



8-1977

A Behavioral Research-Supervising System for Masters Level Research

Michael J. Dillon
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses



Part of the Psychology Commons

Recommended Citation

Dillon, Michael J., "A Behavioral Research-Supervising System for Masters Level Research" (1977).
Masters Theses. 2241.

https://scholarworks.wmich.edu/masters_theses/2241

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



A BEHAVIORAL RESEARCH-SUPERVISING SYSTEM
FOR MASTERS LEVEL RESEARCH

by

Michael J. Dillon

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Master of Arts

Western Michigan University
Kalamazoo, Michigan
August 1977

ACKNOWLEDGEMENTS

I have truly "stood on the shoulders of giants" for the supervision I have received during the preparation of this research. Special appreciation goes to Dr. Richard W. Malott, who has provided the opportunity, the incentive, and the guidance for the entire Behavioral Research-Supervising System. The world needs more system builders like him. To Harry M. Kent, my doctoral supervisor, whose consistent and reasoned suggestions are evident throughout this system. To Dr. R. Wayne Fuqua whose practical and theoretical analyses were most helpful. And to Dr. Paul T. Mountjoy for his recommendations. Lavish praise and hearty approbation go to Ms. Sue Dickerman for her constant, helpful, and cheerful assistance during the many stages of preparation. To these people I am indebted. For the accuracy of the contents herein, however, I assume sole responsibility.

Michael J. Dillon

INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.
2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again – beginning below the first row and continuing on until complete.
4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.
5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International

300 North Zeeb Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR

MASTERS THESIS

13-10,462

DILLON, Michael Joseph, 1952-
A BEHAVIORAL RESEARCH-SUPERVISING
SYSTEM FOR MASTERS LEVEL RESEARCH.

Western Michigan University, M.A.,
1977
Psychology, general

Xerox University Microfilms, Ann Arbor, Michigan 48106

© Copyright by
Michael Joseph Dillon
1977

TABLE OF CONTENTS

CHAPTER		PAGE
I	STUDY I	1
	Introduction	1
	Method	4
	Subjects	4
	Doctoral supervisors	4
	Setting and weekly program	5
	Experimental design	9
	Experimental procedures	10
	General procedures	11
	Dependent variables	17
	Observation and reliability	23
	Results	25
	Reliability	25
	Dependent variables	26
	Discussion	42
II	STUDY II	48
	Introduction	48
	Method	51
	Subjects	51
	Doctoral supervisors	51
	Setting and weekly programs	51
	Experimental procedures	52
	General procedures	53
	Dependent variables	54
	Observation and reliability	59
	Results	60
	Reliability	60
	Dependent variables	60
	Discussion	78

TABLE OF CONTENTS
(Continued)

CHAPTER		PAGE
III	STUDY III	82
	Introduction	82
	Method	83
	Subjects and setting	83
	General procedures	84
	Results and Discussion	86
	Advising system	86
	Research tasks	86
	Conclusions	96
IV	GENERAL CONCLUSIONS	97
V	RECOMMENDATIONS	99
VI	REFERENCES	107
VII	REFERENCE NOTES	109
VIII	APPENDICES	110
	Appendix A	111
	Appendix B	112
	Appendix C	113
	Appendix D	115
	Appendix E	116
	Appendix F	117

TABLE OF FIGURES

FIGURE		PAGE
1	Percent of required points completed and not completed for the total group	28
2	Frequency of students and total number of negative points earned in the baseline condition relative to total earned in the experimental condition	30
3	Percent of required points completed and not completed for review article, log, self-reported hours, group meeting attendance	33
4	Percent of required points completed and not completed for writing, self- editing, data presentation, individual meeting attendance	36
5	Percent of required points completed and not completed for Group G, Group W, Group I	38
6	Percent of required points completed and not completed for the total group	62
7	Frequency of students and total number of negative points earned in the baseline condition relative to total earned in the experimental condition	65
8	Percent of required points completed and not completed for review article, log, writing, self-editing	67
9	Percent of required points completed and not completed for individual meeting attendance, data presentation, self-reported hours	69
10	Percent of required points completed and not completed for Group G and Group I	72
11	Percent of students and mean supervision time per week from chief faculty sponsor in individual and group format	88

TABLE OF FIGURES
(Continued)

FIGURE		PAGE
12	Percent of students and time in months from entrance to graduate program to start of: literature review activities, pilot study or formal implementation, formal thesis write-up	92

TABLE OF TABLES

TABLE		PAGE
I	Weekly thesis program	8
II	Occurrence of specific events in Study I	14
III	Three groups (representing three levels of thesis progress) in Study I and II and their required research tasks, and tasks that could earn extra positive points	22
IV	Occurrence of specific events in Study II	56
V	Record of tasks that earned extra posi- tive points and number of students contributing for each phase of Study I and II	75
VI	An analysis of students from the Standard Supervising System who have not received any research supervision	90
VII	An analysis of students from the Behavioral Research-Supervising System and Standard Supervising System who have not started the research tasks	95

STUDY I

Introduction

Master's thesis research is a long term endeavor with probably three levels of progress before the final thesis document is produced. Research activities extend over many months and in this time the master's student will review the published literature and generate a research design, implement a pilot study or the formal thesis design and collect data, and write and edit thesis drafts in order to produce the final document. Steady, maintained research behavior across these levels is difficult because there are usually no deadlines for completion of these tasks. In many instances, the initial start of research activities is delayed, and then progress is interrupted with long pauses, because there are no rewarding consequences for completion and no aversive consequences for non-completion. A Behavioral Research-Supervising (BRS) System was designed so that graduate students maintained steady rates of completion of research tasks at the three levels of thesis progress.

A system, as distinguished from a non-system, has these characteristics (Knezevich, 1973, p. 4): the setting of goals and objectives; specification of activities and the clustering of these activities related to the goals and objectives; empirically based measurement of outcomes; evaluation of the outcomes; recycling through the earlier components to modify the system.

A behavioral system, as distinguished from a non-behavioral system,

has these characteristics (Malott, 1974, p. 325): reliance on functional relationships to explain behavior; specification of the behavior, consequences and contingencies; observation of the behavior; consequence of the behavior.

There are four distinctive features of the Behavioral Research-Supervising System:

1) Research tasks at the three levels of thesis progress were clearly specified.

2) Deadlines for completion of the separate research tasks were set. The approach here is similar to the one recommended for pacing student work in Personalized System of Instruction courses (Keller, 1968). Under self-pacing, students usually do not complete the majority of course work (if at all) until the final weeks of the semester (Lloyd and Knutzen, 1969). Target dates for completion (Miller, Weaver and Semb, 1974) are then used to maintain steady progress. In the present system weekly deadlines for research-task completion were set.

3) Extrinsic consequences (Malott, 1974, p. 327) were presented for completion and non-completion of the research tasks. The two independent variables -- points toward letters of recommendation and weekly feedback on performance -- comprised a treatment package. Historically, college courses have used points earned during the term as a way of determining a letter grade for the student. In this supervising system, research-task completion earned a positive point and non-completion earned a negative point, but at the end of the semester the point totals were given to the graduate students' chief faculty sponsor for use in letters of recommendation.

The second component of the treatment package was feedback on the student's performance. Feedback on performance has been used with a variety of behaviors in a variety of settings -- increasing writing efficiency in a college classroom (Childers and Haas, 1970); reducing dormitory noise in a community setting (Meyers, Artz and Craighead, 1976); increasing the use of behavior modification techniques by staff members in an institution (Panyon, Boozer and Morris, 1970). In this system the feedback was in the form of cumulative positive and negative points earned on research activities for the week and for previous weeks.

4) Observation of research task completion occurred in weekly meetings with doctoral supervisors. Direct research supervision was provided by students who had just completed their own thesis research. This is a similar approach to the use of proctors in PSI courses to assist the students (the proctor having successfully completed the course too).

It is interesting that the above reference citations are mainly from the fields of behavioral systems analysis and individualized instruction. The defining characteristics of applied behavior analysis research exist (Baer, Wolf and Risley, 1968), and yet, there is a paucity of studies that demonstrate how one is trained to do effective applied behavior analysis research. The Behavioral Research-Supervising System was designed in order to develop and maintain effective research behavior for graduate students doing masters level research.

Method

Subjects

Fifteen Master of Arts (MA) candidates in the Department of Psychology at Western Michigan University participated in this experiment. Thirteen of the 15 were full time graduate students; all students were enrolled in the Applied Behavior Analysis curriculum. I categorized them into three groups based on the progress of the student in completing his or her thesis research. Five students, who implemented their thesis research during this study, made up Group I; four students, who wrote the final draft, made up Group W; six students, who entered the department at the start of the study, composed Group G. Members of Group G generated thesis proposals that could become thesis implementations. The experimenter served as a subject in this experiment (a member of Group I).

The group total consisted of eight females and seven males; Group I - three and two, Group W - two and two, Group G - three and three, respectively. The mean age of the MA students was 24 years and one month at the start of the study. All of the subjects had earned their undergraduate degrees within two years of admittance to the graduate program at Western Michigan.

Doctoral supervisors

The primary research supervisors in this system were the doctoral candidates. They conducted the research meetings and recorded all

data. Originally, four doctoral candidates served as research supervisors to the 15 MA students, but during Week 5 of this study one of the supervisors dropped his advising duties. The three remaining supervisors continued throughout the experiment; the reassigned MA students gave these advising totals: Supervisor A - 5 MA students; Supervisor B - 6; Supervisor C - 4. In general, the research interest of the MA student and/or the setting of the thesis implementation determined which supervisor worked with which student. Each doctoral candidate supervised a mix of subjects from the three levels of thesis progress.

