To the Graduate Council:

I am submitting herewith a thesis written by Carleton Bryan Southworth entitled "A Study of Performance Contrasting Established Groups with Zero-History Groups and Individuals." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Industrial and Organizational Psychology.

John M. Larsen, Jr., Major Professor

We have read this thesis and recommend its acceptance:

Accepted for the Council:

The Graduate School

A STUDY OF PERFORMANCE CONTRASTING ESTABLISHED GROUPS WITH ZERO-HISTORY GROUPS AND INDIVIDUALS

A Thesis

Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Carleton Bryan Southworth

June 1984

DEDICATION

I dedicate this thesis to the memory of my father, with the words on the following page. They were written for a ceremony establishing a memorial grove at the University of Massachusetts Medical School in Worcester with a tree in his memory, by a young man who was a graduate student intern of his at the time of his death.

Infinite living Presence in which all creation lives and moves and has being, common source of all us humans, and trees, too,- we dedicate this copper beech to the fine spirit which Al Southworth has embodied in his life.

As surely as he has enjoyed and shared laughter with those around him, may this tree grow into a place of laughter and ease.

As surely as he has lived his life lending a wise and caring hand to those in need of one, may this be a place of reflection and refreshment.

As surely as he has shown us all the strength and dignity that come from a spirit of good will, may this tree become a place of discovery of inner strength.

As surely as he has worked with great care and diligence in his counseling work, may this tree be a resting place for those who work here in their own areas of expertise.

As surely as he has blessed us all with the freshness of his inquisitive mind and loving heart and with his delight in the whole width and depth of life, - from the sea and sailing, to the lives of the students he counseled, - may this be a place of fresh inquiry and widened love.

So that just as surely as this tree will take from the earth and the air and the rain the nutrients it will need to grow and flourish and become a growing source of beauty for all who come upon it, may we, too, take what is most valuable to each of us from the spirit of this man and incorporate it in our own way in our lives in order to grow and flourish and become sources of beauty.

In this will the spirit of Al Southworth, and this tree, and all of us be blessed.

Terry Allen
September 27, 1983

ACKNOWLEDGEMENTS

I wish to acknowledge the help and guidance of Dr. John M. Larsen, Jr., whose input has contributed to each section of this document. I also wish to thank Dr. Michael Rush, who suggested many pertinent references, and often left photocopies of them in my mailbox. Dr. Eric Sundstrom offered insight and suggestions in connection with the design and analysis of results, and Dr. R. Thomas Ladd helped with the power analysis.

Special thanks are due to Dr. George Levinger for bringing me from a general research topic to a more specific one, and to David Lewis for his suggestions regarding the proposal for this thesis, and for several word processor lessons.

Finally, I wish to acknowledge those whose care and warmth sustained me as I worked on this project: my mother, Dr. Barbara Southworth; my aunt, Lois Southworth; and my friends, Linda Sammataro, George Yancey, and Mark Carlisle.

ABSTRACT

Previous research has shown that newly formed groups can decrease individual performance when group members are engaged in the same task, and when the individuals' contributions are anonymous. This phenomenon has been called the "Ringelmann effect" or "social loafing." This study was designed to determine whether these decreases also occur in well established groups. Three treatment groups were created to form a between-subjects design. Subjects were tested on a brainstorming task either (1) individually, (2) in a newly formed group, or (3) in an established group with a history of cooperative activity. The perfomance decrease found by other researchers in the newly formed groups was not replicated here. No significant differences were found on the brainstorming task between the three treatment groups. (Unexpectedly, a negative correlation was found between feelings of responsibility toward the zero-history group and performance on the brainstorming task.) It was concluded that social loafing may occur in fewer situations than previously believed. Varying the degree to which subjects' performance may be observed is suggested for further research.

TABLE OF CONTENTS

SECTION		PAGE
I.	INTRODUCTION	1
II.	REVIEW OF LITERATURE	4
	Performance Decreases in Groups	4
	Rationale for Research Project	8
III.	METHOD	19
	Subjects	19
	Ideal Design	20
	First Attempted Design	25
	Actual Design	28
IV.	ANALYSIS	32
	Pilot Study and Power Analysis	32
	Reporting of Type I Error	35
	Statistical Tests	37
V.	RESULTS	41
	Outliers	41
	Statistical Results	43
VI.	DISCUSSION AND POST-HOC ANALYSIS	49
	Introduction	49
	Possible Causes for the Absence of Social Loafing	49
	Closing Observations	57
VII.	SUMMARY	61

SECTION PAGE

REFERENCES 63

APPENDIXES 67

APPENDIX A. QUESTIONNAIRES 68

APPENDIX B. FIGURES OF EXPERIMENTAL DESIGNS 74

APPENDIX C. SUPPLEMENTARY STATISTICS 77

VITA 86

LIST OF TABLES

TABLE		PAGE
1.	Pilot Study Means and Standard Deviations for Brainstorming Task by Treatment Group	33
2.	Pilot Study Means and Standard Deviations for Brainstorming Task by Task Type	33
3.	Means and Standard Deviations for Brainstorming Task with Outliers	42
4.	Means and Standard Deviations for Brainstorming Task by Treatment Group without Outliers	44
5.	Correlation of Cohesion with Brainstorming Task	46
6.	Correlation of Task Effort with Brainstorming Task	47
7.	Correlation of Task Responsibility with Brainstorming Task	48
8.	Class Section Means and Standard Deviations for Brainstorming Task without Outliers	78
9.	Treatment Group Means and Standard Deviations for Task Effort without Outliers	78
10.	Class Section Means and Standard Deviations for Task Effort without Outliers	79
11.	Treatment Group Means and Standard Deviations for Cohesion without Outliers	79
12.	Class Section Means and Standard Deviations for Cohesion without Outliers	80
13.	Treatment Group Means and Standard Deviations for Task Responsibility without Outliers	80
14.	Class Section Means and Standard Deviations for Task Responsibility without Outliers	81
15.	Grand Correlation Matrix without Outliers	81

TABLE PAGE

16. Correlation Matrix for Continuing Group Treatment 82

17. Correlation Matrix for Individual Treatment without Outliers 82

18. Correlation Matrix for the Zero-History Group Treatment without Outliers 83

19. Correlations Between Brainstorming Task Z-Scores and Other Variables by Treatment Group 84

20. Correlations Between Brainstorming Task Z-Scores and Other Variables by Class Section 85

I. INTRODUCTION

In the course of scientific inquiry sometimes results contradict our intuitive grasp of how people behave. Ringelmann, A German psychology student of W. Moede, obtained such a result. He discovered that group participation can reduce an individual's performance. His simple experiment was never published, but his teacher reported it in 1927 (Moede, 1927). In his experiment Ringelmann found that an individual decreases his pull on a rope as the number of people pulling with him increases. He found this to be a linear effect. Groups of eight exerted only about 50% as much force as the combined total of the group members' individual performances (reported by Dashiell, 1935). Latané and his research associates, who have contributed extensively to research in this area, refer to this phenomenon as "social loafing." The purpose of the research reported here was to determine if this effect is generalizable to cohesive, developed groups. Heretofore, the only type of group examined has been a collection of individuals assembled for a single experimental session. Such groups cannot develop the affective bonds characteristic of groups which have been involved in cooperative activities over time.

The literature review below summarizes what is known about the generalizability of social loafing, and

outlines the constructs that have been developed to explain and predict it. Following this review, the possible moderating effect of long-term group membership on social loafing is discussed, and an experiment developed to test the hypotheses generated in this discussion is outlined.

Those responsible for managing organizations, as well as society as a whole, have an interest in a clear understanding of how groups influence an individual's productivity. Accordingly, the questions addressed by this research have practical implications for those currently involved in motivating employees.

Predicting the domain of circumstances that elicit social loafing has theoretical importance, too, particularly because the effect runs contrary to our notions of "team spirit" as a means to increase motivation. In reference to this it is interesting to consider that Harkins, Latané, and Williams (1980) found during one experiment that individuals reported no awareness of decreasing their effort in a group setting.

In addition to being against what common sense may tell us about the motivational impact of groups, social loafing also seems to contradict the prediction of social facilitation theory, at least at first glance. Zajonc's (1965) theory of social facilitation can be interpreted as

outlining a mechanism that enhances individual performance on a well-learned behavior, like rope pulling, as group size increases.

Before examining the experiments that have elicited social loafing, note that only additive tasks have been shown to exhibit the social-loafing effect. Steiner (1972, p. 32) explains how additive tasks differ from other types of tasks:

> When tasks are conjunctive or disjunctive, the group product reflects the contribution of a single member. In the case of conjunctive tasks, it is the least proficient member of the group (or the least proficient member who is involved in the conjunctive operations) who determines the quality of the group output. When the tasks are disjunctive, the group may accept any member's outcome as its product. Additive tasks are strikingly different in that they permit the contributions of various members to be summed. Consequently, the group product is a combination of individual outcomes rather than the outcome generated by a single member.

While it might be argued that conjunction, disjunction, and additivity exist on a continuum, tasks that have produced social loafing are all primarily additive in nature according to Steiner's typology of tasks.

II. REVIEW OF LITERATURE

Performance Decreases in Groups

Experiments that have been conducted on social loafing so far have been aimed primarily at determining the causes of the effect. An early hypothesis put forth by Steiner (1972) suggested that a lack of coordination between group members may account for at least part of the decrement in performance. In a rope pulling task, for example, members may pull at slightly different angles, or they may not exert their full efforts simultaneously. Ingham, Levinger, Graves and Peckham (1974) examined this possibility in the first published attempt to replicate Ringelmann's study. In their second experiment, these authors devised a means to convince individuals that they were pulling in groups when they were actually pulling alone. Social loafing still occurred. Their results showed progressive performance decreases in groups of two and three. Interestingly, however, no further decreases occurred with the addition of a fourth, fifth, or subsequent member to the group (the effect was observed to be curvilinear rather than linear). Thus, lack of coordination cannot account for social loafing; rather, the effect is caused by some kind of motivational loss.

