TRAINING, ASSESSMENT, AND TRANSFER OF PRODUCTIVE THINKING

Thesis for the Degree of M. A.
MICHIGAN STATE UNIVERSITY
RONALD CURTIS KIDDER
1972

LIBRARY
Michigan State
University

BINDING BY
HOAG & SONS'
BOOK BINDERY INC.

FEB 25 1985

ABSTRACT

TRAINING, ASSESSMENT, AND TRANSFER
OF PRODUCTIVE THINKING

By

Ronald Curtis Kidder

The present study sought to determine the merits of three distinct methods for enhancing productive thinking on each of two problem types. Criteria included to assess the efficacy of each method included (1) quality of solutions generated, (2) transfer of training to a different type of problem, and (3) persistence as measured three weeks after the training. A further ambition of the study was to determine if training intended to promote productive thinking would also influence judgment ability as measured by each subject's self-evaluations.

Four ten-item productive thinking tests were constructed. Two of the tests employed only sentence and consequence problems. The other two tests employed six different types of problems including hypotheses, sensitivity, titles, sentences, consequences, and graph conclusions.

Ninety-six subjects were randomly assigned to six experimental and to two control groups, providing twelve subjects per group. Within each experimental group one of

three methods proposed for stimulating productive thinking was examined (unassisted-group brainstorming, assisted-group brainstorming, and assisted-individual brainstorming), utilizing one of two problem types, sentence problems or consequence problems.

Upon completion of the training procedure each subject was administered two productive thinking tests, afterwhich the subject was instructed to return for the administration of the second set of tests three weeks later. The tests were coded and randomized, afterwhich two judges independently rated each of the solutions.

In general, only two of the training methods satisfied all of the criteria. When used in conjunction with consequence training problems, both assisted-individual and assisted-group brainstorming were found to be superior methods for enhancing productivity. In general, the assisted brainstorming groups better satisfied the criteria than the unassisted or traditional brainstorming groups.

The conditions necessary for transfer were found to be a function of both the nature of the training method and the nature of the training problem. Transfer of training from Problem A to Problem B did not necessitate transfer of training from Problem B to Problem A. These results were interpreted in terms of the demand characteristics operating and suggestions for future considerations were entertained.

Persistence of training as measured by other investigators was demonstrated. However, when the experimental groups were compared to an experienced control group, only a single instance of persistence was found to exist. Apparently after a three week period, simply having had practice with productive thinking problems was as good a training method as those used with the experimental groups. An explanation was suggested to account for this finding.

And finally, the ability to produce a creative solution was found to be distinct from the ability to accurately judge one's own solutions.

For each test-item, an interjudge reliability coefficient was determined, and for each test, Coefficient Alpha and the test-retest reliability coefficient were determined. These were all considered satisfactory for research purposes. Factor analyses were performed on each test to examine the credibility of the distinction between problem types. The tests using two item-types reduced to two factors, however, the tests using six item-types did not reduce to six factors. An explanation was offered to account for this finding.

TRAINING, ASSESSMENT, AND TRANSFER OF PRODUCTIVE THINKING

By

Ronald Curtis Kidder

A THESIS

Submitted to
Michigan State University
in partial fulfillment
for the degree of

MASTER OF ARTS

Department of Psychology

1972

ACKNOWLEDGMENTS

I would like to express my appreciation to Dr. Donald M. Johnson, chairman of my thesis committee, for his guidance and assistance in the preparation of this manuscript. Gratitude is also extended to both Dr. Gordon Wood and Dr. Richard Marshall for their participation and contributions to the thesis.

TABLE OF CONTENTS

Page

LIST OF TABLES

INTRODUCTION

According to Guilford (1967), there are two broad classes of thinking processes, convergent thinking and divergent thinking. Convergent thinking is characterized by problems having a single solution. In contrast, divergent thinking, more commonly referred to as productive thinking, is characterized by problems with many solutions that cannot be dichotomized as right or wrong, but instead are evaluated along such dimensions as remoteness, cleverness, appropriateness, or originality. The present research is focused on productive thinking and on methods useful in facilitating productive thinking.

Brainstorming Instructions

A considerable volume of research has been generated by the proponents of brainstorming. As initially advocated by Osborn (1957), this method of getting good ideas emphasizes the temporal segregation of hypothesis formation from the judicial evaluation of the adequacy of solutions. Typically the subjects are given a set to freely express initially whatever ideas occur to them, and are instructed to attempt to solve whatever problems are presented by recording all tentative solutions, postponing judgment of those solutions to a subsequent time period.

The current situation with regard to brainstorming seems to revolve about two research questions. First, which is of greater value, individual brainstorming or group brainstorming? And second, by emphasizing quantity rather than quality, does brainstorming result in more superior solutions?

The literature contrasting individual and group brainstorming has basically found that when brainstorming activity is fully developed, the individual procedure often results in the best performance.

In a study conducted by Taylor, Berry, and Block (1957), three problems were presented to ninety-six Yale Juniors and Seniors who had previously worked together in small group sessions. Forty-eight of these subjects were divided into twelve groups of four men each; the other forty-eight brainstormed alone. For each of the three problems, the individuals produced an average of twice as many different ideas. Taylor et al. concluded that group participation had an inhibitory influence on creative thinking during brainstorming.

In a study reported by Dunnette, Campbell, and Jaastad (1963), problems were presented for brainstorming to forty-eight research scientists and forty-eight advertising personnel employed at Minnesota Mining and Manufacturing Company. Within a counterbalanced design, each subject brainstormed certain problems individually and

equated problems as a member of a four-man group. It was found that individuals produced more ideas than groups, and their solutions were of greater quality. Under individual conditions, 23 of the 24 groups produced a larger number of different ideas. The superiority of the individual over group brainstorming was relatively greater when it was preceded by group brainstorming, similar to the results of Parnes. It was concluded that group participation was accompanied by certain inhibitory influences even under conditions that place a moratorium on all criticism.

A further investigation examining the relationship between individual and group brainstorming was reported by Voytas (1968), who attempted to investigate the effects of various combinations of group and individual participation on tasks of creative production. The main dependent variable was the total quantity of ideas produced. The superiority of the individuals in producing ideas was demonstrated as compared to ideas produced by groups. And the production of subjects who did both was significantly greater than the production of subjects who spent their entire time either in a group or who performed individually. With respect to quality, additional analyses were made and conflicting results were obtained. Voytas concluded that the mean quality must somehow interact with the specific type of problem and the type of brainstorming group.

A finding by Collaros and Anderson (1969) may explain, in part, why the above studies comparing the creativity of brainstorming individuals or groups endorse the superiority of the individual condition. Collaros and Anderson hypothesized that the perceived expertise of other members may make a brainstorming group less effective than the pooled results of its members working alone. They employed three groups in their study. In an all-expert condition, each member of a brainstorming group was told that the other members had previously worked in such groups. In a one-expert condition, it was stated that only one member (unidentified) had this experience. No information was given in a control condition. It was reported that the subjects felt more inhibited in the all-expert condition than in the one-expert condition, which in turn, was more inhibited than the control condition. And as hypothesized, originality and practicality of ideas varied according to the degree of felt inhibition, with the control condition having the highest originality and practicality scores.

Addressing themselves to the second question, Meadow and Parnes (1959) found that by emphasizing quantity rather than quality in a thirty-hour course in brainstorming, not only were more solutions produced by these subjects, but more of the solutions produced were of superior quality.

In a further study, Meadow, Parnes, and Reese (1959) studied the effectiveness of the brainstorming procedure

using only subjects who were members of the course in brainstorming. Each subject was given two problems which required creative ability. One problem was administered under brainstorming instructions, the other problem was administered under nonbrainstorming instructions. Again it was found that significantly more good solutions were produced under the brainstorming instructions, and further, significantly more good solutions were produced under the brainstorming instructions when they were given first than when they were followed by nonbrainstorming instructions.

In yet another study conducted by Parnes and Meadow (1959), the performance of a group of untrained subjects was compared to the performance of a group of subjects trained in the use of brainstorming. Basically, they found that brainstorming instructions were an effective method for increasing the production of ideas without sacrificing quality, and that they were even more effective if preceded by extensive training in the use of brainstorming.

The persistence of the brainstorming training was also investigated (Parnes and Meadow, 1960). Experimental and control groups matched for vocabulary ability were compared on six creative ability tests. Experimental subjects were those who had completed the course from eight months to four years previous to the experiment. Control subjects were students registered but uninstructed in the brainstorming course. The results indicated that the experimental

subjects significantly outperformed the control subjects on each of the six creativity tests. It seemed apparent, then, that the increased productivity produced by the course had persisted for at least eight months after its completion. However, an alternative interpretation not considered by these investigators is the possibility that the experimental subjects outperformed the control subjects as a result of their prior exposure to the testing situation. This consideration was not properly controlled for, and hence, the issue of persistence remains unclear.

In a study reported by Compton (1968), the brainstorming approach was contrasted to a group discussion approach. On the initial testing situation and ten to fourteen weeks later, the brainstorming groups surpassed the discussion group in terms of individual gain scores on such factors as originality and seeing differences. Persistence over time once the brainstorming ability was developed seemed to be demonstrated, at least on some productive thinking tests.

Further addressing the distinction between quality and quantity of solutions is a study reported by Weisskopf-Joelson and Eliseo (1961). In their research, one group of subjects was instructed to suspend criticism of ideas produced and another group was instructed to employ critical censorship. On each problem the noncritical group was found to produce a greater number of responses than the

critical group similar to the findings reported by Meadow and Parnes (1959), Parnes and Meadow (1959), and Meadow, Parnes, and Reese (1959). However, the preponderance of responses from the noncritical group was found to owe its origin to the large number of responses of relatively low quality. The mean quality of the critical group's responses was found to be greater than the mean quality of the noncritical group's responses.

The relationship between the absolute number of solutions, the absolute number of good solutions, the proportion of good solutions, and the mean number of good solutions as a function of the type of instructions is documented quite well in a study reported by Johnson, Parrott, and Stratton (1968). With multiple-solution instructions more solutions were obtained at each quality level including a larger increase in the number of inferior solutions. Thus, by increasing the number of solutions produced it was found that more solutions were obtained, but with a sacrifice in mean quality. Compared to single-solution instructions, more superior solutions were produced under multiple-solution instructions. Stratton, Parrott, and Johnson (1970) suggested that if someone other than the producer is to evaluate the solutions, more superior solutions will be produced by multiple-solution instructions despite an overall reduction in quality. If overall quality is desired, single-solution instructions will give better quality per

solution. That is, proportionately more superior solutions will be provided by single-solution instructions.

Criteria-Cued Instructions

To provide the problem solver with a better concept of what is expected of him when instructed to be creative, an obvious technique is simply to supply him with the criteria of "creative" performance. It would seem that when complemented with examples of good and poor solutions even further gains could be obtained. And in fact, several studies have shown this to be the case. Christensen, Guilford, and Wilson (1957) and Johnson, Parrott, and Stratton (1968), for example, have found that an emphasis on quality increased the mean quality and decreased the number of solutions over instructions emphasizing quantity alone.

Gerlach, Schutz, Baker, and Mazer (1964) also found that criteria-cued instructions yielded the most good responses when contrasted to other instructions including brainstorming.

When Johnson (1968) employed criteria-cued instructions coupled with instructions for both single and multiple-cued solutions, he found that criteria-cued instructions increased single-solution quality, and when compared to mean multiple-solution quality, single solutions were better. With respect to criteria-cued instructions, then, the best approach would seem to have the subject supply only a single solution rather than many.

"Assisted-Brainstorming" Instructions

In the present study a synthesis of brainstorming and criteria-cued instructions was attempted. This "assisted-brainstorming"approach involved adapting the brainstorming procedure, as typically construed, to work within the limitations imposed by a set of criteria. It was hoped that the normal brainstorming instructions such as those emphasizing quantity, free-wheeling, interaction of ideas, and deferred judgment could operate within a set of criteria. Examples of good and poor solutions and a possible strategy to employ were also provided. The index of performance, as suggested by Johnson (1968), was a single solution considered by the subject as representing his best effort.

To examine the efficacy of this approach after a 45-minute training period, three criteria were decided upon, including (1) the quality of the solutions generated, (2) transfer of training to a different type of problem, and (3) persistence as measured three weeks after the training. To assess how this brainstorming approach contrasts with brainstorming in the traditional sense, a group brainstorming group was added. And to assess the differential effects of group versus individual participation within the "assisted-brainstorming" approach, all of these conditions were employed.

