PERFORMANCE ON DYADIC SPAN AS A FUNCTION OF SUCCESS, FAILURE, AND SELF REPORTS OF TEST ANXIETY AND SOCIAL DESIRABILITY

by

James Donald Hitchcock

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 7 1

THE UNIVERSITY OF ARIZONA

GRADUATE COLLEGE

I hereby recommend that this dissertation prepared under my direction by JAMES DONALD HITCHCOCK

entitled PERFORMANCE ON DYADIC SPAN AS A FUNCTION OF SUCCESS, FAILURE, AND SELF REPORTS OF TEST ANXIETY AND SOCIAL DESIRABILITY

be accepted as fulfilling the dissertation requirement of the degree of DOCTOR OF PHILOSOPHY

William J. MacKinnon — March 4, 1971

Dissertation Director — Date

After inspection of the final copy of the dissertation, the following members of the Final Examination Committee concur in its approval and recommend its acceptance:*

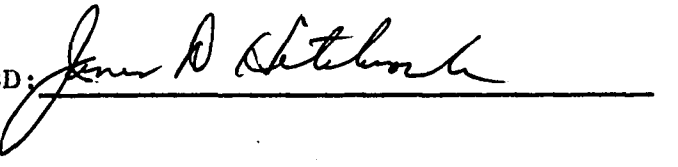
Glenn M. White	March 12, 1971
Ted L. Rosenthal	MARCH 12, 1971
Ralph J. Wetzel	March 12, 1971
Lewis Hertz	March 12, 1971

*This approval and acceptance is contingent on the candidate's adequate performance and defense of this dissertation at the final oral examination. The inclusion of this sheet bound into the library copy of the dissertation is evidence of satisfactory performance at the final examination.

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: 

ACKNOWLEDGMENTS

It is with pleasure that the author expresses his appreciation to Dr. Lewis Hertz, Dr. Theodore L. Rosenthal, Dr. Ralph J. Wetzel, and Dr. Glenn M. White, who critically reviewed the dissertation, and, in particular, to Professor William J. MacKinnon who, as dissertation director, provided invaluable support and guidance during the research for and writing of this dissertation.

In addition, my wife, Nancy, deserves a great deal more reinforcement than she received for her support throughout the graduate program.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

Relatively little attention has been given to the methods for reaching group decisions by pooling individual decisions, as compared, for example, to the attention given to the study of group processes and individual strategies used by group members. Personality variables, nevertheless, do influence group decisions. The present study was designed to (1) provide further understanding of SPAN, a new group decision making method, (2) assess the effects of achievement (test) anxiety and need for social approval on decision making, and (3) examine the effects of success and failure feedback.

With the SPAN technique each member begins with 100 "votes" or points that he may allocate directly to the group's options and/or indirectly to options by allocating points to other members. Group decisions are enhanced when superior group members receive points. As a measure of the tendency for competing responses to interfere with optimal performance, test anxiety was hypothesized to be related to subjects' task or failure avoidance. Similarly, the Crowne and Marlow social desirability scale was hypothesized to be related to subjects' dependence on social feedback and expressed attitudes toward SPAN and the experiment.

Subjects were 64 male college students for whom test anxiety and social desirability scores had been obtained from the administration of the questionnaires under neutral conditions. After being assigned to the three experimental conditions, subjects, participating in pairs, used a dyadic version of SPAN to select the better of five answers to nine successive problems of a factual nature. After each trial, the group and individual performances of pairs were compared with an easy criterion (success condition), difficult criterion (failure), or no criterion (control).

The results indicated that the SPAN method produces better group scores than the mean scores of individual members, but the willingness of subjects to use the techniques in "everyday life" was equivocal. Superior members of pairs were more adaptive than inferior members in that they tended to retain more points on the last trial than on the first, and the relatively inferior members also tended to retain more points on the last as compared with the first trial. Relatively poorer performances were attained by subjects either high or low on both test anxiety and social desirability (defensiveness), indicating that while high anxious, high defensive persons may be "motivationally disturbed," low anxious, low defensive persons may be relatively "unmotivated" under certain conditions.

Irrespective of the adaptive effects of allocating points to superior members, the option to give points away is assumed to be a means of task or failure avoidance. As hypothesized, high anxious subjects tended to give away more points than low anxious subjects with the contrast most evident in the success condition. This supports Atkinson's achievement motivation model, because the probability of success in the success condition was about .50 on a given trial.

As expected, high defensive subjects expressed greater liking for the experiment, but they also expressed a less favorable attitude toward SPAN. These results are interpreted as consistent with Crowne and Marlowe's need approval construct. High defensive subjects express greater liking for the experiment in attempting to gain the experimenter's approval. To explain their relatively less favorable attitude toward SPAN, it is assumed that (1) high defensive subjects expressed a more favorable attitude toward the experiment than was warranted, and (2) their consequent residual negative cognitions were expressed by their relatively more negative evaluation of SPAN. The failure and control conditions were, for different reasons, more threatening to the high defensive subjects than the success condition, producing greater denial and, consequently, relatively more favorable expressed attitudes toward the experiment. The attitudes of the low defensive

subjects are characterized as moderate in contrast to those of the high defensive subjects.

INTRODUCTION

There is considerable interest in group processes within the behavioral sciences. Psychologists have been particularly interested in the dynamics of group behavior and have centered their studies around such issues as the role behavior of leaders and group members, intra-group and inter-group conflict and adjustment, and group psychotherapy. Recently, investigations of the strategies used by individuals to maximize personal gains in group situations have been emphasized, while the study of methods for maximizing group gains has received relatively little attention.

SPAN Decision Making

One of the purposes of this experiment is to elucidate a promising new group decision making method for enhancing group decisions (MacKinnon, 1966a, 1966b; MacKinnon and MacKinnon, 1969). This method, Successive Proportionate Additive Numeration (SPAN), incorporates self ratings, ratings of other group members, and the constant-sum method of requiring a subject to divide 100 points among the set of objects to be judged (Metfessel, 1947; Guilford, 1954, p. 214). SPAN contrasts with the typical one man, one vote, majority rule method of reaching a group decision. Each member begins the decision making process

with 100 "votes" or points instead of just one vote. Additionally, each member has the choice of assigning his points wholly or in part to the options under consideration by the group, or he may assign his points indirectly to the group's options by allocating points to other group members. (Once members make their initial allocations of points a computer completes the process.) Points received by members (rather than options) at the end of a cycle are reallocated on the next cycle to options and members in the same proportions as those characterizing the members' original allocations. This process goes on iteratively, cycle after cycle, until all the points of the group members, but a negligible fraction of a point, are distributed among the options.

Previous studies (Hitchcock, 1967; Willis, 1966; Willis, Hitchcock, and MacKinnon, 1969) have shown that, in comparison with other methods, SPAN can produce better group decisions. These studies, however, have been limited to one trial only. The development of Dyadic SPAN makes it feasible to examine some of the consequences of providing group members with feedback of the results of their allocations.

Dyadic SPAN precludes the necessity for a computer. Analogous to regular SPAN, members of a pair divide 100 points between options and the other member of the pair in multiples of 5 (a minor limitation). Points allocated to

options may be distributed among the options with no limitations. A table (Lillyquist, 1968) has been devised to give immediate knowledge of the number of points (out of 200) which end up being assigned to the respective members' options. For example, suppose member A assigns 65 of his 100 points to the options (because he believes that he has more knowledge about the problem than B) and 35 points to member B. Member B has just a little more faith in A than in himself and gives 55 points to A and 45 points directly to the options. The appropriate row and column of the table informs us that approximately 124.8 points go to options A has chosen (in the same proportions as his original 65 points allocation to options) and 75.2 points go to the options B has chosen. Appendix A describes Dyadic SPAN in detail.

Providing subjects with feedback from their Dyadic SPAN output over several trials on similar problems is expected to result in improved group performance. An interesting aspect of group feedback is that it often misrepresents individual performance. With this in mind Johnston and Nawrocki (1967) have suggested that "evaluative feedback that is some function of te m output may be termed social feedback [p. 145]." Besides this study few others have examined the effects of social feedback on individual and group performance (e.g., Hall, 1957; Zajonc, 1962).

Personality Factors in Decision Making

There is an element of risk in many decisions made by individuals and groups. A decision may be right or wrong, foolish or wise, reckless or conservative. Thus, even in relatively emotion-free areas of cognition, personality variables influence or "moderate" both the kinds of decisions made and the manner in which they are made (Kogan and Wallach, 1964). Kogan and Wallach chose two variables, test anxiety (Alpert and Haber, 1960) and social desirability (Crowne and Marlowe, 1960), as potential moderators of risk taking in a variety of situations. With some success they found that these variables, singly and in combination, extensively moderated individuals' performance on measures of personality, intelligence, cognition-judgment, and decision making.

There are two divergent approaches to the measurement of anxiety (Alpert and Haber, 1960). One approach regards anxiety as a general trait which is presumed to influence behavior in a variety of situations, e.g., social learning (Lucas, 1952; Sarason, 1956a, 1956b); verbal coding (Katchmar, Ross, and Andrews, 1958), high speed color naming (Davidson, Andrews, and Ross, 1956), judgmental confidence and extremity of judgment (Kogan and Wallach, 1965), and arithmetic calculations (Osler, 1954). The second approach takes the position that the function of anxiety is specific to particular situations, such as test taking or academic

achievement (Mandler and Sarason, 1952; Alpert and Haber, 1960). Investigators using test anxiety scales have generally restricted their dependent measures to tasks that are similar to what test anxiety scales are presumed to measure: intelligence or aptitude test-like measures (Mandler and Sarason, 1952; Sarason, Mandler and Craighill, 1952; Waterhouse and Child, 1953), anagrams (Russell and Sarason, 1965; Sarason, 1961), arithmetic problems (Williams, 1955), level of aspiration on a forthcoming exam (Beckwith, Iverson, and Reuder, 1965), and time spent on exams (Smith, 1964).

Most studies of both general anxiety and specific test anxiety have attempted to heighten levels of anxiety by giving subjects the perception of failing or by "ego-threatening" instructions intended to highlight the importance of success versus failure. Under these conditions highly anxious individuals have fairly consistently produced poorer performances as compared with low anxious individuals. The usual explanation for this is that the high anxious person is more likely to engage in task avoidance or irrelevant behavior (Sarason et al., 1952). Atkinson and his co-workers (1957; Atkinson and Feather, 1966) in their work on achievement motivation have developed the idea that test anxiety scales measure a motivational disposition to avoid failure which interacts with the achievement motive in complex ways.

Crowne and Marlowe (1960) have suggested that responding to personality questionnaires in terms of the social desirability content of items reflects a need for social approval. They developed a social desirability scale, "independent of psychopathology," composed of items that are representative of "culturally acceptable and approved behaviors which are, at the same time, relatively unlikely to occur [p. 354]." In The Approval Motive, Crowne and Marlowe (1964) discuss a number of studies of the approval motive construct.

As Ford and Hersen (1967) indicate, however, their ". . . original conception has been changed to include defensiveness as a component of the need for approval [p. 229]." Using Ford's Social Desirability Scale (which correlates "in the .70's and .80's" with the Marlowe-Crowne scale), they provide support for defensiveness as a part of the need for social approval. Other independent investigators have also provided support for Crowne and Marlowe's need approval construct (Fisher and Kramer, 1963; Lichtenstein and Bryan, 1966; Milburn, 1968; Stollack, 1965).

Besides the work by Kogan and Wallach (1964, 1967) no other studies are known to have used both test anxiety and need for social approval as personality variables affecting the decision making process; their latter study is the only one using groups. For this reason the present experiment is somewhat exploratory, and hypotheses about

the influence of test anxiety and need for social approval on decision making in groups are relatively nonspecific. It was expected, however, that subjects high on test anxiety would be especially attentive to the success-failure aspects of problem solving and that subjects scoring high on social desirability would be sensitive to the meaning of social feedback. Those subjects who are high on both characteristics being "motivationally disturbed" (Kogan and Wallach, 1964) were expected to be the least task oriented or conversely the most task avoidant. Subjects low on both test anxiety and social desirability should be the most task oriented and thereby sensitive to the demand characteristics of the situation.