Social loafing depresses performance on different kinds of tasks, and in different kinds of groups. When

individuals were confronted with the task of shouting as loudly as they could, shouting in groups decreased performance (Latané, Williams, and Harkins, 1979). These researchers observed this performance decrement whether individuals shouted in actual groups or in a pseudo-group situation in which participants believed they were shouting with others, but actually were not, and could not see or hear others shouting.

Furthermore, these researchers observed social loafing in both the actual group and pseudo-group experiments despite the fact that those in the real groups were given a different rationale for the experimental purpose than those in the pseudo-group experiment. In the real groups, the experimenters told participants that their primary interest was in judgements of how much noise people make in social settings whereas in the pseudo-group experiment they claimed to be studying the effects of sensory feedback on the production of sound in social groups (Latané et al., 1979).

Adding to evidence that the social-loafing effect is robust in nature, Harkins et al. (1980) obtained decreases in effort when they assigned clapping as the task in a group setting.

Is the social-loafing effect generalizable to cognitive tasks? Evidence indicates that the answer to

this question is yes. Harkins and Petty (1982) found performance decreases on two cognitive tasks when individuals were placed in group settings. In their first experiment, fewer "potential uses for a knife" were generated by individuals when they were grouped. In their second experiment more errors occurred on a vigilance task; when subjects were in a group and their individual performances were not identifiable, performance decreased.

Two additional studies show that social loafing occurs in cognitive tasks. Petty, Harkins, Williams and Latané (1977) conducted an experiment involving the evaluation of poems and editorials. This study showed that group membership reduced motivation and performance. In a task involving the evaluation of psychotherapy, Petty, Harkins and Williams (1980) obtained a similar finding. Individuals who were not part of a group were better able to differentiate between videotapes of high and low quality psychotherapy.

These last two studies may be used to illustrate the importance of social loafing in practical affairs. For example, should performance appraisal be done by combining individually obtained identifiable ratings rather than by combining unidentifiable ratings of group members? Results of these experiments appear to support an affirmative answer. Likewise, if my company were

looking for an innovative way to market an excess inventory of plastic containers, should I delegate this responsibility to identifiable individuals rather than a group? Results of Harkins and Petty (1982) concerning "potential uses for a knife" indicate that the answer to this question, too, is yes. A case will be made, however, that formal, continuing groups differ systematically from zero-history groups such that formal groups with a history of cooperative activity do not exhibit social loafing. If this is true, then an alternative answer to these two questions may be more appropriate.

Latané et al. (1979) put forth a hypothesis for social loafing they called "lessened contingency between input and outcome." Patterned after expectancy theory, this explanation suggests that because an individual's response can not be individually monitored in a group setting, the individual perceives that he or she has less control over his or her rewards. The negative consequences of "slacking off," as well as the positive consequences of high performance, are psychologically divided between group members under this hypothesis. Approval or disapproval of the experimenter (or whatever reward is being offered) loses its quality of individual ownership in group settings. Latané et al. (1979, p. 830) suggest that:

> . . . if inputs were made identifiable and rewards contingent on them, even when in groups, it would be impossible for performers to get a free ride and they would have an incentive to work equally hard in groups of different sizes.

To test this hypothesis Williams, Harkins and Latané (1981) designed an experiment in which group size, actual groups versus pseudo-groups, and identifiability of individual efforts were manipulated as independent variables. In this study identifiability eliminated social loafing, supporting the "lessened contingency between input and outcome" hypothesis. Again, the task was to shout as loudly as one could (see also Kerr and Bruun, 1981). This theory does not, however, offer a comprehensive explanation of social loafing. This inadequacy is explored below.

Rationale for Research Project

Kerr and Bruun (1983) have pointed out that, in all probability, many factors influence an individual's choice of how much effort to exert on a group task. These authors studied non-additive tasks and demonstrated that groups can suppress individual performance even when subjects are identifiable if they are assigned conjunctive or disjunctive tasks. Furthermore, Harkins and Petty (1982) found evidence that social loafing is reduced or eliminated when individuals are assigned especially

difficult tasks or when each group member is assigned a unique task.

Although the "lessened contingency between input and outcome" hypothesis may explain the impact of identifiability on social loafing, it does not predict the absence or reduction of the phenomenon seen when group members are simultaneously assigned unique tasks and individually unidentifiable performances; nor does it predict a reduction or elimination of social loafing in difficult tasks. If the hypothesis that a psychological division of goals has any validity as a cause of social loafing, then it is masked or superseded by other processes in difficult tasks and in situations where group members are assigned unique tasks. A broader paradigm is needed to encompass what is known about social loafing. What is needed is a theory that explains or predicts the impact of these variables--task difficulty, identifiability, and uniqueness of task--under one unifying concept. Petty et al. (1980) have expressed the view that a stimulus, such as a task demand, is processed less thoroughly in groups that exhibit social loafing. This view will be clarified below. It is the construct upon which the rationale for this study is based. Before examining the value of this theory, let us review the known circumstances required to produce the social-loafing

effect. They include the following three factors. First, the task must be primarily additive in nature, with each individual assigned a similar or identical task. Second, the task must be of moderate or low difficulty. We also know that social loafing occurs whether real or pseudo-groups are created. And finally, subjects must remain unidentifiable in the group condition.

Petty et al.'s (1980) stimulus-processing theory asserts that individuals in a social-loafing group do not focus on the task as well as identifiable individuals, perhaps because recognition or other rewards for performance are psychologically divided (accordingly, it subsumes the "lessened contingency between input and outcome" hypothesis). It is quite plausible, too, that a difficult task in itself would heighten arousal because a difficult task requires more concentration: and, individuals may process a stimulus more thoroughly in a group when they are assigned a unique task, because they feel their contribution will be more critical to the group's success. Therefore, the elimination of social loafing observed by Harkins and Petty (1982) in difficult and unique tasks is consistent with a stimulus-processing theory.

Harkins and Petty (1982) found that individuals engaging in difficult or unique tasks felt that their

contributions were more unique than the contributions of those engaged in simple tasks. This result suggests that these two conditions are similar in some respects; both unique and difficult tasks may heighten one's sense of individuality because each condition leads one to feel that his or her contribution to the group is individual. Thus, there may be a relationship between the cause of performance increase seen when subjects are tested as individuals and the cause of performance increase observed when unique or difficult tasks are employed. Kerr and Bruun's (1983) results, while not pertaining directly to social loafing, are also consistent with a stimulus-processing theory. An assumption behind the experiment outlined below is that any variable that affects an individual's arousal to focus upon an additive task will moderate social loafing.

An information-processing theory which gives arousal a central function is the most parsimonious approach to explaining the range of findings relevant to social loafing yet to be advanced in the literature. Before integrating the role of long-term group membership into this theory, let us reexamine the apparent conflict between the findings of social facilitation theory and social loafing. Recall that, according to social facilitation theory, coactors generally enhance

performance on simple tasks. This seems to contradict the results of social-loafing research. However, a relationship between these two phenomena can be shown by focusing on the role social variables play in influencing arousal.

Geen and Gange (1977) review the difficulty early researchers encountered in explaining social facilitation. A satisfactory theory for explaining why coactors or evaluative audiences may help or hinder performance was not achieved until Zajonc (1965) looked at the problem in drive theory terms. Initially, it appeared that the social facilitation phenomenon could be explained by the simple idea that increases in arousal increase dominant responses (Zajonc, 1966). If a task is well-learned, then the "correct" response is dominant. The arousing effects of other people increase the rate of correct responses. Conversely, incorrect responses are dominant when tasks are being learned, especially in regard to difficult tasks. In the case of learning difficult tasks, then, the arousing effects of other people increase the rate of incorrect responses. In conclusion, arousal increases the dominant reponse whether one is learning responses or performing learned responses.

This early distinction between the effects of an audience or coactor on learning versus learned responses

has been elucidated further by recent research. Particularly, audiences or coactors may improve the rate of correct responses during learning if the task is simple, as shown below. This point has implications regarding social loafing, because social loafing occurs only in simple or moderately difficult tasks.

Keep in mind that social facilitation research supports the notion that an increase in drive produced by social influence increases dominant responses. Thus, when one is learning simple responses, social influences may increase the rate of correct responses if correct responses are dominant, even in the initial phases of learning. Cottrell, Rittle, and Wack (1967) found this result in an experiment involving paired-associative lists. This social facilitation study is consistent with Harkins and Pettys' (1982) finding that social loafing occurs only in simple or moderately difficult tasks, as previously mentioned, but not in difficult tasks.

If the same arousal-related mechanism that accounts for social facilitation is operating to produce social loafing, then it would be instructive to ask: what is reducing arousal in social-loafing situations? Why might individuals be more aroused alone than they are in groups producing the social-loafing effect?

Evaluation apprehension appears to be a mechanism that can explain the results examined in this review. Geen and Gange (1977, p. 1284) state:

> . . . Audience-induced anxiety, by producing a narrow range of cue utilization, may help focus the subject's attention on central cues that provide most or all of the information necessary to the solution of simple problems yet prompt the elimination of peripheral cues that may be necessary to the solution of more complex ones.

Remembering that identifiability of performance eliminates social loafing (Kerr and Bruun, 1981; Williams et al., 1981), it is reasonable to propose that evaluation apprehension may account for social loafing. Cottrell (1972) asserts that the impact other people have on drives is learned, not innate. From this vantage point, we can hypothesize that the presence of other people can increase or decrease arousal, depending upon the circumstances. This proposition is supported by Kerr's (1983) conception that social norms account for motivation losses in small groups. Kerr found that to coact with a failing partner causes one to reduce effort when the partner's failure is perceived to be caused by low effort, but does not cause one to reduce effort when the failure is perceived to be caused by low ability. With social norms put in this perspecive, the posssible impact of continuing group membership on social loafing can now be addressed.

Ingham et al. (1974) suggested almost a decade ago that continuing groups may not exhibit social loafing, and this opinion has been reitterated by Kerr and Bruun (1983) in their recent article. What evidence is there to suggest that the presence of a group history may be important in determining the presence or absence of the social-loafing effect?

Two studies have been published that show continuing groups to be detrimental to productivity (Scofield, 1960; Sypher, 1977). In these two studies time spent in social interaction between group members detracted from time spent on the task in cohesive groups, thus decreasing performance. What is the consequence of group cohesion when social interaction is not permitted, or when there is no time factor? Hoogstraten and Vorst (1978) performed one study meeting these conditions.