A working hypothesis was that the assisted-group brainstorming approach is significantly better than either the unassisted-group brainstorming approach, or the control

group; the unassisted-group brainstorming approach is significantly better than either the assisted-individual brainstorming approach, or the control group; and finally, the assisted-individual brainstorming approach is significantly better than the control group.

The author felt that because of the relatively brief training period imposed in the present study (45 minutes), the group brainstorming approach would better benefit the subjects than would the individual brainstorming approach. Previous investigators, of course, have shown that individual brainstorming is generally superior to group brainstorming. However, in these investigations the brainstorming activity was fairly well advanced.

The problem entertained with respect to group versus individual brainstorming was as follows: Within group brainstorming a greater range of solutions seems to be generated, but within that range will the solutions be as penetrating, as well thought-out as might be provided for in individual brainstorming?

Questions of Theoretical Interest

A question of theoretical interest was to determine if training intended to promote productive thinking would also influence judgment ability as measured by each subject's self-evaluations. In a study reported by Johnson et al. (1968), it was found that making subjects evaluate their own

solutions after production to each of a series of problems produced no increase in quality even with criteria-cued instructions. Johnson and Zerbolio (1964) found that allowing subjects to practice evaluating 25 solutions, or to read the evaluations of experts was ineffective in increasing solution quality or judgment accuracy on a transfer task.

The problem used to construct the productive thinking tests employed in the present study were similar to those used by Guilford (1967). However, an essential feature of this study lacking in most other studies was the nature of the productive thinking tests. The tests were constructed of items of psychological content relevant to introductory psychology courses at Michigan State University. The training items were of a similar nature. Two particular problem types were emphasized in the present study, both of which were chosen out of convenience, sentence problems and consequence problems.

Related to the issue of productive thinking tests are several questions of theoretical interest. Since two of the tests constructed were designed to assess transfer effects, and hence, were composed entirely of two distinct item-types, when factor analyzed each test should reduce to two factors, productivity as measured by sentence problems and productivity as measured by consequence problems. Similarly, since two of the tests were composed of six distinctly different problem-types, when factor analyzed,

each of these tests should reduce to six factors, productivity as measured by each of the six problems. And when the four productive thinking tests are factor analyzed, the two sets of parallel tests should each emerge as separate factors, since productivity as measured by different item-types measures essentially different aspects of productivity.

Also of some interest are the values of the reliability coefficients for each of the tests and an examination of how each of the tests are intercorrelated. In a study reported by Cave (1970), creativity tests correlated as highly with each other as they did with intelligence scales. The average correlation among the creativity scales was .37, while the average correlation between the creativity and the intelligence scales was .38. The average correlation among the intelligence scales was .51.

METHOD

Subjects

A sample of ninety-six students (54 female; 42 male) enrolled an introductory psychology courses during the Spring Term of 1971 at Michigan State University volunteered to serve as subjects for this study.

Materials

Four ten-item tests were constructed to measure productive thinking using problems similar to those employed by Guilford (1967). In contrast to the problems employed by Guilford, however, the content of the problems developed in these tests included concepts and issues familiar to introductory psychology students at Michigan State University. These tests may be found in Appendix A.

On two of the tests, A1 and B1, only two kinds of problems were employed. Five sentence problems and five consequence problems alternating in the form SCSC . . . SC for each test were used. On the other two tests, A2 and B2, each of six types of problem was employed: sentences, consequences, graphs, hypothesis, titles, and

sensitivity to problems. Each of the items used on the latter two tests had been used previously (Johnson and Kidder, 1972), and were found to be the best items in terms of item-total correlations. The latter two tests were constructed such that they had parallel item types and their item-total correlations, on the average, were the same.

Two "productive thinking guides" were constructed, one of which emphasized the criteria and examples of good and poor solutions to consequence problems; the other emphasized the criteria and examples of good and poor solutions to sentence problems. (The examples for the sentence problems were as proposed by Johnson, 1968.) The productive thinking guides may be found in Appendix B.

Two sets of training problems, one consisting entirely of sentence problems, the other consisting entirely of consequence problems were developed. Each set emphasized psychological terms and concepts. In developing the sentence items, it was intended to employ terms of variable meaning and which were versatile semantically, e.g. sample, graph, structure, bit. Totally concrete terms were given lesser consideration but were unavoidable. Previous experience (Johnson and Kidder, 1972) had indicated that terms of variable meaning had the greatest potential and were most conducive to creative performance. Less versatile words often could only be combined in at most very mundane combinations. Refer to Appendix C for a listing of training problems.

Procedure

The ninety-six subjects were randomly assigned to six experimental and to two control conditions, providing twelve subjects per condition. Within each experimental condition, one of three methods of stimulating productive thinking was studied, utilizing one of two problem types, sentences or consequences. The methods of stimulating productive thinking included individual brainstorming with exposure to criteria and examples of good and poor solutions (assisted-individual brainstorming); group brainstorming without the examples and criteria of good and poor solutions (unassisted-group brainstorming); and group brainstorming with exposure to criteria and examples of good and poor solutions (assisted-group brainstorming). Each group brainstorming condition was subdivided into four groups of three subjects each, hence, only three subjects brainstormed at a time.

Unassisted-Group Brainstorming Conditions.--These groups were instructed as to the general nature of brainstorming and were then instructed that their task was to produce as many creative solutions to each of ten training problems as possible through the use of Osborn's deferred judgment principle (either sentence or consequence problems) via the following passage:

> What I would like each of you to do is to produce as many creative solutions as you can think of to problem number one, and to verbally express your solutions regardless of how simple or silly you may think they are, because while they may seem unimportant to you, they may help another person in the group by allowing him to see the problem in a new perspective. Again, say everything that comes to mind and do not be critical of your ideas or the ideas of others. When you as a group feel you have exhausted yourselves with a particular problem, you may go on to the next problem. After 45 minutes you will complete your practice, afterwhich each of you will perform individually on a similar task. Introduce yourselves and begin.

It was sometimes necessary for the experimenter to initiate the brainstorming activity, and on occasion, to act as a catalyst. Following the completion of each problem, the experimenter advised the group as to their progress, and when necessary suggested creative solutions not considered by the group. The next problem was then entertained.

Upon completion of their training, the subjects were administered two ten-item tests, Test A1 followed by Test A2. The subjects worked independently and were instructed to be creative in producing solutions to the test problems and to write only their best solution to each problem. They were further instructed to rate their own solutions as they worked on a scale from 1 to 7, with a rating of 7 representing a very superior solution, and a rating of 1 representing a very poor solution. They were told of the forty-five minute time limit and were instructed to complete all twenty items.

Upon completion of tests A1 and A2, the subjects were instructed to return in three weeks for a retest, during which they were administered the parallel forms, B1 and B2,

in that order. The same test conditions were imposed.

Assisted-Individual Brainstorming Conditions.--These groups were given the training guide to read emphasizing sentence or consequence problems, afterwhich their questions were answered. Subjects who were to brainstorm with consequence problems were informed as to the nature of brainstorming and were then instructed to:

> Attempt to think of consequences of consequences, to consider your initial response, often of relative low creativity, as a starting point, and to branch out from there such that the end product is more remote and far reaching, perhaps bearing little relationship to the initial premise when casually observed.

Subjects who were to brainstorm with sentence problems were informed as to the nature of brainstorming and were then instructed to:

> Use all of the words in a sentence; consider as many possible meanings and figures of speech as possible for each word; combine the highly unusual interpretations into a sentence; disregard item 2c in the rating guide, and instead use modifiers given in disparate locations in the sentence; do not modify other words given by the modifiers given; be clever, but no superficially so.

The training problems were then administered and the subjects were instructed that they would have 45 minutes with which to practice, afterwhich they would begin working on a similar task but without help. During the training session, the subjects were instructed to record their final solution and allow the experimenter to advise them of their progress. Solutions not considered by the subjects were suggested.

The identical post-training procedure as mentioned above for the unassisted-group brainstorming conditions was implemented.

Assisted-Group Brainstorming Conditions.--These groups were given the appropriate rating guide and verbal instructions, as in the individual brainstorming condition. They were then informed as to the nature of brainstorming and were given the oral instructions with respect to group brainstorming. From here, the same procedure as mentioned above in the other brainstorming conditions was implemented.

Control Conditions.--There were two control groups. The first control group (AB) was administered tests A1 and A2 under the same conditions and instructions as were each of the experimental groups, however, this group had not had any previous training. These control subjects were instructed to return in three weeks for a retest, wherein tests B1 and B2 were a administered under identical conditions. The second control group (B) was only administered tests B1 and B2.

To summarize, ninety-six subjects were randomly assigned to one of six experimental groups or to one of two control groups. Subjects in each of the experimental groups trained on a specific productive thinking problem (C or S) in a specific manner for a period of 45 minutes. Subjects in the control groups received no training. After the training procedure the subjects were administered tests A1

and A2. A 45-minute time limit was imposted for the completion of those tests.

Test A1 was constructed to assess productive thinking on the specific problem for which the subjects had trained (C or S), and to assess the transfer effect on a specific problem for which the subjects had not trained (C or S). Test A2 was constructed to assess productive thinking using several different types of problems previously found to be good in terms of item-total correlations.

Thus, during the administration of Tests A1 and A2 there were six experimental groups and one control group consisting of the following:

a. Assisted-individual brainstorming with consequences.

b. Unassisted-group brainstorming with consequences.

c. Assisted-group brainstorming with consequences.

d. Assisted-individual brainstorming with sentences.

e. Unassisted-group brainstorming with sentences.

f. Assisted-group brainstorming with sentences.

g. Control (AB)

After a three-week interval the subjects returned and were administered under identical conditions the parallel forms of Tests A1 and A2, Tests B1 and B2, respectively. A second control group was added to assess the exposure effects of the initial testing situation.

Subsequently, each subject's test was coded according to the group he was in and the tests were randomized. The responses to each item were independently rated by two experienced judges, neither of which had knowledge of the code. Tests A1 and B1 were rated by one set of judges; Tests A2 and B2 were rated by another set of judges.[1] Following the rating, the tests were regrouped into their initial conditions and the appropriate statistical analyses were conducted.

---

[1]I would like to express my sincere appreciation to both Richard Hoffman and Len Sawisch for their assistance in rating what at times seemed an endless number of responses.

RESULTS AND DISCUSSION

Interjudge Reliabilities

Two judges working independently rated each subject's solutions on a scale from 1 to 7, with a rating of 7 representing a very superior solution, and a rating of 1 representing a very poor solution. Table 1 illustrates the interjudge reliabilities obtained for each item on each of the four tests.

TABLE 1

Interjudge Correlations for Each Item on Each of Four Tests

	Tests			
Items	Test A1	Test B1	Test A2	Test B2
1	.70	.86	.97	.84
2	.85	.73	.92	.96
3	.76	.70	.94	.94
4	.71	.80	.97	.94
5	.72	.74	.94	.93
6	.76	.72	.90	.92
7	.75	.73	.88	.83
8	.77	.64	.93	.97
9	.84	.75	.96	.84
10	.76	.71	.89	.92

Note.--For tests A1 and A2, N=84; for tests B1 and B2, N=88.

The judges rating the solutions to Tests A1 and B1 were not the same as those judges rating Tests A2 and B2. As is apparent, the second set of judges did better which is reflected in their higher correlations. These judges spent more time both in terms of deciding upon criteria for each problem, and in terms of actual time spent rating the solutions. All correlations were considered satisfactory.

Test Reliabilities

Table 2 indicates the item-total correlations obtained for each item for each of the four tests administered. (For tests A1 and A2, N=84; for tests B1 and B2, N=88 rather than N=96 since some of the subjects did not return for the second administration of the tests.)

TABLE 2

Item-total Correlations for Each Item on Each of Four Tests

	Tests			
Items	Test A	Test B1	Test A2	Test B2
1	.44	.60	.17	.31
2	.60	.71	.28	.49
3	.42	.64	.51	.46
4	.56	.57	.61	.34
5	.48	.66	.60	.58
6	.61	.54	.54	.42
7	.33	.41	.39	.50
8	.47	.54	.63	.28
9	.52	.62	.70	.61
10	.69	.56	.59	.60

Note.--For tests A1 and A2, N=84, for tests B1 and B2, N=88.

If one considers items having an item-test correlation below .40 as poor, only one item need be eliminated from Test A1; three items should be eliminated from Test A2; no items from Test B1 should be eliminated; and three items from Test B2 should be eliminated. Eliminating these items and replacing them with better items would enhance the overall reliability of each test. With respect to quality, it was found that the best items were constructed so as to allow a highly variable response.

