THE RELATION BETWEEN BEHAVIOR IN THE LEVEL OF ASPIRATION TEST SITUATION AND DEVIATIONS FROM EXPECTANCY IN GRADE POINT AVERAGES

A Thesis Presented for the Degree of Master of Arts

by

Jack D. Adams, B.S.

THE OHIO STATE UNIVERSITY

1947

Approved by:

[Signature]
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chapter I</strong> Introduction</td>
<td>1</td>
</tr>
<tr>
<td><strong>Chapter II</strong> Review of Literature</td>
<td>6</td>
</tr>
<tr>
<td><strong>Chapter III</strong> Procedure</td>
<td>14</td>
</tr>
<tr>
<td>A. The Subject</td>
<td>14</td>
</tr>
<tr>
<td>B. The Task</td>
<td>16</td>
</tr>
<tr>
<td>C. The Procedure</td>
<td>17</td>
</tr>
<tr>
<td>D. The Data Collected</td>
<td>18</td>
</tr>
<tr>
<td>E. The Patterns</td>
<td>19</td>
</tr>
<tr>
<td><strong>Chapter IV</strong> Results</td>
<td>23</td>
</tr>
<tr>
<td><strong>Chapter V</strong> Summary and Conclusion</td>
<td>37</td>
</tr>
<tr>
<td><strong>Bibliography</strong></td>
<td>39</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table I  Average Performance Scores for the Group for the First Ten Trials and the Last Ten Trials 24

Table II  Comparison of the Assignment of Behavior Patterns for Raters A and R. 26

Table III  Percentages in Each Category of Acceptable and Unacceptable Patterns for Raters A and R 27

Table IV  Means and Sigmas for Point Hour Ratios and OSPE's by Groups 28

Table V  Distribution of Cases by Group on the Basis of the Nine Level of Aspiration Behavior Categories Raters A and R 29

Table VI  Distribution in Percentages of Acceptable and Unacceptable Patterns for All Groups Rater A 30

Table VII  Discrepancy Between the Expected and Obtained Frequencies in Each Grade Category of Acceptable and Unacceptable Patterns Rater A 30

Table VIII  Chi-Square Between Groups AB, BC-, B64, AC-, AC4, and C4-C for Rater A 31

Table IX  Distribution in Percentages of Acceptable and Unacceptable Patterns for All Groups Rater R 33

Table X  Chi-Squares Between Groups AB, BC-, B64, AC-, AC4, and C4-C 34
Table XI  Means and Sigmas Between Groups for D-Scores, Shifts and Unusual Shifts

Table XII Critical Ratios Between Groups for D-Scores, Shifts and Unusual Shifts
The author wishes to thank Dr. Julian B. Rotter for his indispensable aid in planning and carrying out this study. He is indebted to Dr. Kinzer and the instructors of the 411 Psychology staff at Ohio State University for their cooperation in providing the subjects for the experiment, and to the subjects who gave of their time and patience.
CHAPTER I
INTRODUCTION

The increased enrollment in our colleges and the increased use of vocational guidance centers by college students has made the prediction of college success a vital matter to vocational counselors. Although the intelligence tests used to predict college success are useful there is a dissatisfaction among counselors as to the extent to which they are useful in individual cases. Many students who have high Ohio State Psychological Test scores come in for vocational guidance because they are not doing passing school work. Others who have low OSPE scores are doing above average work. These things have led to consideration of the personality variables that make for success in those of low ability and failure in those of high ability.

The prediction of academic success has been one of the most widely investigated problems in the field of education. Through the use of intelligence tests it has been possible to predict, fairly well, success and failure in school. Derflinger (4) compared the published studies of prediction of college success since 1934 with the earlier studies. He found that the more recent studies showed a higher correlation between intelligence test scores and college grades. The earlier studies would indicate that the median correlation is about .45 while the more recent studies show that the median correlation is significantly higher at about .52.

Since correlations between intelligence test scores and grades rarely go above .60 we can say that there is still a
large measure of error in our predictions. Obviously there are two factors which may account for this; the fact of variability of grades and the fact that intelligence is a narrow aspect of school success. Both of these contain a certain degree of truth. Many writers have emphasized that we have not taken into consideration all of the factors: Toops (27) P. 118) gives moods and emotions as some of the factors that we are failing to measure which do influence school work. According to Feder (6) P. 26) "we can never reach, and in fact should never aspire to reach 100% accuracy in educational prediction. Such an end is impossible of achievement because among other reasons (1) our measuring instruments do not have perfect validity and reliability and (2) we have no way of controlling or directly evaluating the effects of motivation; (3) there are subtle personality factors which preclude the possibility of achieving true measures at each tested performance."

There have been efforts to measure the "subtle personality characteristics" that may determine school success. Emme (5) in a review of the literature on the use of various types of personality tests reports that "there is little evidence that these methods correlate highly with school marks but it is agreed that personal factors are important and that combined with other scores should make for better prediction." In the studies that are reported by Emme the personality tests were paper and pencil questionnaires and trait ratings.
In Durflinger's (4) review of the literature he reports that very few studies attempting to predict college success include any type of personality test. The studies reported have used some type of questionnaire such as the Bernrueter and the correlations found have been quite low from -.21 to .33.

Anderson (1) recently reviewed attempts to show the relationship between personality factors and school success. The Korsbach test has been used but no definite data as to its predictive value has been presented. It is felt, however, that it offers promise along this line. Other studies using questionnaires and standardized measures indicate that there is a positive relationship between emotional maladjustment and lack of school success. Specific information as to validity figures is lacking, however. It seems apparent then that although the importance of personality factors is generally recognized we have failed, so far, to find a means of measuring the personality factors that are important to school success. What is obviously needed then is some other approach to the measurement of these personality variables.