These researchers studied the relationship between cohesiveness and performance on additive and disjunctive tasks, and found high cohesive groups to perform better in eight comparisons (although only four reached statistical significance). Hoogstraten and Vorst manipulated group cohesion in this study by administering personality tests and suggesting, ostensibly on the basis of these tests, that those in high or low cohesion conditions could be expected to like one another (subjects were actually

assigned to conditions at random). This study involved no continuing group membership, nor were any groups created in which cohesion was not manipulated. Moreover, no control over identifiability existed. Still, the pattern of consistent superior performance by the high cohesive groups raises the possibility that continuing group membership may eliminate social loafing. Cohesion has been found to be positively related to performance in a simulation game (Norris and Niebuhr, 1980), in ice hockey teams (Ball and Carron, 1976), and in tank crews (Tziner and Vardi, 1983). These last three studies preclude one from making causal inferences, however, because they are correlational in nature.

The proposition that interpersonal attraction will emerge between group members engaging in cooperative tasks with shared goals is widely supported (see Lott and Lott, 1971). As interpersonal attraction increases, it is likely that a greater concern for the well-being of fellow group members will develop, too. Thus, as the responsibility for an additive task is psychologically divided in established groups, creating a force that lessens the stimulus impact of a task, the countervailing force of interpersonal attraction may increase an individual's motivation to produce a favorable contribution to group performance. Like task difficulty,

this interpersonal attraction may increase the degree to which people process a stimulus, and cause them to focus on the task to a degree that social loafing is eliminated. Performance on a brainstorming task in continuing groups with a history of performance on cooperative tasks was measured to see whether the social-loafing effect would be absent, as predicted by hypothesis 1. Hypothesis 2 predicted that group cohesion correlates with performance.

It was tested by determining whether a relationship existed between a measure of group cohesion (expressed by individuals toward continuing groups) and individual performance on a brainstorming task. Cohesion is defined as interpersonal attraction to group members and to the group itself.

A further test of this hypothesis involved two items designed to measure feelings of responsibility toward subjects' experimental groups. Those who felt more responsibity for their group performance were expected to attain higher scores on the brainstorming task. Finally, hypothesis 3 asserted that zero-history group members will score lower on measures of task effort and involvement than individuals tested as members of continuing groups and subjects assigned to a task individually. This hypothesis was tested by measuring task effort and involvement and comparing treatment means. These

hypotheses are designed to test whether the social-loafing effect is moderated by continuing group membership, and to determine the efficacy of the stimulus-processing theory in regard to explaning this phenomenon.

In one study utilizing cognitive tasks, individuals reported lower effort in unidentifiable group conditions, supporting the stimulus-processing theory (Petty et al., 1977). Stimulus-processing theory is supported by additional evidence (Petty et al. 1980). These experimenters found that individuals evaluating high-quality psychotherapy had more thoughts that were favorable about the stimulus than unidentifiable group members did. Likewise, individuals evaluating low-quality psychotherapy had more thoughts that were unfavorable about the therapy they saw than group members who evaluated low quality psychotherapy did. These results are consistent with a stimulus-processing theory. Therefore, in the research proposed here, continuing group members should rate their effort and involvement in the task higher than zero-history group members if continuing group membership eliminates social loafing and the stimulus-processing theory is correct.

III. METHOD

Three designs are presented below. The ideal design is not feasible to execute, but is given first to represent the conceptual starting point from which the actual design emerged. This ideal design is the design that would minimize variance in the error terms if it were feasible to execute. The design first attempted, like the ideal design, required a high degree of coordination among existing groups. It was altered after results from preliminary inquiries related to scheduling indicated that it would not be feasible to execute, either. Additional alterations that were necessary are elucidated under the heading "actual design."

To review, the three hypotheses were: (1) individual performance in continuing groups will show no decline compared to individual performance alone; (2) individual measurements of cohesion will be positively related to performance in continuing groups; and (3) zero-history groups will score lower on measures of task effort and involvement than either individuals working alone or those working in continuing groups.

Subjects

Ninety-one male and female undergraduate students participated in the study. These participants were drawn

from three sections of a management course employing experiential group exercises as a primary teaching method. The task and other questionnaires that were employed were part of the experiential group exercises required for the course. Students in each course section were assigned to a permanent group that assembled periodically for the duration of the course. Most of the exercises these groups engaged in were cooperative in nature.

Ideal Design

Procedure and treatments. At the outset of this experiment, it had been hoped that four class sections would be included, yielding an N of 160. The ideal design is based on this figure.

A repeated-measures design would be created with three treatment groups. Treatment 1 would be the continuing group treatment, in which individuals would be tested with members of their permanent groups. Treatment 2 would be the zero-history group treatment, in which individuals would be randomly assigned to a group, with other group members coming from other sections of the same course. In these first two treatments, group members would be asked to inscribe their group's name on experimental questionnaires, but would be unidentifiable as individuals. Treatment 3 would be the individual

treatment, where subjects would be asked to sign their names to experimental questionnaires making them fully identifiable.

Because there are three treatment conditions in this design, there are six orders in which the treatments could be administered. Approximately 48 permanent groups of three persons each would participate (based on a final N of 144 after attrition). After stratifying by course section, these groups would be randomly assigned to one of the six treatment orders, so that two permanent groups from each course section would be assigned to each treatment order. Each treatment order would contain eight permanent groups (24 individuals in each treatment order). Combining across all treatment orders yields 144 observations for each of the three treatments. This design is shown in Figure 1 (see Appendix B).

Dependent variables. Three dependent variables would be examined. One dependent variable would be a task used to determine treatment effects regarding social loafing. This task would exist in three parallel forms. After individuals listened to and read instructions, they would procede to write down either "potential uses for a knife," "potential uses for a piece of string," or "potential uses for a piece of wood." One of these parallel tasks has already elicited social loafing

(potential uses for a knife; Harkins and Petty, 1982). The other two have produced similar means and standard deviations in a pilot study. Subjects would be told to stop when 10 minutes had passed.

The first dependent variable would be the total number of uses generated in the brainstorming tasks. This variable would be tested to determine whether social loafing occurred and whether the effect is moderated by continuing group membership (hypothesis 1). Because there would be three parallel forms of the brainstorming task, there would be six possible orders of administration. Each of the six individuals assigned to a particular treatment order section would be randomly assigned to one of these parallel task orders to distribute any systematic order effects that may be present, similar to the way in which these six individuals from each class section would be randomly assigned to treatment administration orders.

A second dependent variable would be a questionnaire designed for this experiment to measure task involvement and effort. This cognitive effort measure would be adapted from a three item questionnaire used successfully in similar contexts by Petty et al. (1977), and by Petty et al. (1980). The first study cited here reported an average correlation of .82 between the three items. This dependent variable would be tested to

determine whether zero-history groups report lower levels of task involvement than in the other two treatment conditions, as predicted by hypotheis 3.

After the administration of other questionnaires and instructions, the subjects would complete a measure of group cohesion designed to determine their orientation toward their permanent group. This cohesion questionnaire would be patterned after one developed by Moran (1964) which Hoogstraten and Vorst (1978) found to have a reliability of .82 (K-R 20). This measure would be tested to determine whether it correlates with performance in the continuing group treatment, as predicted by hypothesis 2. Both the group cohesion and task involvement questionnaires would utilize Likert-type scales.

Constraints rendering the ideal design infeasible. Thirty-six separate conditions would be required merely to randomize treatment-order x parallel-task order effects in this design. Furthermore, the coordination problems involved in bringing the randomly assigned subjects together at the proper time might be insurmountable.

The ideal design would also require subjects in the continuing group treatment to believe they were working on the same task when they actually were not. This manipulation might be accomplished by either giving instructions pertaining to the task in written form only,

or by isolating group members from one another completely (creating a pseudo-group situation). In either case the results would be impaired if the manipulation were not successful. It is noted above that Harkins and Petty (1982) eliminated social loafing by having group members work on unique tasks.

In addition, subjects might compare their performance on one parallel task with their performance on subsequent tasks, setting up a situation in which they competed with themselves. Repeated-measures designs have successfully elicited social loafing in the past (e.g. Latané et al., 1979), but only when the task is such that self-feedback is not possible. A pilot study conducted by this author using the three parallel tasks outlined above did not elicit social loafing, perhaps because the results were confounded by self-feedback. Elimination of this possible confounding factor in the ideal design would require trials widely spaced in time, which would be extremely problematic.

A final problem with the ideal design involves the difficulty of collating the responses from the three treatments for each individual. Any coding system used, or the suspicion that handwriting might be used to collate responses, might contaminate identifiability. It has been

shown that identifiability eliminates the social-loafing effect.

First Attempted Design

To reduce implementation problems associated with the ideal design, a simpler between-subjects design was attempted. This design is shown in Figure 2 (see Appendix B).

Independent variables. According to plan, subjects were randomly assigned to one of three treatment conditions after having been stratified by class section. Subsequently, subjects were assigned to permanent groups within their treatment assignments. This outcome is shown to the right of the heading "groups within class section" in Figure 2. Five or six individuals were assigned to each permanent group. The three groups represented by a rectangle in Figure 2 (beneath each treatment block) depict one of two permanent groups from each class section that were created from those assigned to a particular treatment.

The bottom-most part of Figure 2 represents the way in which the task would be presented in each treatment. Those in the continuing group treatment would be tested together in their permanent groups. Those who had been assigned to the zero-history treatment would be placed in

a new group. In this new group one individual would come from each of the class sections. The group to which a zero-history subject would be assigned would be determined by a random process. Experimental confederates would be used to "fill out" the groups. Those randomly assigned to the individual treatment were to be tested as individuals.

Dependent variables. The dependent variables were identical to those described for the ideal design except that subjects responded to only one of the parallel tasks (potential uses for a knife). An additional dependent variable was included. Those tested in groups answered two questions concerning the amount of responsibility they felt toward their experimental group. Thus, those in the zero-history treatment were asked how much responsibility they felt toward their new group, whereas those in the continuing group treatment were asked how much responsibility they felt toward their permanent group in regard to the group output on the brainstorming task. Acting as an additional method, this variable also tested the stimulus-processing theory. This responsibility measure should be positively related to task performance for each of these treatment groups.