Coefficient Alpha, the reliability in terms of internal consistency, was calculated for each of the tests. For Test A1 the value obtained was .71; for Test A2, .68; and combining Test A1 and A2 resulted in a value of .74. For Test B1 the value obtained was .80; for Test 2, .66; and combining tests B1 and B2 resulted in a value of .80. These results were considered quite satisfactory for research purposes. Since the sample from which these data were drawn was trained to think more productively, and hence, was performing at an enhanced level, a restriction of range was improved through the training procedure. The test reliabilities reported are thus lower than what would be found with a sample of untrained subjects.

On the basis of these results, two things are apparent. First, employing the best items from previous research when constructing tests A2 and B2 resulted in much higher reliabilities, from a previous average of .37 (Johnson and Kidder, 1972), to these results of .68 and .66, respectively.

Secondly, the more homogenous a test, the higher its reliability; tests A1 and B1 were each constructed of only two kinds of problems, whereas, tests A2 and B2 employed six different item types.

Since parallel forms were constructed and the tests were administered three weeks apart, alternative-form reliability coefficients were determined. Table 3 indicates the intertest correlations obtained. (N=71, subjects who completed all four tests.)

TABLE 3

Intertest Correlations

Tests	Test A1	Test A2	Test B1	Test B2
Test A1	1.00			
Test A2	.28	1.00		
Test B1	.47	.32	1.00	
Test B2	.36	.51	.40	1.00

Note.--N=71, subjects who were administered all four tests.

Of the greatest concern are the correlations between the parallel forms, tests A1 and B1; tests A2 and B2, which we find are .47 and .51, respectively. Since there appears to be a marked drop in reliability as compared to Coefficient Alpha, three interpretations are possible; there was a systematic difference in content; subjectivity of scoring; or

variations in people over this time interval. Each of these factors probably contributed to some of the difference, but it is thought that the third factor was the largest since all subjects, including the control subjects, improved over time. (Causing a further restriction of range.)

Of some interest is the fact that the parallel tests had higher correlations with one another than did the non-parallel forms. It has been frequently cited in the literature that tests measuring productive thinking have only moderate correlations with one another, ranging on the average between .30 and .40. The correlations between the nonparallel tests agree with these findings.[2]

Factor Analysis

To examine the credibility of the distinction between problem types factor analyses were performed using principal axes solutions and were interpreted via quartimax and varimax rotations. Two, three, four, and five factor solutions were generated. The following solutions to be reported were found to have both the best simple structure and were the most psychologically interpretable.

---

[2] A possible explanation for the higher correlations between the parallel forms is that they were essentially sampling from the same domain or item pool. However, in this particular study, an alternative explanation is also in order for it is possible that the higher correlations were the product of an order effect, at least with respect to tests A2 and B2. In Test Session One, Test A1 was administered first, A2 second. In Test Session Two, Test B1 was administered first, B2 second. It is entirely possible that the subjects simply spent more time on the first test, and as a result, did less well on the second test.

A working hypothesis was that for tests A1 and B1, two factors would emerge since two problem types were employed. For tests A2 and B2 it was hypothesized that six factors would emerge, since six problem types were employed.

Test A1.--Table 4 summarizes both the quartimax and varimax factor analytic rotations. (N=84). As was predicted two factors emerged, productivity as measured by consequence problems representing one factor, and productivity as measured by sentence problems representing the other factor. For each item, factor loadings for the alternative factor were less than .20, except for item 10, a consequence problem, where the loading was .32 for the sentence factor. The proportion of variance accounted for by this two-factor solution was .47. Both quartimax and varimax solutions were essentially the same.

TABLE 4

Quartimax and Varimax
Rotational Analysis for Test A1

Problem Types	Rotations			
	Quartimax Loadings		Varimax Loadings	
	Factor 1	Factor 2	Factor 1	Factor 2
Sentence	.1407	.5657	.1414	.5655
Consequence	.7177	.1330	.7179	.1322
Sentence	-.0301	.7579	-.0292	.7599
Consequence	.8095	-.0119	.8095	-.0129
Sentence	.0529	.7690	.0538	.7689
Consequence	.8017	.0758	.8088	.0749
Sentence	-.0674	.4514	-.0669	.4515
Consequence	.5044	-.0887	.5043	-.0894
Sentence	.1836	.5492	.1842	.5490
Consequence	.7160	.3167	.7164	.3158

Note.--N=84.

Test B1.--Table 5 summarizes both the quartimax and varimax factor analytic rotations for Test B1. (N=88 since not all subjects returned.) This solution was not as "clean." Psychologically, the two factor solution seemed to be the best. Subsequent solutions divided the items primarily within each of the two major factors, sentence problems and consequence problems. For example, items 1, 3, 5, 7 and 9 represented one factor. A subsequent four-factor solution divided this entry into two factors, 1, 3, and 5 representing one factor; 7 and 9 representing the other. This was not found to be interpretable in terms of content, or any other way.

Using the two-factor solution, there were three inconsistencies. Item 2, a consequence problem, had a loading of .42 on the sentence factor; item 3, a sentence problem, had a loading of .31 on the consequence factor; and item 5, a sentence problem, had a loading of .39 on the consequence factor. All other alternative loadings were less than .30. Both quartimax and varimax solutions were essentially the same. The proportion of variance accounted for by the two-factor solution was .51.

Test A2.--Table 6 summarizes both the quartimax and varimax factor analytic rotations for Test A2 (N=82, two subjects were not included since they completed less than half of the ten items.) A very "clean" two-factor solution emerged, but not along problem types. The first factor

TABLE 5

Quartimax and Varimax
Rotational Analysis for Test B1

Problem Types	Rotations			
	Quartimax Loadings		Varimax Loadings	
	Factor 1	Factor 2	Factor 1	Factor 2
Sentence	.2661	-.5738	.2482	-.5818
Consequence	.6334	-.4021	.6206	-.4215
Sentence	.3356	-.6683	.3147	-.6784
Consequence	.7062	-.0979	.7029	-.1198
Sentence	.4151	-.5565	.3977	-.5691
Consequence	.7054	-.0279	.7042	-.0497
Sentence	-.1439	-.7097	-.1659	-.7049
Consequence	.5303	-.2574	.5221	-.2737
Sentence	.1040	-.7914	.0795	-.7942
Consequence	.7528	.0285	.7534	.0052

Note.--N=88.

might most appropriately be termed "sufficient time for test completion," and the second factor might be termed "insufficient time for test completion." Factor one consisted of the first three items, factor two of the remaining seven.

An alternative hypothesis, however, would be that the first three items measure one factor with respect to productive thinking, the latter seven measure another. The problem with this interpretation is that it does not account for the fact that problems similar to the first three were present in the remainder of the test. The first three problems were also of different problem types. There was only one inconsistency with this two-factor solution, the third item had a loading of .33 on the second factor, insufficient

TABLE 6

Quartimax and Varimax
Rotational Analysis for Test A2

Problem Types	Rotations			
	Quartimax Loadings		Varimax Loadings	
	Factor 1	Factor 2	Factor 1	Factor 2
Sensitivity	-.0652	-.5317	-.0811	-.5295
Consequence	.0071	-.7694	-.0159	-.7693
Hypothesis	.3261	-.7018	.3050	-.7112
Sensitivity	.5961	-.1123	.5925	-.1301
Title	.6373	.1007	.6400	.0816
Sentence	.5626	-.1361	.5583	-.1529
Sensitivity	.3868	-.1312	.3827	-.1427
Sentence	.6758	.0995	.6785	.0792
Graph Conclusion	.7352	-.0437	.7335	-.0657
Consequence	.6920	.2921	.7004	.2713

Note.--N=82.

time. This presents no problem to the initial interpretation. All other alternative loadings were less than .30. Both quartimax and varimax solutions were essentially the same. The proportion of variance accounted for by the two-factor solution was .43.

Test B2.--Table 7 summarizes both the quartimax and varimax factor analytic rotations for Test B2 (N=87, one subject was not included since he completed less than half of the ten items). Again, a two-factor solution best describes the data, since in all subsequent solutions several inconsistencies were to be found, all of which did not lend themselves to psychological interpretation. On the basis

of the two-factor solution, there was one general factor which contained all of the items with the exception of two sentence items, which comprised the alternative factor. Only one inconsistency was found, item 3, a hypothesis, had a loading of .32 on the sentence factor. Both quartimax and varimax solutions were essentially the same. The proportion of variance accounted for by this two-factor solution was .40.

TABLE 7

Quartimax and Varimax
Rotational Analysis for Test B2

Problem Types	Rotations			
	Quartimax Loadings		Varimax Loadings	
	Factor 1	Factor 2	Factor 1	Factor 2
Sensitivity	.5143	-.4619	.6108	-.3238
Sensitivity	.6509	-.0624	.6467	.0969
Hypothesis	.5020	.3191	.4099	.4310
Sensitivity	.3016	.2180	.2399	.2845
Graph Conclusion	.5900	.1850	.5277	.3222
Sentence	.1996	.8526	-.0125	.8755
Title	.5423	.0689	.5096	.1980
Sentence	.1425	.5137	.0141	.5330
Graph Conclusion	.6780	-.2459	.7173	-.0746
Consequence	.5866	.0709	.5520	.2107

Note.--N=87.

Tests A1, B1, A2, and B2.--Since tests A1 and B1 were parallel forms, as were tests A2 and B2, it was hypothesized that the parallel forms would each emerge as separate factors. This prediction was confirmed as is illustrated in Table 8,

(N=71). Only a minor inconsistency was found, wherein a loading of .32 was reported for Test B2 on the factor related to creative performance as measured by tests A1 and B1. The proportion of variance accounted for by this solution was .75. Both quartimax and varimax solutions were essentially the same.

On the basis of these results it is obvious that parallel tests essentially measure the same dimension of creative ability, and nonparallel tests, constructed of different item-types, measure different dimensions of creative ability. One should, thus, be cautious in making generalizations about creativity and creativity tests.

TABLE 8

Quartimax and Varimax
Rotational Analysis for Tests A1, A2, B1 and B2

Tests	Rotations			
	Quartimax Loadings		Varimax Loadings	
	Factor 1	Factor 2	Factor 1	Factor 2
A1	.8666	-.1333	.8655	-.1406
B1	.8116	-.2466	.8095	-.2535
A2	.1169	-.8918	.1094	-.8927
B2	.3228	-.7854	.3161	-.7881

Note.--N=71.

Group Analyses

In general, for each test, sex was not found to be a factor influencing performance, as determined by t tests. As

a result, males and females were pooled on subsequent analyses.

Test A1.--A simple one-way analysis of variance was performed using the total scores of subjects in their respective groups as a function of their performance on Test A1. These results are summarized in Table 9. An F=4.82 was found to be significant at the .001 significance level.

TABLE 9

Analysis of Variance: Test

Source	SS	df	MS	F
Between groups	2749.56	6	458.26	4.82 [a]
Within groups	7320.68	77	95.07	
Totals	10070.24	83		

[a] $p<.001$

A further analysis which employed t tests, again using total scores, is represented in Table 10. Here it was found that each group that trained with consequence problems performed significantly better than the control group. The assisted-group brainstorming condition resulted in the best improvement ($t=4.53$, $p<.001$), followed by the assisted-individual brainstorming condition ($t=2.59$, $p<.01$), followed by the unassisted-group brainstorming condition ($t=2.16$, $p<.025$).

When these experimental groups were compared, subjects in the assisted-group brainstorming condition performed significantly better than subjects in the unassisted-group brainstorming condition ($t=2.59$, $p<.01$), but no other significant differences were found.

TABLE 10

Performance of the Experimental and Control Groups on Test A1

Groups	N	$\bar{X}$	s^2	t
Assisted-Individual Brainstorming with Consequences	12	96.08	201.54	2.59[b]
Unassisted-Group Brainstorming with Consequences	12	90.92	61.00	2.16[a]
Assisted-Group Brainstorming with Consequences	12	99.41	67.54	4.53[d]
Assisted-Individual Brainstorming with Sentences	12	88.00	73.27	
Unassisted-Group Brainstorming with Sentences	12	84.66	23.33	
Assisted-Group Brainstorming with Sentences	12	96.25	116.26	3.17[c]
Control	12	83.50	80.47	

[a] $p < .025$
[b] $p < .01$
[c] $p < .005$
[d] $p < .001$

When training with sentence problems, only subjects in the assisted-group brainstorming condition performed significantly better than the control group ($t=3.17$, $p < .005$). With

sentence practice, neither knowledge of the criteria nor brainstorming alone was sufficient to stimulate creativity, however, the combination of the two was.