When we attempt to predict what an individual will do in a future situation, on the basis of a present performance, we assume that there are general "traits" or behavior patterns which are measurable. Studies already mentioned make this assumption but there is little evidence that the methods used measure general behavior patterns. What is needed, if
we are to predict future behavior is some task which will "call forth" the traits or behavior patterns that we are attempting to predict. A new approach to the problem of the measurement of general traits has been made through the use of level of aspiration tests. The basic procedure in the level of aspiration test as defined by Rotter (20) is this: "a subject is confronted with a task and either before or after practice he is asked to make a statement of how well he will do on the task. After failure or success in reaching this explicitly set goal he is asked to make another estimate. This is repeated several times. Through this procedure it is possible to study, fairly objectively, the effect of success and failure on the explicitly set goals of an individual where success and failure are defined as reaching or not reaching the previously set goal. Studies by Rotter (20, 21, 22,23) have indicated that individual differences are obtained in this situation, characteristic modes of behavior enter into the differences found and that the pattern of responses in such a situation are related to other objective criteria about the individual. Because of its unique character and because it has certain similarities to the school situation it is felt that this technique may have value for the prediction of school success.

If, through this procedure, an objective measure of general traits of an individual can be obtained it would be reasonable to suppose that this measure would be related to other evidences of an individual's behavior.
The primary purpose of this study is to investigate the relationship between behavior in the level of aspiration test situation and success and failure in school work. Success and failure in this study will be considered in relationship to an individual's ability; that is individuals will be selected on the basis of the relationship between their ability to do college work (as measured by the Ohio State Psychological Examination) and the actual grades that they receive. Each subject will be classified into one of nine patterns of behavior on the basis of the Rotter Level of Aspiration Board. It will be possible to determine then if those in the success and failure group can be distinguished on the basis of their behavior in the Level of Aspiration situation.

The secondary purpose of this study will be to determine the reliability of the assignment of behavior patterns as described by Rotter (23).

A check will also be made to determine whether the task as used in this situation, meets the a priori conditions, as set up by Rotter (21), required for the study of individual personality variables.
CHAPTER II

REVIEW OF LITERATURE

An attempt will be made to cover the various definitions of the term "level of aspiration," and the general trend in research studies rather than to attempt an exhaustive review of the voluminous literature. Excellent reviews of the literature have been made by Rotter (20) and Frank (8). Recent studies and a study of the use of the level of aspiration to predict school success will be reviewed.

Definitions and General Trend

The concept of "level of aspiration" was developed by Lewin and his followers, who were interested in it chiefly as it related to the phenomena of success and failure. It is well known that experiences of success and failure are important in an individual's development.

The first experimental study of "level of aspiration" was made by Hoppe in 1937(9). The chief importance of his work was the demonstration of marked individual differences in the interpretation of and reaction to success and failure and his hypothesis that they may be diagnostic of personality differences. His definition was rather ill defined, i.e., he referred to it as the totality of the inner strivings of the individual.

The first American investigation of "level of aspiration" was made by Frank (7). He provided a quantitative technique for measurement of aspiration level. The subject was asked to perform a task, such as quoits, which was subject to quan-
tative measurement, a number of times in succession, and after each performance was requested to give an estimate of what he intended to do the next time. Frank found the most useful measure to be the "difference score" or the difference between the subject's estimate and his immediately preceding performance. His definition was more precise than previous ones; he defined level of aspiration as "the level of future performance in a familiar task which an individual, knowing his level of past performance in that task, explicitly undertakes to reach."

In a more recent study Gould employed a method similar to Frank's where a "Difference score" was obtained (12). Through interviews with the subjects she found that the explicit goals of Frank were quite different than the implicit goals of Hoppe.

Gardner (10) in a discussion of the use of the term "level of aspiration" summarized the trend as follows:

"Hoppe thought of the level of aspiration as the totality of certain highly subjective aims. Later investigators rejected his methods as inadequate for getting at these aims and set up more precise and objective methods. These methods were measuring or precisely what the individual was willing to make public about his aims and no more."

The author's main point is that there is no level of aspiration, but levels of aspiration. The recent experiments in this area do not get at inner goals and aspirations but simply overtly expressed goals (in a social situation) to a specific task. In spite of this he feels that interesting personality variables have been uncovered with the method.
"Used in this very narrow sense, level of aspiration is a truly quantitative concept. Used in any other sense, it is not only pseudo-quantitative but mythical. It may be objected that in narrowing the concept down to the point where it means something, we have narrowed it down to the point where it means nothing. The only convincing answer to this objection is to be found in the recent experimental literature on the subject which indicates that even in its strictest meaning the concept represents one of the most interesting personality variables that has been uncovered in recent years."

The more recent studies have accepted an operational definition of the term "level of aspiration" and attempts have been made to determine whether individuals reveal in their responses to this situation stable, specific personality variables, and whether or not measures obtained in this situation are related to "personality traits" as measured by other means.

Studies by Gardner (9) and Gould (13) attempted with little success to relate "personality traits" directly to a single derived score, one representing the difference between the subject's stated aspiration and performance.

Rotter, critical of the methods used in previous studies, made attempts to avoid the difficulties found in early studies and to relate the measures obtained with his test to validation criterion. On the basis of experimental results on the development of a controlled method he concluded (21 p. 421).