Procedure. A problem that made the first attempted design infeasible involved the difficulty of bringing

three permanent groups, one from each course section, together simultaneously. This plan would have required about 15 people, each with separate schedules, to meet simultaneously. This requirement could not be met. The design that would have been followed if this had been possible is outlined in the two paragraphs that follow.

About 15 individuals would be present at each experimental administration. To avoid the possibility that subjects would guess the purpose of the experiment before hand, subjects would have signed up for an administration on a group by group basis, according to their permanent group assignment, regardless of which treatment they were to be in. Only one of the three treatments would be administered during each experimental administration. Accordingly, each permanent group would have a choice of three or four time slots in which to participate. Although there would be some attrition due to this process, all permanent groups that remained in the continuing group treatment would have at least three individuals present, which is a sufficient number to elicit social loafing (Ingham et al., 1974). Because the number of group participants in the continuing group treatment would vary somewhat, the number of individuals in each zero-history group would also be varied by adding

or subtracting experimental confederates. Each treatment condition would have approximatley 30 participants.

Upon arrival to the experimental session, those in the individual treatment condition would be seated separately. Those in the continuing group treatment would be seated in a semi-circle with their other permanent group members, and those in the zero-history treatment would be seated similarly with people from other course sections and experimental confederates.

Actual Design

The coordination problems described above required the creation of pseudo-groups in the zero-history group treatment. This was undesirable because the stimulus cue--the idea that one's responses would be combined with those of other unknown persons--was not as concrete as the corresponding cue for the continuing group treatment. Recall, however, that Latané et al. (1977) obtained the social-loafing effect with pseudo-groups. The design originally planned would have resulted in an attrition rate which would have offset any theoretical advantage it might have. An additional arguement for this change involves the possible negative impact attrition could have on the generalizability of results, if attrition were non-random.

In order to work around these constraints, each of the three classes were tested independently. Each class was divided into three treatment conditions, which were created according to the description above. Each treatment condition was then tested in a separate room, with one male experimenter present for each condition. The instructions are elucidated in the following section.

Instructions. All subjects were read the following instructions, which are similar to those used by Harkins and Petty (1982).

> Today you will take part in what is called a brainstorming task. You will be asked to come up with as many uses for a knife as you can. Don't be concerned about the quality of your reactions. The uses can be ordinary or unusual. Simply list as many as you can.

In the individual treatment, this statement was followed by further oral instructions:

> In this experiment the interest is in the number of uses you can come up with as an individual. Therefore, please sign your name here _____. In appreciation for your participation, one of you will be awarded a coupon for a free 'all you can eat' pizza lunch following this session. This coupon will be awarded on a random basis. Remember, list as many uses for a knife as you can.

In the continuing group treatment, subjects heard:

> Your responses are anonymous. In this experiment, the interest is in the number of uses your class work group can come up with. Therefore, only your class work group is marked on this questionnaire.

These instructions were followed by the "free lunch" message, and a final encouragement to list as many uses as possible.

In the zero-history treatment, subjects heard:

> Your responses are anonymous. In this experiment, the interest is in the number of uses a new group can come up with. Your responses will be combined, on a random basis, with three or four other people's responses. The other members of your group will be drawn from other classes.

These instructions were followed by the "free lunch" message and an additional encouragement to "do one's best." The anonymity of responses was emphasized a second time for those in the group treatments.

Conditions. Ten minutes was the allotted time span for the brainstorming task (as described in the ideal design section). Finally, subjects were asked to answer the questionnaires designed to measure (1) task involvement and (2) cohesion in their permanent groups. Subjects in the zero-history and continuing group treatments responded to an additional two-item questionnaire designed to measure their feelings of

responsibility toward their experimental group (which was either their permanent group or a new group, as indicated above). Copies of these questionnaires are included in Appendix A.

IV. ANALYSIS

Pilot Study and Power Analysis

A pilot study using pseudo-groups was conducted. Three parallel brainstorming tasks described earlier (potential uses for a knife, piece of string, and piece of wood) were administered. This experiment employed a repeated measures randomized design with an N of 35. No significant differences were found between task means. Possible reasons for this result were considered. A summary of means and standard deviations is shown below in Table 1.

The mean of each task is similar to one reported by Harkins and Petty (1982) for "potential uses for a knife" in their fourth experiment. These researchers obtained a mean of 21.37 in an unidentifiable group condition. They did not report standard deviations. In the same experiment, Harkins and Petty reported a mean of 30.67 when individuals believed they alone were generating uses for a knife. The assignment of a supposedly unique object to each individual eliminates social loafing, just as identifiability does. Thus, a resonable benchmark for performing a power analysis is: 30.67 minus 21.37 equals 9.30 (the difference between the two conditions just described).

This power analysis is conducted as though multiple t-tests were being conducted. It is based only upon number of responses to the task. Because the Wilks' lambda analysis used more information than this (by combining the task effort data with the number of responses to the task), this analysis is conservative. Additional reasons are given later in the discussion to show that this analysis is conservative.

Table 1

Pilot Study Means and Standard Deviations for Brainstorming Task by Treatment Group

	Continuing Group	Zero-History	Individual
$\bar{M}$	21.765	23.206	21.941
$\sigma n-1$	7.836	8.082	9.566

Table 2 shows the same data separated by task:

Table 2

Pilot Study Means and Standard Deviations for Brainstorming Task by Task Type

	Knife	String	Wood
$\bar{M}$	23.382	21.008	22.412
$\sigma n-1$	8.521	8.317	8.627

Dividing the benchmark difference-between-means by the standard deviation for potential uses for a knife obtained in the pilot study yields a measure of the effect size that is free of the original unit of measurement. This statistic is defined in the following manner by Cohen (1969, p. 18): $d = (\mu 1 - \mu 2)/\sigma$ where $\mu 1$ and $\mu 2$ are independent sample means and σ is the standard deviation common to both samples. This statistic is referred to by Cohen as the effect-size index. Once this measurement is calculated it is a simple matter to find the power associated with t-tests of different sample sizes in an accompanying power table. In this case, a two-tailed test determined power.

Using the benchmark difference of 9.30 and the standard deviation for the "knife" task from the pilot study yields the following effect size: 9.30/8.52 = 1.09. Cohen (1969) indicates a power in excess of 92% for a sample size of 33 and an α level of .01 for this effect size. Because three comparisons are to be made in the main analysis of this data, it is necessary to calculate a family α level. Bonferroni's t-statistic, used here, yields a conservative measure of the family α when used for non-independent tests (see Myers, 1979):

Family α = p(one or more sig. results, all Ho true)

$$= 1 - (1 - \alpha)^{k}$$

$$= 1 - (1 - .01)^{3}$$

$$= .0297 \text{ for the family of three comparisons.}$$

This corresponds to a family power of $.92^{3} = 77.87\%$ probability of finding three true differences of the effect size under scrutiny. However, the family α will be set at .05 in practice, making this a conservative estimate (along with the other factors just discussed). Although sample sizes actually obtained were slightly smaller than 33, this analysis should be conservative enough to offset this deficit.

Reporting of Type I Error

Conventions followed by most researchers have dictated that Type I error probabilities should not be reported for non-significant results. When results are significant, these probabilities have typically been reported in increments such as "exceeding .05, .01, .005," and so on; they have not been reported precisely.

Computers have eliminated the difficulties formerly inherent in establishing precise Type I probabilities; if modern computers had been available when statistical tests currently in use were developed, then alternative

conventions might have emerged. The arguments for full and accurate disclosure of Type I error probabilities are persuasive.

Obviously, information is lost when Type I error probabilities are reported by approximation. Those advocating the traditional conventions cited above may argue that misleading conclusions will acrue if these conventions are ignored. According to this line of reasoning, misperceptions may occur due to the temptation to attribute undue importance to results approaching critical α benchmarks. Such a position may be tenable in regard to journals intended for statistically naive audiences, but this orientation appears to be unnecessary, and even harmful, in connection with material aimed at professionals who are familiar with probability theory. These conventions deprive qualified professionals of information, while those who are inclined to impute more practical importance to an insignificant result (or to a significant result) than rational grounds permit are unlikely to be deterred by conventions intended to thwart such inclinations.

Let us look, then, at some specific uses to which precise Type I probabilities may be put. As mentioned above, the ease with which Type I probabilities can be obtained is one argument for full and precise discloser of

statistical results. Such reports may be particularly valuable in exploratory studies, such as this one, indicating areas where more statistical power may be called for. Moreover, the growing popularity of meta-analysis increases the usefulness of full disclosure of probabilities. Plainly, a Type I error probability of .25 may be of little importance in and of itself, and yet, in combination with results from other studies, such a result may yield useful information. Therefore, insignificant Type I probabilities are reported here as well as significant probabilities.

Statistical Tests

A MANOVA for fixed effects with the three treatments constituting the independent variable was the primary analysis. Dependent variables for this procedure were (1) number of uses for a knife (per individual), and (2) scores on the effort questionnare (per individual). Wilks' lambda was used to test the results for significance (see Tatsuoka, 1971, p. 84). The permanent group cohesion data were not included in this part of the analysis because no treatment differences were expected.

Next, the three treatments were to be considered as the independent variable in two one-way ANOVA designs with fixed effects. In one of these ANOVAs, brainstorming task scores were to be the dependent variable. In the other,

task effort would have been the dependent variable. These ANOVA's were not carried out due to insignificant results in the MANOVA just described, which meant that the ANOVA's would be non-significant, too, (see Hays, 1973, p. 457 for an outline of this analysis). F-tests, followed by Tukey's pairwise comparisons were to have been employed following these ANOVAs.

In a third ANOVA, scores on the effort questionnaire were used as a blocking factor. Again, this analysis considered brainstorming task scores to be the dependent variable. In this ANOVA, the treatments were considered as a fixed factor. This procedure was performed by splitting each treatment group at its median on the task effort measure (see Myers, 1978). This allowed a test for interaction effects between task effort and treatment type (as well as main effects for treatment and task effort).