Irrespective of the adaptive effects of allocating points to one's (presumably superior) partner, the opportunity to give points away may be used by some types of possibly task avoidant subjects to minimize or avoid failure on the problems. Thus, the number of points allocated to partner may be a sensitive measure of task avoidance. It may be noted that "task" is being used here in the restricted sense of selecting better answers to problems. In a broader sense, the full SPAN involves the assessing of group members, including one's self, in relation to the goal of selecting better options.

On any given trial a subject's best guess as to which member of the pair may be the superior problem solver

is a function of the preceding trials. Thus, the ability of particular personality types to shift adaptively over trials is assessed. The possibility of differences in expressed attitudes toward the experiment were deemed likely to obtain among the personality types with subjects high on need for social approval most likely to express more favorable attitudes.

METHOD

Subjects and Design

Adapted versions of the Alpert-Haber Achievement Anxiety Test and the Marlowe-Crowne Social Desirability (SD) scale were administered to approximately 950 subjects in introductory psychology discussion sections under neutral, non-experimental conditions, i.e., discussion section leaders administered the questionnaires rather than the experimenter, completing the questionnaire was voluntary, and anonymity was assured.

The original Alpert-Haber Achievement Anxiety Test, permitting five levels of agreement-disagreement with each item, was modified to permit only dichotomous choices: "true," the item "applies to me" or "false," the item "does not apply to me." In addition, the scale was treated as uni-dimensional rather than bi-dimensional when being scored as done by Herron (1964) and by Kogan and Wallach (1964). The scale in this form will be referred to as the test anxiety (TA) scale.

One week following the administration of the questionnaires, volunteers were drawn from the same discussion sections to participate at $1.40 per session in a "social psychological experiment in the area of problem solving in groups." It was unlikely that any volunteers

associated the questionnaires with this phase of the experiment. A minimum number (60) of male subjects was desired, but additional subjects were scheduled to ensure an adequate number of subjects. Four additional subjects were subsequently run; thus, a total of 64 subjects participated in the experiment.

The questionnaires of the subjects were scored and ranked. The subjects were then assigned to one of the three experimental conditions (success, failure, and control) alternately, according to their test anxiety scores, i.e., the subject with the rank of "1" was assigned on the throw of a die to a condition, the second ranked subject was assigned on the throw of a die to one of the remaining conditions, the third ranked subject was assigned to the remaining condition, the fourth ranked subject was assigned to the same condition as the subject ranked "1," and the remaining subjects were assigned alternately in the established sequence. Such a procedure resulted in essentially similar distributions of text anxiety scores, because of the relatively narrow range of obtained scores (1-18) and the large number of duplicated scores. (Someone other than the experimenter scored the questionnaires and made the assignments so that the experimenter had no knowledge of the subjects' test scores during the experiment.)

Because of the nearly zero correlation between the TA and SD scales, r = .05, obtained by Kogan and Wallach (1964), it was assumed that equal numbers of low and high subjects on the SD scale would also obtain in the experimental conditions. "High" or "low" on both scales was defined as above or below the group median on the respective scales. Equal numbers of high and low subjects on the two scales did result in the experimental conditions, although equal numbers of the four personality subtypes did not result. The numbers of personality subtypes that resulted were as follows: success condition, low TA-low SD = 8, low TA-high SD = 3, high TA-low SD = 4, high TA-high SD = 7; failure condition, 4, 7, 8, 3, respectively; control condition, 4, 6, 4, 6, respectively. Over all conditions equal numbers, 16, of each subtype were obtained.

The unequal numbers of the personality subtypes appears to have occurred by chance, because there was very little association between the TA and SD scores (r = .088), and, therefore, are not believed to have biased the results. In addition, the method for analyzing most of the results, analysis of variance by the method of unweighted means (Winer, 1962), allows for the presence of unequal n's.

The experimental design is a four factor between-within nested design with three experimental conditions, two levels of test anxiety, two levels of social desirability, and eight repeated measures.

Task

Subjects in all conditions worked in pairs (pairings made on the basis of time subjects had available for the experiment) on nine similar problems of a factual nature selected from The World Almanac. (See Appendix C for a copy of the practice problem and Appendix D for copies of the eight "test" problems. Appendix F provides the answers to the problems.) The eight problems were divided into two blocks of four problems each, and the order of the blocks was counter-balanced with half the dyads receiving them in the sequence A-B, and half receiving B-A. Using Dyadic SPAN, subjects attempted to achieve the maximum group score by allocating more points to the better of five possible answers to each problem. After each of the first eight problems (including the practice problem), all subjects received feedback individually--their individual scores and the group scores--in weighted averages from 0 to 100. No feedback was provided for the last problem. (Appendix A, the Instructions For Problem Solving Pairs, describes in detail how the scores were calculated.) Subjects in the Success and Failure conditions also received a third score which compared their group score with fictitious group norms.

Procedure

As the subjects arrived in the experimental room, they were introduced to each other and informed, "You will be working together as partners later on." They were then seated at a built-in table which ran the length of the room with opaque partitions at regular intervals forming booths. (Thus, the subjects were physically seated next to each other, but were unable to observe each other's performance.) Subjects were then given the three pages of Instructions for Problem Solving Pairs with the instructions, "Please read along as I read the instructions aloud. This may seem a little awkward, but I have to be sure that everyone gets the same instructions." Instructions for all conditions explained in detail Dyadic SPAN, the subjects' task, and the method for scoring subjects' responses. When the section of the instructions referring to the Illustrative Problem was begun, subjects were given copies of the Illustrative Problem. (See Appendix B for a copy of the Illustrative Problem.) Where reference was made to the Dyadic SPAN table, subjects were shown the table. After the subjects and experimenter read section 4 of the instructions, any questions were answered before the experimenter gave subjects the Practice Problem. Upon completion of the Practice Problem, subjects' individual and group scores were calculated while the subjects observed how their scores were determined for illustrative

purposes. Any questions subjects had were answered by reading again the relevant section of the instructions. The last section of the instructions was read with the different endings appropriate to the different experimental conditions:

> (Control) You will receive two scores: your individual score and your group score.
>
> We don't know as yet how well people may be expected to do on these problems. Therefore, one of the things we are concerned with is how difficult these problems are for college students, like yourself and your partner.
>
> (Success) You will receive three scores: your individual score, your group score, and a score that compares your group with other groups.
>
> It is known that on problems of this kind, average performance results in a score of about 50, because, as you will see, these are pretty hard problems. For college students, like yourself and your partner, we estimate that if you do an adequate job, you should be able to earn a score of about 55. As a guide to help you know how well you are doing, you will receive a score comparing your group score with the score that we expect from college students.
>
> (Failure) You will receive three scores: your individual score, your group score, and a score that compares your group with other groups.
>
> It is known that on problems of this kind, average performance results in a score of about 70, because, as you will see, these are pretty easy problems. For college students, like yourself and your partner, we estimate that if you do an adequate job, you should be able to earn a score of about 75. As a guide to help you know how well you are doing, you will receive a score comparing your group score with the score that we expect from college students.

The subjects were then given Problem 1 with the instructions: "Go ahead and begin." After the subjects

completed Problem 1, the experimenter picked up the problem sheets and gave the subjects Problem 2 while reminding them: "You may rank the answers, but postpone any decisions about how you will make your allocations on the next problem until you receive the results of the last problem." The subjects had about 2-1/2 to 3 minutes to study the problem while the experimenter computed their scores and filled in the blanks of the feedback slips of paper. The feedback from Problem 1 was given to the subjects with the instructions: "Now you may go ahead and make your allocations." The procedure for each Problem, 2-8, was the same as for Problem 1 except that subjects were not given the "reminders" and the feedback from Problem 8 was not computed for the subjects.

After Problem 8, subjects were given a short questionnaire to assess their reactions to the experiment with the instructions: "There's just one thing more--please fill these out." (See Appendix E for a copy of the questionnaire.) The questionnaire included two items that assessed the subjects' attitudes toward SPAN and the experiment and a third item intended to reveal whether the subjects were aware of (or believed there had been) any deception in the experiment. Two subjects who indicated a belief that some deception had been employed were interviewed briefly. Both subjects were inaccurate as to the nature of the deception and both maintained that they still

attempted to perform the tasks as instructed; their results were included in the subsequent analyses.

Finally, the purpose of the experiment was explained to the subjects, and they were cautioned not to discuss the experiment.

RESULTS

Outcomes of the SPAN Method

The results pertaining to the efficacy of SPAN as a group decision making method are considered first. The SPAN scores for the 32 pairs of subjects were compared, by t-tests for matched pairs, with the "simple" means of the pairs' individual scores over blocks 1 and 2. The results indicated that the SPAN scores exceeded the pair mean scores ($t = 1.6922$, $p = .10$[1] with 63 d.f.). Similar

---

1. The customary practice for evaluating the significance of a sample statistic is to use the null hypothesis test, i.e., ". . . one rejects or accepts a null hypotheses according to whether or not the value of a sample statistic falls within a certain predetermined 'rejection region' of its possible values [Rozeboom, 1960, p. 417]." In this regard, Rozeboom discusses several reasons why traditional practice is changing or should change. He points out, for example, that, in practice, beliefs and decisions reflect the likelihood of outcomes and do not require that the probabilities of given events arbitrarily be .05 or .01. A compromise, between tradition and practices advocated by Rozeboom, is attempted in this study. The probabilities for all outcomes of $p < .25$ are reported in the summary tables; where no probability is given, it may be assumed to exceed .25. The "exact" probabilities are given, based on linear interpolation, a conservative means of estimating the actual probability under the null hypothesis. That is, the true probability is less than the one reported, though the difference might appear only if additional decimal places were used. With this understanding, equal signs, rather than approximation signs, will be used to report results of the analyses. With few exceptions, however, results greater than .10 are not discussed. As noted in the tables, some of the results, for which probabilities greater than .10 were obtained, have been graphed and are appended to the text. All

t-tests were made to assess whether or not (matched) SPAN scores indicated improvement in block 2 over block 1. No differences were obtained, either for overall changes or for specific treatment conditions. Actually, the mean score of block 1, 55.598, was nearly the same as that of block 2, 55.358. The possibility of differences among the treatment means for SPAN scores over blocks 1 and 2 was also assessed. No differences were obtained among treatment conditions or between blocks, and there was no interaction of treatment conditions and blocks of trials.

Effects of Personality Variables

Individual performances on the problems as a function of personality are considered next. Analysis of variance by the method of unweighted means was carried out on the scores (blocks 1 and 2) of the 64 subjects. (A FORTRAN IV program for this analysis was obtained from Shoemaker [1969].) As shown in Table 1 and Figure 1, the results indicate that social desirability (defensiveness) and test anxiety interact ($p = .04$). Better performance appears to result from the combination of either high defensiveness and low anxiety or low defensiveness and high anxiety.

In addition, individual performance scores for the eight problems were essentially unrelated to social

---

probabilities in this investigation are based on either 2-tailed tests or 2-tailed probability estimates.