"These instructions were fairly successful in eliminating the possibility of misinterpretation upon the part of the subjects without involving judgment as a factor in the subject's reaction or reducing the extent of individual differences."
The task creates a great deal of interest, does not appear to be influenced by performance level, shows little learning after a short initial practice period, and appears to be free of attitudes and standards resulting from previous contact with this or a similar task. The task also allows for a large number of trials, providing adequate quantitative results, in a relatively short time. Several scores which attempt to represent the subject's full pattern of response appear to be reliable measurable."

Rotter found (22) that in the "level of aspiration" situation stable modes of behavior are found. However, he found that these stable modes of behavior do not bear a linear relationship to any single score. The analysis of extreme scores suggested the presence of several patterns of response.

Rotter then (23) studied the relationship between the patterns of response previously found and individual clinical findings. He concluded that "the full value of this technique for clinical and some experimental purposes lies in the consideration of the whole pattern of response rather than of the D-score or any other single score."

Recent Studies

Since the time of the reviews by Rotter and Frank there have been several studies made. These will be briefly reviewed. In all of the recent studies the level of aspiration has been used as an overtly expressed goal to a specific task.

Studies by Preston and Bayton (18, 19) continue the investigation of the problem of whether the behavior measured is general from one situation to another and whether or not information about the performance of a super ordinate group
will have the same effect upon each of three levels of aspiration. They refine the nature of the task somewhat by introducing three levels of aspiration instead of one; that is, (1) the best that the S's expect to do, (2) the performance actually expected, and (3) the least they expect to do. From these levels they calculate three ranges. As subjects Negro College students were used. As a source of pressure upon the estimates they used the device of telling the S's that white students had reached a certain level of proficiency in the tasks used. They found that "the reliability of the estimates is very high and that the degree of generality observed is very extensive." They state that these definite conclusions are undoubtedly related to the refined definition of the level and to the introduction of rivalry. On the basis of their results they say that the concept of the level of aspiration must be expanded to include a variety of tendencies. They include two:

1. Level of aspiration is affected by consideration of self preservation and

2. Level of aspiration is a form of self motivation.

Bayton (3) studied the hypothesis that the needs which operate in the determination of the level of aspiration continue to exert their influence upon later performance. Subjects used were matched for ability to perform the tasks. The experiment was designed so as to permit a trial by trial test of the hypothesis in an ego-involved task and a non-ego involved task. The results support the view that the influ-
ence which find expression in the level of aspiration con-
tinue to be operative upon the events which follow the sta-
tement of the aspiration.

Gruen (14) studied the level of aspiration behavior in
relation to personality adjustment in adolescent, as deter-
mined by the Rogers Test of Personality Adjustment. The ex-
perimental situation consisted of a modified symbol-substi-
tution task. The results indicate that the differences in
D-scores and average deviations are significant at the 1% level between the well adjusted and maladjusted groups. He
concludes "Generally, it appears that the level of aspiration
behavior seems to reflect underlying personality needs and
wants, and that it might be used as another measure of per-
sonality adjustment."

These recent studies support the evidence found in ear-
lier studies that generality of behavior is found through the
level of aspiration situation and that this behavior is rela-
ted to other facts about the individual.

Use to Predict School Success

Sears (24) based her study on the hypothesis that one
factor in the level of aspiration pattern for a given task
is the characteristic past experience of success or failure
which the individual associates with the task. Subjects
were selected on the basis of their previous experience with
the task. Through the use of tests similar to ones on which
they had previously failed or succeeded the question of ge-
nerality from task to task was eliminated. She had three
groups matched for age, sex and intelligence. Group I unsuccessful in all school subjects, Group II successful in all, Group III successful in reading and unsuccessful in arithmetic. Her results indicated that the children with a history of failure in a subject had higher difference scores, on the average, than successful ones. She also found that they differed markedly in the spread of their difference score, the unsuccessful children making up almost all of the cases at the extremes of the distribution. Experimentally induced success brought the reaction of all subjects into a more homogeneous but still differential distribution than did the neutral conditions of stimulation.

The study by Sears does not provide evidence as to whether experiences of success and failure in a specific task will influence behavior in other tasks. That is, she did not deal with the question of the generality of the behavior observed.

Anderson and Brandt (2) made a study of level of aspiration in 53 school children. Their purpose was to determine the effects of knowledge of one's own achievement plus the achievement of the rest of the class on estimates of future performance. A control group of 24 children who did not know their scores were used. They found that the performance of the experimental group improved significantly over the control group. Also they found that in the experimental group those with poor achievement set their goals considerably higher than their past performance and
those with high achievement set their goals relatively lower. Knowledge of performance of the group tended to bring the estimates of future performance closer together.
CHAPTER III
PROCEDURE

A. The Subjects:

Subjects for the experiment were selected on the basis of their Ohio State Psychological Examination Test Scores and their Point Hour Ratios only. Other factors about the individual were not taken into consideration.

The Ohio State Psychological Examination is the entrance examination used at Ohio State. Its reliability has been reported as .98. An unpublished study by Toops (27) of 3024 students in 1940-41 was used as the basis for the statistics in this report. The correlation with school grades was found to be .54.

The figure used for school grades is the cumulative point hour ratio of each student for the number of quarters that he has been in school.

The point hour ratio is the ratio of the total number of credit hours by the total number of hours taken. Credits are figured on the basis of 4 points for an A, three points for a B, two points for a C, one point for a D, and 0 points for an E. Grades and point hour ratios were verified from the official records of the University.