One female and two males served as doctoral supervisors. Two supervisors completed their Masters' degrees in an experimental analysis of behavior curriculum; the third did his in an applied area. All three supervisors presently worked in applied systems, either as an instructor or as a manager. The mean age of the supervisors at the start of this study was 26 years one month; each had finished his or her thesis within three years of the start of the study.

Setting and weekly program

Master of Arts students in the Applied Behavior Analysis program take six required semester hours of Master's Thesis credit. While participating in this study the subjects earned course credit -- either in terms of Master's Thesis credit or in the course Research in Behavior Analysis. Members of Group G earned Research in Behavior Analysis credit until the semester in which they implemented their thesis design.

The requirements of this study were only one component of the

thesis requirements under the subjects' chief faculty sponsor. (The chief faculty sponsor was the chairman of the student's Orals Committee.) Specific overlaps were: 1) individual meetings with the sponsor and another faculty member ran concurrently with this study. These meetings occurred at triweekly intervals throughout the semester with each professor and lasted one half hour. The MA student discussed research issues about his or her thesis in these meetings, but this did not involve any of the behavioral contingency relations from this study. The sponsor also discussed the student's performance (for example, the student should increase his number of comments in the group meeting) in academic and self-management areas. This too was independent of the present study. 2) The sponsor attended, on a triweekly basis, the small group meeting between the doctoral supervisor and research advisees. The subjects knew, in advance, the schedule of when the sponsor would be present. 3) The sponsor moderated the weekly large group meeting in which MA student research was the main topic, but the doctoral supervisors did not record any data in this meeting. (All meetings in which the supervisors recorded data, the supervisors moderated.)

A week of activities under this system (see Table I) ran as follows: on Thursday afternoon the small group meeting between the doctoral supervisor and his or her supervisees lasted for one hour. Following a short refreshment break, the large group meeting convened in which the two faculty members, the doctoral supervisors and the MA students discussed topics related directly to a particular thesis, or general research issues, plus a business item agenda. This meeting lasted for one and three-quarter hours. On Monday or Tuesday of the next calendar

Table I: Example of the weekly thesis schedule showing research task due dates and receipt of the feedback form.

TABLE I

EXAMPLE OF THE WEEKLY THESIS SCHEDULE SHOWING RESEARCH
TASK DUE DATES AND RECEIPT OF THE FEEDBACK FORM

Week N

Wed.

Thur. ————— Tasks due: group meeting attendance
review article
data presentation

Fri.

Sat.

Sun.

Mon.

Tues. ————— Tasks due: individual meeting attend-
ance
log
self-reported hours
writing (and research
proposal)
self-editing

Week N + 1

Wed.

Thur. ————— Student receives written feedback
form showing Week N performance

Fri.

(etc.)

week the supervisor met individually for half an hour with each MA supervisee. The doctoral supervisor recorded data only in the small group and individual meetings. A 'thesis' week lasted seven days and ran from the student's individual meeting to his or her next individual meeting.

Experimental design

I used a BAB reversal design in this study. This type of design starts with the experimental condition in effect, follows with the baseline condition and then reintroduces the experimental condition. The three phases in this study were as follows: (B) Points Toward Recommendation and Feedback; (A) Baseline; (B) Points Toward Recommendation and Feedback.

I used the BAB design rather than an ABAB or ABA design for three reasons: 1) the length of the semester did not allow a full ABAB reversal design because I expected three to five data points to establish a trend in each phase; 2) ending the semester in the experimental condition was probably a much stronger test of the treatment package than the ABA design. Typically, other courses provide poor contingency management of a student's academic behavior. The student must complete a large amount of work at the close of the semester in order to finish the course. The immediacy of the deadlines and grade consequences in these concurrent activities work against steady, maintained progress on thesis research. Because the student's thesis research seldom has these immediate demands it is often the first to drop out. Good performance by the subjects at this time of the semester would show the value of this

system; and 3) if this system was effective in controlling research behavior then it would be profitable for the subjects and the group to use it (which a BAB does more than an ABA). Similarly, I discounted a between group design -- using an experimental and control group -- because withholding the benefits of this system from the control group subjects was of doubtful ethical merit if another equally powerful design was available.

From preliminary observations of masters level research (an earlier pilot study) I believed the dependent variables would reverse under the BAB design. There are few intrinsic rewards for doing research -- the tasks are hard work and extend over a number of months. The long initial delay in starting research and its "on-again-off-again" variability in maintenance for most MA students indicated that the research tasks would reverse upon removal of the contingency relationship of this system.

A multiple-baseline design across individuals or groups would have presented two problems: 1) administration time expanded when there were two or more groups; 2) interaction between members of the groups was unavoidable as the supervising system now functioned. This verbal contact among differentially treated subjects was experimentally unacceptable (Kazdin, 1973, pp. 519-520).

Experimental procedures

Points Toward Recommendation and Feedback. In the experimental condition, positive and negative points and written weekly feedback were contingent on the MA student's performance. At the end of the

study the student's chief faculty sponsor received these point data for use in letters of recommendation. The feedback was in the form of cumulative points earned for the week and for previous weeks.

Baseline. The doctoral supervisors continued to record the positive and negative points the students earned; the sponsor still "expected" the students to keep up their research work (verbal directive). But the students did not receive the written weekly feedback on their behavior and the points of identified, individual students, collected during this condition, were never shown to the sponsor.

The two independent variables constituted a treatment package in this study -- no studies in educational technology have used points in this manner. The feedback component was added in order to help guide the subjects' future research behavior based on their past performance. At the start of the study the faculty sponsor announced that a total of zero negative points earned by a student in the experimental condition would indicate good performance. The weekly feedback would presumably make the point toward recommendation contingencies more effective by showing the student his or her current point totals.

General procedures

I obtained informed consent signatures from each participant in this study (MA students and doctoral supervisors). This form clearly described the behavioral contingency relationships that affected a participant, the option that each participant could

withdraw at anytime from the study and that the data collected in this experiment remained completely anonymous in whatever form of publication or presentation the author attempted.

Subjects received handouts describing the dependent variables, general procedures and a schedule of meeting dates with the two faculty members at the start of this study. I gave out a questionnaire to all 15 of the MA students and the three doctoral supervisors as a check of their understanding of the requirements of the advising system. This occurred in the Thursday large group meeting of Week 7 (see Table II). No positive or negative points were contingent on the scores of the MA students or supervisors; the supervisors answered all questions on the form while MA students answered only those questions applicable to their level of thesis progress. I told the system participants a week in advance to review the operations of the advising system as described in the handouts. I provided answer keys to the system members immediately after each completed the questionnaire. At the next large group meeting I discussed with the group members the most frequently missed questions.

I encouraged the doctoral supervisors to schedule their weekly individual meeting times with MA students on Monday or Tuesday of the thesis week. This was important for two reasons: 1) the feedback on the students' performance should occur as soon as possible after the completed week's research behavior in order to effectively guide the following week's behavior. With this arrangement of meeting times the students completed a week's worth of tasks and then received a written record of their performance at the Thursday large

Table II: Occurrence of specific events in Study I by experimental condition and week.

TABLE II

OCCURRENCE OF SPECIFIC EVENTS IN STUDY I BY EXPERIMENTAL
CONDITION AND WEEK

<u>Conditions</u>	<u>Weeks</u>	<u>Specific Occurrences</u>
B	1	
	2	
	3	
	4	Research proposal due for Group G
A	5	Formal announcement -- sponsor still expects tasks in baseline to be completed
	6	
	7	List distributed -- small group meeting attendance by sponsor
	8	
B	9	
	10	
	11	System evaluation by participants
	12	

group meeting. This was within two or three days (depending on the exact day of the individual meeting) of the previous week. 2) I wanted the subjects' research behavior that earned points in one week to have actually occurred in that week. The MA student turned in the more difficult and time consuming tasks at the individual meeting. If this meeting was at the end of the 'thesis' week then work towards completing the tasks more likely occurred in that week.

The doctoral supervisor assigned two kinds of points to the MA student's performance: positive and negative. The MA student earned a positive point if he or she met criteria on the required research task. The MA student earned a negative point if he or she did not meet criteria on the required research task. It was possible for the student to do extra tasks each week and thus he or she could earn some extra positive points. But, if the student missed a deadline and earned a negative point, then that task always showed a negative point.

The doctoral supervisors turned in their data sheets (Appendix A) at the end of each thesis week. I calculated the percentage of points completed (for a group or separate task) by dividing the number of positive points earned by the total number of required points (x 100) for that week. I calculated the percentage of points not completed by dividing the number of negative points earned by the total number of required points (x 100) for that week. Doing extra tasks (that is, earning extra positive points) could inflate the percent completed value; however, extra positive points could not similarly affect the percent of required points not completed. If

a student failed to completed a required task this would increase the percent not completed, even if the student completed extra (non-required) tasks, because the number of negative points in the numerator remained the same.

During the experimental condition I prepared written weekly feedback and distributed the individual feedback forms (Appendix B) to the MA students at the break between the small and large group meeting on Thursday. The feedback form showed the number of points required for the student, the number of positive and negative points earned on each of the required tasks for that week and that week's total summed. The additional category of the total number of positive and negative points earned under each required task indicated cumulative performance of the MA student across the semester in the experimental condition. Positive points indicated the 'quantity' of thesis work completed while negative points indicated the 'timing' of the work -- non-completion because the student missed a deadline.