Table 1. Analysis of Variance: Individual Scores on Problems

Source	df	MS	F	Probability
Between Subjects	63			
Success-Failure-Control (A)	2	593.345	.402	
Test Anxiety (B)	1	3747.967	2.542	p = .12
Social Desirability (C)	1	855.308	.580	
A x B	2	281.747	.191	
A x C	2	1878.197	1.274	
B x C	1	6938.294	4.705	p = .04
A x B x C	2	1885.418	1.279	
Error (between Ss)	52	1474.673		
Within Subjects	64			
Problem Blocks (D)	1	1484.647	.489	
D x A	2	952.345	.313	
D x B	1	6.831	.002	
D x C	1	49.047	.016	
D x A x B	2	1449.389	.478	
D x A x C	2	3136.395	1.034	
D x B x C	1	1093.996	.361	
D x A x B x C	2	2119.462	.699	
Error (within Ss)	52	3033.582		
Total	127			

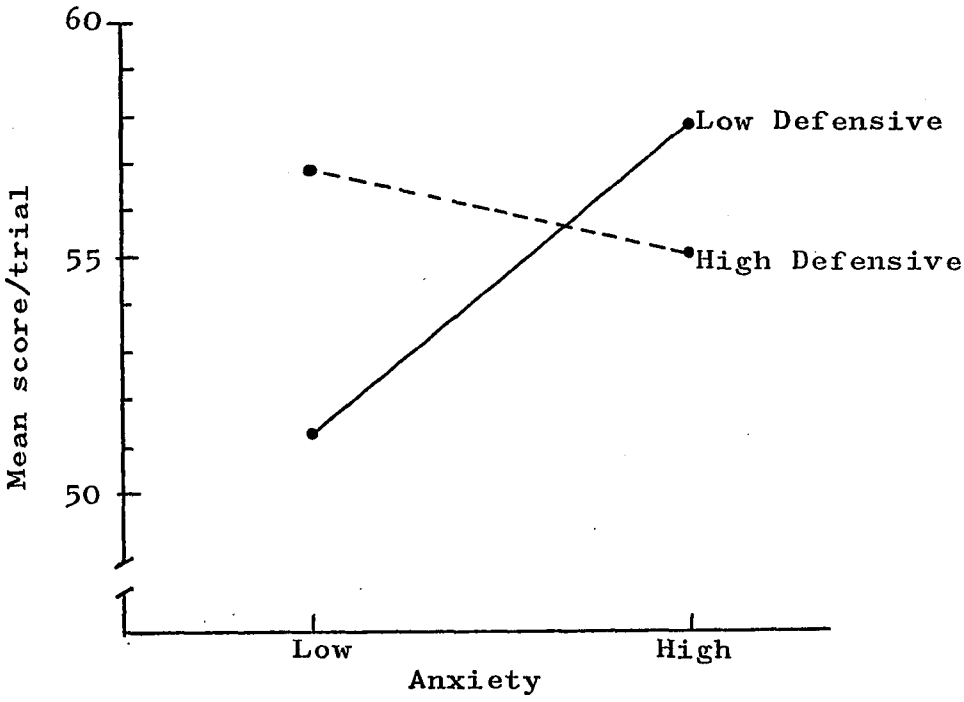


Figure 1. Effects of the interaction of anxiety and defensiveness on individual performance.

desirability scores (r = .0538) and test anxiety scores (r = .0173). In view of the nearly zero correlation obtained between individual performance and test anxiety, the result (p = .12, Table 1) that differential performance was slightly associated with the two levels of anxiety appeared to require some further explanation. A plot of the performance-anxiety scores was made which indicated the possibility of a nonlinear relationship between anxiety and performance, with better performance associated with intermediate levels of test anxiety. Subsequent analysis, however, did not support the impression that a nonlinear relationship between anxiety and performance obtained.

The general tendency to direct points to one's partner, irrespective of relative performance, was also analyzed with the results shown in Table 2. High anxiety appears to be related to giving away more points, low anxiety mean/trial = 32.44 vs. high anxiety mean/trial = 38.57 (p = .05). While the interaction of test anxiety and treatment conditions is not significant (p = .14), a graph (Figure 2) of the interaction indicates that the overall difference between high and low anxious subjects is primarily attributable to the success condition. It may also be noted that the main effect of trials appears to be to give away fewer and fewer points over trials (Figure 3). Additional tests were made of the linear and quadratic trends with the following results: linear component,

Table 2. Analysis of Variance: Points Directed to Partner

Source	df	MS	F	Probability
Between Subjects	63			
Success-Failure-Control (A)	2	2298.712	1.911	p = .17[a]
Test Anxiety (B)	1	5070.448	4.214	p = .05
Defensiveness (C)	1	1770.668	1.471	p = .24
A x B	2	2615.453	2.173	p = .14
A x C	2	2898.859	2.409	p = .11[b]
B x C	1	153.018	.127	
A x B x C	2	251.896	.209	
Error (between Ss)	52	1203.383		
Within Subjects	448			
Problems (D)	7	793.428	2.690	p = .02
D x A	14	230.051	.780	
D x B	7	141.349	.479	
D x C	7	312.568	1.060	
D x A x B	14	264.630	.897	
D x A x C	14	228.730	.776	
D x B x C	7	187.250	.635	
D x A x B x C	14	162.567	.551	
Error (within Ss)	364	294.940		
Total	511			

[a]See Appendix G, Figure 7, for a graph of this result.

[b]See Appendix G, Figure 8.

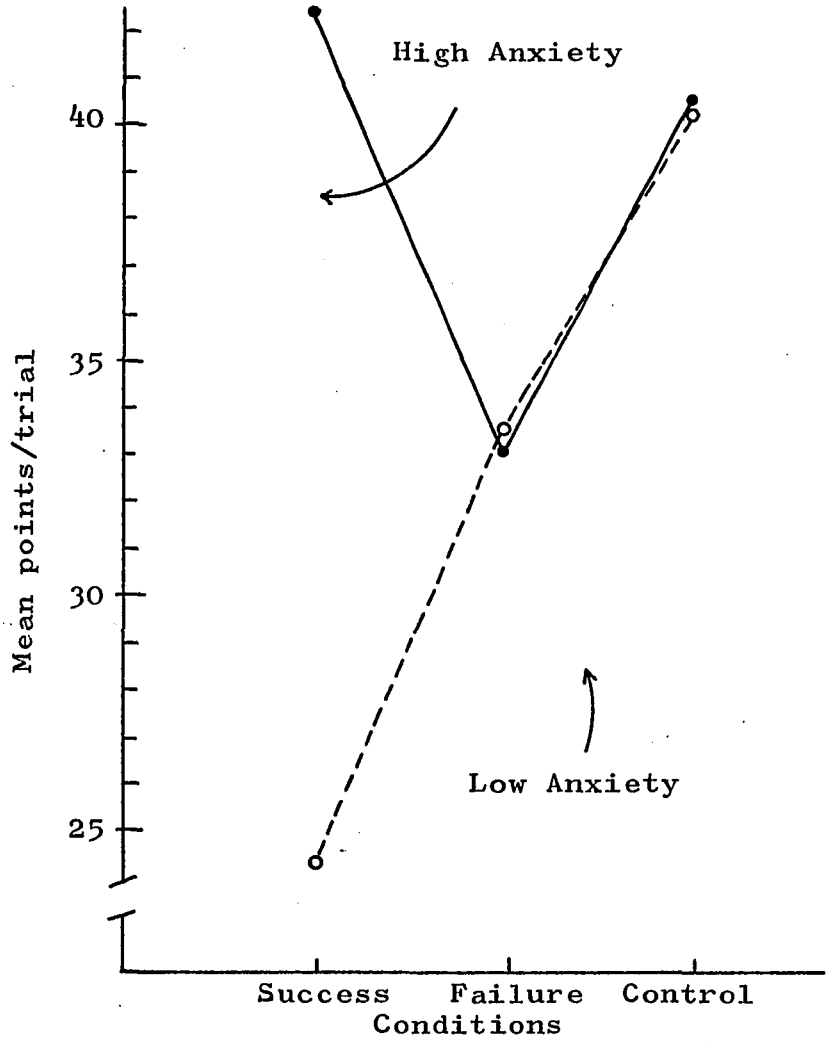


Figure 2. Effects of the anxiety x conditions interaction on points directed to partners.

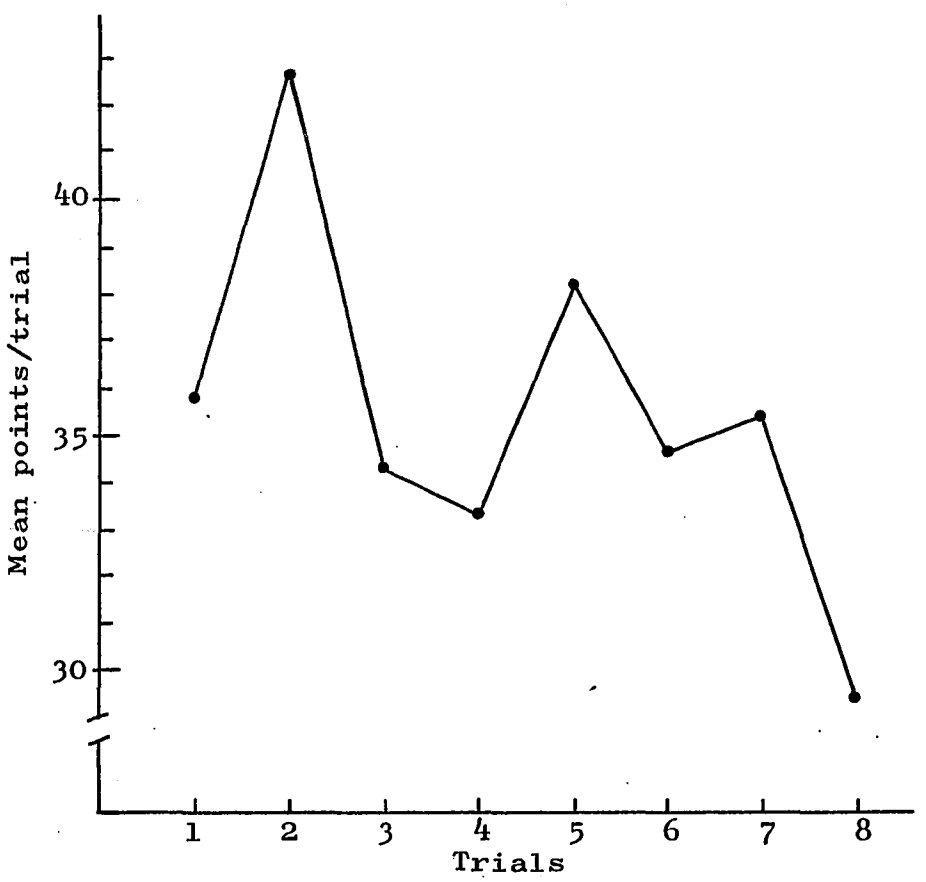


Figure 3. Main effect of trials on mean points directed to partners.

$F = 7.1448$, $p = .01$ with 1/364 d.f., quadratic trend, $F = 1.2114$, $p > .25$.

The trend toward giving fewer points away over trials (or conversely, to retain more points over trials) is clearly adaptive for the relatively superior members of the pairs, but is non-adaptive for the relatively inferior members. Therefore, separate analyses were made of the superior members' (members of pairs who performed better on the practice through the 7th trials) and inferior members' tendencies to shift adaptively from the practice to the 8th trial. The number of points that subjects retain or give away on the practice trial is regarded as an approximate base line of initial response tendencies; the responses on the 8th trial are a function of experience accumulated over the intervening trials.