Subjects were selected on the basis of the relationship between their O.S.P.E. score and their point hour ratios. Four groups of students were selected:

1. Group A. Those whose grades were more than one sigma below what you would expect. (based on the regression formula given by Garrett (22) Y = .015 x + 1.02)
2. Group B. Those whose grades were more than one sigma above what you would expect on the basis of their O.S.P.E.

3. Group C. Those whose grades were within one-half a sigma plus or minus of what you would predict on the basis of their O.S.P.E.'s. This group is divided into two parts: 
   C+ those whose O.S.P.E. and Point Hour Ratios are high and 
   C- those whose O.S.P.E. and Point Hour Ratios are low.

An effort was made to obtain all of the cases for this study from the Psychology 411 classes.

Psychology 411 at Ohio State University is a course in remedial techniques for college students. Remedial training at Ohio State University is directed toward the development of better study skills on the assumption that even the best students work at a low level of efficiency. In a recent study (16) it was found that students in the 411 classes range in scores on the Ohio State Psychological Test from the first percentile to the 98th and the median percentile score is 41. They also range in grades from very low to very high. It was felt, therefore, that an adequate sample of cases could be obtained from these classes.

However, after exhausting all possible cases from the 411 classes from January through April 1947 it was found impossible to obtain enough students who had grades significantly below what you would expect on the basis of their Ohio State Psychological Test Scores. It was apparent that
this type of case was no longer available in the 411 classes. Therefore 15 of the 109 cases were selected from throughout the University. These cases were selected on the basis described above and were called by phone. Two out of three of those called showed up for testing. It is significant that it was difficult to find these cases in any of the colleges. It is evident that the distribution of grades at the lower end of the scale has changed since the time of the study by Toops. Possibly some teachers are more lenient towards the returned veteran. This may be a source of negative results in this study.

B. The Task:

The level of aspiration board used was the one developed by Rotter (21). It is a board made of pine wood, 38 inches long with a square groove down the center. A steel ball is hit along the groove by a stick resembling a miniature billiard cue. The subject attempts to bring the ball to rest in a certain depression valued at ten points. On each side of the ten are depressions numbered down to one. Regularly spaced depressions preceding the numbered units and also one placed in the center of each numbered unit slow down the speed of the ball and provide a resting place for it when it comes to a stop. The score is dependent upon how closely to the central unit the ball comes to rest regardless of the direction. The central unit, painted in white with the black number ten on it, counts 10 points. The ones on either side count nine points and so on. The other units decrease to a
value of 1 point.

C. General Procedure

The procedure for testing in this study was the same as that used in the level of aspiration test developed by Rotter (21) and described below.

1. The subject was told to simply try to hit the ball so that it will stop at ten and to go ahead and practice for a while.

2. After 10 to 20 practice trials the subject was given the following instructions to read:

"This is a test of motor control. The idea is always to aim for the ten. Your score will depend on how close to the ten you come. You will be given a series of trials in which you should try to get as high a total score as possible. Before you start each trial, however, you will have to tell me the score you expect to get and you will not be credited with anything over that score. If your score is lower than your bid, then the score you will be credited with will be two points off your bid for every point you fall below in your actual score. For example, if you say you will score 15 and score 20, for the five trials, you will only get credit for 15; if you say 15 and score 10, then you will only get credit for five. You can see that once your bid is made it is always to your advantage to score as high as possible."

After the subject read them the experimenter went over the instructions verbally, giving as many examples of the scoring as was necessary to be sure that the subject fully comprehended.

3. If the subject inquired, he was told that he could change his estimate as many times as he wanted or he could keep it the same.

4. When the subject made his estimate it was written
down. The score for the five trials was added and the number of credits was calculated. The credit score, not the actual performance score was then written in large numberals where the subject could see it.

5. The instructions were not repeated after the first trial, the experimenter said at the first regular trial "You tell me each time what you are going to do." If the subject failed to make an estimate before hitting the ball, he was stopped, reminded and the fact of his forgetting noted down. All spontaneous comments and exclamations were noted.

6. In all cases a short rest was given midway through the test.

D. Data Collected

The following information was obtained for each subject.

1. His Ohio State Psychological Examination Score.

2. His point hour ratio

3. Data from the Level of Aspiration Test.
   A. D-score (mean of the differences i.e., each estimate minus the preceding performance)
   B. Frequency in shifts, absolute number of shifts out of 19.
   C. Number and nature of responses following success and failure and including unusual responses.
   D. Forgetting to state one's bid.
   E. Incidental behavior and comments.
E. Classification into Patterns

The cases selected were classified, on the basis of their responses, into one of the nine Patterns described by Rotter (23).

The classifications are as follows:

1. Low Positive D-Score Pattern. In this response estimates are, on the average, higher than past performances but adequate adjustments are made to both success and failure. There is an average number of shifts and generally an absence of unusual shifts. If there is one unusual shift, it occurs under somewhat excusable circumstances.

2. Low Negative or Very Slightly Positive D-Score Pattern. This is similar to pattern number one but the tendency here is in the direction of cautiousness and protection. An average number of shifts; usually no, or only one, unusual shifts are present.

3. Medium High D-Score. This is characterized by high D-scores, although usually not as high as pattern No. 7. The subjects set high goals, try hard to reach them, and show exceptionally deep involvement in the task. They are responsive to success and failure, however, and do not leave the reality of the situation to the extent of those subjects falling in pattern 6. This pattern is continuous with the low positive D-score pattern representing the tendencies there to a greater degree. None, or only one, unusual shift is present.
4. **Achievement Followers.** In this group there is a constant change of estimate to a score exactly the same or quite close to the previous achievement. No stability of estimates is present; consequently, there are a large number of shifts, an absence of unusual shifts, and generally a D-score quite close to zero.

