparent that those concerned with measuring human behavior with testing devices must be concerned not only with the effect of various response sets in influencing the responses of particular individuals in particular situations, but also they must be concerned with the effect of social acquiescence response set in influencing the responses of groups of subjects in any situation which may contain cues concerning group norms relevant to the test items.