Analysis of variance by the method of unweighted means was performed on the 32 superior subjects' tendencies to retain points on the practice and 8th trials with the results shown in Table 3. Because the restricted number of subjects resulted in a n of 0 in one cell and because the program would not run without at least one subject per cell, the analysis was performed with an extra "dummy" subject, whose scores were the same as the means of all the subjects. The dummy subject did not substantially affect the results of the F ratio for the main effect of "Practice-8th-Trial" as indicated by a separate analysis with the

Table 3. Analysis of Variance: Superior Members-Adaptive Shift[a]

Source	df	MS	F	Probability
Between Subjects	32[a]			
Success-Failure-Control (A)	2	143.626	.366	
Anxiety (B)	1	138.895	.353	
Defensiveness (C)	1	12.465	.031	
A x B	2	61.771	.157	
A x C	2	472.322	1.202	
B x C	1	101.902	.259	
A x B x C	2	88.168	.224	
Error (between Ss)	21	392.877		
Within Subjects	33			
Practice--8th Problem (D)	1	8029.683	49.407	p = .005[b]
D x A	2	99.778	.614	
D x B	1	4.787	.029	
D x C	1	25.727	.158	
D x A x B	2	112.122	.689	
D x A x C	2	14.924	.091	
D x B x C	1	364.629	2.244	p = .166[c]
D x A x B x C	2	26.007	.160	
Error (within Ss)	21	162.520		
Total	65			

[a]Includes dummy subject.

[b]Based on 1/20 d.f.

[c]See Appendix G, Figure 9.

dummy subject's scores being 50-50. With the 50-50 scores being counter to the general trend, the F ratio was still high ($F = 41.8265$, $p = .005$ with 1 and 20 d.f.). In any case, it seems clear that, as with the overall trend, superior subjects retain more points on the 8th trial than on the practice trial.

Analysis of variance was also performed on the adaptive shifts of the 32 inferior subjects (number of points directed to the superior member on the practice and 8th trials) with the results shown in Table 4. While the overall trend was to give fewer points to partner on the 8th trial as compared with the practice trial (practice mean = 46.875; 8th trial mean = 37.188; $p = .12$), this result was not significant and is noted only because it is similar to that of the superior subjects.

The question of whether the amount that superior members shifted might be related to their relative superiority over their partners was also assessed. For given pairs, the differences between the superior members' and the inferior members' individual performances on the practice through 7th trials were correlated with the differences between the respective superior members' points retained on the practice and 8th trials. The degree of superiority of the superior members over inferior members was essentially unrelated to the amount that the superior members shifted from the practice to 8th trials ($r = .1539$,

Table 4. Analysis of Variance: Inferior Members-Adaptive Shift

Source	df	MS	F	Probability
Between Subjects	31			
Success-Failure-Control (A)	2	122.468	.292	
Test Anxiety (B)	1	508.518	1.214	
Social Desirability (C)	1	101.519	.242	
A x B	2	78.959	.188	
A x C	2	236.878	.565	
B x C	1	298.343	.712	
A x B x C	2	367.403	.877	
Error (between Ss)	20	418.958		
Within Subjects	32			
Practice--8th Problem (D)	1	928.234	2.799	p = .12
D x A	2	67.392	.203	
D x B	1	415.803	1.254	
D x C	1	106.411	.321	
D x A x B	2	645.776	1.947	p = .18[a]
D x A x C	2	342.599	1.033	
D x B x C	1	20.776	.063	
D x A x B x C	2	411.314	1.240	
Error (within Ss)	20	331.667		
Total	63			

[a]See Appendix G, Figure 10, for a graph of this result.

p = .22). The analogous adaptive shift in the number of points that inferior subjects directed to the superior members on the practice and 8th trials showed no relationship between their relative degree of inferiority and their tendencies to give points to the superior members (r = -.0469).

A chi-square test of the frequency of adaptive and non-adaptive shifts of the superior and inferior members indicated that superior subjects were more inclined to shift adaptively than were inferior subjects (p = .00002, two-tailed test). Of the superior subjects, 90.62% were adaptive (i.e., retained more points on the 8th trial than on the practice trial), while only 9.37% were non-adaptive (i.e., did not shift). No superior subject retained fewer points on the 8th trial than on the practice trial. While 37.5% of the inferior subjects were adaptive (i.e., directed more points to their partners on the 8th than on the practice trial), 62.5% were non-adaptive (i.e., 12.5% did not shift, and 50% retained more points on the 8th than on the practice trial). It may be noted that the definition of an adaptive shift for a superior member, i.e., that he retains more points on the 8th trial than on the practice trial, involves a response measure different from the definition of an adaptive shift for an inferior member, i.e., that he gives away more points on the 8th than on the practice trial. Consequently, the result that the inferior

subjects were less adaptive is, no doubt, somewhat artificial, and the discussion below will be speculative.

The expressed attitudes of subjects toward SPAN and the experiment are considered next. Analysis of variance was performed on the attitudes of subjects toward SPAN and toward the experiment with the results shown in Table 5. The high F-ratio ($F = 75.412$, $p = .005$) associated with the differences between the means (in mm.'s) of expressed attitudes toward SPAN and the experiment must be interpreted cautiously, because the two items are worded somewhat differently. Nevertheless, the mean of the liking for SPAN scores of 75.75 (corresponding to the neutral point of the scale, i.e., "Maybe, I'd like to use the method . . . in everyday life.") appears to reflect a somewhat less favorable attitude than the mean of the liking for experiment scores (scored in the positive direction) of 115.53, which corresponds to a point between "I liked the experiment a little" and "I liked the experiment very much."

The tendency ($p = .06$) of high and low levels of defensiveness to interact with attitudes toward SPAN and toward the experiment is graphed in Figure 4. The possible criticism that, because the two items are worded differently, differences obtained in expressed attitudes may be spurious, becomes less tenable when one bears in mind that the differences reflect <u>relative</u> preferences. In any case, it appears that subjects with stronger needs for social

Table 5. Analysis of Variance: Subjects' Attitudes Toward SPAN and Toward Experiment

Source	df	MS	F	Probability
Between Subjects	63			
Success-Failure-Control (A)	2	686.496	.647	
Anxiety (B)	1	27.834	.026	
Defensiveness (C)	1	174.939	.165	
A x B	2	122.872	.116	
A x C	2	869.031	.819	
B x C	1	190.951	.180	
A x B x C	2	1865.848	1.759	p = .20
Error (between Ss)	52	1060.487		
Within Subjects				
Span-Experiment (D)	1	41018.593	75.412	p = .005
D x A	2	277.940	.511	
D x B	1	3.117	.006	
D x C	1	2136.640	3.928	p = .06
D x A x B	2	1007.987	1.853	p = .20
D x A x C	2	1701.860	3.129	p = .06
D x B x C	1	1359.174	2.499	p = .12[a]
D x A x B x C	2	726.199	1.335	
Error (within Ss)	52	543.925		
Total	127			

[a]See Appendix G, Figure 11, for a graph of this result.

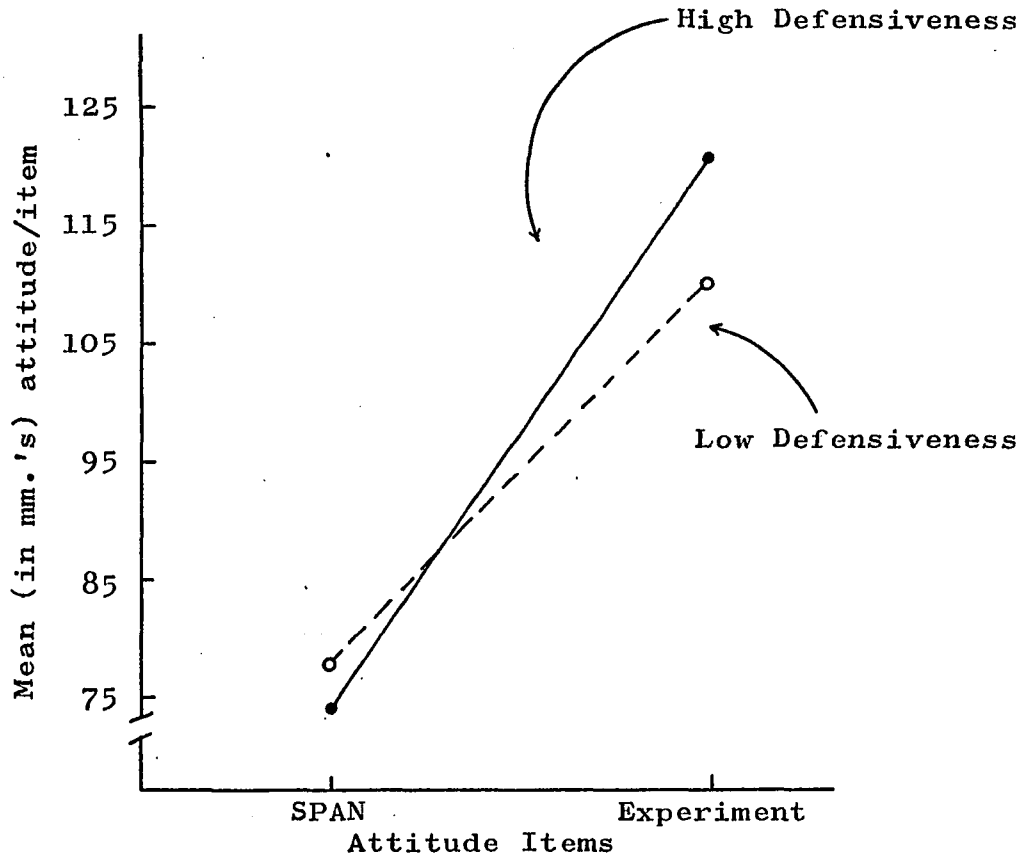


Figure 4. Effects of the interaction of defensiveness x SPAN-experiment on attitudes toward SPAN and the experiment.

approval tend to express more favorable attitudes toward the experiment (than do subjects with less need for social approval) and to express less favorable attitudes toward SPAN as compared with relatively non-defensive subjects.

Further analysis of the Defensiveness x SPAN-EXP. interaction supports the finding above that high and low defensive subjects tend to express their attitudes toward SPAN and the experiment in different directions. A Fisher's exact probability test was made on the 22 subjects who were inconsistent in their expressions of attitudes toward SPAN and the experiment, i.e., their scores were above the median on one item and below the median on the other (Figure 5), with the result, $p = .02$ (two-tailed test). If it is assumed that the inconsistent subjects contributed most to the interaction, then this result is support for the tendency of defensiveness and attitudes toward SPAN and toward the experiment to interact.

Most of the subjects (42) were consistent, i.e., expressed attitudes below, or above, the median on both items, and, consequently contributed little, if any, to the interaction. The median score, 10.30 (i.e., the median between the first and second quartiles for all 64 subjects), for the 32 low defensive subjects indicates that the inconsistent low defensive subjects are comparable to the other low defensive subjects. Similarly, the median score, 18.25 (i.e., median between the third and fourth quartiles),

	High Def.	Low Def.
Above median on SPAN, below on Exp.	3 Median def. = 22.75	8 Median def. = 13.00
Below median on SPAN, above on Exp.	10 Median def. = 18.5	1 Median def. = 12.00

Figure 5. Fisher exact test of inconsistent subjects with cell median scores.

for the 32 high defensive subjects indicates that the inconsistent high defensive subjects are representative of the high defensive subjects.

The interaction of Defensiveness and SPAN-EXP. also tends to have interacted with the treatment conditions ($p = .06$, Figure 6). High defensiveness is related to greater expressed liking toward the experiment in the control and failure conditions, but relatively less liking for the experiment in the success condition.