5. **The Step Pattern.** In this reaction the subject refuses to lower his estimate. There are usually few shifts but there may be as many as eight. The only shifts are up and usually no adjustment is made even to repeated failures. D-scores may range from low positive to high positive but are usually fairly high. The characteristic of this pattern is that the subject feels that once he has made an estimate he must reach it without retreating. Behavior in the situation is commonly called stubborn and persistent. Unusual shifts are infrequent but may occur with some of the higher D-scores.

6. **Very High Positive D-score Pattern.** In this response there is an average amount of shifts and frequently one or more shifts upward after failure. In this case the response is largely a phantasy response. The subject leaves the reality of the situation and gains his satisfaction merely from the statement of high goals itself or by implying by the statement that he expects to do that well and, in fact, is surprised that he is not reaching his goal. This lack of contact with reality is often emphasized by a large number of unusual shift, up
after failure, and by the repression of failures. The presence of this pattern implies little as to whether the subject is aggressive or retiring in his reaction to daily problems but merely points to the strong tendency toward unreal solutions when they are possible under conditions of frustration.

7. **High Negative D Score Pattern.** This response is characterized by an average number of shifts and frequent shifts down after success. Predominant in the individual's behavior is the desire not to take a chance, to avoid failure at all cost. This is the extreme of the low negative pattern.

8. **Rigid Pattern.** This is characterized by the absence of shifts. Occasionally the individual will shift once and then return to the original level or, if early in the test, stay at the new level. The essential criterion in this pattern is the avoidance of the problem situation by maintaining the original estimate regardless of the achievement. Occasionally there is some confusion between this and the Step Pattern where the individual's first commitment is so high that he has little or no opportunity to shift up and refuses to shift down. The comments and behavior in the situation are the means of separating these two types. In the case of the Rigid Pattern the interest and involvement are not as deep. The subjects frequently comment that they "are taking 25 for the rest of the trials," in spite of instructions
to state their estimate before each trial, or they tell the experimenter that they are not going to change their estimate. Sometimes the subject will shift around uncertainly in the first few trials and then lapse into the Rigid Pattern or he may attempt twice to leave it and immediately return. In this way it is possible to be classified in this pattern when there are 4 or 5 shifts present. However, the great majority of subjects showing this pattern have no more than two shifts. In this rigid pattern the D-score can range from high negative to high positive; with the Step Pattern it ranges from low positive to high positive. The Rigid Pattern may be combined with the High Negative or High Positive Pattern.

9. The Confused or Breakdown Pattern. In this pattern there is a very high frequency of shifts. The D-score can be any size but usually is low positive or negative. Unusual shifts of both kinds are frequent. The pattern is actually characterized by the breakdown of the individual in the problem situation. Unable to provide any adequate method of response, he is impulsive and unpredictable and his behavior lacks consistency.

Classification into patterns was done independently by the writer and by Dr. J. B. Rotter.
CHAPTER IV

Results

The results will be presented in three parts (1) Evaluation of the task. (2) Reliability of the Assignment of Behavior Pattern numbers and (3) Relationship between Behavior on the Level of Aspiration Test and Success and Failure in School Work.

1. Evaluation of the task.

On the basis of analysis of the literature Rotter set forth several requirements which must be met if the level of aspiration task is to provide optimum conditions for the study of individual personality variables operating in the situation with the Aspiration Board. He found (21) that:

A. The absolute level of performance was not influencing behavior in the situation. (correlation of absolute level of performance and difference score.)

B. Following the practice trials learning was at a minimum. (Average performance for the group for the first ten compared with average for the last ten.)

C. There was little difference in the height of performance level for the subjects. (Average standard deviation and range for total score.)

D. There was for each individual considerable variability from trial to trial. (Median standard deviation of individual standard deviations.)

E. He found reliability for the scores used (Odd vs. Even D-scores.)
In order for this study to be a valid one the same conditions must obtain. The same calculations were therefore repeated and are presented below.

The correlation between the total performance score and the mean of the differences between each estimate and the preceding performance, for each individual was -.36 with a standard error of .08. The correlation obtained indicates that the size of the D-score bears a slightly negative relationship to the total performance score.

Table I shows the relationship between the average performance for the Group for the first ten trials compared with the average for the last ten. Although the mean of the last ten is two points higher than the mean of the first ten the increase in learning is so slight as to have little effect on performance.

**TABLE I**

<table>
<thead>
<tr>
<th></th>
<th>Average Performance Scores for the Group for the First Ten Trials and the Last Ten Trials</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Ten</td>
<td>Average score 26.9</td>
<td></td>
</tr>
<tr>
<td>Last Ten</td>
<td>Average score 29.0</td>
<td>2.1</td>
</tr>
</tbody>
</table>

The average performance for the group of 109 cases was 27.92. The Standard Deviation of the average performances for all subjects was 3.48 and the range was 18 to 38. Eighty six percent of the cases were included between the performance levels 23 and 32. It is evident that there was little
difference in the height of performance level for the subjects of this investigation. These scores are almost exactly similar to the group of 209 cases previously tested by Rotter (21).

For each individual subject there was considerable variability from trial to trial, so that the possibilities of having to change his estimate always existed. The median Standard Deviation of the individual Standard Deviations for each of the 109 subjects was 6.47.

The odd-even reliability of the difference scores (corrected by the Spearman-Brown prophecy formula) (15) for this group was .69. Although this is not as high as the reliability obtained by Rotter (20) it is high enough to indicate a certain amount of consistency.