When these experimental groups were compared, subjects in the assisted-group brainstorming condition performed significantly better than both subjects in the assisted-individual brainstorming condition ($t=2.08$, $p < .025$), and subjects in the unassisted-group brainstorming ($t=3.4$, $p < .001$). No other significant differences were found.

On the basis of these results, it would appear that the success of each of the brainstorming techniques was determined, in part, by the particular problem utilized in the training procedure. Evidence has been found for the assertion that the assisted-group brainstorming condition will facilitate performance to a greater extent than either assisted-individual brainstorming activity or unassisted-group brainstorming activity. But this assertion must be qualified as this result was a function of the particular training problem used. No evidence has been found for the assertion that unassisted-group brainstorming activity will enhance performance significantly better than assisted-individual brainstorming activity.

A more detailed analysis of the above data was undertaken by dividing each subject's total score into total score for sentence problems and total score for consequence problems; t tests were employed, the results of which are summarized in Table 11.

TABLE 11

Performance of the Experimental and Control Groups on either Sentence or Consequence Problems for Test A1

Groups	N	$\bar{X}$	s^2	t
Assisted-Individual Brainstorming with Consequences				
Consequences	12	48.75	49.84	2.736[b]
Sentences	12	47.33	89.15	1.736[a]
Unassisted-Group Brainstorming with Consequences				
Consequences	12	46.50	38.09	2.033[a]
Sentences	12	44.42	26.48	
Assisted-Group Brainstorming with Consequences				
Consequences	12	52.88	33.90	4.45[d]
Sentences	12	47.33	11.52	3.10[c]
Assisted-Individual Brainstorming with Sentences				
Consequences	12	40.58	42.45	
Sentences	12	47.42	29.72	2.569[b]
Unassisted-Group Brainstorming with Sentences				
Consequences	12	42.33	69.87	
Sentences	12	42.33	18.61	
Assisted-Group Brainstorming with Sentences				
Consequences	12	44.42	56.27	
Sentences	12	51.83	22.88	4.99[d]
Control				
Consequences	12	41.50	34.27	
Sentences	12	42.00	23.82	

[a] $p < .05$
[b] $p < .01$
[c] $p < .005$
[d] $p < .001$

On the basis of these results, it is evident that all subjects who practiced with consequence problems, regard-

less of method, did significantly better on consequence problems than did the control group. The best performance was produced under the assisted-group brainstorming condition ($t=4.45$, $p < .001$), followed by the assisted-individual brainstorming condition ($t=2.736$, $p < .01$), followed by the unassisted-group brainstorming condition ($t=2.033$, $p < .05$). Among these experimental groups, the only significant difference found was between the assisted-group brainstorming condition and the unassisted-group brainstorming condition ($t=2.61$, $p < .01$).

But not only did these groups do significantly better on consequence problems. Subjects in the assisted-individual brainstorming condition and the assisted-group brainstorming condition also did significantly better on sentence problems ($t=1.736$, $p < .05$: $t=3.1$, $p < .005$), suggesting a certain amount of transfer occurred to a different problem, which measures a different dimension of creativity. When these experimental groups were contrasted, only a single significant difference was found, wherein subjects in the assisted-group brainstorming condition performed better than the subjects in the unassisted-group brainstorming condition. ($t=2.605$, $p < .01$) on the consequence problems.

In no case did transfer occur to consequence problems under conditions characterized by sentence training. However, in two cases, performance was significantly better on sentence problems. The assisted-individual brainstorming condition ($t=2.569$, $p < .01$), and the assisted-group brainstorming condition

(t=4.99, p <.001) were significantly better than the control group. Unassisted-group brainstorming in groups with sentence problems resulted in no gains on either problem type.

When these experimental groups were contrasted on sentence scores, subjects in the assisted-group brainstorming condition performed signficantly better than either subjects in the assisted-individual condition (t=2.11, p < .025), or the unassisted-group brainstorming condition (t=5.11, p < .001). Contrary to expectation, however, subjects in the assisted-individual brainstorming condition were found to perform significantly better than subjects in the unassisted-group brainstorming condition (t=2.54, p < .01). No other significant differences were found among these groups.

It appears then, that for groups which practiced with sentence problems, knowledge of the criteria for a good solution was all that was necessary for enhanced performance on that type of problem. When coupled with brainstorming in a group situation performance was facilitated to an even greater extent, but unassisted-group brainstorming, by itself, was found insufficient to stimulate later performance. Furthermore, it was only because of the greatly facilitated performance on sentence problems that the assisted-group brainstorming did significantly better than the control group, as measured by total scores.

For groups that practiced with consequence problems, brainstorming, criteria, or both significantly facilitated performance on consequence problems, and in order for transfer to occur, exposure to the criteria was all that was found necessary. Brainstorming, by itself, was found to be insufficient to enhance performance on a new kind of problem.

Refer to Table 12 for a summary of the above results. On the basis of the results reported from Test A1, it would appear that assisted-individual brainstorming activity or unassisted-group brainstorming activity, as training techniques, may either facilitate performance on subsequent similar problems, or they may have no effect, depending upon the nature of the problem. Assisted-group brainstorming was found to have a positive effect regardless of the training problem.

With respect to solution quality, group brainstorming was not found to have an inhibitory effect as some investigators have suggested. When group brainstorming was conducted after exposure to both criteria and examples of good solutions, performance was greatly enhanced.

With respect to the experimental conditions, some evidence was found supporting the hypothesis that subjects in the assisted-group brainstorming condition would perform significantly better than either subjects in the unassisted-group brainstorming condition or subjects in the assisted-individual brainstorming condition. But again,

this assertion must be qualified as an important consideration was the particular problem used during the training. For sentence problems the hypothesis was supported, however, for consequence problems only partial support was found.

We are left uncertain at this point with respect to the hypothesis that unassisted-group brainstorming activity will result in significantly better performance than the assisted-individual brainstorming activity. When contrasting total scores, no significant differences were found between the two conditions, regardless of problem type. However, when the total scores were separated into consequence and sentence problem scores, a significant difference was obtained for sentence problems (though not for consequence problems), but opposite in direction to that which was predicted: Assisted-individual brainstorming performance was found to be significantly better on Test A1 than unassisted-group brainstorming performance.

The conditions necessary for transfer were found to vary as a function of both the nature of the training and the nature of the problem. For transfer to sentence problems, the essential condition apparently, was exposure to the criteria and examples of good solutions to consequence problems, inherent in both the assisted-individual brainstorming condition and the assisted-group brainstorming condition. Unassisted-group brainstorming with consequence

problems was found to be insufficient as a training technique to enhance performance on sentence problems.

And to reiterate, in no case did transfer occur to consequence problems under conditions characterized by sentence problem training.

A possible explanation for the unidirectionality of the transfer reported is suggested by the generalizability of the training techniques. It is entirely possible that the sentence problem training was simply too 'problem specific,' whereas the consequence problem training was not. Adapting from suggestions to be remote, far reaching, unusual, etc., could have been more readily generalized to sentence problems than suggestions of how to combine words in novel ways, etc., could have been generalized to consequence problems.

Let us consider the above interpretation in terms of the demand characteristics operating in each of the groups. The set of demand characteristics operating in each of the brainstorming groups was the suggestion for quantity, deferred judgment, free-wheeling, and interaction. But perhaps the key demand characteristic of the assisted groups training with consequence problems was the suggestion to penetrate from an original idea along many logical associations to a diverse consequence. Whereas, the key demand feature of the assisted-groups training with sentence problems was the suggestion to consider the greatest range of possible

TABLE 12

Differences Between Groups
on Tests A1, B1, A2, and B2

Training Problems	Tests	Training Techniques: Assisted-Group Brainstorming Compared to UGB	Assisted-Group Brainstorming Compared to AIB	Unassisted-Group Brainstorming Compared to AIB	Control Group (AB) Compared to UGB	Control Group (AB) Compared to AGB	Control Group (AB) Compared to AIB	Control Group (AB) Compared to CG(B)	Control Group (B) Compared to UGB	Control Group (B) Compared to AGB	Control Group (B) Compared to AIB
Consequence Problems	A1	S= + C+	S= = C=	S= = C=	S= - C-	S- - C-	S- - C-				
	A2	=	=	=	=	=	=				
	B1	S= = C=	S= = C=	S= = C=	S= = C=	S= = C=	S= = C=		S- - C-	S- - C-	S- - C-
	B2	=	=	=	=	=	-	S= + C+	=	=	-
Sentence Problems	A1	S+ + C=	S+ + C=	S- = C=	S= = C=	S- - C=	S- = C=				
	A2	+	+	+	-	-	=				
	B1	S+ + C=	S+ = C=	S= - C-	S= = C=	S- - C=	S= = C=		S= = C=	S- - C-	S- - C-
	B2	+	+	+	=	-	=		=	-	=

Note.-- (+) indicates a positive significant difference was found; (=) indicates no significant difference was found; (-) indicates a negative significant difference was found.

solutions, and then to associate the most novel ramifications within the range.

It would appear, then, that the ability to penetrate to a solution is more successfully transferred to solving sentence problems than the ability to consider a wide range of solutions is successfully transferred to solving consequence problems.

On the basis of this interpretation, caution should be exerted in the future when constructing 'training guides.' One should not make them so general as to be vague, nor so specific as to limit their generalizability.

Optimally, instructions should provide a subject with sufficient direction or information about the nature of a desired solution but the instructions should not be so narrow as to inhibit the imagination. Simply instructing the subjects to be creative was found insufficient to obtain the experimenter's desired behavior. And yet, the more specific a set of instructions, the less generalizable the resulting behavior was found to be.

An alternative approach might be to construct several 'specific' training guides and to allow the subject to form a more intuitive model of creativity in general. That is, by extracting the salient aspects from each of the specific guides, the subject would be allowed to blend his innate capacities with the strategies that would best maximize his creative potential. More research is indicated.

From this analysis, it would seem that before a productive thinking task can be successfully terminated, two conditions are necessary. First, the final solution must be cognitively consistent with basic, inherent premises held by the subject. The greater the cognitive complexity, the more the solution must be analyzed to ascertain its validity. Furthermore, the greater the cognitive complexity, the fewer the solutions that will be accepted as creative, since there are many concepts to compare resulting in a greater chance of rejection.

Second, the solution must be novel to the person. Again, the greater the cognitive complexity, the fewer the solutions that will be accepted as creative, since exposure to many diverse points of view allows less of a chance for forming novel ramification not previously considered. What is considered to be creative by one person at one level of development may not be considered to be creative by another person at another level of development.

Test A1, Subject's Ratings.--Of further importance were the subject's own ratings of their performance. Each subject's total score was determined as a function of his own rating and was analyzed through t tests. Table 13 summarizes the basic findings obtained.

Only two significant differences were found. Subjects in the assisted-group brainstorming condition rated

TABLE 13

Subject's Ratings of Own Solutions on Test A1

Groups	N	X	s^2	t
Assisted-Individual Brainstorming with Consequences	11	41.68	57.56	
Unassisted-Group Brainstorming with Consequences	12	42.50	42.64	
Assisted-Group Brainstorming with Consequences	12	42.67	43.52	1.74[a]
Assisted-Individual Brainstorming with Sentences	11	35.18	76.56	
Unassisted-Group Brainstorming with Sentences	11	41.73	32.62	
Assisted-Group Brainstorming with Sentences	12	46.83	47.61	2.82[b]
Control	12	36.16	124.33	

Note.--Some subjects, although present, did not rate their own solutions.

[a] $p < .05$
[b] $p < .005$

themselves significantly higher than control subjects rating themselves, regardless of the type of problem used in training ($t=1.74$, $p < .05$ for consequence problems; $t=2.82$, $p < .005$ for sentence problems). Subjects who simply brainstormed in groups, or subjects who brainstormed individually with assistance, did not rate themselves higher than control subjects who rated themselves, inspite of the evidence as reported above where performance as rated by experienced judges was found to be significantly better in some instances.

When the experimental groups were compared, it was found that subjects in the assisted-group brainstorming condition that trained with sentence problems rated themselves significantly higher than did either subjects in the assisted-individual brainstorming condition ($t=3.56$, $p < .001$), or the unassisted-group brainstorming condition ($t=2.08$, $p < .025$). No other significant differences were found. Refer to Table 15 for a summary of the above results.

Of considerable interest, then, was the finding that while some of the various methods employed to facilitate creativity were successful, they were not successful in promoting within each subject the ability to accurately judge or assess his own solutions. In each condition, the mean rating of the solutions as judged by the subjects was less than that reported by the experienced judges. Apparently, the ability to produce a solution is distinct from the ability to judge a solution, and training in one does not necessitate enhanced performance in the other.