The possibility that there might be differences among the expressed attitudes of superior and inferior members of pairs was analyzed by the method of unweighted means. The treatment conditions were combined to form a 2 x 2 x 2 x 2 design with two levels of performance (superior-inferior), two levels of anxiety, two levels of defensiveness, and the two attitude items. No differences

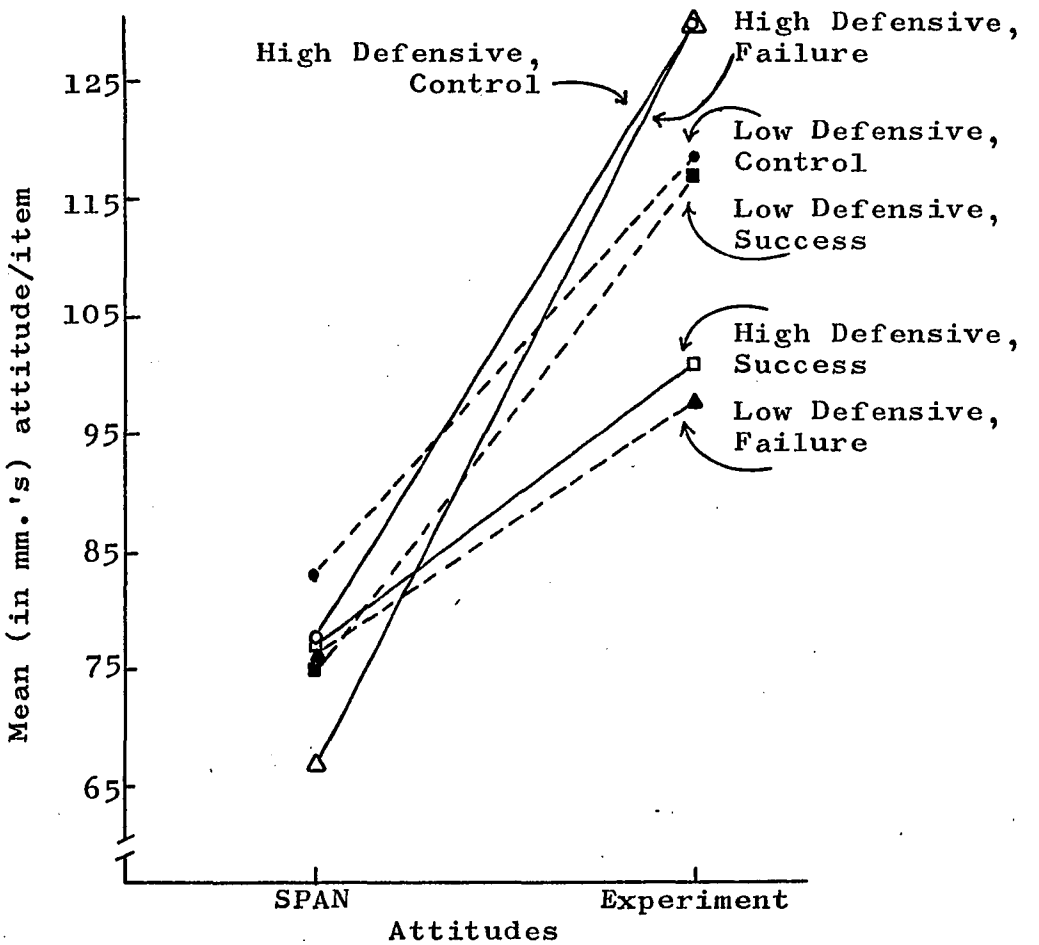


Figure 6. Effects of the interaction of defensiveness x conditions x SPAN-experiment on attitudes toward SPAN and the experiment.

were obtained between superior and inferior subjects at any levels (or combination) of anxiety and defensiveness. Expressed attitudes toward SPAN and the experiment, therefore, appear to be unrelated to how well subjects performed relative to their partners.

DISCUSSION

The SPAN Decision Making Method

The results of this study are consistent with previous studies (Hitchcock, 1967; Kelly, 1968; Willis, 1966) in respect to the relative superiority of the SPAN group decision making method over other methods. In the current study, SPAN performances exceeded the mean performances of pair members ($p = .10$). The hypothesis that increasing familiarity with SPAN and feedback from successive trials over a series of similar problems would result in enhancing the effectiveness of the SPAN method was not confirmed. In other words, the number of points received by the superior members of pairs in block 2 did not substantially exceed the number of points received in block 1. Apparently, the relative superiority of one member of a pair may not have been particularly obvious to the members or, at least, not obvious to the inferior members. Because the superior members were relatively more adaptive, it is possible that the superior members tended to be aware of their relative superiority, whereas the inferior members were unaware of their relative inferiority. Thus, the relative differences in performance may have been large enough to be apparent to the superior members, but not great enough to be apparent to inferior members. It would not be surprising if it were

easier to recognize one's superiority than one's inferiority.

A more likely explanation for the failure of subjects to respond to differences in relative performances may be based on the finding that over trials subjects directed fewer points to their partners ($p = .01$). If the use of SPAN over several trials resulted in increasing reliance on the subjects' own judgments, then the greater adaptability of the superior members was, at least partly, an artifact of the way in which adaptability was defined. Nevertheless, because 37.5% of the inferior members did direct more points to their superior partners on the 8th trial than on the practice trial, and because none of the superior members directed more points to their inferior partners on the 8th trial than on the practice trial, it appears that at least some of the inferior members responded to relative differences in performance rather than conforming to the overall trend.

In any case, the results are not clear enough to allow for a clear choice among alternatives in accounting for the failure of SPAN outcomes on block 2 to exceed those of block 1. The most plausible explanation is a combination of the alternatives discussed above. That is, because of individual and inter-subject variability of performance on the problems, it was difficult for many subjects to determine who was actually superior. Under such ambiguous

conditions, cultural norms favoring self-reliant behavior may tend to predominate with the resulting tendency to retain more points over trials.

A concern in the development of the SPAN technique has been whether it would be accepted in place of traditional methods, even if it were demonstrated to be more effective. Previous studies, therefore, have given some consideration to whether subjects were "satisfied" with the SPAN methods (Hitchcock, 1967; Willis, 1966), or whether they would prefer to use a SPAN method rather than the more traditional one-vote method in reaching group decisions (Kelly, 1968). In the present study, subjects were asked if they would ". . . like to use the method to reach some group decision in 'everyday' life?" As their mean response was a neutral "maybe," in contrast to a more favorable evaluation of the overall experiment, it is difficult to say whether or not they would favor using SPAN in other situations. The results of Kelly's (1968) study are somewhat more definitive in this regard. He showed that subjects believed that a SPAN method would select better answers to problems like the one used in his experiment, but that they would prefer to use the traditional one-vote method when voting for organizational "policy proposals" or "candidates for office." It is likely that the subjects in the current study would share similar attitudes toward SPAN and traditional methods. The question of whether people

would accept SPAN (or analogous methods) for reaching group decisions, even having had the superiority of SPAN demonstrated, is as yet unanswered.

Influences of Personality Variables

The effects of test anxiety on the measures considered in this study appear to be consistent with the results of previous studies. It was hypothesized that "motivationally disturbed" subjects, i.e., subjects scoring high on test anxiety and social desirability, would tend to be the most task avoidant. Irrespective of the adaptive effects of allocating points to superior partners, the option to give points away is regarded as a possible means for task avoidant subjects to minimize the negative self evaluative effects of failure. Thus, a relationship between high anxiety and/or high defensiveness and directing points to partners would be expected. It was found that high anxious subjects directed more points to their partners than did low anxious subjects ($p = .05$), but that this effect was mainly attributable to differences obtained in the success condition ($p = .14$, Figure 2). This latter result may best be interpreted in terms of Atkinson's theory of achievement motivation (Atkinson, 1957; Atkinson and Feather, 1966).

Briefly, Atkinson proposed that the motivation to perform an achievement related task is an additive function

of the motivation to achieve success (Ms) and the motivation to avoid failure (Maf). When Maf is greater than Ms, the subject will decline to perform the task in question unless he is constrained by incentives extrinsic to the task, e.g., the desire to please the experimenter. Theoretically, at least, both Maf and Ms are greatest at intermediate probabilities of success, i.e., around .50. This hypothesis has been supported in studies in which subjects with high Ms expressed a preference for intermediate probabilities, whereas subjects with low Ms (and presumably relatively high Maf) preferred either high or low probabilities of success (Atkinson and Feather, 1966). Atkinson and Feather also provide support for the assumption that test anxiety questionnaires measure motivational dispositions to avoid failure in achievement situations.

If the high anxious (higher Maf relative to Ms) subjects in the success condition perceived the problems to be of intermediate difficulty, then their tendency to give more points to partners would provide support for the Atkinson model. In fact, it does seem plausible that subjects in the success condition perceived the problems as intermediate in difficulty because, although the fictitious norms were set at a level (55) that was intended to be low, the actual performance (mean/trial = 54.127) of subjects in the success condition resulted in success on only about half the trials. Presumably, therefore, motivation to

avoid failure was higher in the success condition than in the failure condition (where performance rarely exceeded the norms) and control condition (where no norms were used). The high anxious subjects in the success condition tended to give more points to their partners in an attempt to minimize the effects of possible failure.

Kogan and Wallach (1964) regard high defensive, high anxious individuals as "motivationally disturbed," while low defensive, low anxious individuals are assumed to be "free from motivational disturbance." Results obtained in the current study suggest that this may be too simple a characterization of low defensive, low anxious individuals. It was indicated that better performance on the problems was achieved by the low defensive, high anxious and high defensive, low anxious subjects, while subjects high or low on both variables attained relatively poorer performances ($p = .04$, Figure 1). Evidently, subjects high on both variables may indeed be motivationally disturbed, but low defensive, low anxious subjects may be relatively unmotivated to do their best on the problems rather than merely being "free from motivational disturbance." The low defensive, low anxious subjects in Kogan and Wallach's study received various sums of money contingent upon their performance of various tasks. The corresponding subjects in the present study had no such motivation to perform well on the task as the money they received in payment was

contingent upon little more than their physical presence and a minimal degree of cooperation. In the Kogan and Wallach study, money may have been a sufficient reinforcer to motivate the low defensive, low anxious subjects, whereas in the present study the reinforcers available to the low defensive, low anxious subjects may have been relatively weak.

The results of the analysis of expressed attitudes toward SPAN and the experiment appear to be consistent with prior research on the approval motive (Crowne and Marlowe, 1964). As expected, the high defensive subjects expressed relatively greater liking for the experiment than low defensive subjects ($p = .06$, Figure 4), but they also expressed relatively less liking for SPAN. If it is assumed that the subjects believed the experiment to be important to the experimenter, then the high defensive subjects' higher rating of the experiment is readily explained. Being dependent on social approval, the high defensive subjects tended to rate the experiment in the direction that would most likely win the experimenter's approval.

The explanation for the tendency of high defensive subjects to express relatively less liking for SPAN, than did the low defensive subjects (Figure 4), is more complicated, but is similar to the explanation used by Crowne and Marlowe to account for the results of a study

they report (Crowne and Marlowe, 1964, pp. 133-149). In their study, after high and low defensive subjects were provoked to anger by an experimental accomplice who cheated them, the high defensive subjects were more susceptible to the accomplice's modeling of euphoric behavior and subsequently rated the accomplice more favorably. The explanation, given by Crowne and Marlowe, was that the high defensive subjects "repressed" their hostile feelings, but, in order to account for their emotional arousal, they more readily adopted the alternative cognition of euphoria. Applying this explanation to the current study, the assumption is made that the high defensive subjects expressed a more favorable attitude toward the experiment than was objectively warranted. As a consequence, their residual negative cognitions were interpreted or expressed by a relatively more negative evaluation of SPAN.

Additional assumptions are required to account for the tendencies of high defensive subjects to more favorably rate the experiment in the failure and control conditions than in the success condition ($p = .06$, Figure 6). First, it is assumed that outright failure in the failure condition constituted a greater threat (than the relative success of the success condition) which results in greater denial and, consequently, greater expressed liking for the experiment and the most negative evaluation of SPAN. Second, it is assumed that the lack of criteria for judging

the adequacy of performances in the control condition is in itself threatening to high defenseive subjects, because of defensive individuals' dependence on the favorable evaluation of others and their desire for "normative anchoring" (Crowne and Marlowe, 1964). Thus, the control condition also tended to produce more denial and a more favorable rating of the experiment. The relative success of the high defensive subjects in the success condition resulted in their relatively less favorable rating of the experiment, because they were relatively less threatened and thereby less motivated to deny. The low defensive subjects' ratings of SPAN and the experiment are interpreted as moderate in contrast to the ratings of high defensive subjects.