These results indicate that the task as administered to this group meets the requirements set forth by Rotter as being optimum for the study of individual personality variables operating in the situation.

2. Reliability of the Assignment of Behavior Pattern Numbers.

Table II shows a scatter plot of the ratings obtained independently by the two raters.

The coefficient of contingency obtained for this relationship was .86. According to Wert (28) a figure of .90 would be the maximum for a 9 x 9-fold classification. It is apparent, then, that there is a definite tendency for classifications in the one distribution to be associated with
### TABLE II
Comparison of the Assignment of Behavior Patterns for Raters A and R.

<table>
<thead>
<tr>
<th>Rater A</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior Patterns</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td>7</td>
<td></td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
<td></td>
<td>4</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>2</td>
<td>1</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>6</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>3</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>37</td>
<td>23</td>
<td>20</td>
<td>11</td>
<td>5</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

those in the other distribution.

Since the cases are widely distributed over the nine categories it was decided to break them into two more meaningful groups for the purpose of analysis. The breakdown described previously by Rotter is used. He describes (23) the numbering as roughly the order of the social acceptability of that type of response. Patterns 1-3 represent culturally acceptable methods of solution of problems of self-evaluation. Patterns 4-9 are considered socially un-
acceptable as they demonstrate tendencies towards solutions of problems that either limit the achievement of the individual, show little contact with reality, or indicate attempts to escape the problem situation.

It is evident from Table II that rater A assigned many more of the low ratings (4-9) than did rater R. Of those that rater A assigned in the 4-9 group rater R disagreed on 36 percent while of those that rater K assigned in the 4-9 group rater A disagreed on only one case or two tenths of one percent.

Table III gives a breakdown in terms of percent between those assigned in the socially acceptable and socially unacceptable groups.

Table III Percentages in Each Category of Acceptable and Unacceptable Patterns for Raters A and R.

<table>
<thead>
<tr>
<th></th>
<th>Acceptable (1-3)</th>
<th>Unacceptable (4-9)</th>
<th>Total Rater A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unacceptable</td>
<td>15%</td>
<td>26%</td>
<td>41%</td>
</tr>
<tr>
<td>Acceptable</td>
<td>58%</td>
<td>1%</td>
<td>59%</td>
</tr>
<tr>
<td>Total Rater K</td>
<td>73%</td>
<td>27%</td>
<td></td>
</tr>
</tbody>
</table>

With this breakdown there is perfect agreement in eighty-four percent of the cases and lack of agreement in sixteen percent of the cases. The sixteen percent disagreement is due to the fact, as previously shown, that rater R did not agree on the many cases assigned the unacceptable patterns (4-9) by rater A. Since rater A was inexperienced it seems reasonable to suppose that his ratings were not as
valid as those of rater R; also, however, it indicates that the source of the lack of perfect reliability lies in the difficulty (for an inexperienced rater) of distinguishing, in some cases, between the socially acceptable and unacceptable patterns.

3. Relationship Between Behavior on the Level of Aspiration Test and Success and Failure in School Work:

Nature of the Groups Selected.

Table IV shows the average Ohio State Psychological test Scores and Point Hour Ratios for the four groups selected. It is apparent that those in group A have grades below what you would expect on the basis of their Psychological Test Scores, those in Group B above what you would expect and those in group C at what you would expect. There is a sizable difference in O.S.P.E.'s and Point Hour Ratios between groups.

<table>
<thead>
<tr>
<th></th>
<th>A Grades Below Expectancy</th>
<th>B Grades Above Expectancy</th>
<th>C At Expectancy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean OSPE 53.6</td>
<td>Mean OSPE 34.53</td>
<td>Mean OSPE 67.7</td>
</tr>
<tr>
<td></td>
<td>Mean PHR 1.32</td>
<td>Mean PHR 2.82</td>
<td>Mean PHR 2.18</td>
</tr>
</tbody>
</table>

Table V shows the distribution of cases by groups on the basis of the nine Level of Aspiration Behavior Categories.
as scored by rater A and rater R.

Table V. Distribution of Cases by Group on the Basis of the Nine Level of Aspiration Behavior Categories (Rater A and rater R)

Academic Group

<table>
<thead>
<tr>
<th>Behavior Categories</th>
<th>A Below Expectancy</th>
<th>B Above Expectancy A</th>
<th>C- At Expectancy low grades</th>
<th>C- At Expectancy high grades</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rater A Rater R</td>
<td>Rater A Rater R</td>
<td>Rater A Rater R</td>
<td>Rater A Rater R</td>
</tr>
<tr>
<td>1</td>
<td>7 7</td>
<td>5 12</td>
<td>5 9</td>
<td>8 9</td>
</tr>
<tr>
<td>2</td>
<td>3 5</td>
<td>8 7</td>
<td>3 2</td>
<td>9 9</td>
</tr>
<tr>
<td>3</td>
<td>4 5</td>
<td>6 5</td>
<td>3 5</td>
<td>4 5</td>
</tr>
<tr>
<td>4</td>
<td>4 3</td>
<td>2 2</td>
<td>5 4</td>
<td>2 2</td>
</tr>
<tr>
<td>5</td>
<td>4 2</td>
<td>3 2</td>
<td>0 0</td>
<td>2 1</td>
</tr>
<tr>
<td>6</td>
<td>1 2</td>
<td>2 2</td>
<td>7 4</td>
<td>1 1</td>
</tr>
<tr>
<td>7</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>8</td>
<td>2 1</td>
<td>2 0</td>
<td>1 0</td>
<td>1 0</td>
</tr>
<tr>
<td>9</td>
<td>0 0</td>
<td>2 0</td>
<td>3 3</td>
<td>0 0</td>
</tr>
</tbody>
</table>

| Number of Cases     | 25                 | 30                   | 27                         | 87                         |
Table VI shows the distribution in percentages of acceptable and unacceptable patterns for all groups for rater A.