If the MANOVA had been significant, then the results of the first ANOVA, on the brainstorming task, would have determined whether social loafing occurred in the zero-history group condition as predicted. This test also would have shown whether the effect occurred in the continuing group treatment. A second ANOVA would have been performed on the task involvement data, and would have shown whether those in the zero-history treatment

perceived themselves to be oriented toward the task to a lesser degree than those in the other two treatments, as predicted. The third ANOVA showed whether a main effect existed for task effort and whether an interaction between task effort and treatment was present.

To determine whether group cohesion is related to task performance, the group cohesion measure was correlated with results from the brainstorming task in the continuing group treatment using the Pearson-r. The other treatments were also tested this way. The Fisher's-z transformation was used to determine whether these correlations were significant. These tests indicated whether "high cohesion" people work harder in general, or work harder only in groups toward which they feel cohesive.

As indicated, the zero-history and continuing group treatments completed two additional items. These items measured the subjects' sense of responsibility toward their experimental group (this experiment used a pseudo-group in the zero-history treatment). These items were averaged and correlated with brainstorming task scores for each treatment individually, as well as with both treatments combined. Positive correlations were expected, because the continuing group treatment subjects were expected to turn in superior performances on the

brainstorming task due to their increased concern for their group's performance compared to those in the zero-history treatment.

Following the planned analysis a post-hoc investigation of the data is presented, in keeping with the exploratory nature of this study.

V. RESULTS

Outliers

The data array created from experimental results was systematically checked, datum by datum, against the original questionnaires. Three errors were detected and corrected. Upon initial examination of the raw "number of uses for a knife" data it appeared that three outliers were present. A Shapiro-Wilk W-statistic of normality was calculated for each treatment sample (Shapiro and Wilk, 1965). Individual and zero-history treatment samples were found to have a significant probability of being from non-normal populations when these outliers were included. Table 3 shows the means and standard deviations of these data on a treatment group by treatment group basis. Reexamination of the pilot study data, which contained 105 observations from 35 individuals (as opposed to 91 observations from 91 individuals in this study) indicated no such extreme scores were present in the pilot study results.

These outliers were approached in the following manner. Some individuals generated groups of uses which were highly similar. For example, many individual lists

Table 3

Means and Standard Deviations for
Brainstorming Task
with Outliers

Treatment	Mean	Standard Deviation
I[a] (n=31)	22.77	10.66
Z (n=30)	23.23	11.91
C (n=30)	25.40	7.86

[a] I = individual treatment; Z = zero-history group treatment; C = continuing group treatment.

of "uses for a knife" contained more than one food item, along with an indication that each one could be cut up with a knife.

To determine whether outliers could have resulted from individuals "going overboard" in this respect, without generating new uses that were different in any significant way from those they had already written, categories were created. These intuitively generated categories included "cutting food," "peeling food," "carving or cutting wood," "cutting rope, string, etc.," and so on. Those uses for a knife that could be subsumed under one category were then counted as one use. This process did not eliminate the outliers, nor did it change

which individuals were outlying. Accordingly, this approach was abandoned.

Instead, analyses were conducted with and without the outliers. A z-score with an absolute value greater than plus or minus 2.58 ($p<.01$) constituted the exclusionary rule. Z-scores were based on treatment means and standard deviations (as opposed to the mean and standard deviation of individual measures from all treatments combined). Three excessively large scores were rejected. Two were deleted from the zero-history group, whereas one was rejected from the individual treatment. Not one was rejected from the continuing group treatment. Performing the Shapiro-Wilk W-statistic on these revised data yielded no significant results for any of the treatment groups. The rationale for this procedure is explained in the discussion section below. Means and standard deviations for the brainstorming task when the outliers are not included are shown below in Table 4.

Statistical Results

Reliabilities. Reliabilities were established for the scales measuring group cohesion, task effort, and task responsibility using Cronbach's coefficient α (Cronbach, 1951). When these reliabilities are rounded to two places

Table 4

Means and Standard Deviations for Brainstorming Task by Treatment Group without Outliers

Treatment	Mean	Standard Deviation
I[a] (n=30)	21.58	8.42
Z (n=28)	20.93	8.34
C (n=30)	25.40	7.86

[a] I = individual treatment; Z = zero-history group treatment; C = continuing group treatment.

following the decimal, no differences were seen between reliabilities with and without outliers on any of the scales. Reliability for the cohesion scale equaled .83 (n=91 or 88), whereas the reliability of the task effort scale equaled .90 (n=91 or 88) and the task responsibility scale had a reliability of .87 (n=60 or 58).

Multivariate analysis. The planned MANOVA was conducted twice--with and without outliers. Recall that (1) brainstorming task scores and (2) task effort were the dependent variables for this procedure. When outliers were included, the procedure yielded a Wilks' lambda of 1.11 ($p<.36$, with 4 and 174 df) whereas this value changed to 2.25 ($p<.07$ with 4 and 168 df) when the outliers were excluded. Neither of these results is significant.

Hence, the separate ANOVA's for each of these dependent variables were not carried out.

Cohesion correlations. Table 5 presents the correlations between cohesion and "number of uses for a knife" for each treatment group. These statistics were predicted to be significant. As seen here, the only correlation that is significant (the correlation between cohesion and task performance for the individual treatment group) is significant only when the outliers are included. Results of this analysis were expected to be significant for the continuing group treatment only, and significant in a positive direction.

Analysis of task effort. Next, an ANOVA was performed with performance on the brainstorming task as the dependent variable. This analysis considered treatment as a factor and used task effort as a blocking variable (two blocks, each treatment divided at the treatment median of task effort scores). A significant main effect for task effort was found. No interaction effect occurred between task effort and treatment, and, of course, no main effect for treatment group was observed (which is consistent with previous findings). These results were obtained with and without outliers. The Pearson-r product moment correlations of task effort with

performance on the brainstorming task are shown in Table 6.

Table 5

Correlation of Cohesion with Brainstorming Task

Treatment	With Outliers	Without Outliers
I[a]	-.37 (p<.04) n=31	-.23 (p<.23) n=30
Z	-.14 (p<.47) n=30	-.06 (p<.77) n=28
C	-.09 (p<.62) n=30	-.09 (p<.62) n=30
I+Z+C	-.20 (p<.06) n=91	-.12 (p<.26) n=88

[a] I = individual treatment; Z = zero-history group treatment; C = continuing group treatment.

Task responsibility correlations. Correlations between the task responsibility measure and brainstorming task scores are shown in Table 7. Positive correlations were expected here. Because subjects in the individual treatment group were tested as individuals, and because this measure corresponds to responsibility felt for the experimental work group, no measure of this variable was taken in the individual treatment. An unexpected negative

correlation between task responsibility and brainstorming task scores was obtained for the zero-history treatment. Other correlations are not significant.

Table 6

Correlation of Task Effort with Brainstorming Task

Treatment	With Outliers	Without Outliers
I[a]	.30 ($p<.11$) n=31	.27 ($p<.15$) n=30
Z	.36 ($p<.05$) n=30	.32 ($p<.11$) n=28
C	.54 ($p<.002$) n=30	.54 ($p<.002$) n=30
I+Z+C	.37 ($p<.0004$) n=91	.35 ($p<.0007$) n=88

[a] I = individual treatment; Z = zero-history group treatment; C = continuing group treatment.

Table 7

Correlation of Task Responsibility with Brainstorming Task

Treatment	With Outliers	Without Outliers
Z[a]	-.45 (p<.02) n=30	-.45 (p<.02) n=28
C	.22 (p<.25) n=30	.22 (p<.25) n=30
Z+C	-.18 (p<.18) n=60	-.08 (p<.57) n=58

[a] Z = zero-history group treatment; C = continuing group treatment.

VI. DISCUSSION AND POST-HOC ANALYSIS

Introduction

Possible reasons for the absence of social loafing are discussed first. In this section, the rationale for analyzing results without outliers is given. Next, after possible reasons for the absence of social loafing are discussed, some closing observations are made. In particular, reasons are considered for the negative correlations unexpectedly encountered for the brainstorming task performance with the (1) cohesion and (2) task responsibility measures.

Possible Causes for the Absence of Social Loafing

No difference on the brainstorming task was observed between the individual and zero-history group treatments. Perhaps this is the most curious finding of this project. The Ringelmann effect, (social loafiing), did not occur. This surprising outcome duplicates results of the pilot study. Recall that the pilot study utilized a repeated measures design with each of the three treatments occurring sequentially. An absence of social loafing in this case was hypothesized to have occurred because subjects compared their performance from one session to another.

This study raises the possibility that more complicated or (at present) obscure contingencies

predicate whether or not social loafing will occur. Two possibilities will be considered here. The first concerns the properties of the sample employed; the second focuses upon the experimental manipulations. Thus, the first possibility to be considered is that the treatment groups may have been non-equivalent on important dimensions. Random assignment increases the probability of equivalency, but it does not guarantee it. The second possibility considered is that social loafing may be an artifact of particular experimental designs. Designs from related published studies will be compared to the design this study used in exploring this possibility.

Outliers and non-equivalent groups. Note that the 91 individuals in the sample were not randomly chosen from a larger population. It is true that subjects in the total sample were assigned at random to the three treatments; however, creating the permanent groups to which all participants belonged prior to data collection required the use of existing classes. Thus, in effect, subjects selected themselves to participate in this experiment.

Differences on the dependent variables were found between the three classes from which participants were drawn. Possible sources of these differences are examined below. The possibility that self-selection to classes

(and hence to the experiment) produced a non-normal grand sample is discussed, along with the implications such an event may have had on experimental outcomes.

Remember that three sections of the same course participated in this experiment. A one-way ANOVA with class section as a factor and brainstorming task scores as the dependent variable indicated a significant difference between the three class sections ($p<.0001$, with or without outliers included in the analysis). One of the classes participating met in the evening, for three hours, once a week. The other two met two or three times a week during ordinary class hours. More uses for a knife were produced by the evening class than the other two classes, and this of course is responsible for the high significance level reached by the F-ratio in the analysis just described.

Do students who give up an evening of their time differ from other students? It seems probable that older students who have jobs and families, and possibly more direction in their lives, would be over-represented in an evening class. Perhaps students who were more serious about their academic pursuits would give up an evening more willingly, as well. Of course, students who simply could not find an alternative class section would be represented in the evening class, too.