I used a criterion-referenced system in providing feedback to each MA student. I announced at the beginning of the study a goal of zero negative points earned for each student. Generally, members of each group worked on the same number of required research tasks throughout this experiment. Students did not receive formal feedback on the performance of their group peers during this experiment.

I announced the change in contingencies for each phase at the Thursday large group meeting. This was in the form of a vocal statement to the group. The actual change occurred after the subjects' individual meeting following that large group meeting; so the formal

announcement preceded the change by four or five days.

The student could postpone a deadline (tacitly stated as -- 'if unavoidable circumstances occurred') if he or she had the consent of his or her doctoral supervisor in advance of the deadline. This was not the same as merely notifying the supervisor in advance -- for example, leaving a note on the supervisor's desk involved no consent. There had to be prior agreement with the supervisor.

At the end of this study the MA students filled out a detailed evaluation that asked them to rate various aspects of the advising system. Areas included the value of current requirements to the student, possible changes for the next semester, and other social validity ratings by the participants.

Dependent variables

I gave the following descriptions to the MA students and the doctoral supervisors at the start of the experiment. The doctoral supervisors checked these tasks on Thursday of the thesis week:

Group meeting attendance. This included both small and large group meetings. The students were recorded in attendance if they arrived before 4:00 P.M. Thursday.

Review article. MA students completed the 10 categories on the literature review form on each article read (Appendix C). Subjects read articles relevant to their thesis topic. They wrote a minimum of 100 words with an entry under each category (even if it was to say this category did not apply). They summed and circled the total number of words at the top of the page. The doctoral

supervisor checked the completed form in the small group meeting. The MA student discussed the article in this meeting but with no points contingent.

Data presentation. MA students who implemented their thesis design and collected weekly data presented these new data each week in the small group meeting. The student who collected data at intervals greater than one week's time contracted a week in advance with the supervisor for presentation. The MA students complied the data in graphic, tabular or statistical form.

The doctoral supervisor checked these tasks on Tuesday of the thesis week:

Individual meeting attendance. The MA student met weekly for half an hour on an individual basis with his or her doctoral supervisor. The student was recorded in attendance if he or she arrived no later than the scheduled time of the meeting.

Log. The log was the student's intellectual diary for the week. It listed concepts, ideas, procedures and procedure changes, statements that came from the student's research meetings, and things in the environment that affected the study. The log contained a minimum of 200 words; entries were by calendar dates and it ran from one individual meeting to the next one. The MA student summed and circled the total number of words written at the beginning of the new week.

Self-reported hours. The MA student summed the total number of hours worked on thesis activities for the week -- this included meeting, reading, writing and research time but excluded paid assistantship hours in an applied setting and hours worked in an applied

setting for practicum credit. Students enrolled for Master's Thesis credit worked a minimum of 12 hours per week on thesis activities in order to earn a positive point for this task; students not currently enrolled for this credit recorded their hours for a positive point but there was no minimum time requirement. Students summed and circled the reported hours worked at the start of each new thesis week.

Writing. This requirement was normally in terms of the formal thesis write-up -- a thesis outline, 750 new words on a section (in either the American Psychological Association or Systems Analysis (Malott, 1974, page 324) format), rewriting an entire section. However, the student substituted, with the prior consent of the supervisor, written materials used in the implementation of the thesis design (for example, handouts, training package).

Self-editing. MA students edited their own writing as a way to improve their writing style. The students edited the 750 words or section (written in the previous task) in either of two ways: they submitted the original draft along with the completed work, or on the original rough draft the student made corrections in a different color ink.

Research proposal. This was neither a weekly task, nor was it required for all 15 subjects. The six new Group G students completed a research proposal (turned in at the individual meeting) at four week intervals. This task provided the opportunity for the new MA students to generate thesis designs by analyzing their applied work setting as a place containing many applied behavior analysis thesis

topics. The students provided a statement of their general topic and the setting they worked in and specific recommendations under each of these four categories as examples of possible thesis implementations: dependent variables, independent variables or setting, design (and other methodological issues), and subjects. The analysis was in terms of articles read by the MA students, suggestions logged from research meetings and actual work done in the setting. The proposal required a minimum of 200 words which the students summed and circled at the top of the page, an entry under each category, and complete sentences (I defined as a clause having a subject and a verb). On the Thursday small group meeting, following submission of the proposal, the MA students in Group G did not do a review article. The students discussed their proposals with their doctoral supervisors at this time. The supervisor required a proposal rewrite if he or she determined the proposal was unsatisfactory.

The students were required to complete those tasks relevant to their level of thesis progress. The students in Group G (see Table III) worked on these tasks: attendance at the group and individual meetings, review article, data presentation (when implementation began), log, self-reported hours, research proposal (due at four week intervals); students in Group W worked on these tasks: individual meeting with doctoral supervisor, writing, self-editing; students in Group I worked on all eight weekly tasks. Students earned extra positive points when they completed additional tasks over the weekly requirement. Such tasks were: review article, writing (based on 750 word increments -- that is 1500 words earned +2 positive points,

Table III: The three groups in this study and their required research tasks; also shown are the tasks that earned positive points when they were completed over the weekly requirement.

TABLE III

THE THREE GROUPS IN THIS STUDY AND THEIR REQUIRED RESEARCH TASKS; ALSO SHOWN ARE THE TASKS THAT EARNED POSITIVE POINTS WHEN THEY WERE COMPLETED OVER THE WEEKLY REQUIREMENT

<u>Tasks</u>	<u>Group G</u>	<u>Group W</u>	<u>Group I</u>
1) Group meeting attendance	X		X
2) Review article*	X		X
3) Data presentation	X		X
4) Individual meeting attendance*	X	X	X
5) Log	X		X
6) Self-reported hours	X		X
7) Writing*		X	X
8) Self-editing*		X	X
9) Research proposal*	X		

* -- Tasks that could earn extra positive points

2250 words earned +3 positive points, etc.), self-editing these words (on the same point scale), individual meeting with doctoral supervisor and research proposal.

Observation and reliability

The three doctoral supervisors served as primary observers. The supervisor assigned a positive point if the MA student correctly met criteria on each of his or her weekly required tasks; a negative point indicated the response did not meet criteria; the supervisor recorded a dash when the behavior did not apply to the student. The supervisor also recorded the student's quantitative output on certain tasks -- the number of words written on the log, '750 words' requirement, self-editing, review article and research proposal and the number of hours recorded for the self-reported hours task.

The experimenter served as the secondary reliability observer. On an unannounced, though frequent basis across the semester, the experimenter attended individual and small group meetings and inspected the work of the MA student(s). Interaction between the secondary observer and the participants in the meeting was brief (two to four minutes in individual meetings; eight to ten minutes for small group meetings) and specific contact with the supervisor and student(s) was avoided.

I used a reliability calculation (Bijou, Peterson and Ault, 1968) in which the reliability percentage equalled the number of agreements divided by the total number of agreements and disagreements (x 100). I calculated these percentages on the eight research

tasks (plus the research proposal) and the quantitative outputs and also figured percentages across supervisors. The secondary observer assigned positive and negative points to the MA student's behavior (Reliability Sheet found in Appendix D) and then added to the total number of agreements and disagreements after comparison with the supervisor's data sheet. In the case of disagreements the secondary observer discussed these specific instances with the doctoral supervisor within one week of the reliability check.

In summary, 15 MA students were supervised on their thesis research by three doctoral candidates. Eight research tasks with weekly deadlines were specified; completion of the task by the deadline earned a positive point, non-completion earned a negative point. Subjects in Group I worked on all eight tasks per week; subjects in Group W did writing, editing, individual meeting attendance; subjects in Group G did all eight weekly tasks but writing and editing, plus the research proposal due at four week intervals. The two independent variables -- points toward recommendation and weekly written feedback on performance -- comprised a treatment package. At the end of the semester the points earned on research activities were given to the MA students' chief faculty sponsor for use in letters of recommendation; the weekly feedback form showed the individual MA student's performance (in terms of positive and negative points) for the previous week and cumulative totals. The design was a BAB reversal, in which the 'B' stands for the experimental condition and the 'A' for the baseline condition. The doctoral supervisors recorded all data during this study; the experimenter served as a secondary reliability observer. A Type II reliability calculation was used.

Results

Reliability

Mean reliability for the three doctoral supervisors for the entire study is 90%. Individual percentages are: Supervisor A -- 70%; Supervisor B -- 96%; Supervisor C -- 90%. I sampled Supervisors B and C on all eight tasks in each of the three phases of the study. Scheduling differences between the secondary observer and Supervisor A's individual meetings with MA students resulted in a low number of reliability checks. On Supervisor A I checked only group meeting attendance, review article and data presentation in the two experimental phases; in the baseline phase reliability checks included all research tasks but self-editing.

I sampled approximately 30% of the total possible occurrences of research-task completion in this experiment. Separate reliability percentages on the dependent variables are: group meeting attendance -- 93%; review article -- 88%; data presentation -- 87%; individual meeting attendance -- 97%; log -- 96%; self-reported hours -- 87%; writing -- 94%; self-editing -- 87%.

Reliability percentages on the quantitative output of the MA students' work range from 92 to 100% with a mean of 95%. These are: review article words -- 92%; log words -- 96%; words written -- 100%; words self-edited -- 100% (± 10 words); self-reported hours -- 92% (± $\frac{1}{2}$ hour).