Table VI. Distribution in Percentages of Acceptable and Unacceptable Patterns for All Groups (Rater A)

<table>
<thead>
<tr>
<th>Groups</th>
<th>Acceptable 1-3</th>
<th>Unacceptable 4-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Below Expectancy</td>
<td>56.0</td>
<td>44.0</td>
</tr>
<tr>
<td>B Above Expectancy</td>
<td>63.3</td>
<td>36.7</td>
</tr>
<tr>
<td>C+ High OSPE and High Grades</td>
<td>77.8</td>
<td>22.2</td>
</tr>
<tr>
<td>At Expectancy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C- Low OSPE and Low Grades</td>
<td>40.8</td>
<td>59.2</td>
</tr>
</tbody>
</table>

In order to determine accurately whether or not any relationship existed, a test of Chi-Square was used. Table VII gives the discrepancy between the expected and the obtained frequencies.

Table VII. Discrepancy Between the Expected and Obtained Frequencies in Each Grade Category of Acceptable and Unacceptable Patterns (Rater A)

<table>
<thead>
<tr>
<th>Acceptable</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below Expectancy</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above Expectancy</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expectancy</td>
<td></td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>At Expectancy</td>
<td></td>
<td></td>
<td>5.2</td>
</tr>
<tr>
<td>C+ High OSPE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C- Low OSPE</td>
<td></td>
<td></td>
<td>4.8</td>
</tr>
</tbody>
</table>

Unacceptable 1.0 1.0 5.2 4.8

Total Cases 25 30 27 27
The null hypothesis of this problem is stated thus: There are no factors (other than chance) at work to distinguish (on the basis of the Level of Aspiration Test) between those performing at expectancy and those performing significantly above or below expectancy in grades. The value of chi-square is equal to 5.07. According to the chi-square tables in Snedecor (25) this kind of difference could occur by chance 20 times in 100. Since five chances in 100 is usually accepted as a significant difference we must accept the null hypotheses in this case that there are no significant differences between the groups. Chi-Squares computed between individual groups are presented in Table VIII.

<table>
<thead>
<tr>
<th>Table VIII. Chi-Square Between Groups AB, BC−, BC+, AC−AC+, and C+C− for Rater A.</th>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Below Expectancy, B Above Expectancy</td>
<td>.74</td>
</tr>
<tr>
<td>B Above Expectancy, C− Low Grades and Low OSPE</td>
<td>1.04</td>
</tr>
<tr>
<td>B Above Expectancy, C+ High Grades and High OSPE</td>
<td>.40</td>
</tr>
<tr>
<td>A Below Expectancy, C− Low Grades and Low OSPE</td>
<td>.68</td>
</tr>
<tr>
<td>A Below Expectancy, C+ High Grades and High OSPE</td>
<td>1.90</td>
</tr>
<tr>
<td>C+ High Grades and C− Low Grades and High OSPE</td>
<td>6.0</td>
</tr>
</tbody>
</table>

There are no differences large enough for us to reject the null hypothesis except in the case of the relationship between the C+ and C− groups. Chi-Square for these groups was 6.0 which means that there is less than one chance in one hundred of no significant differences. In this case we can
reject the null hypotheses that there are no differences between the groups. Behavior in the Level of Aspiration Test Situation does seem to differentiate somewhat between those of high ability and high grades and those of low ability and low grades. Although both of these groups are functioning at expectancy the individuals in the low group constantly face their inability to measure up to those in the high group. Instead of accepting their position they react in such ways as to deny the problem or to set high goals in spite of their achievement.

Evidence presented by Rotter (21) indicates that individual differences found in the Level of Aspiration Situation are not due to poor judgment but to basic personality trends. He found that there was little correlation between several measures of ability on judgment tests and performance on the Level of Aspiration Test. McGhee also presents evidence (17) to show that there is a difference between a simple judgment and a measure of aspiration. He concluded that "the difference seems to lie in the fact that the ego-level of the subjects is more involved in the erection of levels of aspiration than in making a judgment." It would seem then that the difference obtained in this study is probably related to personality trends rather than to low college ability.

Although the reliability of the separate ratings was fairly high there was some difference and so it was thought that perhaps some relationship might exist between the ratings
made by rater R and the four groups.

Table IX gives the percentages of acceptable and unacceptable patterns for all groups made by rater R.

Table IX. Distribution in Percentages of Acceptable and Unacceptable Patterns for All Groups. Rater R.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-3 Acceptable</td>
</tr>
<tr>
<td>A Below Expectancy</td>
<td>68</td>
</tr>
<tr>
<td>B Above Expectancy</td>
<td>80</td>
</tr>
<tr>
<td>C High OSPE and High Grades</td>
<td>85.2</td>
</tr>
<tr>
<td>C Low OSPE and Low Grades</td>
<td>59.3</td>
</tr>
</tbody>
</table>

Again there are no great differences between groups. Chi-Square with this distribution of patterns was 3.39, which means that there are 30 chances in 100 of no significant differences. We must therefore again accept the null hypotheses that there are no factors at work to distinguish (on the basis of the Level of Aspiration Test) between those performing at expectancy and those performing significantly above or below expectancy in grades. Chi-Squares computed between individual groups for Rater R are presented in table X.
Table X Chi-Squares Between Groups AB, BC-, BC^4, AC-, AC^4, and C^4C-, AB Below Expectancy, Above Expectancy Rater R