Thus, it is possible that the evening class drew a handful of atypical students, who, because they were few in number, could have been placed disproportionately in one treatment group merely by chance. Previously, it was mentioned that the treatment groups were chosen after the grand sample had been stratified by class section (to distribute any effects class section could have had on the final outcome). What is being suggested here is the possible existence of a small group of atypical students in the evening class which should have been stratified as well (or deleted from the analysis). In short, a possible contamination of the experiment has just been described, as well as an explanation for the significant Shapiro-Wilk W-statistics described in the results section.

To explore this possibility further, the grand sample was tested using the Shapiro-Wilk W-statistic on brainstorming task scores. First, it was tested with the outliers (as defined in the analysis section). This test showed the unabridged grand sample to have a probability of $p<.01$ of being drawn from a normal population. Next, when outliers were excluded, this figure increased to $p>.12$. A more detailed breakdown showed that both the morning and afternoon classes, when treated as samples, exhibited significant W-statistics when outliers were included; however, no class exhibited a significant

W-statistic when the outliers were excluded. These results support the argument that class effects may have been deleterious in this experiment.

Another point is worth mentioning in regard to the issue of class-section effects. Another significant difference was found. A one-way ANOVA with class section as a fixed factor and task effort as the dependent variable yielded a significant F-ratio ($p<.05$, with or without outliers). Given the class section differences on brainstorming task scores, this is not surprising, because these two measures were shown to be correlated in the results section. In future experiments similar to this one it may behoove the experimenter to consider class section effects more carefully than was done in this case.

To determine whether the significant class section effects still existed after the outliers were deleted, an additional analysis was performed. An ANOVA was conducted with class section as a factor and with brainstorming task scores as the dependent variable with the outliers excluded. A significant difference still existed between the evening class mean and the means of other class sections. These means are shown in Appendix C. To reduce class section effects as much as possible all analyses were repeated using brainstorming z-scores. These z-scores were calculated individually for each class

section from class section means and standard deviations. The results of the MANOVA were virtually identical to the results obtained with raw scores when outliers were excluded. The two-way ANOVA described earlier was also repeated on these z-scores and results were virtually identical to those obtained on raw brainstorming task scores when outliers were excluded. Correlations based on these z-scores are shown in Appendix C, and are reported by class section and by treatment group.

Returning to the issue of outliers on the brainstorming task, removing the outliers is justified according to Huber (1977, p. 3) because lengthening the tails of a distribution explodes the variability of the standard deviation and ". . . any conscientious statistician will do something about [outliers] before calculating the mean square error." Huber goes on to recommend the Shapiro-Wilk W-statistic as a criterion for rejecting outliers. Barnett and Lewis (1978) also recommend using the W-statistic on samples coming from a population suspected of having a long, flat tail (as in this case).

Artifacts of experimental design. In reexamining previous social-loafing experiments, several facts were discovered regarding the handling of participants in the group treatments (which were, of course, zero-history

group treatments in all cases). In the Ingham et al. (1974) study those in the zero-history group treatments were truly submerged in a group inasmuch as no one (as far as participants could tell) could see, hear or otherwise perceive their individual performance on the task, even in the most cursory way.

In later experiments, the task is such that this would not have been true. In those experiments conducted by Latane and his associates, the tasks were such that subjects could have gotten an idea of how they were doing in relation to others, if they could have seen or heard the other participants. However, with one exception, the experiments were conducted in a way that ameliorated the affects of even the most casual interaction between subjects in the group treatment. Thus, subjects were tested individually using pseudo-groups (Petty et al., 1980) or they were physically isolated from one another (Williams et al., 1981; Harkins and Petty, 1982), or groups, instead of individuals, were used as the level of analysis (Latané et al., 1979).

Because all the researchers just cited were cooperating, they may have learned in pilot research that even subtle interaction between subjects eliminated the social-loafing effect. Clearly, these researchers took pains to avoid even limited interaction between subjects,

except in the study where groups were used as the level of analysis (which may have been the point at which they discovered a problem). Notice that the study which used this style of analysis, by Latané et al., was published first in this series.

Perhaps interplay between individuals caused the absence of social loafing in the present experiment, since subjects were seated side by side in each experimental session. Unfortunately it was necessary to conduct sessions this way because it was not feasible to test participants in smaller groups. Too few groups existed for groups to be used as the level of analysis. Furthermore, the use of pseudo-groups in the zero-history group treatment may have weakened the social-loafing effect; and the presence of the zero-history participants' permanent groups was an unfortunate concession to feasibility as well. In summary, several mechanisms which have just been described may have interfered with the study of social loafing in this experiment.

An additional point may be made in connection with the posssible contamination of these findings through unintended group influences. Perhaps social loafing occurs only when an individual is "in a group," and yet isolated from others or other groups to an extent that is unusual in ordinary practical affairs. Further

investigations regarding varying degrees of identifiability may illuminate this possibility. One way to accomplish this might be to vary the physical distance between participants so that those who were close to other participants could get a more accurate "gist" of how others were doing compared to those seated more widely apart. No doubt, some participants tend to glance at other students' work.

Closing Observations

A prior relationship between task effort and performance on a brainstorming task has been established (Petty et al., 1977; Petty et al., 1980). The positive relationship between task effort and performance on the brainstorming task fulfilled expectations. There is certainly no surprise in observing that those who believe they are working hard actually are. This correlation is interesting nonetheless, because one experiment found no such relationship (Harkins et al., 1980). In the Harkins study, however, the experiment involved clapping with ear plugs on, which restricted self-feedback. Obviously, self performance could be more accurately gauged in the task studied here.

Concerning the relationship between cohesion and task performance, it should be noted that those correlations that approached or attained significance when

outliers were included moved away from significance when outliers were deleted. Because of this finding, and because the use of a family α would lead to a conclusion of no significance in this sequence of statistics, it is appropriate to say that no relationship exists between these variables in this experiment. Tests for significant differences between these treatment group correlations would have been carried out if any of the correlations had been significantly different from zero with outliers excluded. In support of the approach taken here, Chang (1976, pp. 70-71) concluded:

> The significance test investigated in the present study [for diffference between two Pearson-r product-moment correlation coefficients] provides acceptable control over Type I error when sampling from normal bivariate populations. The significance test is less effective in controlling Type I error with increasing non-normality and magnitude of ρ values in the parent populations.

The absence of a relationship between cohesion and task performance in the continuing group treatment is surprising because Hoogstraten and Vorst (1978) found a positive correlation between performance and cohesion measured by a questionnaire nearly identical to the one used here. It is possible that the group influences described above masked the differences that would have created a positive correlation.

Finally, the significant negative correlation in the zero-history treatment group between responsibility felt toward the experimental group and performance on the brainstorming task is unexpected. One possible accounting for this is that those who felt little responsibility toward their zero-history group had a hightened sense of their own individuality. They may have been thinking more in terms of their own performance than the performance of their group (which was a pseudo-group in this treatment). This may not have occurred in the continuing group treatment because the participants believed their scores would be combined with group participants who were present, immediate and salient.

On the other hand, the presence of familiar individuals in the zero-history treatment group may have caused the psuedo-group manipulation to seem unreal. In a sense, then, this outcome may be an artifact associated with a lack of success in manipulating zero-history group members to conceptually imbed themselves in a pseudo-group. That is, many of the subjects in the zero-history treatment may not have felt as though they were part of a group at all.

In conclusion, the results of this experiment indicate that social loafing may be a phenomenon that is less robust than previously published experiments would

lead one to believe. Additional research aimed at more clearly understanding variables already known to influence social loafing may be useful before investigating the role that continuing group membership may play. Equivalent-group research on continuing groups is problematical, and it may make sense to wait until more is known about social loafing before proceeding to investigate whether it occurs in continuing groups. One manipulation has been described that is intended to determine whether a linear effect exists between an individual's performance and his or her degree of identifiability to other group members. In future experiments it would be helpful to create actual zero-history groups instead of pseudo-groups, and thus remove a possible source of experimental contamination.

VII. SUMMARY

For many years the social-loafing effect has been unknown or ignored by the majority of behavioral scientists, despite its importance for motivational strategies within organizations. Recent studies have shown the effect to be robust in groups of unacquainted people. The mediating effects of long-term group membership, however, have not been studied. A plausible rationale that such effects may exist has been established. Results of this study were uncertain because Seashore (1954) and others have found cohesion to have varying effects on group performance.

This research project gave credence to the observation that the motivational impact of groups upon individual performance is complex. Continuing groups did not perform significantly better than individuals or zero-history groups. Additional research into the impact of propinquity upon the tasks studied here may be illuminating. Finally, this research has highlighted the importance of considering class section idiosyncracies in regard to experimental "noise." Differences between the motivational impact of new groups versus the motivational impact of established groups are difficult to study experimentally. Because so much of our experimentally derived knowledge of groups is coming from zero-history

groups, the importance of understanding any differences that may exist between zero-history and continuing groups is growing. Hopefully, this will lead researchers to overcome methodological stumbling blocks hindering this endeavor.

REFERENCES

REFERENCES

Ball, J. R., & Carron, A. V. The influence of team cohesion and participation motivation upon performance success in inter-collegiate ice hockey. Canadian Journal of Applied Sport Sciences, 1976, 1, 271-275.

Barnett, V., & Lewis, T. Outliers in Statistical Data. New York: John Wiley & Sons, 1978.

Chang, C. (1977). Effects of non-normality and sample size on effectiveness of significance test of difference between two independent correlation coefficients (Doctoral dissertation, University of Iowa, 1976). Dissertation Abstracts International, 37, 6217B.

Cohen, J. Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press, 1969.

Cottrell, N. B. Social facilitation. In C. G. McClintock (Ed.), Experimental Social Psychology. New York: Holt, Rinehart & Winston, 1972.

Cottrell, N. B., Rittle, R. H., & Wack, D. L. The presence of an audience and list type (competitional or noncompetitional) as joint determinants of performance in paired-associates learning. Journal of Personality, 1967, 35, 425-434.

Cronbach, L. J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

Dashiell, J. F. Experimental studies of the influence of social situations on the behavior of individual human adults. In C. Murchison (Ed.), A Handbook of Social Psychology. Worchester, Mass.: Clark University Press, 1935.