Dependent variables

Figure 1 shows the percent of required points completed and not completed in each week of the study for all 15 subjects. The MA students showed a high rate of completion in the first experimental phase (median 90% completed), dropped substantially in the baseline phase (68%), and nearly recovered the original performance level in the second experimental phase (86%). Tasks done extra slightly inflated the percent of points completed curves. But the percent of required points not completed showed a similar pattern: a median of 14% in the initial experimental phase, a rise to 34% in the baseline phase, and a return to 14% in the second experimental phase. (All other graphs in Study I (Figures 3-5) are analyzed in terms of the percent completed curves. Analysis of the graphs in terms of the percent not completed curves -- which are not affected by tasks done extra -- does not change the interpretation of the results.)

Figure 2 shows that individual performance of the 15 subjects is closely related to the average group performance. To obtain a subject's score on this bar graph I used the following formula:

$$X_{\text{score}} = \frac{(\# \text{ of negative points earned in baseline phase})}{2} - \frac{(\# \text{ of negative points earned in the experimental phases})}{2}$$

I first equated the number of opportunities to earn negative points in the baseline and experimental phases for the three groups. This involved arithmetically increasing the mean number of point opportunities in the two experimental phases in order to equal the point opportunities in the baseline phase. This produced proportional or equal

Figure 1: Mean percent of required points completed (dots connected with a solid line) and not completed (circles connected with a dashed line) in each week of the study for the total group. There was a mean of 70 required points per week (range 55-84); N = 15. The horizontal solid lines indicate median percent completed and not completed in each phase. During the baseline phase no point or feedback contingencies were in effect. (These two features -- horizontal median lines in each phase, no point or feedback contingencies in baseline -- apply to Figures 3-6 and 8-10.)

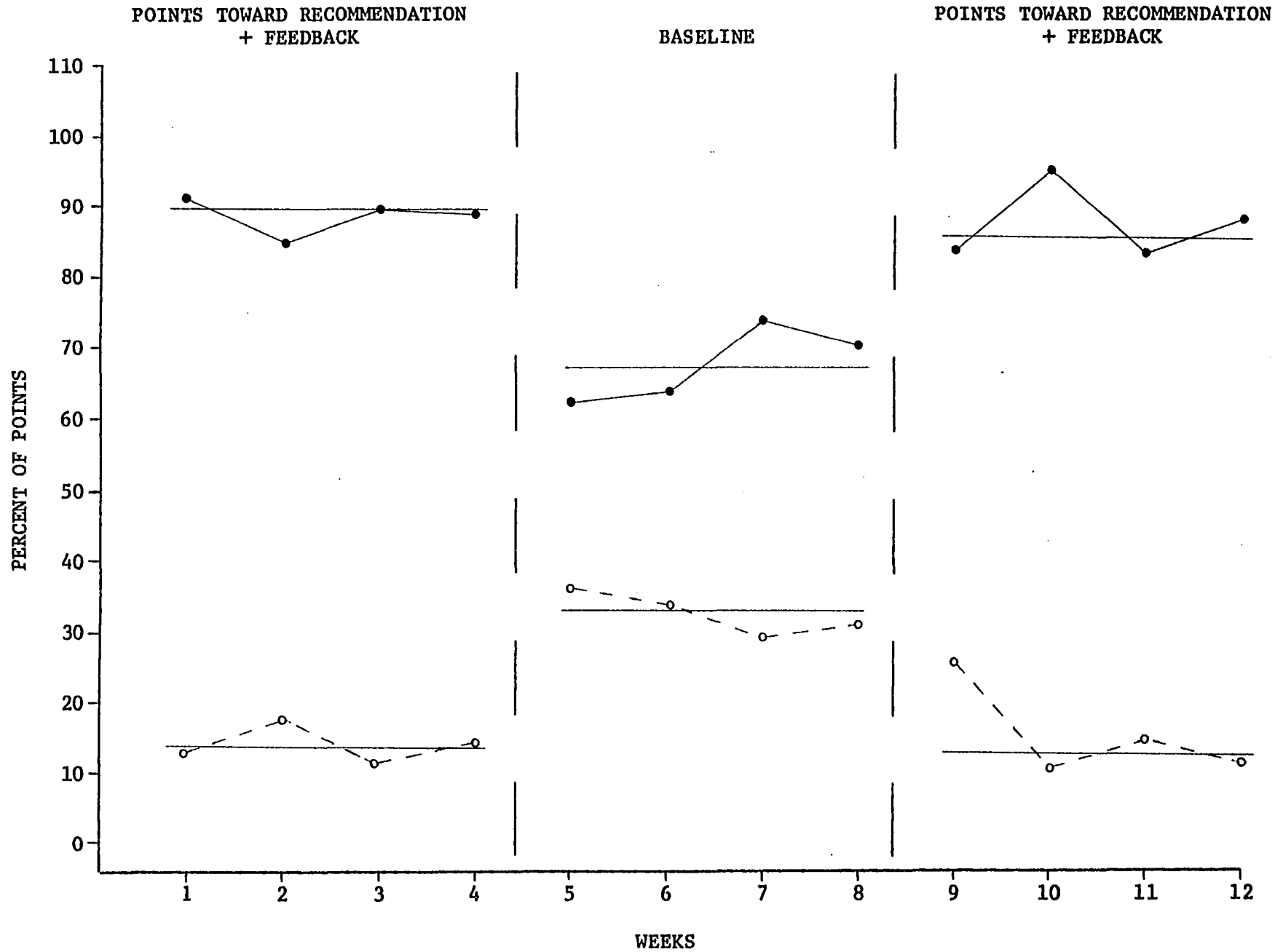
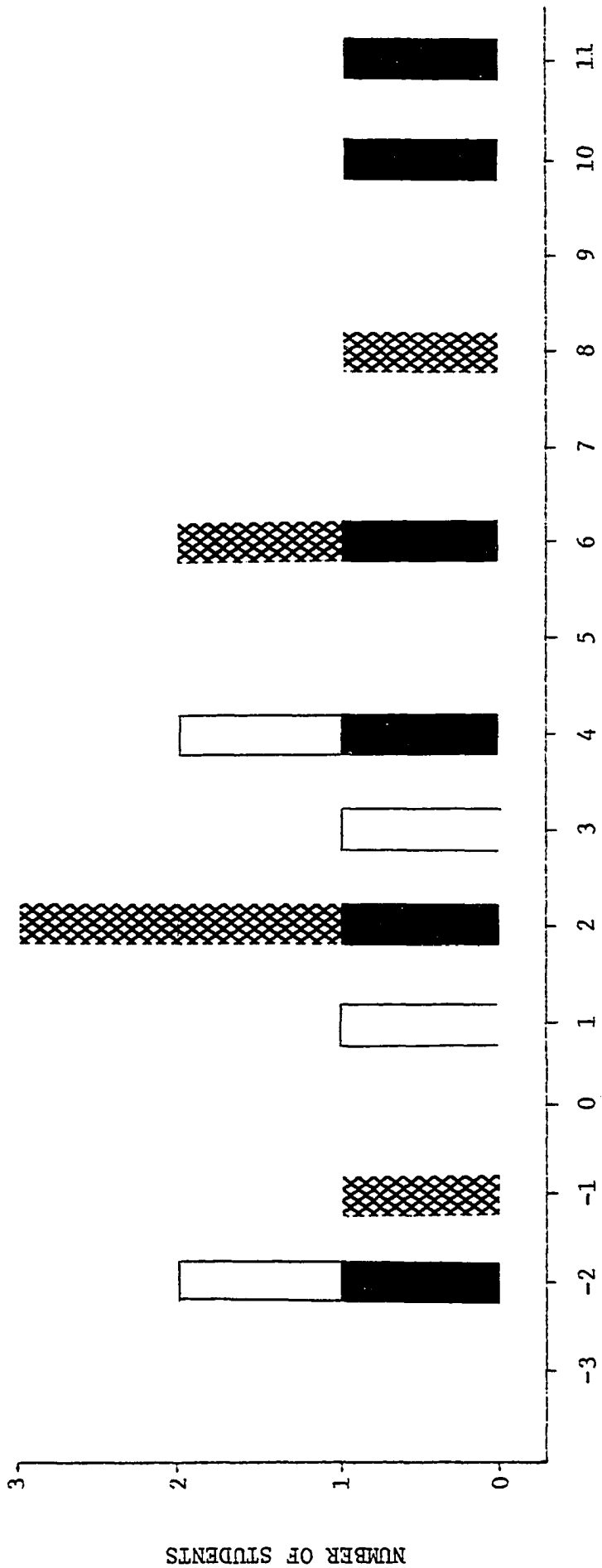
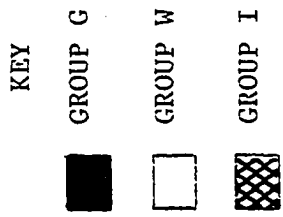


Figure 2: Frequency bar graph shows the relative performance of the 15 subjects for the entire experiment. Subjects who exhibited scores greater than zero earned more negative points in the baseline condition than they earned in the experimental condition. Subjects who exhibited scores less than zero earned more negative points in the experimental condition. "G" = member of the generating group; "W" = member of the writing group; "I" = member of the implementing group.