<table>
<thead>
<tr>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB Below Expectancy, Above Expectancy</td>
</tr>
<tr>
<td>B Above Expectancy, C^4, Low Grades and Low OSPE</td>
</tr>
<tr>
<td>B Above Expectancy, C^4, High Grades and High OSPE</td>
</tr>
<tr>
<td>A Below Expectancy, C^4, Low Grades and Low OSPE</td>
</tr>
<tr>
<td>A Below Expectancy, C^4, High Grades and High OSPE</td>
</tr>
<tr>
<td>C^4 High Grades and High OSPE</td>
</tr>
</tbody>
</table>

Chi-Squares obtained here are similar to those obtained for rater A. The only significant difference was that between groups C^4 and C-. In this case there were less than two chances in one hundred of no significant differences.

Since these particular groupings were not fruitful (except in the case of C^4 C- group) a new breakdown of pattern numbers was made. It was thought that there might be some relationship between those in the more cautious patterns 2 and 4 against those in the more positive patterns 1, 3, and 5. The Chi-square in this case was .35, too low to reject the null hypotheses.

Although no relationship between the total pattern and success and failure was obtained it was felt that perhaps one of the single measures (D score, shifts and unusual shifts) obtained in the test situation might give significant data.
Critical ratios were obtained for each of these measures between group A and B, Group A and C- and group B and C+. Table XI and XII present the data. The size of the critical ratios obtained indicates that there is no significant difference between any of the groups A, B, and C on the basis of any one of these single measures of behavior.
Table XI  Means and Sigmas Between Groups for D-Scores,  
Shifts and Unusual Shifts

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C-</th>
<th>C+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Below</td>
<td>Above</td>
<td>Low OSPE</td>
<td>High OSPE</td>
</tr>
<tr>
<td>M</td>
<td>2.47</td>
<td>1.7</td>
<td>2.15</td>
<td>2.01</td>
</tr>
<tr>
<td>D Score</td>
<td>2.48</td>
<td>2.17</td>
<td>2.64</td>
<td>2.18</td>
</tr>
<tr>
<td>N</td>
<td>25</td>
<td>30</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Shifts</td>
<td>8.56</td>
<td>9.03</td>
<td>1.18</td>
<td>9.25</td>
</tr>
<tr>
<td></td>
<td>4.07</td>
<td>3.28</td>
<td>3.54</td>
<td>3.64</td>
</tr>
<tr>
<td>N</td>
<td>25</td>
<td>30</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Unusual Shifts</td>
<td>.40</td>
<td>.60</td>
<td>1.18</td>
<td>.51</td>
</tr>
<tr>
<td></td>
<td>.63</td>
<td>1.11</td>
<td>1.19</td>
<td>.79</td>
</tr>
<tr>
<td>N</td>
<td>25</td>
<td>30</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>

Table XII  Critical Ratios Between Groups for D-Scores,  
Shifts and Unusual Shifts

<table>
<thead>
<tr>
<th></th>
<th>AB</th>
<th>AC-</th>
<th>AC+</th>
<th>BC+</th>
<th>BC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>D Score</td>
<td>.23</td>
<td>.08</td>
<td>.13</td>
<td>.11</td>
<td>.13</td>
</tr>
<tr>
<td>Shifts</td>
<td>.09</td>
<td>.49</td>
<td>.12</td>
<td>.04</td>
<td>.44</td>
</tr>
<tr>
<td>Unusual Shifts</td>
<td>.16</td>
<td>.53</td>
<td>.11</td>
<td>.07</td>
<td>.36</td>
</tr>
</tbody>
</table>
CHAPTER V

Summary and Conclusions

A. Summary:

This study attempted to distinguish between students who have high ability and low grades and those who have low ability and high grades on the basis of their behavior in the Level of Aspiration Test Situation. One hundred and nine students, some selected from all Psychology Classes and a few selected from all colleges of the University were tested. They were assigned one of nine behavior patterns independently by the author and by Dr. J. B. Rotter on the basis of their behavior in the test situation. These behavior patterns were then compared with their scholastic standing as described above. The results indicate there is considerable consistency between the assignment of behavior patterns, but there is little relationship between the level of aspiration behavior ratings and discrepancy between ability and grades. There is, however, some indication that the Level of Aspiration Categories distinguish between those of high ability-high grades and those of low ability-low grades.

Further evidence was obtained in this study that the Level of Aspiration Task meets the a priori conditions set forth by Rotter, on the basis of his study of the literature, as being optimum for the study of individual personality variables.
B. Conclusions:

1. There is considerable consistency in the assignment of Level of Aspiration Behavior Patterns between two independent raters. The chief source of disagreement is due to the difficulty of determining whether or not an individual falls into the extreme socially unacceptable pattern.

2. A significant difference was found between those of high ability-high grades and those of low ability-low grades on the basis of their behavior in the Level of Aspiration Test Situation. Those in the low group tend to have more socially unacceptable patterns.

3. A lack of relationship between behavior in the Level of Aspiration Test Situation and deviation from expectancy in grades was found. This result may be due, in part at least, to the change in grade distributions from the time of the original study (on which the basis for selection of cases was made) and the time of the actual selection of cases.
BIBLIOGRAPHY


18. Preston, M. G. and Bayton, J. A. Differential effect of a social variable upon three levels of aspiration. J. Exp. Psychol. 1941, 29, 351-369.


27. Toops, H. A. Correlation Between Ohio State Psychological Test Scores and Point Hour Ratios. 1941 Ohio State University.