Geen, R. G. & Gange, J. J. Drive theory of social facilitation: Twelve years of theory and research. Psychological Bulletin, 1977, 84, 1267-1288.

Harkins, S. G., Latané, B., & Williams, K. Social loafing: Allocating effort or taking it easy? Journal of Experimental Social Psychology, 1980, 16, 457-465.

Harkins, S. G., & Petty, R. E. Effects of task difficulty and task uniqueness on social loafing. Journal of Personality and Social Psychology, 1982, 43, 1214-1229.

Hays, W. L. Statistics for the Social Sciences. New York: Holt, Rinehart and Winston, Inc., 1973.

Hoogstraten, J., & Vorst, H. C. M. Group cohesion, task performance, and the experimenter expectancy effect. Human Relations, 1978, 31, 939-956.

Huber, P. J. Robust Statistical Procedures. Philadelphia: Society for Industrial and Applied Mathematics, 1977.

Ingham, A. G., Levinger, G., Graves, J., & Peckham, V. The Ringelmann effect: Studies of group size and group performance. Journal of Experimental Social Psychology, 1974, 10, 371-384.

Kerr, N. L. Motivation Losses in small groups: A social dilemma analysis. Journal of Personality and Social Psychology, 1983, 45, 819-828.

Kerr, N. L., & Bruun, S. E. Ringelmann revisited: Alternative explanations for the social loafing effect. Personality and Social Psychology Bulletin, 1981, 7, 224-231.

Kerr, N. L., & Bruun, S. E. The dispensability of member effort and group motivation losses: Free-rider effects. Journal of Personality and Social Psychology, 1983, 44, 78-94.

Latané, B., Williams, K., & Harkins, S. Many hands make light the work: The causes and consequences of social loafing. Journal of Personality and Social Psychology, 1979, 37, 822-832.

Lott, A. J., & Lott, B. E. Group cohesiveness as interpersonal attraction: Antecedents of liking. In Hinton, B. L., and Reitz, H. J. (Eds.), Groups and organizations. Belmont, California: Wadsworth Publishing Company, Inc., 1971.

Moede, W. Die Richtlinien der leistungs-psychologie. Industrielle Psychotechnik, 1927, 4, 193-207.

Moran, G. The Cohesion Problem. Sittard, The Netherlands: Albert Drukkerijen, 1964.

Myers, J. L. Fundamentals of Experimental Design. Boston: Allyn and Bacon, Inc., 1979.

Norris, D. R., & Niebuhr, R. E. Group variables and gaming success. Simulation and Games, 1980, 11, 301-312.

Petty, R. E., Harkins, S. G., & Williams, K. D. The effects of group diffusion of cognitive effort on attitudes: An information-processing view. Journal of Personality and Social Psychology, 1980, 38, 81-92.

Petty, R. E., Harkins, S. G., Williams, K. D., & Latané, B. The effects of group size on cognitive effort and evaluation. Personality and Social Psychology Bulletin, 1977, 3, 579-582.

Scofield, R. W. Task productivity of groups of friends and non-friends. Psychological Reports, 1960, 6, 459-460.

Shapiro, S. S, and Wilk, M. B. An analysis of variance test for normality (complete samples), Biometrika, 1965, 52, 591-611.

Steiner, I. D. Group Process and Productivity. New York: Academic Press, 1972.

Sypher, H. E. Task productivity of continuing and zero-history groups in two communication networks. Perception and Motor Skills, 1977, 44, 1333-1334.

Tatsuoka, M. M. Multivariate Analysis. New York: John Wiley & Sons, Inc., 1971.

Tziner, A., & Vardi, Y. Ability as a moderator between cohesiveness and tank crews performance. Journal of Occupational Behavior, 1983, 4, 137-143.

Williams, K., Harkins, S., and Latané, B. Identifiability as a deterrent to social loafing: Two cheering experiments. Journal of Personality and Social Psychology, 1981, 40, 303-311.

Zajonc, R. B. Social facilitation. Science, 1965, 149, 269-274.

Zajonc, R. B. Social Psychology: an Experimental Approach. Belmont, California: Wadsworth Publishing Co., Inc., 1966.

APPENDIXES

APPENDIX A

QUESTIONNAIRES

Questionnaire 1

Please answer the following questions concerning the brainstorming task you just completed.

1. To what extent were you trying hard to generate as many uses for a knife as you could?

1	2	3	4	5	6	7	8	9	10	11
Trying very hard										Not trying hard

2. How much effort did you put into generating as many uses for a knife as you could?

1	2	3	4	5	6	7	8	9	10	11
A great deal of effort										Very little effort

3. How involved were you in this task?

1	2	3	4	5	6	7	8	9	10	11
Very involved										Not involved

[The zero-history and continuing group treatments answered the next two questions, which were designed to measure the importance and responsibility individuals felt toward their experimental group. Those in the individual condition did not see these two questions.]

4. How important was it to you for your group to do well?

1	2	3	4	5	6	7	8	9	10	11
Very important										Not important

5. How much responsibility did you feel for your group's performance?

1	2	3	4	5	6	7	8	9	10	11
A great deal of responsibility										Very little responsibility

Questionnaire 2

Please respond to the following questions concerning the group with whom you have been working for the past academic quarter.

1. How attractive is your group to you?

   (1) Very Attractive
   (2) Somewhat Attractive
   (3) Neither Attractive Nor Unattractive
   (4) Somewhat Attractive
   (5) Very Unattractive

2. Would you like to stay a member of this group?

   (1) Yes, Certainly
   (2) Probably
   (3) Not Sure
   (4) Probably Not
   (5) Certainly Not

3. If it were up to you how often would this group meet?

   (5) Quite Infrequently
   (4) Infrequently
   (3) Occasionally
   (2) Frequently
   (1) Very Frequently

4. If another course or work group were to begin, would you choose the same people for that group as are in your present group if you could do the choosing?

   (5) Certainly Not
   (4) Probably Not
   (3) Not Sure
   (2) Probably
   (1) Yes, Certainly

5. Would you advise a friend to join your group if it were possible?

   (1) Yes, Certainly
   (2) Probably
   (3) Not Sure
   (4) Probably Not
   (5) Certainly Not

6. Do the people in your group show interest in the viewpoint of others in the group?

   (1) Yes, Certainly
   (2) Usually
   (3) Not Sure
   (4) Only Occasionally
   (5) Almost Never

7. In your opinion, do the people in group have the same conception of the purpose of group activities?
   (1) Yes, Certainly
   (2) Probably
   (3) Not Sure
   (4) Probably Not
   (5) Certainly Not

8. Would most of the people in your work group be willing to accept additional work in the group's behalf if this would help the group?

   (5) Certainly Not
   (4) Probably Not
   (3) Not Sure
   (2) Probably
   (1) Yes, Certainly

9. If people from other groups were to criticize or belittle your group, would most of the people from your group contradict them and defend your group?

   (1) Yes, Certainly
   (2) Probably
   (3) Not Sure
   (4) Probably Not
   (5) Certainly Not

10. Do most of the people from the group have the same opinion of the purpose of your group?

    (1) Yes, Certainly
    (2) Probably
    (3) Not Sure
    (4) Probably Not
    (5) Certainly Not

11. If your group should make up several rules on its own authority, would most people gladly observe them?

    (5) Certainly Not
    (4) Probably Not
    (3) Not Sure
    (2) Probably
    (1) Yes, Certainly

12. If something has to be decided in your group, do most of the people weigh the opinions of others in the group?

    (1) Yes, Certainly
    (2) Probably
    (3) Not Sure
    (4) Probably Not
    (5) Certainly Not

13. Most people in my group are aware that they belong to a very congenial group.

    (1) Yes, Certainly
    (2) Probably
    (3) Not Sure
    (4) Probably Not
    (5) Certainly Not

14. Is it more important for most of the people in the group to achieve a better discussion than to "score points" on their own?

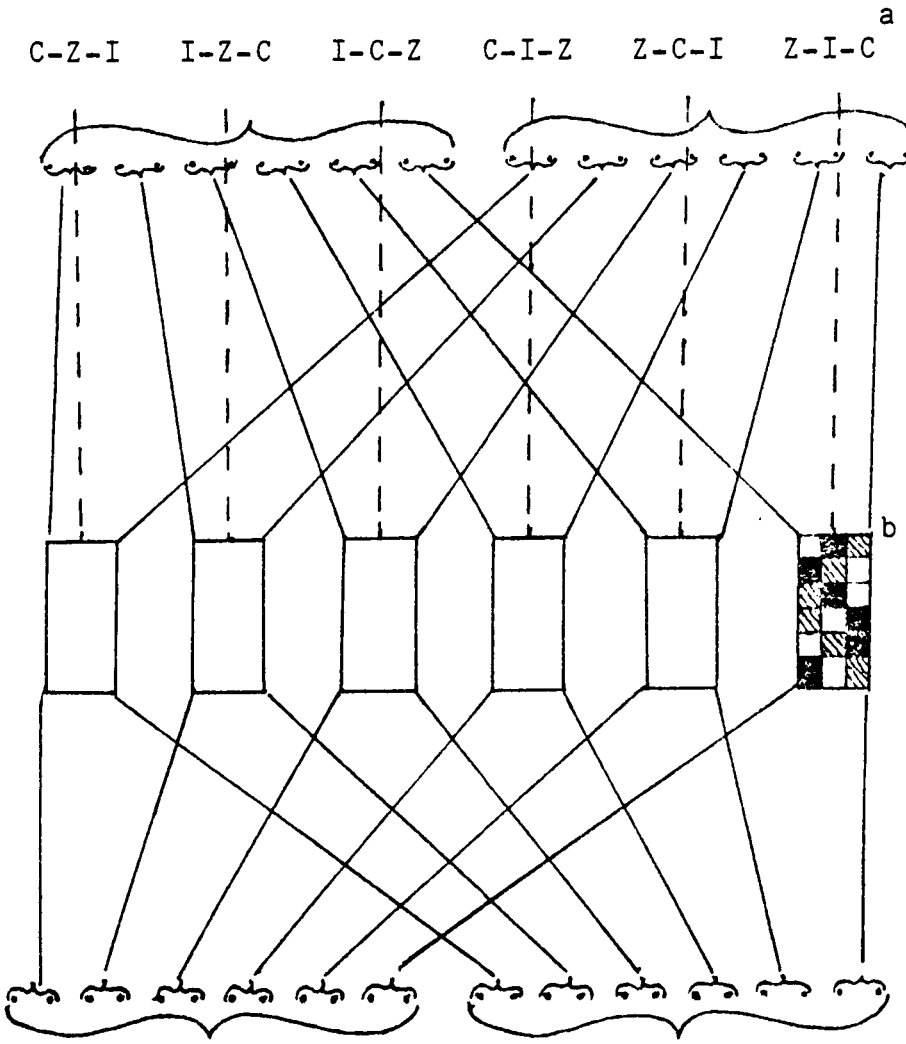
    (1) Yes, Certainly
    (2) Probably
    (3) Not Sure
    (4) Probably Not
    (5) Certainly Not

15. My group is successful in handling the tasks assigned to it.

(5) Certainly Not
(4) Probably Not
(3) Not Sure
(2) Probably
(1) Yes, Certainly

APPENDIX B

FIGURES OF EXPERIMENTAL DESIGNS



[a] C=continuing group treatment; Z=zero-history group treatment; I=individual treatment. These six large rectangular blocks represent each possible order of administration for these three treatments.