The experimenter feels, however, that this interpretation should be considered cautiously. It must be emphasized that the subjects were rating their own solutions, and not solutions in general. An alternative explanation is that the subjects' scores simply reflect the confidence they had in their own performances. At any rate, more research is indicated to analyze each subject's concept of his performance as a function of training.

When the subject's total scores were reduced to total sentence and total consequence scores, it was found that subjects in the unassisted-group brainstorming condition with consequence problems for training rated themselves significantly higher than did the control group on sentence problems ($t=1.976$, $p < .05$), and the assisted-group brainstorming that trained with sentence problems rated themselves significantly higher than did the control subjects on both sentence and consequence problems ($t=1.92$, $p < .05$; $t=3.184$, $p < .005$, respectively). Refer to Table 14 and Table 15 for summaries.

On the basis of these results, it would appear that subjects who train for only 45 minutes, while performing significantly better than control subjects, were unaware of their enhanced performance as was measured by their own ratings. Of course, this conclusion is speculative, as no pretest was administered to determine what the subjects considered their base level to be.

Test A2.--A simple one-way analysis of variance was performed using total scores as a function of each subject's performance on Test A2. The results are summarized in Table 16. A significant finding reflected in an $F=2.61$ indicated that the null hypothesis may be rejected at the .05 significance level.

A subsequent analysis utilizing t tests revealed only two significant findings. Refer to table 17. Unassisted-group brainstorming and assisted-group brainstorming with

TABLE 14

Subject's Ratings of Own Solutions on either Sentence or Consequence Problems for Test A1

Groups	N	$\bar{X}$	s^2	t
Assisted-Individual Brainstorming storming with Consequences				
Consequences	11	20.50	18.65	
Sentences	11	21.18	28.15	
Unassisted-Group Brainstorming with Consequences				
Consequences	12	20.50	7.25	
Sentences	12	22.00	19.64	1.976[a]
Assisted-Group Brainstorming with Consequences				
Consequences	12	21.50	12.42	
Sentences	12	21.17	14.15	
Assisted-Individual Brainstorming with Sentences				
Consequences	11	16.82	23.24	
Sentences	11	18.36	18.41	
Unassisted-Group Brainstorming with Sentences				
Consequences	11	21.09	8.81	
Sentences	11	20.64	17.50	
Assisted-Group Brainstorming with Sentences				
Consequences	12	22.25	11.84	1.922[a]
Sentences	12	24.58	20.08	3.184[b]
Control				
Consequences	12	18.33	38.25	
Sentences	12	17.83	33.98	

Note.--Some subjects, although present, did not rate their own solutions.

[a] $p < .05$
[b] $p < .005$

sentence problems were both found to enhance performance when compared to the control condition ($t=2.305$, $p < .025$;

TABLE 15

Differences Between Groups
as Determined by Subject's Ratings of
Own Solutions on Tests A1, B1, A2, and B2

Training Problems	Tests	Training Techniques									
		Assisted-Group Brainstorming Compared to		Unassisted-Group Brainstorming Compared to	Control Group (AB) Compared to				Control Group (B) Compared to		
		UGB	AIB	AIB	UGB	AGB	AIB	CG(B)	UGB	AGB	AIB
Consequence Problems	A1	S= = C=	S= = C=	S= = C=	S- = C=	S= - C=	S= =C=				
	A2	=	=	=	=	=	=				
	B1	S= = C=	S= = C=	S= = C=	S= = C=	S= = C=	S= =C=		S= = C=	S= = C=	S= = C=
	B2	=	-	=	=	=	=	= S= C=	=	=	=
Sentence Problems	A1	S+ + C=	S+ + C=	S= + C+	S= = C=	S- - C-	S= =C=				
	A2	=	=	=	-	-	=				
	B1	S= = C=	S+ + C=	S+ + C=	S= - C-	S= = C=	S= =C=		S= =C=	S= = C=	S= = C=
	B2	=	=	=	=	=	=		=	=	=

Note.-- (+) indicates a positive significant difference was found; (=) indicates no significant difference was found; (-) indicates a negative significant difference was found.

t=3.697, p .001). Of some importance is the fact that the former group did not perform significantly better on Test A1, nor in this group was there any indication on Test A1 of transfer to other problems.

TABLE 16

Analysis of Variance: Test A2

Source	SS	df	MS	F
Between groups	2446	6	407.7	2.61[a]
Within groups	11873	76	156.2	
Totals	14319	82		

[a] $p < .05$

A possible explanation for the superior performance by most of the subjects in the experimental groups on Test A1, but not on Test A2, was that more time was spent on the initial test by these subjects. This seemed to be indicated by the factor analytic results where the two emergent clusters consisted of the first three problems followed by the remaining seven problems.

It is not really understood why only the two groups that had successfully trained with sentence problems were able to spend some time on the second test, than those groups that had successfully trained with consequence problems. But perhaps the sentence problems required more time to successfully complete in the absence of training than did the consequence problems.

TABLE 17

Performance of the Experimental and Control Groups on Test A2

Groups	N	$\bar{X}$	s^2	t
Assisted-Individual Brainstorming with Consequences	12	83.08	100.81	
Unassisted-Group Brainstorming with Consequences	12	82.25	117.52	
Assisted-Group Brainstorming with Consequences	12	79.00	223.82	
Assisted-Individual Brainstorming with Sentences	12	71.75	392.52	
Unassisted-Group Brainstorming with Sentences	12	83.67	46.06	2.305[a]
Assisted-Group Brainstorming with Sentences	12	90.33	102.79	3.697[b]
Control	12	77.17	49.24	

[a] $p < .025$
[b] $p < .001$

Two interpretations, other than total time spent per test, are plausible explanations for finding few significant differences on Test A2. Possibly transfer of training from one type of problem to several types was not great enough to be exposed by this short ten-item test. Or perhaps transfer simply did not occur. It seems highly likely that transfer of training may vary from problem type to problem type, that finding transfer from consequence problems to sentence problems does not imply transfer from sentence problems, or for that matter, to any other kind of problem.

Another interpretation is that the test (A2) did not allow for widely variable responding. Both of these interpretations are plausible, and when taken together, suggest why the trend of the means was in some instances toward significance.

Further analysis of the experimental groups revealed that the three groups that trained with sentence problems differed significantly from one another in the predicted direction. The assisted-group brainstorming condition performed significantly better than either the assisted-individual brainstorming condition ($t=2.89$, $p < .005$), or the unassisted-group brainstorming condition ($t=1.89$, $p < .05$). Furthermore, the unassisted-group brainstorming condition performed significantly better than the assisted-individual brainstorming condition ($t=1.97$, $p < .05$). No other significant differences were found between the experimental groups that practiced with consequence problems. The trend, thus far, seems to indicate that the predicted differences in the experimental groups are plausible only for those groups that trained with sentence problems. Refer to Table **12** for a summary of the above results.

Test A2, Subject's Ratings.--Subjects also rated their own solutions on Test A2, and when compared with the control subject's ratings of their solutions, two experimental groups were found to significantly differ. Refer to Table 18.

TABLE 18

Subject's Ratings of Own Solutions on Test A2

Groups	N	$\bar{X}$	s^2	t
Assisted-Individual Brainstorming with Consequences	10	33.00	55.00	
Unassisted-Group Brainstorming with Consequences	12	37.58	76.63	1.795[a]
Assisted-Group Brainstorming with Consequences	9	36.22	67.51	
Assisted-Individual Brainstorming with Sentences	8	32.38	150.36	
Unassisted-Group Brainstorming with Sentences	10	37.10	60.10	
Assisted-Group Brainstorming with Sentences	10	39.70	54.01	2.399[b]
Control	12	30.92	88.63	

Note.--Some subjects, although present, did not rate their own solutions.

[a] $p < .05$
[b] $p < .025$

Subjects in the assisted-group brainstorming condition with sentence training were found to rate themselves significantly higher than the control subjects rated themselves ($t=2.399$, $p < .025$), and subjects in the unassisted-group brainstorming condition were found to rate themselves significantly higher than did the control subjects ($t=1.795$, $p < .05$).

No significant differences between the experimental groups were found. Again, the distinction between the ability to produce superior solutions and the ability to judge

solutions is evident. Refer to Table 15 for a summary of the above results.

Test B1.--Scores on tests B1 and B2 reflect performance after a three-week interval from the time of training to the administration of these tests.

A simple one-way analysis of variance was performed using total scores of subjects as a function of their performance on Test B1. These results are summarized in Table 19 where it is evident that the null hypothesis may be rejected since an obtained F=3.67 was found to be significant at the .001 level of significance.

TABLE 19

Analysis of Variance: Test B1

Source	SS	df	MS	F
Between groups	2388.86	7	341.266	3.67[a]
Within groups	7438.23	50	92.978	
Totals	9827.09	87		

[a] $p < .001$

When the experimental groups were compared with the experienced control group (AB), only one significant difference was found to persist. The assisted-group brainstorming condition with sentence training performed significantly better than did the control condition ($t=2.56$, $p < .01$).

Of considerable consequence was the finding that the experienced control group (AB) performed significantly better than the new, inexperienced control group (B) ($t=3.00$, $p < .005$). Apparently, exposure to the testing situation and some funtion of time combined somehow to facilitate performance.

When comparing the experimental groups to the inexperienced control condition (B), all of the groups were found to perform significantly better, with the exception of the unassisted-group brainstorming with sentence training condition.

These results indicate that each of the approaches initially found to promote productivity was also found to promote productivity as measured three weeks later. And of considerable interest was the finding that except for a single instance, the experience of the control group was as effective a procedure for enhancing productivity as each of the experimental approaches when measured three weeks later. Considering the brevity of the training sessions, this is somewhat understandable.

When the experimental groups were compared, only two significant differences were found. For the groups that trained with sentence problems, subjects in the assisted-group brainstorming condition did significantly better than subjects in the unassisted-group brainstorming condition ($t=2.82$, $p < .005$), and subjects in the assisted-individual brainstorming condition also did significantly better than

subjects in the unassisted-group brainstorming condition ($t=2.02$, $p < .05$). The first result was in the predicted direction, whereas the second finding was not. No other significant differences were found.

TABLE 20

Performance of the Experimental and Control Groups on Test B1

Groups	N	$\bar{X}$	s^2	t_{AB}	t_B
Assisted-Individual Brainstorming with Consequences	9	91.56	37.28		3.82 [c]
Unassisted-Group Brainstorming with Consequences	11	92.18	139.36		2.84 [b]
Assisted-Group Brainstorming with Consequences	9	93.99	119.15		3.41 [b]
Assisted-Individual Brainstorming with Sentences	11	95.28	128.82		3.74 [c]
Unassisted-Group Brainstorming with Sentences	12	86.58	86.08		
Assisted-Group Brainstorming with Sentences	12	98.34	122.24	2.56[a]	4.69 [c]
Control (AB)	12	88.91	41.36		3.00 [b]
Control (B)	12	81.17	38.70		

Note.--Not all subjects returned for the second administration of the tests.

[a] $p < .01$
[b] $p < .005$
[c] $p < .001$

A more detailed analysis of the above data was undertaken by dividing each subject's total score into total score

for sentence problems and total score for consequence problems. Refer to Table 21.

TABLE 21

Performance of the Experimental and Control Groups on either Sentence or Consequence Problems Problems for Tests B1

Groups	N	$\bar{X}$	s^2	t_{AB}	t_B
Assisted-Individual Brainstorming with Consequences					
Consequences	9	47.11	21.86		2.913[d]
Sentences	9	44.44	17.28		2.561[c]
Unassisted-Group Brainstorming with Consequences					
Consequence	11	47.91	39.09		2.864[d]
Sentences	11	44.27	43.82		2.046[a]
Assisted-Group Brainstorming with Consequences					
Consequences	9	47.88	63.11		2.423[b]
Sentences	9	46.11	18.36	1.812[a]	3.378[d]
Assisted-Individual Brainstorming with Sentences					
Consequences	11	49.73	25.01		4.337[e]
Sentences	11	45.55	65.07		2.262[b]
Unassisted-Group Brainstorming with Sentences					
Consequences	12	44.16	37.97		
Sentences	12	42.42	33.90		
Assisted-Group Brainstorming with Sentence					
Consequences	12	47.42	47.72		2.498[b]
Sentences	12	50.92	33.99	3.853[*]	5.349[e]
Control (AB)					
Consequences	12	46.66	17.15		3.031[d]
Sentences	12	42.25	26.75	2.309[b]	
Control (B)					
Consequences	12	41.75	14.20		
Sentences	12	39.42	21.54		

Note.--Not all subjects returned for the second administration of the tests.