Summary of Major Results and Implications

SPAN performances of pair members exceeded their unweighted means ($p = .10$). While a probability of .10 may be regarded as relatively low, this result is noteworthy because it is consistent with the results of previous studies (Hitchcock, 1967; Kelly, 1968; Willis, 1966).

High anxious submects directed more points to their partners than did the low anxious subjects ($p = .05$), but this effect was mainly attributable to differences obtained in the success condition (Figure 2). This result is interpreted as support for Atkinson's theory of achievement motivation (Atkinson, 1957; Atkinson and Feather, 1966).

That is, persons with relatively greater dispositions to avoid failure (high TA) are relatively more motivated to minimize the effects of possible failure when engaged in tasks of intermediate difficulty.

Low defensive, high anxious and high defensive, low anxious subjects attained higher overall performance scores than did subjects high or low on both variables ($p = .04$, Figure 1). Thus, the Kogan and Wallach (1964) characterization of low defensive, low anxious individuals as being "free from motivational disturbance," implying that low defensive, low anxious subjects would be expected to attain superior performances, may be somewhat misleading. In some situations, a relatively high level of defensiveness, in combination with low anxiety, or a high level of anxiety, with low defensiveness, may function to facilitate better performances than low or high levels of both defensiveness and anxiety.

A tendency was obtained for the high defensive subjects to express relatively greater liking for the experiment than did the low defensive subjects, but relatively less liking for SPAN ($p = .06$, Figure 4). The interaction of defensiveness and expressed attitudes toward SPAN and the experiment also interacted with the treatment conditions in that the high defensive subjects tended to more favorably rate the experiment in the failure and

control conditions than in the success condition as compared to the low defensive subjects ($p = .06$, Figure 6).

These results appear to be consistent with the need for social approval construct (Crowne and Marlowe, 1964). The high defensive subjects, being dependent on the approval of others, tended to more favorably rate the experiment in order to gain the experimenter's approval. To explain their relatively less favorable attitude toward SPAN, it is assumed that (1) high defensive subjects expressed a more favorable attitude toward the experiment than was warranted, and (2) their consequent residual negative cognitions were expressed by their relatively more negative evaluation of SPAN. The failure and control conditions were, for different reasons, more threatening to the high defensive subjects than the control condition, producing greater denial and, consequently, relatively more favorable expressed attitudes toward the experiment.

APPENDIX A

INSTRUCTIONS FOR PROBLEM SOLVING PAIRS

This is a study of problem solving in pairs using a new decision-making method that we are interested in studying. You will be using the method, called SPAN, to select the best solutions to 8 successive problems and to do your best as a pair. After each problem, you will get feedback on how you did on the problem. This is explained in detail below. However, don't hesitate to ask questions at any time.

1. The SPAN method. In some respects, the method is similar to the familiar one-man, one-vote majority rule method of reaching a group decision. One difference is that instead of one "vote," you will have 100 votes, or points, for each problem. Therefore, instead of having only one point to give to only one answer, you will have 100 points to allocate among the 5 answers to each problem. Presumably, in order to do your best, you will want to give more points to the better answers. Another difference from the one-man, one-vote method is that you have the choice of splitting your 100 points in some way between the answers and your partner. In other words, if you think that your partner may have more knowledge about

a particular problem than you, and/or you are relatively unsure of your own knowledge about the problem, then you have the option of allocating some portion of your 100 points to him and some points directly to the answers. This should become clearer as we look at the Illustrative Problem.

2. The Illustrative Problem. At the top of the sheet, you will note the problem and five possible solutions. Below the problem are two lines for the first part of the SPAN method, that is, the splitting of points. In this example, 50 points have already been indicated as the number an individual might allocate to his PARTNER. Fifty points have also been allocated to ANSWERS, because an individual has 100 points to start with. He did not have to split his points 50-50; he could have split them in any other way, in multiples of 5, just so the total equalled 100. The next line illustrates how the number of points allocated to ANSWERS could be assigned among the answers. This individual has assigned an equal number of points to each answer, even to answers which, in this case, are obviously poor answers. But this will help illustrate how the individual and group scores are calculated. This is all you will have to do when you actually begin working on the problems. The experimenter will calculate your scores, but we want you to understand how the individual and group scores are obtained.

3. The individual score. Before we can calculate the individuals' scores for the Illustrative Problems, let's assume that the two individuals in the two Illustrative Problems you and your partner have, are also partners. They have each allocated 50 points to PARTNER and 50 points to ANSWERS. The points they have allocated to ANSWERS are proportioned equally among the answers. But what happens to the 50 points each has assigned to his partner? That is, each partner has received 50 points from his partner, but what does each do with the points he receives? Well, what happens (in the SPAN method) is that any points received are <u>automatically</u> re-allocated in the same proportions as the original allocation of 100 points. Thus the 50 points each has received are split into 50-50 proportions, that is, 25 points for PARTNER and 25 points for ANSWERS. The 25 points for ANSWERS are assigned in the same proportions among the answers as in the first assignment, 5 points to each answer, in this case. Again, each partner ends up with some points (now 25) which are re-allocated, again, in the same proportions as originally allocated. This goes on repeatedly, with fewer and fewer points going back and forth between the partners, until all the points end up being assigned to the answers. This is probably confusing, but there is a table for SPAN that is used to determine how many points a particular partner ends up with. After the initial allocations are made, the table

informs us where the points end up. In our example, we don't really need the table, because it's fairly clear that both partners end up with 100 points.

Now we can calculate the individual scores. Both partners end up with 100 points, 20 points going to each of the 5 answers. Besides the 50 points that each partner sent directly to the answers, each ends up with 50 more which, in this case, go equally to each answer. Values or scores have been arbitrarily assigned to the answers: 100 is the score for the best answer, 75 for the second best, 50 for the third, 25 for the fourth, and 0 for the fifth. To obtain an individual's score, we first multiply the number of points that end up going to a particular answer times the score for that answer. We do this for each answer that receives points, then sum these products and divide by the total number of points that the individual ends up with. This gives us the individual score. Thus, in the Illustrative Problem, both individuals end up with scores of 50:

$$\frac{\overset{a}{20(25)} + \overset{b}{20(50)} + \overset{c}{20(100)} + \overset{d}{20(75)} + \overset{e}{20(0)}}{100} = 50.$$

If these individuals had given all their points to "c," the best answer, they would have received scores of 100: $\frac{100(100)}{100} = 100$. If they had given all their points to "e," the worst answer, they would have received scores of 0:

$\frac{100(0)}{100} = 0$. Thus, depending on which answers receive points, the individual (and group) scores may vary between 0 and 100.

An individual's score may be easily calculated from his initial allocation of points: because an individual's score is uninfluenced by the number of points he ends up with. In the Illustrative Problem:

$$\frac{\overset{a}{10(25)} + \overset{b}{10(50)} + \overset{c}{10(100)} + \overset{d}{10(75)} + \overset{e}{10(0)}}{50} = 50$$

4. The group score. The group score for the partners is determined by how many points each partner ends up with and the partners' individual scores. To calculate the group score, the number of points each partner has at the end of the SPAN process is multiplied times his individual score, then the two products obtained this way are added, and the sum is divided by the total number of points (200). In this example, the group score would be:

$$\frac{\overset{A}{100(50)} + \overset{B}{100(50)}}{200} = .50.$$

To take a different example, suppose partner "A" obtained an individual score of 60 and partner "B" obtained a score of 40. Their group score would be determined by how many points each partner received through the SPAN process. If they both received 100, the group score would be:

$$\frac{\overset{A}{100(60)} + \overset{B}{100(40)}}{200} = 50.$$

If A received 150 points (and B, 50), the group score would be:

$$\frac{\overset{A}{150(60)} + \overset{B}{50(40)}}{200} = 55.$$

If A received 50 points (and B, 150), the group score would be:

$$\frac{\overset{A}{50(60)} + \overset{B}{150(40)}}{200} = 45.$$

Thus, the group score is not a simple average of partners' individual scores. Instead, the (weighted) group score tends toward the individual score of the partner who receives the most points.

5. The Practice Problem should help make this information more meaningful. Go ahead and work on the problem individually (without talking to your partner) and make your allocations. Then we'll go over the mechanics of computing individual and group scores.

6. (After completion of Practice Problems) You and your partner are now ready to begin working on the 8 problems. There are no restrictions on how you make your allocations. You may allocate all your 100 points to PARTNER; allocate all your 100 points to ANSWERS; split

your points 50-50, 45-55, 55-45; and so on. You may distribute your points among several answers in any proportions you want, or assign all your points to one answer. Remember, however, that your task on each problem is to do your best as a pair.

When you and your partner are through with each problem, give your problem sheets to the experimenter, and he will compute your scores while you begin studying the next problem. You may rank the answers, but postpone any decisions about how you will make your allocations on the next problem until you receive the results of the last problem.

(The remaining instructions contain the different instructional sets for the three experimental conditions. See the text for these instructions.)

APPENDIX B

ILLUSTRATIVE PROBLEM

Which of these mammals weighs the most?

a) Dog

b) Cow

c) Blue whale

d) Elephant

e) Domestic cat

Number of points allocated to PARTNER: (Circle one number)

0 5 10 15 20 25 30 35 40 45 (50) 55 60 65 70 75 80 85 90 95 100

(Total points allocated to PARTNER & ANSWERS must sum to 100)

Number of points allocated to ANSWERS: (Circle one number)

0 5 10 15 20 25 30 35 40 45 (50) 55 60 65 70 75 80 85 90 95 100

Points assigned among ANSWERS: (Must sum to number allocated to ANSWERS above)

a) 10 b) 10 c) 10 d 10 e) 10

APPENDIX C

PRACTICE PROBLEM

Which of these breeds of dogs was most popular in 1966, according to pure-bred registration figures?

a) St. Bernards

b) English-Springer Spaniels

c) Scottish Terriers

d) Golden Retrievers

e) Doberman Pinschers

Number of points allocated to PARTNER: (Circle one number)

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

(Total points allocated to PARTNER & ANSWERS must sum to 100)

Number of points allocated to ANSWERS: (Circle one number)

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

Points assigned among ANSWERS: (Must sum to number allocated to ANSWERS above)

a)____ b)____ c)____ d)____ e)____

APPENDIX D

TEST PROBLEMS

Which of these U. S. magazines had the highest circulation per issue during 1966?

a) McCall's

b) Saturday Evening Post

c) Look

d) Better Homes & Gardens

e) Life

Note: The format for this problem and those following is to be identical to those of the illustrative and practice problems.

PROBLEM_____

Which of these rivers is the longest, from source to outflow?

a) Mekong

b) Ganges

c) Volga

d) Congo

e) Danube

PROBLEM_____

Which of these U. S. cities has the greatest population (1960 census)?

a) Cleveland

b) Houston

c) Boston

d) Pittsburgh

e) Washington, D. C.

PROBLEM_____

Which of these states had the highest per capita income in 1966?

a) Ohio

b) Wisconsin

c) Hawaii

d) Kansas

e) New Hampshire

PROBLEM_____

Which of these U. S. seaports had the most maritime traffic (vessels arriving and sailing) in the first half of 1967?

a) Seattle

b) San Francisco

c) Houston

d) Baltimore

e) Boston

PROBLEM_____

Which of these causes of accidental deaths has been the most common in recent years (1960-1966)?

a) Machinery

b) Poison gases

c) Firearms

d) Burns

e) Drowning

PROBLEM_____

Which of these companies had the highest total sales or revenues for 1966?

a) Mobil Oil Corp.

b) General Electric Co.

c) U. S. Steel Corp.

d) Eastman Kodak

e) Radio Corporation of America

PROBLEM_____

Which of these states has the greatest population (1960 census)?

a) New Jersey

b) Michigan

c) Minnesota

d) Indiana

e) Ohio

APPENDIX E

QUESTIONNAIRE

Please be frank with your answers to these questions: SPAN may be used by groups larger than two people. Would you ever like to use the method to reach some group decision in "everyday" life? (Place a check anywhere on the line.)