TOTAL NUMBER OF NEGATIVE POINTS EARNED IN
BASELINE CONDITION RELATIVE TO TOTAL EARNED IN EXPERIMENTAL CONDITION

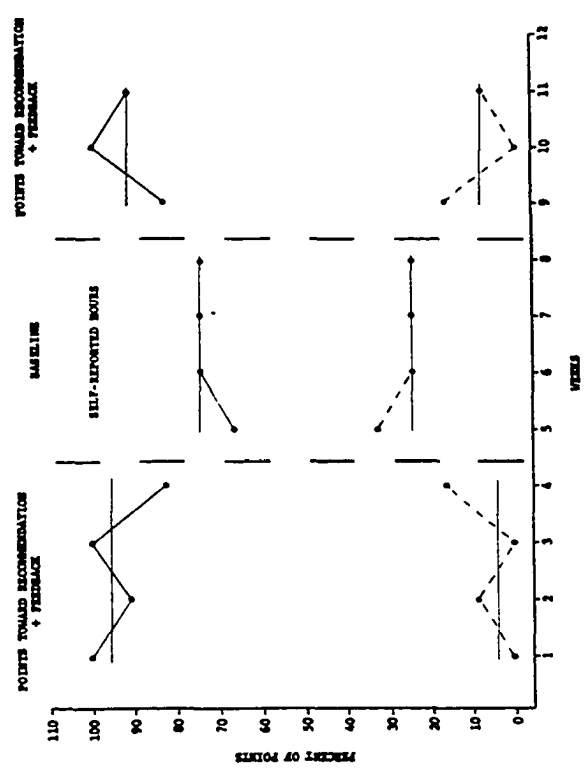
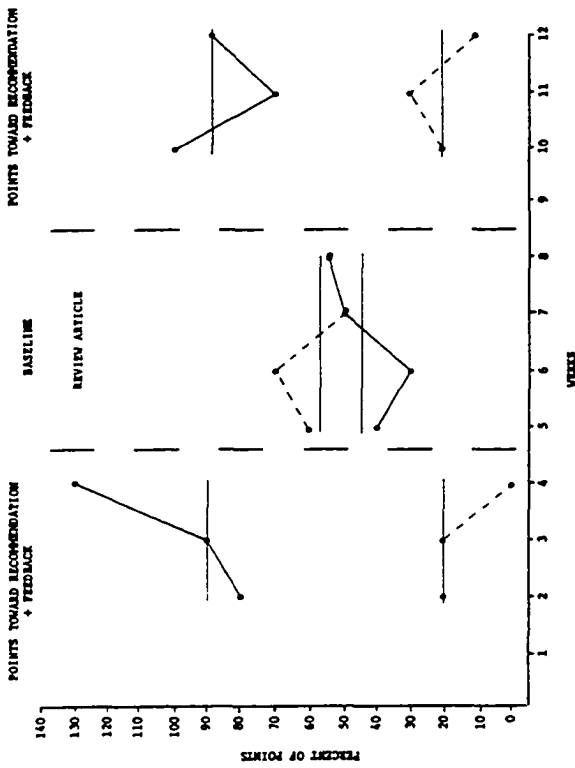
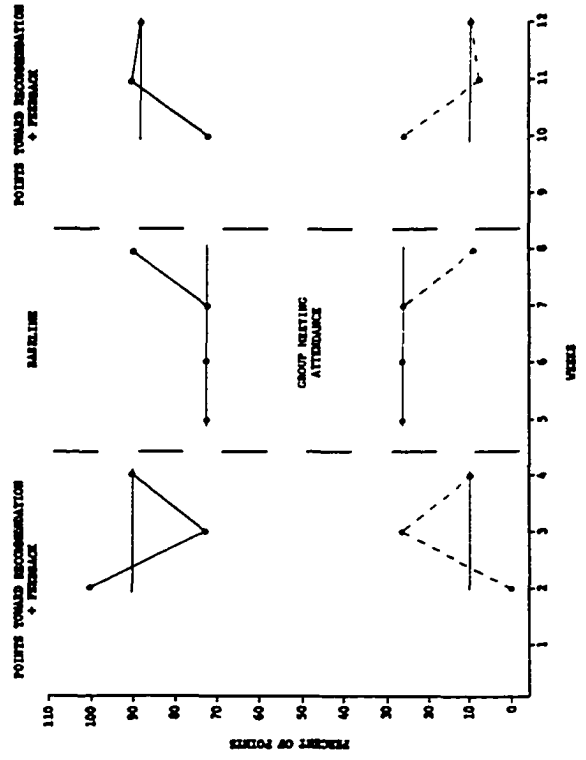
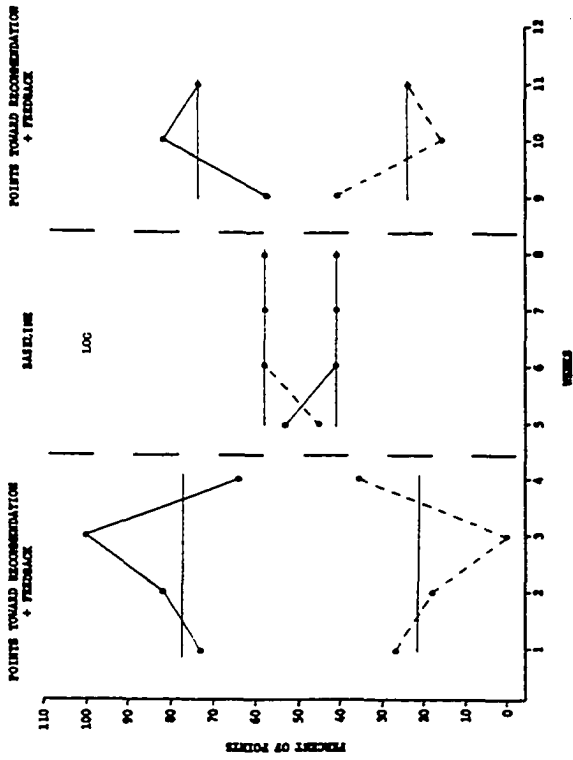
ratios of point opportunities for the experimental and baseline conditions. I then multiplied the number of negative points earned in the experimental phases by this conversion ratio, summed these two numbers and then used the above formula. The 12 resulting positive values in Figure 2 indicate that 80% of the subjects earned proportionately more negative points in the baseline condition than in the experimental condition. The three subjects who earned proportionately more negative points in the baseline condition than in the experimental condition represented each of the three groups.

Figure 3 depicts the results on four of the separate research tasks. Review article, log and self-reported hours show clear behavioral effects -- in both experimental phases the median percents completed are at least 15% higher than the baseline percent completed median; review article is 44% higher, log 33% higher, self-reported hours 17% higher.

Group meeting attendance shows equivocal results because of intra-condition variability. It seemed from inspection of the graphs in this study that variability could be defined as: two or more data points that 1) dropped at least 15% from their experimental condition median percent(s) completed and equalled or overlapped any data points in the baseline condition, or 2) rose at least 15% from their baseline condition median percent completed and equalled or overlapped any data points in the experimental condition.

For group meeting attendance both of the experimental phases median percents completed are a minimum 16% higher than the baseline phase median, but two experimental phase data points drop 17 and 16%

- Figure 3: Percent of required points completed (dots connected with a solid line) and not completed (circles connected with a dashed line) in each week of the study for:
- a) Review Article -- mean of 10 required points per week (range 5-11); N = 10.
 - b) Log -- mean of 12 required points per week (range 11-12); N = 12.
 - c) Self-Reported Hours -- mean of 12 required points per week (range 11-12); N = 12.
 - d) Group Meeting Attendance -- mean of 11 required points per week (range 9-11); N = 11.



and overlap baseline phase data, and a baseline phase data point rises 17% and overlaps experimental phase data.

Figure 4 shows the remaining four research tasks. Writing, self-editing, data presentation and individual meeting attendance show no effect -- the baseline phase median percent completed is equal to or higher than one or both experimental phase medians. Writing and self-editing both have the baseline phase median percent completed higher than the first experimental phase median; for data presentation all three median percents completed are equal; for individual meeting attendance the baseline phase median percent completed is higher than the second experimental phase median.

Figure 5 shows the performance of the MA students at the three levels of progress of thesis research. Group G and Group W show a clear effect -- in both experimental phases for Group G the median percents completed are a minimum 26% higher than the baseline percent completed median; Group W is 15% higher. Group I shows an equivocal effect -- the median percents completed are a minimum 10% higher than the baseline median.

The first research proposal for the members of Group G was due in Week 4 -- 100% of the required points were completed (N = 6); the second proposal was due in Week 8 -- 80% of the required points were completed (N = 5).

The data recorded by the supervisors on the MA students' quantitative output on research tasks generally supported the effects just indicated for completion and non-completion of the tasks. Separate word totals on review article, log, and hours total on self-reported

- Figure 4: Percent of required points completed (dots connected with a solid line) and not completed (circles connected with a dashed line) in each week of the study for:
- a) Writing -- mean of nine required points per week (range 7-12); N = 9.
 - b) Self-Editing -- mean of seven required points per week (range 3-9); N = 7.
 - c) Data Presentation -- mean of three required points per week (range 1-7); mean of three subjects per week (range 1-7).
 - d) Individual Meeting Attendance -- mean of 15 required points per week (range 14-15); N = 15.