[b] Each row of three squares represents one of the six possible orders of administration of the parallel tasks.

Figure 1. Diagram of the ideal design.

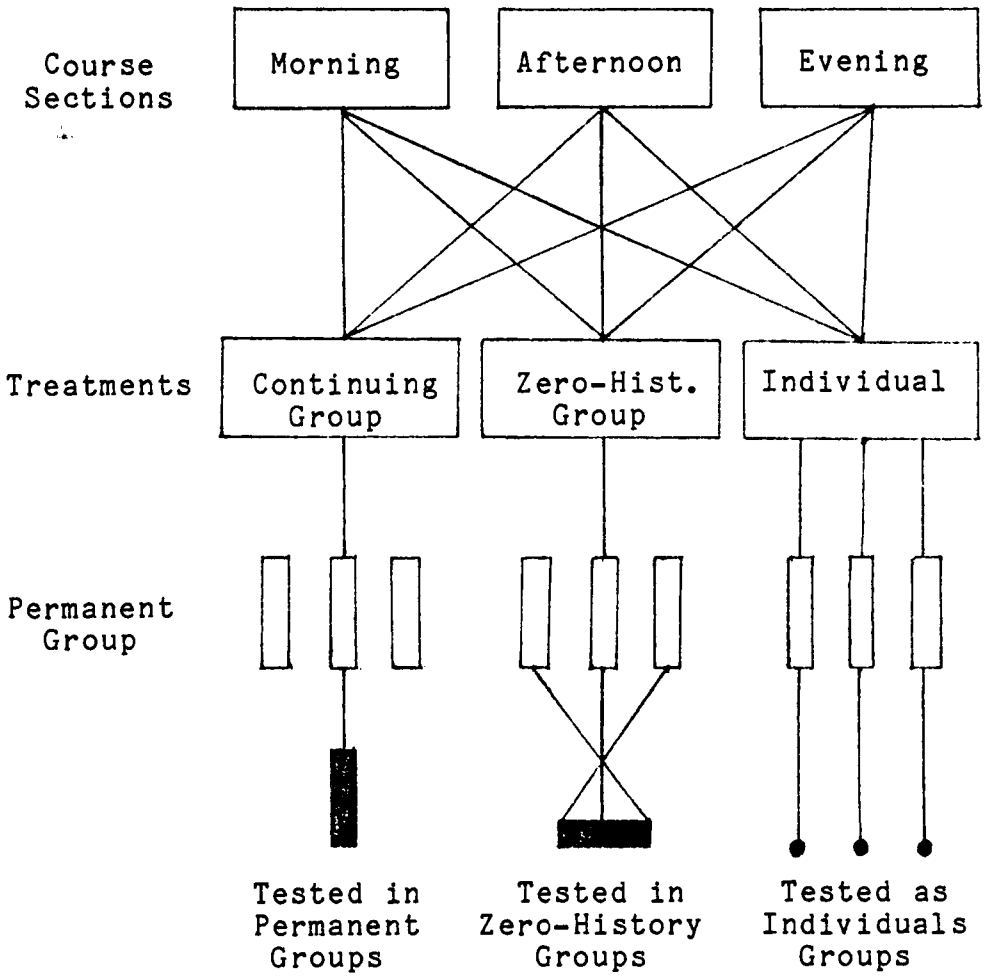


Figure 2. Diagram of the first attempted design.

APPENDIX C

SUPPLEMENTARY STATISTICS

Table 8

Class Section Means and Standard Deviations for Brainstorming Task without Outliers

Class Section		Mean	Standard Deviation
Morning	(n=28)	21.11	7.87
Afternoon	(n=30)	19.03	6.48
Evening	(n=30)	27.77	8.18

Table 9

Treatment Group Means and Standard Deviations for Task Effort without Outliers

Treatment		Mean	Standard Deviation
I[a]	(n=30)	3.03	1.32
Z	(n=28)	3.55	1.70
C	(n=30)	3.51	1.71
I+Z+C	(n=88)	3.56	1.58

[a] I = individual treatment; Z = zero-history group treatment; C = continuing group treatment.

Table 10

Class Section Means and Standard Deviations for Task Effort without Outliers

Class Section		Mean	Standard Deviation
Morning	(n=28)	3.35	1.62
Afternoon	(n=30)	3.87	1.62
Evening	(n=30)	2.87	1.39

Table 11

Treatment Group Means and Standard Deviations for Cohesion without Outliers

Treatment		Mean	Standard Deviation
I[a]	(n=30)	1.92	.44
Z	(n=28)	1.90	.35
C	(n=30)	1.88	.45
I+Z+C	(n=88)	1.90	.41

[a] I = individual treatment; Z = zero-history group teatment; C = continuing group treatment.

Table 12

Class Section Means and Standard Deviations for Cohesion without Outliers

Class Section		Mean	Standard Deviation
Morning	(n=28)	1.99	.31
Afternoon	(n=30)	1.88	.39
Evening	(n=30)	1.83	.50

Table 13

Treatment Group Means and Standard Deviations for Task Responsibility without Outliers

Treatment		Mean	Standard Deviation
Z[a]	(n=28)	2.80	2.13
C	(n=30)	2.20	1.32
Z+C	(n=58)	2.49	1.77

[a] Z=zero-history group; C=continuing-Group Treatment.

Table 14

Class Section Means and Standard Deviations for Task Responsibility without Outliers

Class Section		Mean	Standard Deviation
Morning	(n=18)	2.25	.93
Afternoon	(n=20)	2.75	1.46
Evening	(n=20)	2.45	2.52

Table 15

Grand Correlation Matrix without Outliers

	Task Effort	Task Respons.	Group Cohesion	# of Uses for Knife
Task Effort	1.00	.35	-.13	.35
p<	.00	.008	.24	.0008
n=	88	58	88	88
Task Respons.		1.00	.17	-.08
p<		.00	.21	.56
n=		58	58	58
Group Cohesion			1.00	-.12
p<			.00	.26
n=			88	88
# uses for a Knife				1.00
p<				.00
n=				88

Table 16

Correlation Matrix for Continuing Group Treatment

	Task Effort	Task Respons.	Group Cohesion	# of Uses for Knife
Task Effort[a]	1.00	.41	-.03	.54
p<	.00	.03	.90	.002
Task Respons.		1.00	.31	.22
p<		.00	.10	.25
Group Cohesion			1.00	-.09
p<			.00	.63
# of Uses for Knife				1.00
p<				.00

[a] n=30 for all correlations.

Table 17

Correlation Matrix for Individual Treatment without Outliers

	Task Effort	Group Cohesion	# of Uses for Knife
Task Effort[a]	1.00	.10	.27
p<	.00	.62	.15
Group Cohesion		1.00	-.23
p<		.00	.23
# of Uses for Knife			1.00
p<			.00

[a] n=30 for all correlations.

Table 18

Correlation Matrix for the Zero-History Group Treatment without Outliers

	Task Effort	Task Respons.	Group Cohesion	# of Uses for Knife
Task Effort[a]	1.00	.28	-.30	.32
p<	.00	.15	.12	.10
Task Respons.		1.00	-.02	-.44
p<		.00	.94	.02
Group Cohesion			1.00	-.06
p<			.00	.77
# of Uses for Knife				1.00
p<				.00

[a] n=28 for all correlations.

Table 19

Correlations Between Brainstorming Task Z-Scores and Other Variables by Treatment Group

	Continuing Group	Zero-History Group	Individual
Task Effort[a]	.42	.32	.10
p<	.03	.10	.63
n=	30	28	30
Task Respons.	-.17	-.57	-
p<	.37	.002	
n=	30	28	
Group Cohesion	-.14	-.16	-.30
p<	.47	.43	.11
n=	30	28	30

[a] The brainstorming task z-scores in this analysis are based on class section means and standard deviations without outliers.

Table 20

Correlations Between Brainstorming Task Z-Scores and Other Variables by Class section

	Morning	Afternoon	Evening
Task Effort[a]	.41	.22	.21
p<	.03	.26	.27
n=	28	30	30
Task Respons.	-.08	-.46	-.36
p<	.74	.05	.12
n=	18	20	20
Group Cohesion	-.29	-.23	-.10
p<	.14	.22	.60
n=	28	30	30

[a] The brainstorming task z-scores in this analysis are based on class section means and standard deviations without outliers.

VITA

Carleton B. Southworth was born in Winchester, Massachusetts, on October 4, 1955. He moved to Amherst, Massachusetts, shortly before his sixth birthday. He was graduated from Amherst Regional High School in June, 1973. He entered Vassar College in the fall of 1973 and was graduated with a Bachelor of Arts degree in 1977.

Between college and high school, he was employed as a substitute high school teacher in Amherst, and was also employed designing and building residential housing.

He was awarded the Master of Science degree in Industrial and Organizational Psychology in June, 1984, from The Graduate School of the University of Tennessee, Knoxville.