Definitely Not	No	Maybe	Yes	Definitely Yes

How do you now feel about the experiment?

I liked the exp. very much	I liked the exp. a little	I neither liked nor disliked the exp.	I disliked the exp. a little	I disliked the exp. very much

In many studies, psychologists are interested in the possibility that the results might have been partly determined by the subjects thinking that they had been deceived by the experimenter. By the very nature of psychological research, you were not fully informed about all that we are interested in. But do you think you have been deliberately lied to, deceived, or tricked in any way?

Yes_____ No_____

APPENDIX F

ANSWERS TO PROBLEMS

In descending order:

Most popular breeds of dogs: St. Bernards, Doberman Pinschers, Scottish Terriers, English-Springer Spaniels, Golden Retrievers.

Magazines with highest circulation per issue: McCall's, Look, Life, Better Homes & Gardens, Saturday Evening Post.

Longest rivers: Congo, Mekong, Volga, Danube, Ganges.

Population of cities: Houston, Cleveland, Washington, D. C., Boston, Pittsburgh.

Per capita income: Hawaii, Ohio, Wisconsin, Kansas, New Hampshire.

Maritime traffic: San Francisco, Baltimore, Houston, Seattle, Boston.

Causes of accidental deaths: burns, drowning, firearms, machinery, poison gases.

Highest total sales or revenues: General Electric, Mobil Oil Corp., U. S. Steel Corp., Radio Corporation of America.

Population of states: Ohio, Michigan, New Jersey, Indiana, Minnesota.

APPENDIX G

SUPPLEMENTAL FIGURES

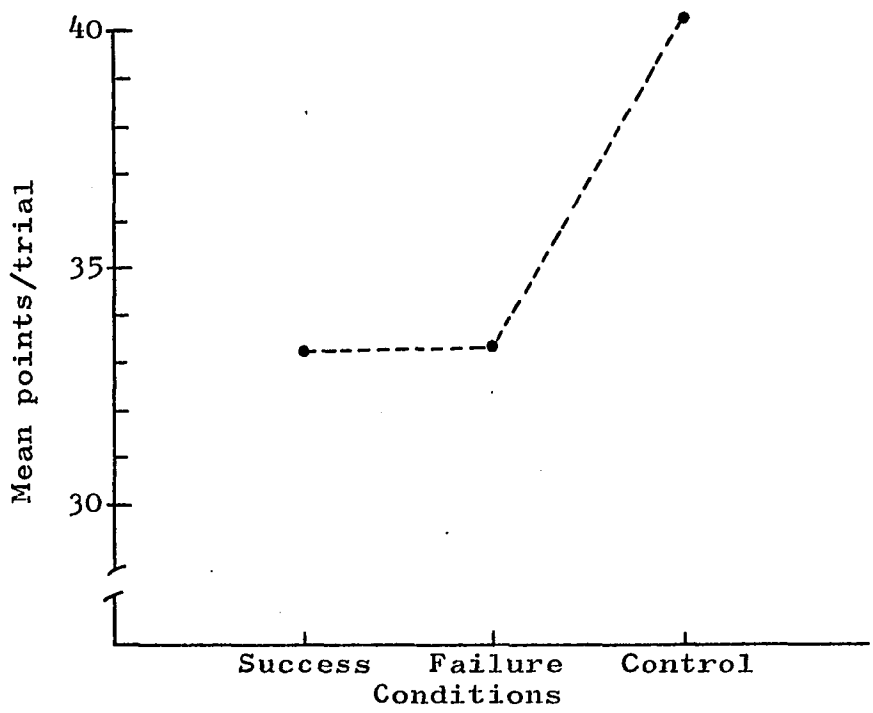


Figure 7. Main effect of conditions on points directed to partners.

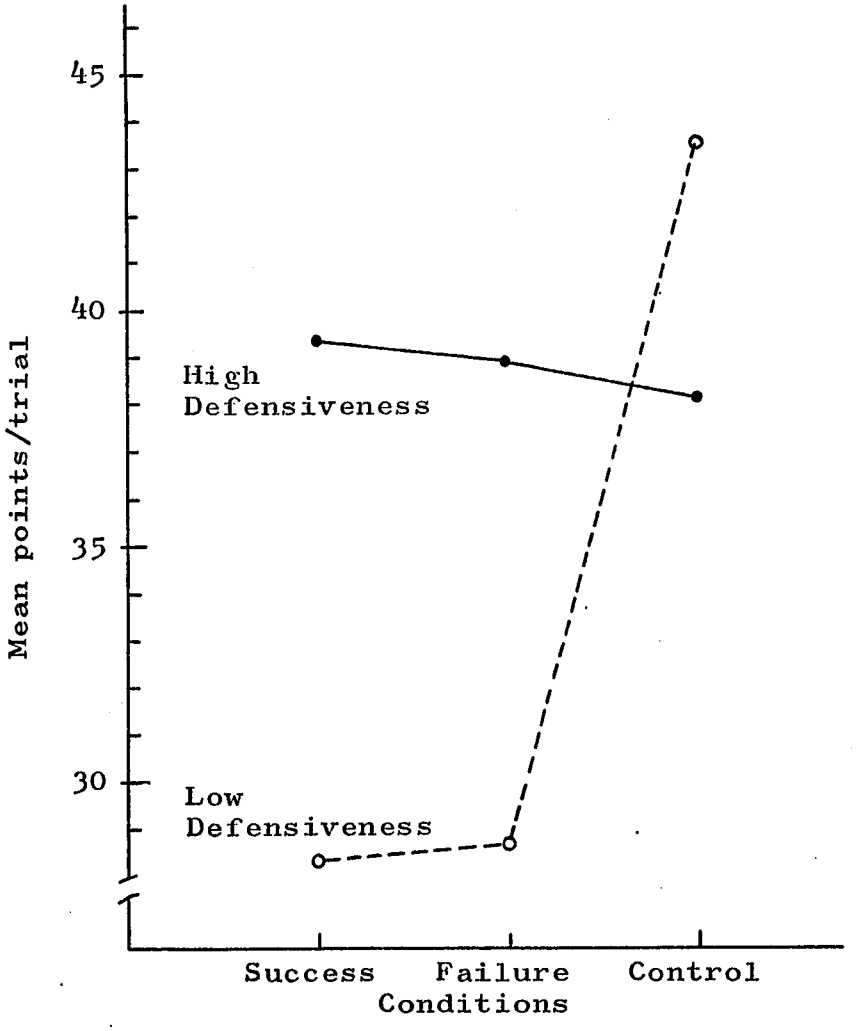


Figure 8. Effects of the interaction of defensiveness and conditions on points directed to partners.

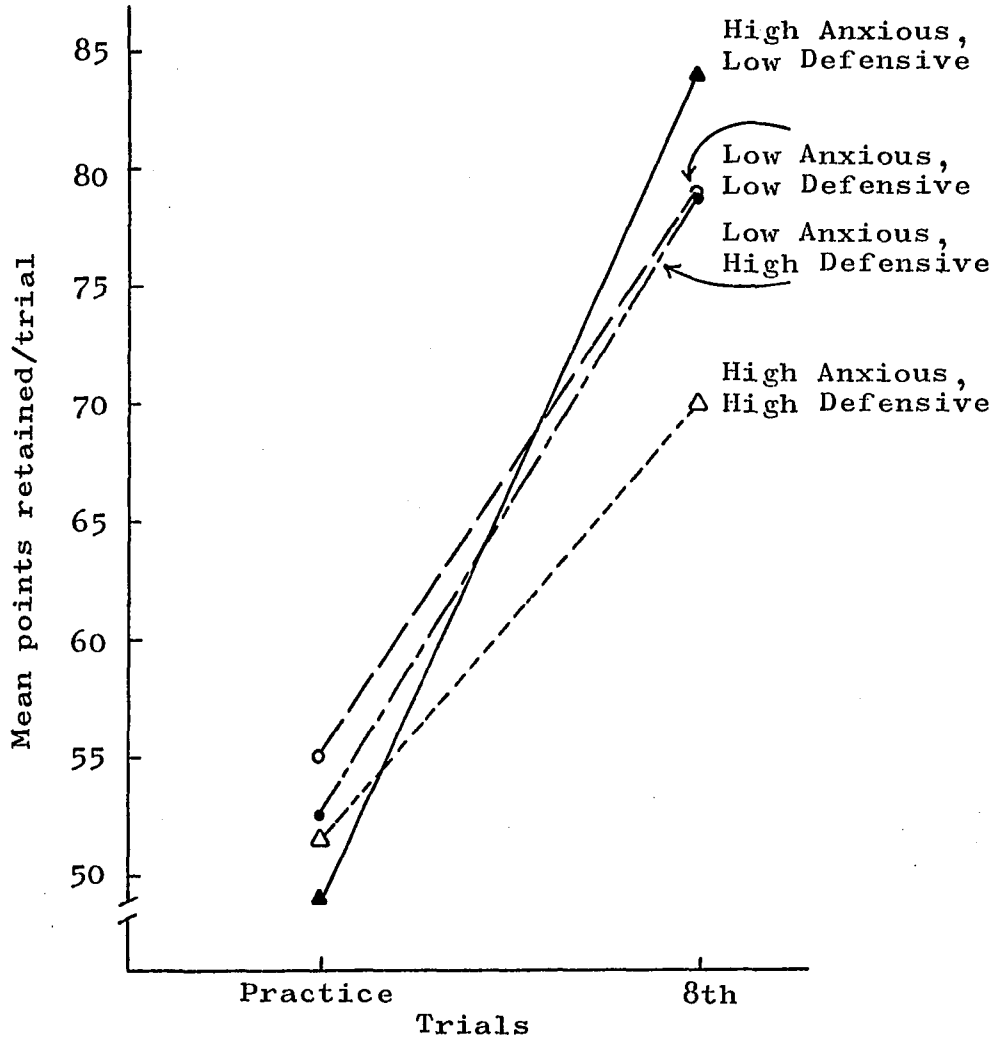


Figure 9. Effects of the interaction of trials x anxiety x defensiveness on the adaptive shifting of superior members.