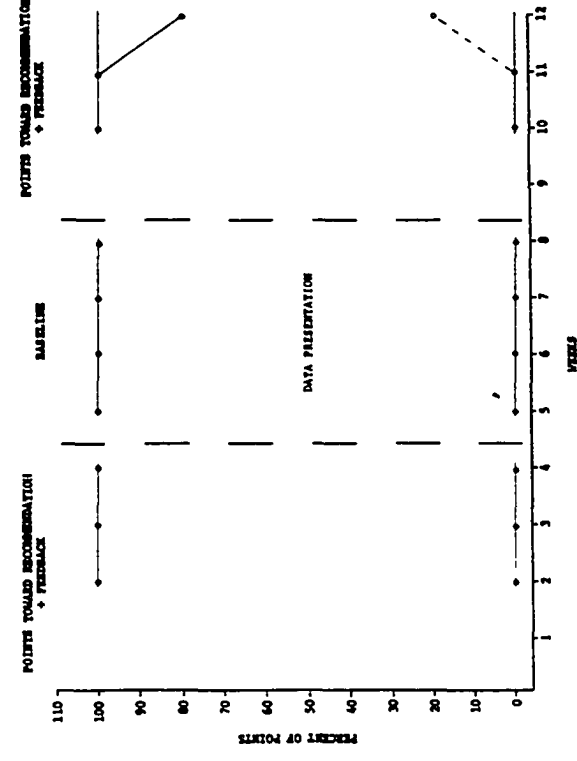
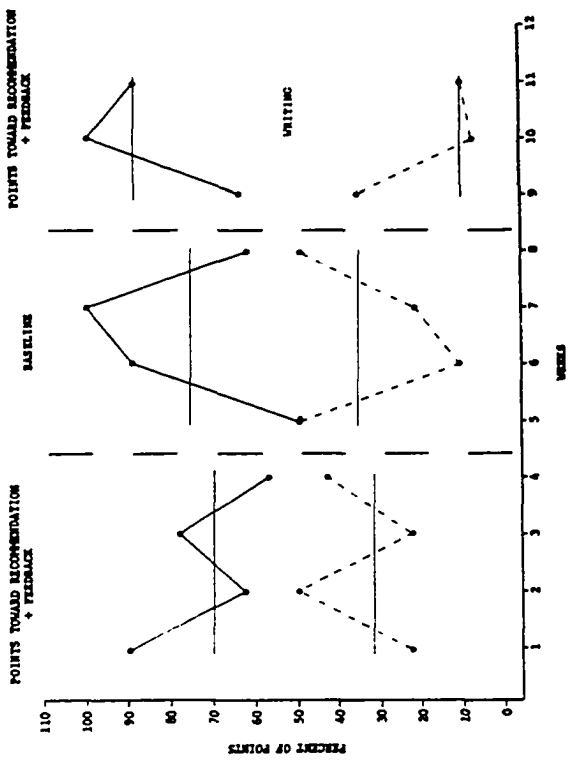
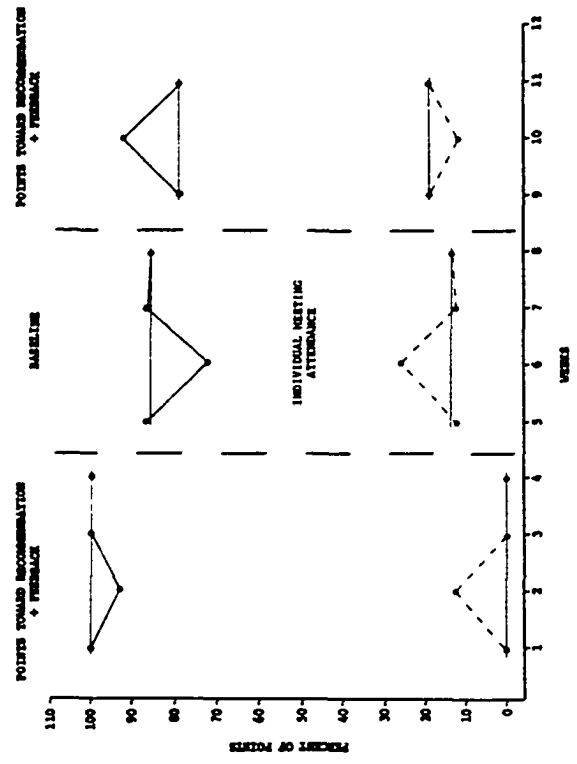
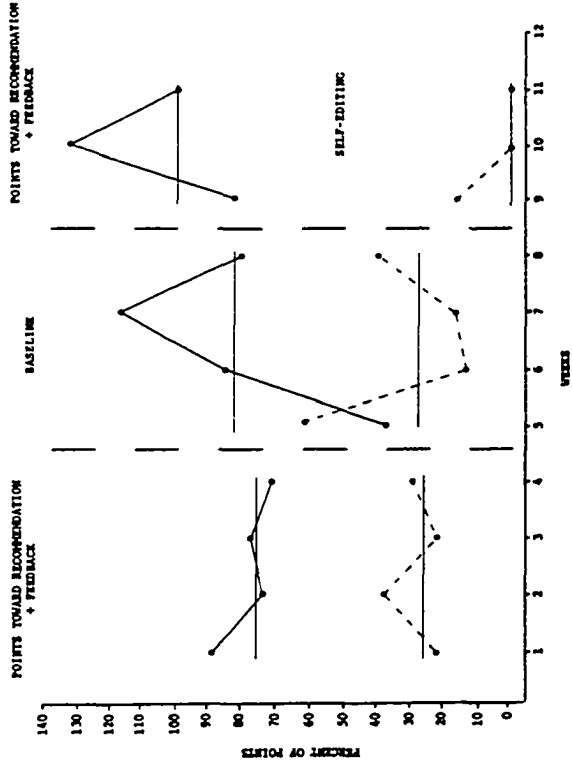
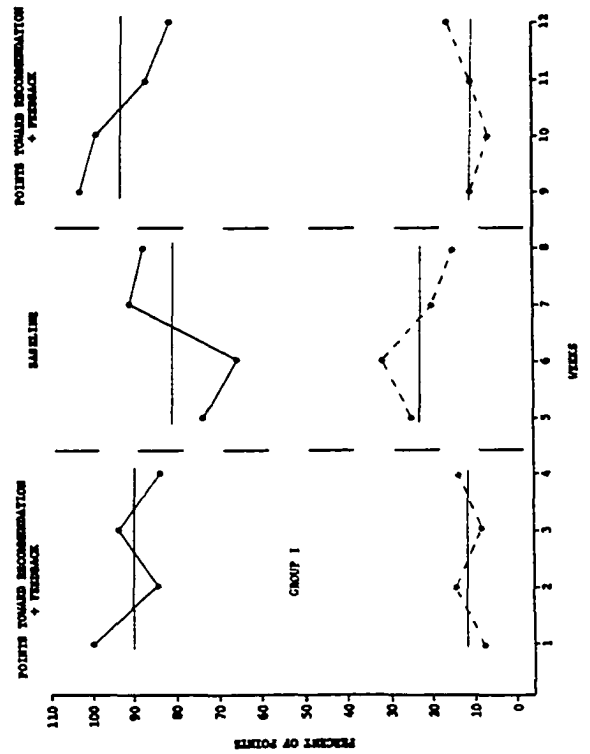
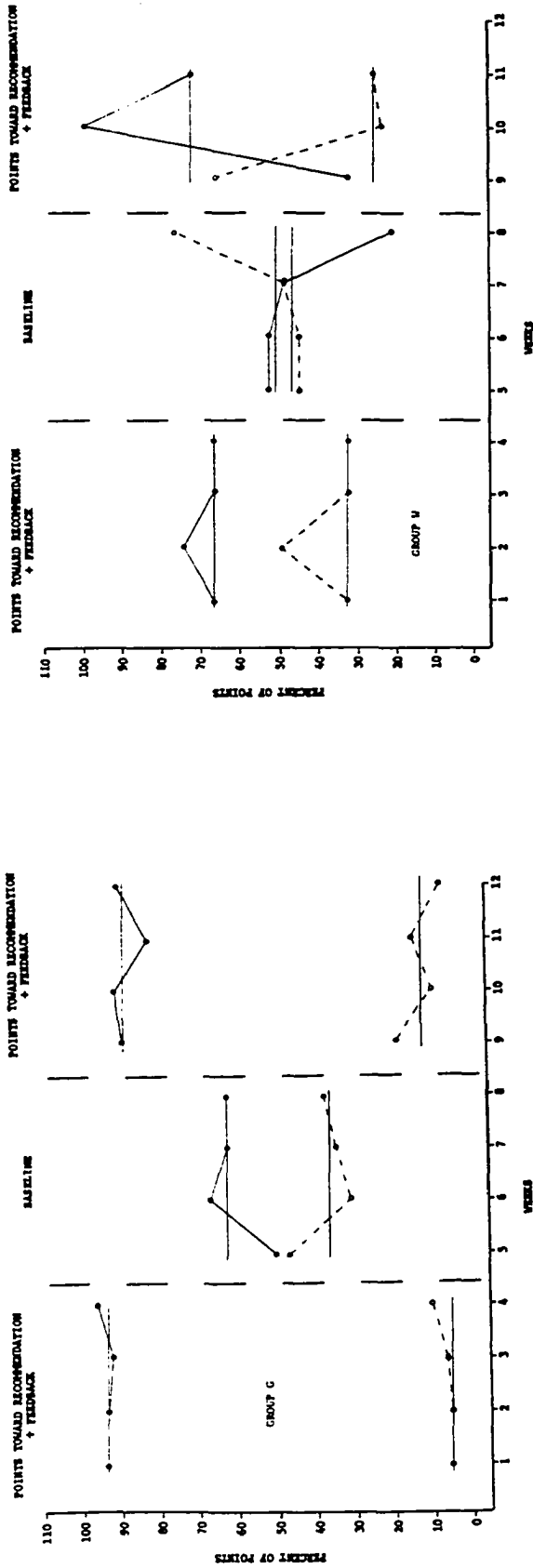


Figure 5: Percent of required points completed (dots connected with a solid line) and not completed (circles connected with a dashed line) in each week of the study for:

- a) Group G -- mean of 28 required points per week (range 18-38); N = 6.
- b) Group W -- mean of 11 required points per week (range 9-12); N = 4.
- c) Group I -- mean of 32 required points per week (range 24-37); N = 5.



hours (students doing 12 hours per week), were consistently over the minimum requirements in the first experimental phase; the totals always dropped lower than these first levels (and often below the minimum requirements) during the baseline phase; in the second experimental phase the totals were again over the minimum but not to the same level as the first experimental phase.