[a] $p < .05$ [d] $p < .005$
[b] $p < .025$ [e] $p < .001$
[c] $p < .01$ [*] $p < .001$

Several things become apparent. Whereas on Test A1 there were only two instances of transfer (assisted-individual brainstorming and assisted-group brainstorming, both with consequence problems), on Test B1 transfer seemingly occurred in each of the experimental groups that had previously performed successfully on the problem for which they were trained.

It is also evident that the experienced control condition (AB) improved not on both problem types, but only on consequence problems ($t=3.03$, $p < .005$).

The fact that the experienced control group improved only on the consequence problems casts some doubt as to whether transfer occurred among the two experimental groups that had trained with sentence problems, the assisted-individual brainstorming condition and the assisted-group brainstorming condition. It may well be that the enhanced performance by these groups on the consequence problems was not so much a function of their previous training, but rather a function of the time imposed between the first and second testing situations.

When comparing the experimental groups, again, the only significant differences occurred among those groups which had practiced with sentence problems. Subjects in the assisted-group brainstorming condition performed significantly better than subjects in the unassisted-group brainstorming condition ($t=2.82$, $p < .01$), and this difference was attributable primarily to enhanced performance on sentence problems ($t=3.574$, $p < .005$). Subjects in the assisted-individual

brainstorming condition also performed significantly better than subjects in the unassisted-group brainstorming condition ($t=2.02$, $p < .05$), but this result was attributable primarily to enhanced performance on consequence problems ($t=2.367$, $p < .025$). No other significant differences were found. Refer to Table 12 for a summary of the above results.

Test B1, Subject's Ratings.--When comparing how the subjects in each condition rated their own solutions, only one significant difference was found. Ironically, subjects who simply brainstormed in groups with sentence problem training rated themselves significantly higher than the experienced control subjects rated themselves ($t=2.04$, $p < .025$). And when the items were broken down into consequence and sentence items, it was found that they rated themselves better only on the consequence problems ($t=2.132$, $p < .025$).

Again more research is indicated to assess each subject's concept of his performance as a function of training, and again, it is quite evident that the ability to produce solutions is distinct from the ability to rate solutions (at least with respect to one's own solutions).

When comparing the experimental groups, significant differences were found only among those groups that had trained with sentence problems. Subjects in the assisted-group brainstorming condition were found to perform significantly better than subjects in the assisted-individual

TABLE 22

Subject's Ratings of Own Solutions on Test B1

Groups	N	$\bar{X}$	s^2	t_{AB}
Assisted-Individual Brainstorming with Consequences	8	36.13	48.90	
Unassisted-Group Brainstorming with Consequences	12	36.67	47.52	
Assisted-Group Brainstorming with Consequences	9	38.67	38.25	
Assisted-Individual Brainstorming with Sentences	10	35.50	98.04	
Unassisted-Group Brainstorming with Sentences	12	43.67	44.61	2.04*
Assisted-Group Brainstorming with Sentences	9	43.22	57.69	
Control (AB)	12	36.75	94.35	
Control (B)	12	42.66	60.97	

Note.--Not all subjects returned for the second administration of the tests, and of those subjects who did return, not all rated their own solutions.

*$p < .05$

brainstorming condition ($t=1.865$, $p < .05$) and this was attributable primarily to enhanced performance on sentence problems ($t=2.726$, $p< .01$). Subjects in the unassisted-group brainstorming condition were also found to perform significantly better than subjects in the assisted-individual brainstorming condition ($t=2.275$, $p < .025$), and this result was attributable primarily to enhanced performance on sentence problems ($t=2.31$, $p < .025$). No other significant

TABLE 23

Subject's Ratings of Own Solutions on either Sentence or Consequence Problems for Test B1

Groups	N	$\bar{X}$	s^2	t_{AB}
Assisted-Individual Brainstorming with Consequences				
Consequences	8	19.38	20.48	
Sentences	8	16.75	12.69	
Unassisted-Group Brainstorming with Consequences				
Consequences	12	17.25	10.02	
Sentences	12	19.42	18.08	
Assisted-Group Brainstorming with Consequences				
Consequences	9	19.56	11.80	
Sentences	9	19.11	9.71	
Assisted-Individual Brainstorming with Sentences				
Consequences	10	18.70	27.61	
Sentences	10	16.90	26.89	
Unassisted-Group Brainstorming with Sentences				
Consequences	12	21.92	26.44	2.132*
Sentences	12	21.75	17.35	
Assisted-Group Brainstorming with Sentences				
Consequences	9	20.44	12.02	
Sentences	9	22.78	16.62	
Control (AB)				
Consequences	12	17.25	31.30	
Sentences	12	19.50	28.45	
Control (B)				
Consequences	12	21.33	20.25	
Sentences	12	21.33	23.33	

Note.--Not all subjects returned for the second administration of the tests, and of those subjects who did return, not all rated their own solutions.

*$p < .025$

differences were found. Refer to Table 15 for a summary of the above results.

Test B2.--A simple one-way analysis of variance was performed using total scores on Test B2. The results are summarized in Table 24, where it is evident that the null hypothesis may be rejected since an obtained F=2.38 was found to be significant at the .05 significance level.

TABLE 24

Analysis of Variance: Test B2

Source	SS	df	MS	F
Between groups	2006	7	286.6	2.38*
Within groups	9598	80	120.0	
Totals	11604	87		

*$p < .05$

A subsequent analysis employing t tests is reported in Table 25. The two control groups were not found to differ significantly, and only two of the experimental groups were found to perform significantly better than either of the control groups. Subjects in the assisted-individual brainstorming condition with consequence training performed significantly better than the control subjects ($t=2.399, p < .025$ for the experienced control group; $t=2.439$, $p < .025$ for the inexperienced control group). And subjects in the assisted-group brainstorming condition with sentence training performed significantly better than the control subjects ($t=2.957$, $p < .005$ for the experienced control group; $t=3.021$, $p < .005$ for the inexperienced control group).

TABLE 25

Performance of the Experimental and Control Groups on Test B2

Groups	N	$\bar{X}$	s^2	t_{AB}	t_B
Assisted-Individual Brainstorming with Consequences	9	86.00	50.50	2.399[a]	2.439[a]
Unassisted-Group Brainstorming with Consequences	11	82.55	68.79		
Assisted-Group Brainstorming with Consequences	9	75.56	458.91		
Assisted-Individual Brainstorming with Sentences	11	74.27	86.02		
Unassisted-Group Brainstorming with Sentences	12	80.58	46.41		
Assisted-Group Brainstorming with Sentences	12	88.58	92.99	2.957[b]	3.021[b]
Control (AB)	12	77.58	72.99		
Control (B)	12	76.83	88.88		

Note.--Not all subjects returned for the second administration of the tests.

[a] $p < .025$
[b] $p < .005$

On the basis of the results from Test B2, it would appear that the persistence of the training effects had survived for only two of the experimental conditions, the assisted-individual brainstorming condition with consequence training, and as on Test B1, the assisted-group brainstorming condition with sentence training.

It is not really understood why the highly significant results between each of the seven treatment conditions did not

occur as was found on Test B1. Some hint is provided when comparing the respective means and variances of the two tests. In all of the experimental groups except one, the variances were much less for Test B2 than for Test B1. And for all of the groups, the respective means were much lower for Test B2 than for Test B1 suggesting that not only was the performance of a lesser quality on Test B2, but it was also of a lesser variety. This may have been the fault of the subjects, in that they spent less time on this test, or they simply were not able to think creativity on these problem types. Or it may have been a fault of the test, in that the items did not lend themselves readily to highly variable responding, hence, the reduction in the mean performance and variance.

When comparing the experimental groups, significant differences were again found only among those groups that had trained with the sentence problems. All differences were in the predicted direction with subjects in the assisted-individual brainstorming condition performing significantly better than either subjects in the unassisted-group brainstorming condition ($t=2.347$, $p < .025$), or subjects in the assisted-individual brainstorming condition ($t=3.621$, $p < .005$). Furthermore, subjects in the unassisted-group brainstorming condition performed significantly better than subjects in the assisted-individual brainstorming condition ($t=1.87$, $p < .05$). Refer to Table 12 for a summary of the above results.

While most of the general implications derived from the comparisons between the experimental groups and the control groups have been discussed, the general trends existing between the experimental groups have not yet been systematically explored. These trends are best summarized in Table 12.

Basically, the efficacy of each of the experimental conditions was found to vary as a function of the particular training problem. In general, for those groups that trained with consequence problems, no significant differences were found to exist. Assisted-group brainstorming, unassisted-group brainstorming, or assisted-individual brainstorming were of about equal value in promoting productive thinking.

However, for those groups that trained with sentence problems, a very different pattern of results was obtained. When tested on diverse item-types, group brainstorming either assisted or unassisted was clearly found to be superior to assisted-individual brainstorming. This conclusion should be considered with caution, however, since it rests on the results of Test A2 and Test B2, previously suspected of having too little time for successful completion. The results of Test A1 and Test B1 also reflected the superiority of group brainstorming, but only when combined with assistance. Inconclusive results were obtained when contrasting unassisted-group brainstorming with assisted-individual brainstorming using sentence problems for training.

Table 27 analyzes each of the brainstorming approaches employed in the present study as a function of the criteria. As is apparent, both assisted-individual and assisted-group brainstorming were found to be superior methods for enhancing productivity, but only when used in conjunction with consequence problems. It would thus appear that the advantage of a particular brainstorming technique is in part determined by the nature of the problem to be brainstormed upon. What is best for one set of problems may not be best for another set of problems.

Test B2, Subject's Ratings.--No significant differences were found when comparing the subject's ratings of their own solutions against either control condition, nor were there any differences among the experimental groups on the basis of the subject's own ratings. Refer to Table 26.

To summarize the results as determined by the subject's self-evaluation, for each test, the data seemed to indicate that successful performance in producing creative solutions was distinct from the ability to accurately judge solutions. Training in production did not necessitate enhanced performance in judgment. In general, the mean ratings of the solutions as judged by the subjects was less than that reported by the experienced judges.

An alternative conclusion was also entertained, however. It seems reasonable that since the subjects had only rated

their own solutions, their objectivity could very well have been impaired. And further, their ratings could very well have simply reflected their confidence.

TABLE 26

Subject's Ratings of Own Solutions on Test B2

Groups	N	$\bar{X}$	s^2
Assisted-Individual Brainstorming with Consequences	7	35.57	63.29
Unassisted-Group Brainstorming with Consequences	12	32.00	69.09
Assisted-Group Brainstorming with Consequences	8	28.00	72.57
Assisted-Individual Brainstorming with Sentences	9	31.78	218.19
Unassisted-Group Brainstorming with Sentences	11	41.27	113.29
Assisted-Group Brainstorming with Sentences	10	35.30	64.01
Control (AB)	11	35.09	62.26
Control (B)	11	36.91	83.17

Note.--Not all subjects returned for the second administration of the tests, and of those who did, not all rated their own solutions.

Table 15 summarizes the general trend of results as determined by the subject's self-evaluations. For those groups that trained with consequence problems, generally, no significant differences were found to exist. Similarly, when these groups were compared to the control groups, relatively few significant differences were found. For those groups that trained with sentence problems, consistent

significant differences in self-evaluation were obtained on Test A1 and Test B1, those tests administered first in each session. As was the case with the experienced judges' ratings, the group brainstorming conditions, either assisted or unassisted, resulted in significantly better self-evaluation than the assisted-individual brainstorming condition.

TABLE 27

Experimental Groups as a Function of Criteria

Groups	Criteria		
	Quality*	Transfer	Persistence*
Assisted-Group Brainstorming			
Sentence Problems	+	0	+
Consequence Problems	+	+	+
Unassisted-Group Brainstorming			
Sentence Problems	0	0	0
Consequence Problems	+	0	+
Assisted-Individual Brainstorming			
Sentence Problems	+	0	+
Consequence Problems	+	+	+

*Training Problem.

In general, it would appear that while the training procedures employed in the present study were in some instances successful in stimulating productive thinking, they were unsuccessful in providing each subject an awareness of their enhanced performance. More research is indicated to analyze a subject's concept of his performance as a function of training.

SUMMARY

Summary of Test Analyses

Two judges working independently can rate solutions on divergent thinking problems with a reasonably high degree of reliability. As a result, subjectivity of rating can be minimized helping to provide for high test reliability.

The reliabilities of the tests reported were considered to be less than would appear in a sample of untrained subjects due to the restriction of range imposed through the training procedures. However, the reliabilities of the tests reported were considered sufficient for research purposes with Coefficient Alpha for the four tests ranging from .66 to .80. Procedures for improving each of the four tests were suggested.