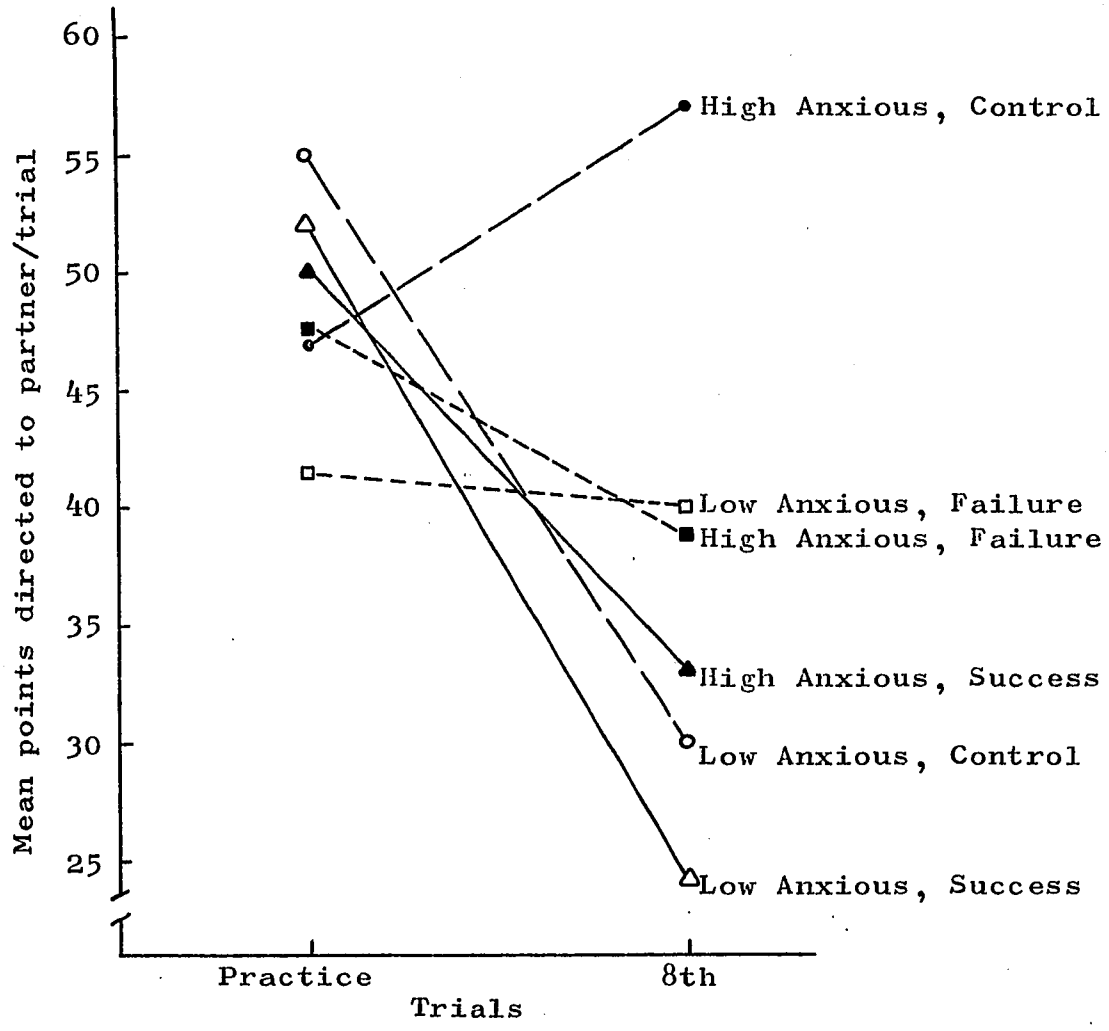


Figure 10. Effects of the interaction of trials x conditions x anxiety on the adaptive shifting of inferior members.

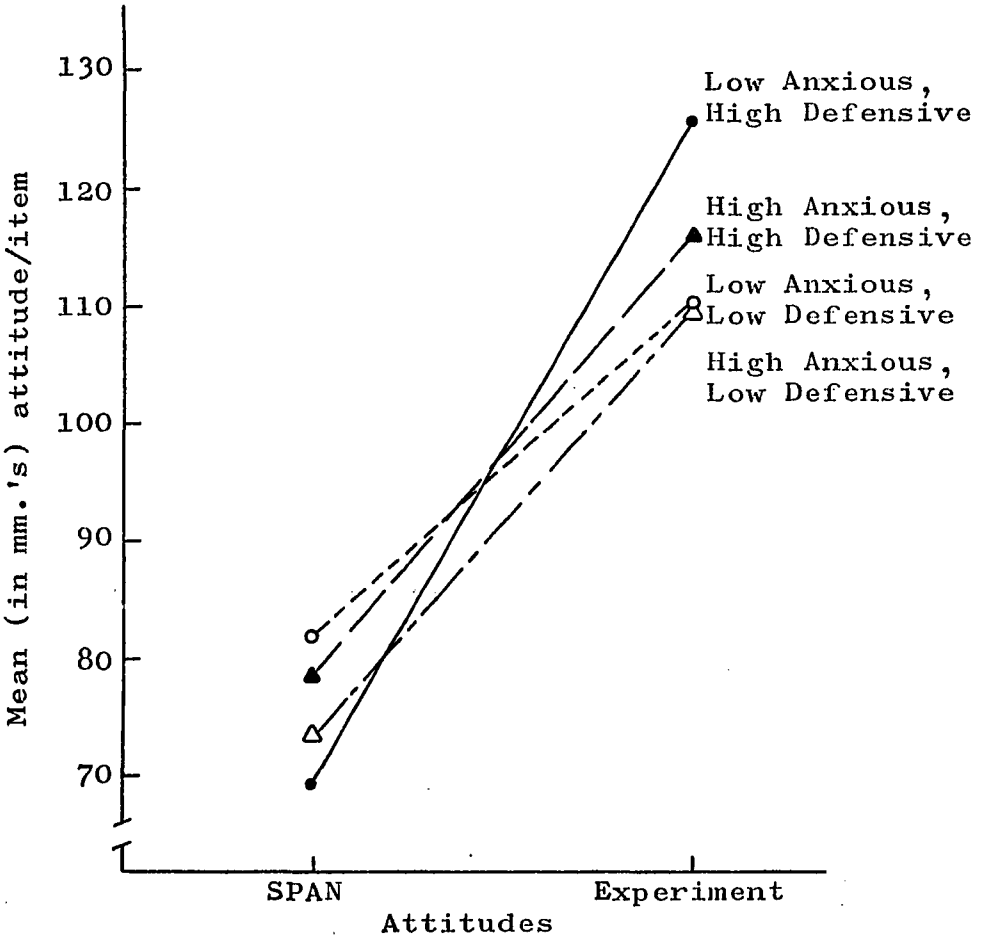


Figure 11. Effects on the interaction of defensiveness x anxiety x SPAN-experiment on attitudes toward SPAN and the experiment.

REFERENCES

Alpert, R., and Haber, R. N. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Atkinson, J. W. Motivational determinants of risk-taking behavior. Psychological Review, 1957, 64, 359-372.

Atkinson, J. W., and Feather, N. T. A theory of achievement motivation. New York: John Wiley & Sons, 1966.

Beckwith, J., Iverson, M. A., and Reuder, M. E. Test anxiety, task relevance of group experience, and change in level of aspiration. Journal of Personality and Social Psychology, 1965, 1, 579-588.

Crowne, D. P., and Marlowe, D. A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 1960, 24, 349-354.

Crowne, D. P., and Marlowe, D. The approval motive: Studies in evaluative dependence. New York: John Wiley & Sons, 1964.

Davidson, W. Z., Andrews, T. G., and Ross, S. Effects of stress and anxiety on continuous high-speed color naming. Journal of Experimental Psychology, 1956, 52, 13-17.

Fisher, G. M., and Kramer, R. A. The relation of the Marlowe-Crowne Social Desirability scale to the Cattell Anxiety scale. Journal of Clinical Psychology, 1963, 19, 204-205.

Ford, L. H., Jr., and Hersen, M. Need approval, defensive denial, and direction of aggression in a failure-frustration situation. Journal of Personality and Social Psychology, 1967, 6, 228-232.

Guilford, J. P. Psychometric methods. New York: McGraw-Hill, 1954.

Hall, R. L. Group performance under feedback that confounds responses of group members. Sociometry, 1957, 20, 297-305.

Herron, E. W. Relationship of experimentally aroused achievement motivation to academic achievement anxiety. Journal of Abnormal and Social Psychology, 1964, 69, 690-694.

Hitchcock, J. D., Jr. The SPAN group decision-making method in fraternities. Unpublished master's thesis, University of Arizona, 1967.

Johnston, W. A., and Nawrocki, L. H. Effect of simulated social feedback on individual tracking performance. Journal of Applied Psychology, 1967, 51, 145-151.

Katchmar, L. T., Ross, S., and Andrews, T. G. Effects of stress and anxiety on performance of a complex verbal-coding task. Journal of Experimental Psychology, 1958, 55, 559-564.

Kelly, J. F. The 3-point SPAN group decision-making method in sororities. Unpublished master's thesis, University of Arizona, 1968.

Kogan, N., and Wallach, M. A. Risk taking: A study in cognition and personality. New York: Holt, Rinehart, and Winston, 1964.

Kogan, N., and Wallach, M. A. Personality and situational determinants of judgmental confidence and extremity. British Journal of Social and Clinical Psychology, 1965, 4, 25-34.

Kogan, N., and Wallach, M. A. Group risk taking as a function of member's anxiety and defensiveness levels. Journal of Personality, 1967, 35, 50-63.

Lichtenstein, E., and Bryan, J. H. CPI correlates of the need for approval. Journal of Clinical Psychology, 1966, 22, 453-455.

Lillyquist, M. J. FORTRAN computation of a table for the SPAN decision-making method in dyads. Unpublished master's thesis, University of Arizona, 1968.

Lucas, J. D. The interactive effects of anxiety, failure, and intra-serial duplication. American Journal of Psychology, 1952, 65, 59-66.

MacKinnon, W. J. Development of the SPAN technique for making decisions in human groups. American Behavioral Scientist, 1966, 9(9), 9-13. (a)

MacKinnon, W. J. Elements of the SPAN technique for making group decisions. Journal of Social Psychology, 1966, 70, 149-164. (b)

MacKinnon, W. J., and MacKinnon, M. M. The decisional design and cyclic computation of SPAN. Behavioral Science, 1969, 14, 244-247.

Mandler, G., and Sarason, S. B. A study of anxiety and learning. Journal of Abnormal and Social Psychology, 1952, 47, 166-173.

Metfessel, M. A proposal for quantitative reporting of comparative judgments. Journal of Psychology, 1947, 24, 229-235.

Milburn, T. W. Note on Goldfried's cross validation of Marlowe-Crowne social desirability items. Psychological Reports, 1968, 22, 653-654.

Osler, S. F. Intellectual performance as a function of two types of psychological stress. Journal of Experimental Psychology, 1954, 47, 115-121.

Rozeboom, William W. The fallacy of the null-hypothesis significance test. Psychological Bulletin, 1960, 57, 416-428.

Russell, D. G., and Sarason, I. G. Test anxiety, sex, and experimental conditions in relation to anagram solution. Journal of Personality and Social Psychology, 1965, 1, 493-496.

Sarason, I. G. The effect of anxiety and two kinds of failure on serial learning. Journal of Personality, 1956, 25, 381-392. (a)

Sarason, I. G. Effect of anxiety, motivational instructions, and failure on serial learning. Journal of Experimental Psychology, 1956, 51, 253-259. (b)

Sarason, I. G. The effects of anxiety and threat on the solution of a difficult task. Journal of Abnormal and Social Psychology, 1961, 62, 165-168.

Sarason, S. B., Mandler, G., and Craighill, P. G. The effect of differential instructions on anxiety and learning. Journal of Abnormal and Social Psychology, 1952, 47, 561-565.

Shoemaker, D. M. A FORTRAN IV program for the analysis of variance by the method of unweighted means. Behavioral Science, 1969, 14, 252.

Smith, C. P. Relationships between achievement-related motives and intelligence, performance level, and persistence. Journal of Abnormal and Social Psychology, 1964, 68, 523-532.

Stollack, G. E. Relationship between EPPS and Marlowe-Crowne SD scale. Psychological Reports, 1965, 16, 731-732.

Waterhouse, I. K., and Child, I. L. Frustration and the quality of performance. Journal of Personality, 1953, 21, 298-311.

Williams, J. E. Mode of failure, interference tendencies, and achievement imagery. Journal of Abnormal and Social Psychology, 1955, 51, 573-580.

Willis, J. E. A comparative study of different quantitative methods of obtaining group decisions by collecting and synthesizing individual decisions. Unpublished doctoral dissertation, University of Arizona, 1966.

Willis, J. E., Hitchcock, J. D., and MacKinnon, W. J. SPAN decision making in established groups. Journal of Social Psychology, 1969, 78, 185-203.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Zajonc, R. B. The effects of feedback and probability of group success on individual and group performance. Human Relations, 1962, 15, 149-161.