But word totals for writing and self-editing never followed this pattern. Usually these tasks alternated on consecutive weeks across the entire study with an oscillation around the minimum requirements regardless of the conditions in effect.

The Thursday tasks of review article, group meeting attendance and data presentation (Figures 3 and 4) show a missing data point for Week 1 because the first experimental phase started with the Tuesday individual meeting; and for Week 9 when the Thanksgiving vacation occurred and there was no Thursday meeting. The Tuesday tasks of log, self-reported hours, writing, editing and individual meeting attendance (Figures 3 and 4) show a missing data point for Week 12 because the final week of the study included the Thursday meeting only. One subject missed this final group meeting when she left for Christmas vacation. I dropped this subject's last week of data from the study because this type of absence was the only such occurrence for any student during the semester. Members of Group W (Figure 5) have no Thursday tasks required and so no data point for Week 12 appears.

The doctoral supervisors and MA students filled out the questionnaire on the dependent variables and general procedures of the supervising system. The 15 subjects scored a mean of 83% correct (range

65-100%); the three doctoral supervisors scored a mean of 93% correct (range 90-95%).

At the end of the study 13 of the 15 subjects completed an evaluation of the supervising system. The most favorably rated aspect of the system was the individual meeting between the MA student and his or her doctoral supervisor. Ninety-two percent of the MA students said that during the baseline phase they always knew what behavioral contingency relationships were in effect; in the experimental phases 77% were always clear. During the actual baseline phase 100% of the MA students were always clear what phase was in effect; during the experimental phases 92% of the students were always clear what phase was in effect. During the entire experiment 92% of the 13 respondents were always clear what research tasks they were required to do each week. The students rated the supervising system as well-organized. The weekly feedback forms and the positive and negative point procedures were clear, while the points toward recommendations were less clearly understood.

The six members of Group G endorsed the research proposal as an effective approach for generating research designs. By the end of the study four of the six students implemented their proposal and collected data for presentation. An added benefit of this research task occurred for the student's doctoral supervisor. The written proposal allowed the supervisor to clearly understand and helpfully critique the student's research designs.

Five of the 13 students indicated that they hoarded; I defined "hoarding" in this way -- the student completed a research task in

Week N but did not turn it in to the supervisor until Week N + n (which was any succeeding week). This occurred for the review article, log and writing. When the hoarding occurred or when the student turned the task in to the supervisor was not limited to any one phase. Six of the 13 students indicated they scheduled an extra meeting, at one time in the semester, with the sponsor, the second faculty member in the advising system, another faculty member or another thesis committee member. Three of the MA students used formal contingency contracting outside of the supervising system -- to graph weekly data, to implement a pilot study, and the third student had a self-management contract covering the eight research tasks.

Of particular importance to the long term effectiveness of this system is the favorable overall rating provided by its participants. Nine of the 13 MA students (69%) said they preferred this type of supervising system with its weekly regularity of specified research tasks, deadlines and contingencies on performance over a less structured approach. One subject (8%) considered the aversive features of the system, like negative points and missed deadlines, as unacceptable. (Three of the students (23%) did not clearly approve or condemn the system in the evaluation.)

Discussion

The present results indicate that points toward recommendation and feedback are effective in controlling research behaviors. For the entire study 12 of the 15 students earned proportionately less negative points in the experimental condition than in the baseline condition.

How is it that the research behaviors of the subjects were controlled by the independent variables? Clear analysis of the point and feedback contingencies is difficult because the controlling relations are very subtle. The treatment package sets up two schedules, with one based on the positive point as a learned reward and the negative point as a learned aversive. The rewarding and aversive properties of these two kinds of points are established through the verbal statements present in the system (for example, handouts, group discussions, rules from the chief faculty sponsor). The positive points are available on a limited hold schedule in which the student must complete a research task within a specified period of time each week; the negative points are part of an avoidance procedure in which the completion of a research task prevents the occurrence of an aversive (the negative point and ultimately, a poor letter of recommendation).

In order to be maximally effective as consequences for the subjects' research behavior, the points should directly follow the behavior. Only for group and individual meeting attendance were the points directly contingent on behavior -- that is, there was no temporal delay between behavior that earned the point and when the

point was assigned by the supervisor. But throughout the study the subject's other research tasks were also checked in these two meetings. And so the repeated pairing (though greatly delayed) of the research output and the points sustained the rewarding and punishing value of the points.

The two kinds of point contingencies -- presenting and preventing -- also provided consequences in the form of self-management statements by the subjects (Malott and Whaley, 1976). Instances of off task behavior, when the student should have been completing a research task, were probably punished by self-given aversive statements; these statements decreased the occurrence of behavior that would eventually earn negative points. Following partial or full completion of a research task, the student supplied him or herself with rewarding statements; these statements tended to increase the occurrence of behaviors that would earn positive points.

An important stimulus function in this supervising system was in the form of rule control. As Skinner (1969) has stated, a rule is a statement that describes:

- 1) the topography of the action involved,
- 2) the setting the action occurs in,
- 3) and, the results of the action.

An example of a rule used in this study was:

- 1) topography -- the response definition of any of the dependent variables;
- 2) setting -- the announcement that the experimental condition was in effect;

3) and, consequences -- the points earned in this condition received by the student's chief faculty sponsor. Each MA student read, heard, and probably repeated this rule (or similar variations) many times during the study. The rule then cued the occurrence of the subject's research behavior.

The feedback component of the treatment package also served a stimulus function. The quantitative measures on the form served as positive feedback when they maintained the research activities or made them more likely to occur; they served as negative feedback when they made the acts less likely to occur (Malott, Tillema and Glenn, in press). The feedback form could have also cued the occurrence of self-given rewarding and punishing statements and appropriate rules.

The sources of behavioral control in this supervising system are quite complex. We have hypothesized three main functional causes: 1) the points as learned rewards and aversives on two separate schedules; 2) self-management techniques, such as self-given rewarding and aversive statements provided by the subject; 3) rule and feedback control. Further speculation on the exact contribution of each is probably not warranted.

Writing (Figure 4) did not show clear behavioral control by the treatment package -- the baseline condition median was higher than the first experimental condition median. Of the eight research tasks in this study, writing probably was the most difficult to complete. But with the present point scale an MA student could not earn more positive points for writing completion, or negative points for non-completion, than for any other of the less effortful research

tasks. In Study II the consequences for the writing task should more accurately reflect the higher response cost -- provide a greater point value for completion as well as non-completion of the task.

Self-editing (Figure 4) was not as effortful as writing, but its completion was wholly dependent on the prior completion of the writing task. It was for this reason that writing and self-editing showed similar percent completed and not completed curves in this study.

The effects of the independent variables on group meeting attendance (Figure 3) were weakened because other powerful contingency relationships were present -- for example, the chief faculty sponsor moderated the large group meeting and low attendance by a student was noticeable; often the sponsor requested students to discuss certain issues related to their thesis design that had general interest to the group or that the sponsor desired group input on. Regardless of the condition in effect these factors may have increased the percent of tasks completed for group meeting attendance. The sponsor also attended the small group meeting on a regular basis and in a similar way may have affected the percent of tasks completed for review article and data presentation.

Data presentation (Figure 4) did show a high rate of percent completion across the entire study. But it is difficult to accurately analyze in terms of the independent variables because only a small number of required points occurred each week. Weeks 2-6 contain only seven required points (mean of 1.4 required points per week) while the remaining five weeks contain a mean of five required points per week. The number of students who collected data increased directly

with the number of implementations -- during Weeks 2-6 only two people implemented; in the remaining five weeks five people implemented -- with three in Week 7 and two in Week 10. However, only two of the implementations lasted longer than five weeks and this distorts the results found in Figure 4.

I am not sure why individual meeting attendance (Figure 4) was not affected by the point and feedback contingencies. The particular requirement of this task might contain an explanation. The MA student was recorded in attendance if he or she arrived on time and stayed for the full half hour. For the entire study only five of the 20 negative points for "non-attendance" were for actually missing a meeting; three-quarters were earned for tardiness. In the first experimental phase there were two negative points (one absence, one tardy); in the baseline phase there were 10 negative points (one absence, nine tardy); in the second experimental phase there were eight negative points (three absences, five tardy). Accordingly, in Figure 4 performance was high in the first experimental phase and much lower in baseline. Late arrival caused the decrease in the baseline phase and not actual missed meetings, but a late student could still show the supervisor his or her completed Tuesday tasks (for positive points) and could also discuss research issues in the remainder of the half hour. The consequences for late arrival, other than one negative point, were not that aversive, and so, with the onset of the second experimental phase, the effectiveness of the point and feedback contingencies was lessened.

The results from Study I show that points earned toward recommendation letters combined with weekly feedback on the subjects' performance

are effective in controlling research behavior. Twelve of the 15 MA students earned proportionately less negative points in the experimental condition than in the baseline condition. For the three levels of thesis progress in this system, Group G and Group W show clear effects, and Group I shows equivocal effects. When all the parts -- the total group average, individual performance, the eight research tasks, the three groups based on thesis progress -- are examined in terms of the whole system the behavioral effects are evident. Yet performance on the separate research tasks is not nearly as distinct -- effects are clear only for review article, log, and self-reported hours, equivocal for group meeting attendance, and non-existent for writing, self-editing, data presentation and individual meeting attendance. Study II is a direct attempt to bring all of the eight research tasks under the control of the point and feedback contingencies.