With respect to the factor analytic results, two of the hypotheses proposed were confirmed. It was found that sentence and consequence items each emerge as separate factors, indicating that each of these problem types contributes something unique with respect to measuring creative ability.

It was further found that parallel tests essentially measure the same dimension of creative ability, and nonparallel tests, constructed of different item types, measure different dimensions of creative ability. This latter finding suggests

that one should be cautious in making generalizations about creativity, and creativity tests.

A third hypothesis was not supported, as it was found that tests A2 and B2 did not reduce to six factors, one for each item type. Three reasons may be suggested for this finding. The first explanation is that there simply were not six factors. But assuming there were, an alternative interpretation is that time played a role. This is certainly indicated. The third explanation is that there were not a sufficient number of items representing each of the problem types for them to emerge as distinct factors. The latter two interpretations seem most plausible.

While the correlations found between the nonparallel tests in this study substantiate previous findings of relatively low correlations between tests of divergent thinking ability, the correlations between the parallel forms were found to be much higher. An order effect was suggested to account for the inflated correlations in one instance. However, in the other instance a different conclusion is merited. It would seem that even in the case of creativity, when two tests are essentially sampling from the same domain, as is the case with many aptitude tests, relatively high correlations may be found.

Summary of Group Analyses

According to the criteria imposed in the present study, both assisted-individual brainstorming and assisted-group

brainstorming were found to significantly facilitate productive thinking. The efficacy of each of these methods, however, was in part a function of the particular training problem used. When used in conjunction with consequence problems, both assisted-individual and assisted-group brainstorming were found to be superior methods for enhancing productivity.

The assisted-individual and assisted-group brainstorming methods were unsuccessful in conjunction with sentence training problems because of their failure to promote transfer. Examination of the transfer phenomenon revealed that the conditions necessary for transfer were a function of the training method and the nature of the training problem. Transfer of training from Problem A to Problem B did not necessitate transfer of training from Problem B to Problem A. Unassisted-group brainstorming was also found inadequate in satisfying the criteria imposed in the present study. An explanation in terms of the specificity of the demand characteristics operating was suggested to account for these findings.

It would thus appear that the brainstorming procedure which emphasizes quantity of solutions, free-wheeling, interaction of ideas, and deferred judgment can operate within a set of criteria-cued instructions and satisfy three pertinent criteria when the brainstorming strategy suggested is of a general nature. In general, the assisted-brainstorming

groups better satisifed the criteria than the unassisted or traditional brainstorming groups.

When the experimental groups were compared, basically it was found that for those groups that had trained with consequence problems, no significant differences existed. However, for those groups that trained with sentence problems, assisted-group brainstorming was found to be superior to either assisted-individual or unassisted-group brainstorming. Inconclusive results were obtained when contrasting unassisted-group brainstorming with assisted-individual brainstorming using sentence problems for training.

It would appear then, at least early in the brainstorming process, that on a task requiring the subject to consider a wide range of solutions (sentence problems), group brainstorming may be of greater value than individual brainstorming. Furthermore, at least early in the brainstorming process, on a task requiring solutions to be more penetrating and well thought-out (consequence problems), group brainstorming may be as equally beneficial as individual brainstorming.

And finally, as other investigators have shown, the ability to produce a creative solution was found to be distinct from the ability to accurately judge one's own solutions. In general, while the training procedures employed in the present study were in some instances successful in stimulating productive thinking, they were unsuccessful in

in providing each subject an awareness of their enhanced performance.

REFERENCES

REFERENCES

Cave, R.L. A combined factor analysis of creativity and intelligence. Multivariate Behavioral Research, 1970, 5, 177-191.

Christensen, P.R., Guilford, J.P., and Wilson, R.C. Relations of creative responses to working time and instructions. Journal of Experimental Psychology, 1957, 53, 82-88.

Collaros, P., and Anderson, L. Effects of perceived expertness upon creativity of members of brainstorming groups. Journal of Applied Psychology, 1969, 53, 159-163.

Compton, M.F. An attempt to foster creative thinking in teachers, Dissertation Abstracts, 1968, 29 (I-A), 164.

Dunnette, M.D., Campbell, J., and Jaastad, K. The effects of group participation on brainstorming effectiveness for two industrial samples. Journal of Applied Psychology, 1963, 47, 30-37.

Gerlach, V.S., Schutz, R. E., Baker, R.L., and Mazer, G.E. Effects of variations in test directions on originality test responses. Journal of Educational Psychology, 1964, 55, 79-83.

Guilford, J.P. The Nature of Human Intelligence. New York: McGraw Hill Inc., 1967.

Johnson, D.M. Improvement of Problem Solving Processes. Final Report. East Lansing; Michigan State University, 1968.

Johnson, D.M., and Kidder, R. C. Productive thinking in the college classroom. Educational Development Program, Michigan State University, 1972.

Johnson, D.M., Parrott, G.L., and Stratton, R.P. Production and judgment of solutions to five problems. Journal of Educational Psychology Monograph Supplement, 1968, 59 (6, pt. 2).

Johnson, D.M., and Zerbolio, D.J. Relations between production and judgment of plot-titles. American Journal of Psychology, 1964, 77, 99-105.

Meadow, A., and Parnes, S. Evaluation of training in creative problem solving. Journal of Applied Psychology, 1959, 43, 189-194.

Meadow, A. Parnes, S., and Reese, H. Influence of 'brainstorming' instructions and problem sequence on a creative problem solving test. Journal of Applied Psychology, 1959, 43, 413-416.

Osborn, A. F. Applied Imagination: Principles and Procedures of Creative Thinking. New York: Scribner's, 1957.

Parnes, S., and Meadow, A. Effects of 'brainstorming' on creative problem solving by trained and untrained subjects. Journal of Educational Psychology, 1959, 50, 171-176.

Parnes, S., and Meadow, A. Evaluation of persistence of effects produced by a creative problem solving course. Psychological Reports, 1960, 1, 357-361.

Stratton, R.P., Parrott, G.L., and Johnson, D.M. Transfer of Judgment training to production and judgment of solutions on a verbal problem. Journal of Educational Psychology, 1970, 61, 16-23.

Taylor, D.W., Berry, P.C. and Block, C.H. Does group participation using brainstorming facilitate or inhibit creative thinking? Technical Report No. 1, 1957, Yale University, Department of Psychology, Office of Naval Research.

Voytas, R.M. Some effects of various combinations of group and individual participation in creative participation, Dissertation Abstracts, 1968, 28 (9-B), 3911.

Weisskopf-Joelson, E., and Eliseo, T. An experimental study of the effectiveness of brainstorming. Journal of Applied Psychology, 1961, 45, 45-49.

APPENDICES

APPENDIX A

PRODUCTIVE THINKING TESTS

NAME_____

TEST A

1. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

   art semantic master

2. What are some of the consequences of racial prejudice in the U.S. today? (One or two sentences). Rate your answer from 1 to 7.

3. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

   culture thinking process

4. Of what consequence to human behavior would result through the elimination of psychological defense mechanism? (One or two sentences). Rate your answer from 1 to 7.

5. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

   perception variable color

6. Suggest some of the psychological consequences of getting old. Rate your answer (one or two sentences) from 1 to 7.

7. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

   interest principle net

8. Suppose statistical procedures have not yet been developed. Of what consequence would this have to psychology? Rate your answer (one or two sentences) from 1 to 7.

9. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

path analysis psycho

10. What are some of the social consequences of being shy and timid? Rate your answer (one or two sentences) from 1 to 7.

TEST A

Rating of answers for this test is on the basis of creative thinking. Rate each of your answers when completed with this test on a scale from 1 to 7. A rating of 7 indicates a very unusual, creative answer. A rating of 1 indicates a very poor answer. Answers should not exceed two sentences.

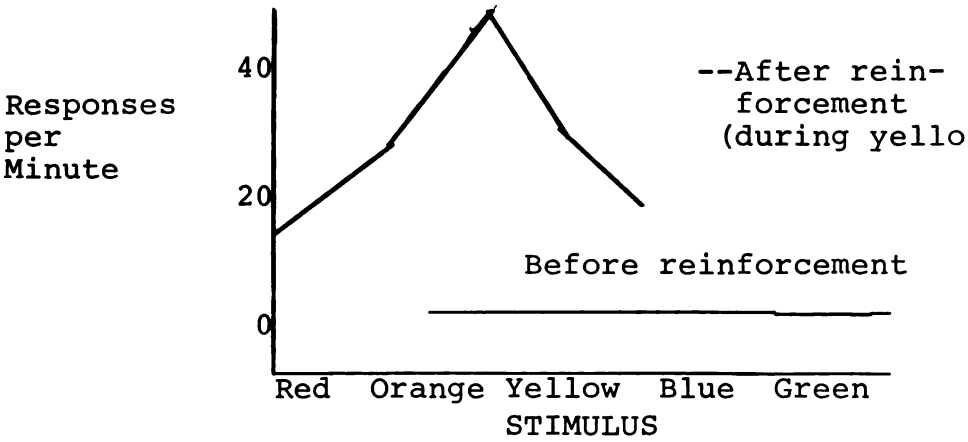
1. In what way is language and intelligence alike?

2. Suppose television had been invented and had become popular before books. What would be the consequences for the learning processes of today's children?

3. Write a hypothesis worth testing about the relation between psychologists, rats, and other living things.

4. What kind of reinforcement occurs in social situations that psychologists have not yet studied in the laboratory.

5. What would be a suitable title for the following graph?



6. Write an imaginative sentence that includes the following words:

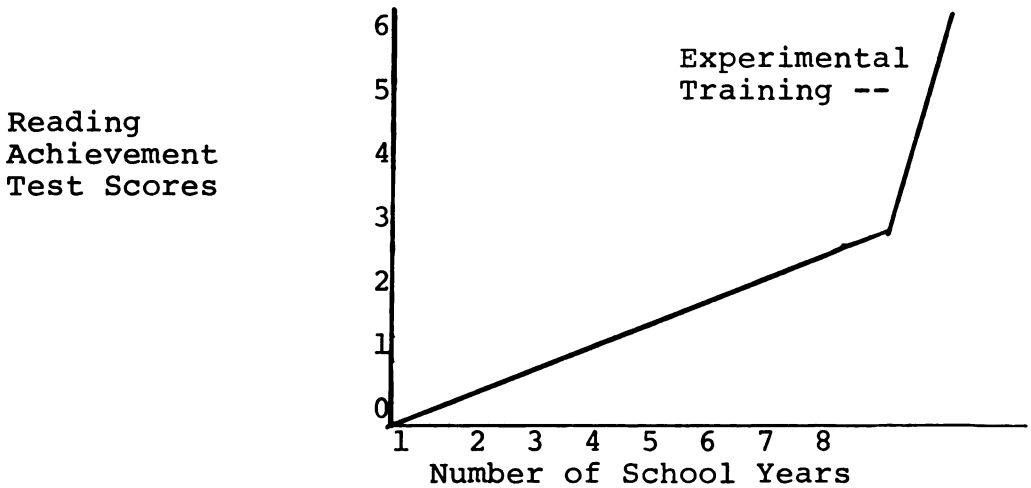
repression interference retrieval

7. We are all aware of discrimination against racial groups, religious groups, women, foreigners, old people, etc., invent a new and different kind of discrimination that you can describe in one sentence.

8. Write an imaginative sentence that includes the following words:

emotion perception satiation

9. Write a general psychological conclusion that can be inferred from the graph at the right.



10. How would your thoughts be organized if you did not have a language?

TEST B

1. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

   identification figure discrimination

2. If people did not organize into groups, suggest some of the possible consequences. Rate your answer (one or two sentences) from 1 to 7.

3. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

   normal hypothesis behavior

4. Suppose two identical twins were separated shortly after birth. One was sent to the Orient, the other to Alaska. In what ways would they differ at the age of 30? Rate your answer (one or two sentences) from 1 to 7.

5. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

   class condition empathy

6. Suggest some of the psychological consequence of marriage. (One or two sentences). Rate your answer from 1 to 7.

7. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

   structure research image

8. What are some of the social consequences for a person who is callous and competitive? (One or two sentences).. Rate your answer from 1 to 7.

9. Write an imaginative sentence that includes the following words and rate your sentence from 1 to 7.

   strategy ship relation

10. Suggest some of the psychological consequence of being an orphan. (One or two sentences). Rate your answer from 1 to 7.

TEST B

Rating of answers for this test is on the basis of creative thinking. Rate each of your answers when completed with this test on a scale from 1 to 7. A rating of 7 indicates a very unusual, creative answer. A rating of 1 indicates a very poor answer. Answers should not exceed two sentences.

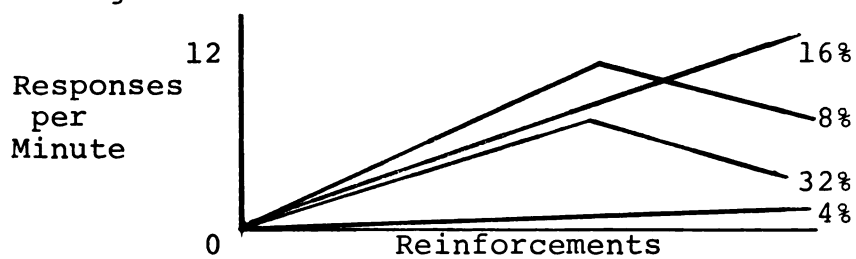
1. What do experimental psychology and the experimental theater have in common?

2. Suppose you want to train an animal in a new, unique way. Invent a reinforcement schedule different from those used by others.

3. Write a hypothesis worth testing relating the achievement motive to innate talent.

4. You have an eccentric roommate. He likes to boat his head against the wall because it feels good when he stops. What question occurs to you as a student of human behavior that you would like to investigate if you had the time?

5. What general psychological conclusion can be drawn from the graph at the right?



6. Write an imaginative sentence using the following words:

   language speech acquisition

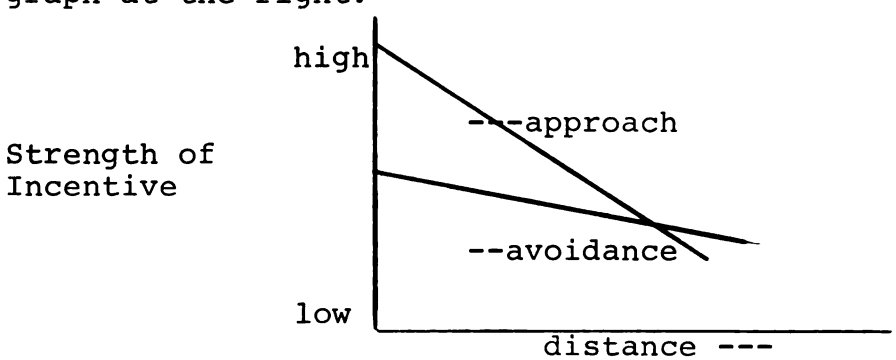
7. The following is a summary of a series of experiments. Write an original title for this series.

   Schachter's ingenius experiments point to the possibility that an affiliative incentive may be broader than physical contact . . . He suggested that the affiliative goal is to reduce confusion about the emotions that a person has in connection with certain anxious situations. College women faced with the prospect of being shocked were given the opportunity of being with other people prior to facing the shock. Schachter argues that "misery seeking company" in this setting is for the sake of establishing communication about emotional responses that the girls feel unsure about.

8. Write an original sentence that includes these words:

   heredity environment retardation

9. What general psychological interpretation can be drawn from the graph at the right?



10. Suppose a pill is discovered which completely inhibits the effects of any kind of frustration. What would be the consequences for the development of personality?

APPENDIX B

PRODUCTIVE THINKING GUIDES

SENTENCE PROBLEM TRAINING

The purpose of this training procedure is to (a). Construct as clear an image as possible as to what is meant by "creative" in producing sentences based upon three words which are given to you.(b). Assist you in producing and judging your productions.

All items are rated on a 7 point scale from least creative (1) to most creative (7). Basically, the items are rated as follows:

For a rating of:	The sentence must have these characteristics.
1	Does not use all three words
2	Lists the words.
3-4	Uses all the words in a sentence which flows well, but the words stand out; a mediocre sentence.
5-7	A creative or unusual sentence.

To ullustrate how the rating guide has been used to judge sentences, we will present some poor, some good, and some superior sentences selected from previous research using the following four words: expensive happy horse lake. The following are examples of <u>poor</u> solutions:

1. The happy expensive horse jumped into the lake.

Explanation: This is a poor solution because it is a simple listing of the adjectives and they modify the noun given, Horse. It does not take much imagination to produce this sentence, and it really doesn't read very smoothly as the words stick out.

2. The expensive horse made the boy near the lake happy.
Explanation: This is fairly good, but "lake" is used as the subject and "expensive" modifies "horse" which makes these words obvious. The given words should be better integrated into the sentence, so that you have to look twice to make sure they are all included.

The following are examples of good sentences, though not superior:

1. A happy scene it was, the children splashing in the lake, a horse basking in the sun; their time not being expensive.

Explanation: This is a good sentence because it uses all the words unobtrusively and uses the adjectives other than to modify the given nouns. It falls short of being superior in that it is so complex it does not read smoothly.

2. I'm happy that our expensive weekend at Horse Lake is over with.

Explanation: The novel use of "Horse Lake," and the fact that it is so smooth reading and well-constructed, makes this a good sentence. A superior solution, however, would have a little something more. This one lacks imagination or humor.

The following are examples of superior sentences:

1. The only thing about the lake that we were not too happy about was spending that expensive time swatting horse flies.

Explanation: This is a clever sentence that reads smoothly and is well constructed. The words fit into the structure unobtrusively, and "horse flies" is a novel use of the word "horse."

2. A beautiful,peaceful lake and a spirited horse, neither of which were expensive in comparison to how happy they make her, proved the best cure for Anne's dejected state of mind.

Explanation: Did you have to read this twice to find all four words? The best sentence will use the words so they are not obvious. In addition it is a compound sentence which reads exceptionally well, and the adjectives are used to modify words other than the given nouns.

For a sentence to be rated as creative it must exemplify some of the following:

1. All words fit unobtrusively in the sentence.
2. Use the critical words in an unusual context.
   a. <u>pun</u>

      "lead" typically will be interpreted as a noun, a more creative usage might be as verb, to lead. "Smelt" may be interpreted as a species of fish or a refining process
      more on puns to come...

   b. <u>verb or modifier</u>

      "torpedo" typically will be interpreted as a noun, a more creative usage might be as a verb, to torpedo. "envy" typically interpreted as a noun, a more creative usage might be as a verb, to envy, and if you're really hep on your Freud, you might envision something like torpedo envy, a novel usage of the word torpedo. "grass" typically interpreted as a noun, a more creative usage might be as a modifier as in Grass Lake; as a pun, grass might be used as in a joint, as in pot. "personality" typically used as a

noun, a more creative usage might be as a modifier, as in personality inventory.

c. compound words Suppose you are given the words

normal sea sub

A creative combination of the words might combine sub-normal thus, you now have only two words to relate into a sentence, and the possibility for a more novel response.

3. Be a clever or witty sentence--your manipulation of the words creates a new perspective or approach at interpreting them.

Are there any questions?

CONSEQUENCES TRAINING

The purpose of this training procedure is to (a). Construct as clear an image as possible as to what is meant by "creative" in producing unusual consequences. (b). Assist you in producing and judging your productions.

All items are rated on a 7 point scale from least creative (1) to most creative (7). Basically the items are rated as follows:

For a rating of:	The consequence must have these characteristics:
1	No answer
2	A denial of the question; an opinion.
3	A trivial consequence or an invalid consequence.
4	A well written valid consequence, but obvious.
5-7	A creative or unusual consequence.

To illustrate how the rating guide has been used to judge consequences, we will present some poor and some good consequences selected from previous research.

1. What would happen if there were no sex drive?

A poor consequence: less people; it would ruin all the fun; or there wouldn't be any love-making.

A good consequence: no need for birth control; changes in art, music, clothing styles, the economy, and family units; less fighting, jealousy, social problems.

2. Suppose you are a monkey and the hairless creatures outside your box are depriving you of any kind of sensory stimulation. What is going through your mind at this time?

A poor consequence: I want out; nothing; or wait til I get out!

A good consequence: Boredum, exploratory fantacies, mind producing stimuli are created via nerve impulses.

3. Suppose children were born without any pain reflexes. Suggest some of the consequences for child development.

A poor consequence: they wouldn't learn; they wouldn't behave; they wouldn't feel any pain.

A good consequence: trial and error learning would be altered; learning via emotional responses; wouldn't completely grasp the meaning of pleasure; vocabulary wouldn't include "pain" words.

For a consequence to be rated as creative, it must exemplify some of the following characteristics:

1. Not obvious, that is, <u>remote</u> and <u>far reaching</u> -- of an idea thought through to its ultimate end.

2. Uncommon, striking, extraordinary--has slight connection or relation to the initial premise or statement, an infrequent response.

3. May be <u>clever</u> or <u>witty</u>.

4. Has breadth, considers several approaches to the solution, that is, economic, political, religious, social, political, educational, psychological.

Are there any questions?

APPENDIX C

TRAINING PROBLEMS

NAME_____

PRACTICE SENTENCES

Write an imaginative sentence that includes the following words and rate your sentence from 1-7.

1. ability stimulus structure

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

_____

2. abnormal stress sample

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

_____

3. rhythm formation concept

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

_____

4. depression inhibition love

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

_____

STOP. HAVE EXPERIMENTER RATE YOUR SENTENCES AND ADVISE.

5. emotion mean graph

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

6. fear subception plans

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

7. frustration perceptual nonsense

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

8. learning response curve

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

STOP: HAVE EXPERIMENTER RATE YOUR SENTENCES AND ADVISE.

9. model goal conformity

S RATING _____
E RATING_____

SENTENCE TO BE RATED:

---

10. unconscious esteem self

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

11. shaping sex behavior

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

12. information logic bit

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

STOP. HAVE EXPERIMENTER RATE YOUR SENTENCES AND ADVISE.

13. heredity environment encounter

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

14. foster performance motivation

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

15. organ environment LSD

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

16. experiment twins test

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

STOP. HAVE EXPERIMENTER RATE YOUR SENTENCES AND ADVISE.

17. birth inference behavior

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

18. language transfer acquisition

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

19. code conflict peers

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

---

20. theory factor aptitude

S RATING_____
E RATING_____

SENTENCE TO BE RATED:

WRITE A REMOTE CONSEQUENCE TO THE FOLLOWING QUESTIONS IN ONE OR TWO SENTENCES AND RATE YOUR ANSWER ON A SCALE FROM 1 to 7.

CONSEQUENCE PRACTICE

1. If the optic nerve and the auditory nerve were cross, we would hear the lightning and see the thunder. What would be the long term consequences of such a crossing?

RATING S
E

2. Consider a hypothetical State off the coast of Zombie-land, Utopia Z. The Utopians are of a socialist origin and have a standard income of seven thousand dollars per each individual over 18. How much such a culture's social attitudes differ from ours?

RATING S
E

3. If during the course of evolution man had developed as a marine animal instead of a land animal, what differences would you expect in his sensory equipment?

RATING S
E

STOP. LET EXPERIMENTER RATE YOUR ANSWERS AND ADVISE.

4. Suppose you are a Peace Corps worker in a lesser developed country. You find that your most immediate problem is to get the villagers to use sanitation facilities so as to reduce the disease in the area. You can speak the native language and are aware of the villagers deep sense of familial pride. How might you approach the task?

RATING S
E

5. If you were studying social conformity in the college classroom, what activities would you observe?

RATING S
E

6. How would human behavior differ if there weren't any "body talk?"

RATING S
E

7. It is possible that recent student movements of college campuses will change the shape of modern psychology. What sort of change will this be?

RATING S
E

8. If it becomes possible to change human abilities by manipulating chemical substances in the chromosomes, what change would have the greatest effect?

RATING S
E

9. How might a person's IQ be limited by his language?

RATING S
E

10. Suggest some of the psychological consequences of losing a limb.

RATING S
E

11. Suggest some of the psychological consequences of being a homosexual.

RATING S
E

STOP. LET EXPERIMENTER RATE YOUR ANSWERS AND ADVISE.

12. Suggest some of the psychological consequences of music.

RATING S
E

13. Suggest some of the psychological consequences of being deaf.

RATING S
E

14. Suggest some of the psychological consequences of being a genius.

RATING S
E

15. Suggest some of the psychological consequences of being ugly.

RATING S
E

16. What are some of the social consequences of being considerate?

RATING S
E

17. Suggest some of the psychological consequences of being Black.

RATING S
E

18. Suggest some of the psychological consequences of being reprimanded.

RATING S
E

19. Suggest some of the psychological consequences of being impotent.

RATING S
E

20. Suggest some of the psychological consequences of being fat.

RATING S
E

STOP. LET EXPERIMENTER RATE YOUR ANSWERS AND ADVISE.

MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 03062 2348