STUDY II

The design of this study included a number of participant suggestions from the end of the semester evaluation as well as experimental considerations generated from an analysis of the empirical results of the first study.

1) The MA students rated the Thursday large group meeting as unstructured and too long while the small group meeting was too short to deal with the students' research issues. So for this study we set the time of the large group meeting at one and one quarter hours (with a business item period if time remained) and the small group meeting expanded to one and one half hours; I also posted a weekly agenda sheet for issues to be discussed. (When two of the doctoral supervisors could not attend the Thursday meeting until 5:00 P.M. the order of the meetings was switched -- large group now followed by small group.)

2) Generally the students rated the log as ineffective. The word total was too high and its ultimate utility questioned. As a result we made these changes:

- a) new descriptions of areas to write on,
- b) the supervisor would read the student's log in the individual meeting,
- c) and a log summary was added to insure reviewing of past entries.

3) The MA students requested a norm referenced measure on the

the weekly feedback form. So during this study they received the percent of points completed and not completed for the group in which the student was a member.

4) I also included a list of Graduate College requirements in this semester's advising system. As requested by the participants these tasks were optional and were to be individually contracted between the supervisor and the student. Completion or non-completion of the contracted task earned a positive or negative point but these data were not included on the weekly feedback form or in the results of Study II. I took this approach for two reasons: 1) the initial description of these tasks (reviewed by the participants towards the end of Study I) was not favorably received by all the subjects. Three or four vigorously objected to the comprehensive scope of the tasks and said I had no business making some particular tasks requirements for points. 2) For these tasks the student received +1 point for completion regardless of the condition in effect, since they were contracted individually. This was a different procedure compared to the regular point procedures on research tasks, and I considered its inclusion as a category on the feedback form to be confusing and possibly misleading to the student.

Specific experimental procedures were: 1) I decided to use the same BAB reversal design in this second study but to provide for longer experimental phases. The missing data points in Study I hindered the analysis of the effects of the treatment package.

2) The point requirement on writing and self-editing was increased so that it more nearly reflected the relative amount of

behavior involved in completing these tasks (without making the point value of other tasks irrelevant). At the same time more comprehensive self-editing requirements were included.

3) And finally, I handed out the questionnaire on dependent variables and general procedures of the advising system before the start of the study; and so all participants knew their score and the errors made at the start of the first week.

Method

Subjects

Fifteen MA candidates in the Department of Psychology at Western Michigan University participated in this study. Fourteen of the 15 were full time graduate students and all were enrolled in the Applied Behavior Analysis program. Thirteen of the 15 served in Study I. The same five students composed Group I; eight students now made up Group G -- the original six plus two new admittees to the supervising system; Group W contained two students. The experimenter was a member of Group I.

The group total consisted of seven females and eight males: Group I -- three and two, Group W -- one and one, Group G -- three and five respectively. All of the subjects had earned their undergraduate degrees within two years of admittance to the graduate program at Western Michigan.

Doctoral supervisors

The same three doctoral candidates continued as research supervisors. Each supervised five MA students; Supervisors A and C had no students from Group W.

Setting and weekly program

The behavioral requirements of Study II were again only one component of the thesis requirements under the subjects' chief faculty

sponsor. The subjects continued to receive course credit for participating in the study. The chief faculty sponsor again announced that all research tasks required for a student should be completed regardless of the condition in effect. The students also continued to meet triweekly with the two faculty members and they discussed research issues in these meetings. The sponsor continued to attend on a triweekly basis each small group meeting; the subjects knew this schedule in advance. The sponsor moderated the weekly large group meeting. These activities ran concurrently with Study II.

The weekly program of thesis activities was similar to the first experiment. The only change was for the group meeting on Thursday -- the large group meeting now preceded the small group meeting and only lasted for one and one-quarter hours.

Experimental procedures

I also used the BAB reversal design in this study. The three phases were the same: (B) Points Toward Recommendation and Feedback; (A) Baseline; (B) Points Toward Recommendation and Feedback.

The weekly feedback form included all the categories used in Study I. Based on student preference in the evaluation at the end of the previous semester, I added a norm referenced feedback category. The subjects received quantitative descriptions of their own performance plus the percent of required points completed and not completed on the group's performance in which they were a member. Unlike the individual feedback categories, the group performance was not cumulative and did not show the separate research tasks or the total

required points.

This was the new point scale for completion and non-completion of the writing and self-editing:

<u>Total Words Written</u>	<u>Points Earned</u>
0-749	-2
750-1499	+2
1500-2249	+4
2250-2999	+6

The scale continued on (if necessary) in multiples of 750. The student earned the same number of points at the exchange rate depicted on this scale, for self-editing these words. (Note that this is twice the point value of Study I.)

General procedures

Subjects received revised handouts describing the dependent variables and general procedures, and a schedule of meeting dates with the two faculty members at the start of the study. I gave out the revised questionnaire on research tasks and system procedures before the study began. The supervisors and subjects filled it out on their own and received feedback on their performance outside of the Thursday large group meeting. The experiment started after the participants knew their score and the correct answers for the questions they missed.

The two kinds of points for performance, the calculation of percent of required points completed and not completed, tasks done extra and the postponement of deadlines were identical to the previous study. I continued to make the announcement of the change in

contingencies for each phase at the Thursday large group meeting. The students continued to receive their weekly feedback forms at this meeting. The subjects again filled out an evaluation at the end of the semester.

Dependent variables

The Thursday tasks -- group meeting attendance, review articles, data presentation -- were identical to Study I; these Tuesday tasks -- individual meeting attendance, self-reported hours, and writing were also the same. I gave out the following revised descriptions to the MA students and the doctoral supervisors at the start of the study:

Log. I added these criteria: ideas from other courses, from faculty members, from articles and books you have read, and self-memo belong in the log. The log entries included a running summary of topics in the student's discussion section and for analyzing data and the type of graphs needed in the thesis write-up. The doctoral supervisor read the log in the individual meeting. Twice during the semester the MA student prepared a log summary based on past log entries of this semester but written since the previous summary. The log summary comprised two or more statements of significance on the two (or more) most important developments that affected the student's thesis. This was a minimum of 400 words long; the total number of words was summed and circled at the top of the page. The log summary precluded the regular weekly log requirement for that week; the supervisor also read this. I announced the summary due date (which was the MA student's next individual meeting) at the group meeting of that week.

Self-editing. This consisted of four requirements introduced cumulatively across the study (see Table IV). These were an attempt to help teach the student a writing style based on readability.

1) Active versus passive voice -- the student underlined the verb(s) in each sentence and placed an "A" or "P" over each verb to indicate the voice. Editing marks were in a different color ink than the text.

2) First versus third person -- the student underlined any personal pronoun used as the subject of a sentence and placed a "1st", "2nd", or "3rd" above the pronoun. These editing marks were in a third color.

3) Content review -- the student underlined the topic sentence in each paragraph and then placed, on each succeeding sentence of the paragraph, a checkmark over the period if the sentence logically related to the topic sentence. The student provided written justification on another sheet of paper for those sentences without checkmarks. A fourth color of ink indicated these editing marks.

4) Cumulative and short sentences -- the MA students were required to complete a training package (developed by Mary Tillema, manuscript in preparation) on identifying and writing cumulative sentences. A cumulative sentence has more free modifiers at the end and beginning of the sentence and not in the middle; a free modifier is any word or words that stand before the subject of a sentence (with the exception of coordinate conjunctions like 'and', or 'but') and medial or final sentence words set off by commas, dashes or parentheses.

The supervisors checked the accuracy of the editing marks during

Table IV: Occurrence of specific events in Study II by experimental condition and week.

TABLE IV

OCCURRENCE OF SPECIFIC EVENTS IN
STUDY II BY EXPERIMENTAL CONDITION AND WEEK

<u>Conditions</u>	<u>Weeks</u>	<u>Specific Occurrences</u>
B	1	Questionnaire on supervising system's dependent variables and procedures Active vs. passive voice self-editing
	2	
	3	
	4	
	5	Log summary due
A	6	
	7	
	8	
B	9	
	10	
	11	
	12	
	13	Log summary due System evaluation by participants

the individual meeting.

Research proposal. Only the two new members of Group G qualified for this task. The supervisor contracted individually with the MA student for completion of this task, but all other requirements remained the same.

Graduate College procedures. I prepared a description of 15 separate tasks related to the MA student's successful completion of the graduate program. These included the English Qualifying Examination, Candidacy Application, forms for registration and departmental requirements, residency and diploma application, fellowships and research stipends open to graduate students, important pamphlets in preparing the thesis document, sequence of submitting the thesis through the Psychology Department and the Graduate College, list of deadlines for submission of papers and sites for conventions in behavior analysis in 1977 tasks related to the supervising system -- performance review with doctoral supervisor, undergraduate advising, individual meeting with Committee members, formal presentation of thesis results to the large group, and Informed Consent and Human Subjects Review Committee procedures. These requirements were optional for each MA student. I provided a checklist which indicated those tasks a student contracted to complete, by what dates, and what points earned. The doctoral supervisor recorded these data on the checklist.

The members of the three groups worked on the same tasks as in the first experiment; the same tasks done extra also applied (see Table III).

Observation and reliability

The exact same procedures occurred in this experiment as in Study I.

Results

Reliability

Mean reliability for the three doctoral supervisors for the entire study is 94%. Individual percentages are: Supervisor A -- 88%, Supervisor B -- 91%, Supervisor C -- 99%. I sampled each supervisor on all eight research tasks in each of the three phases of the study.

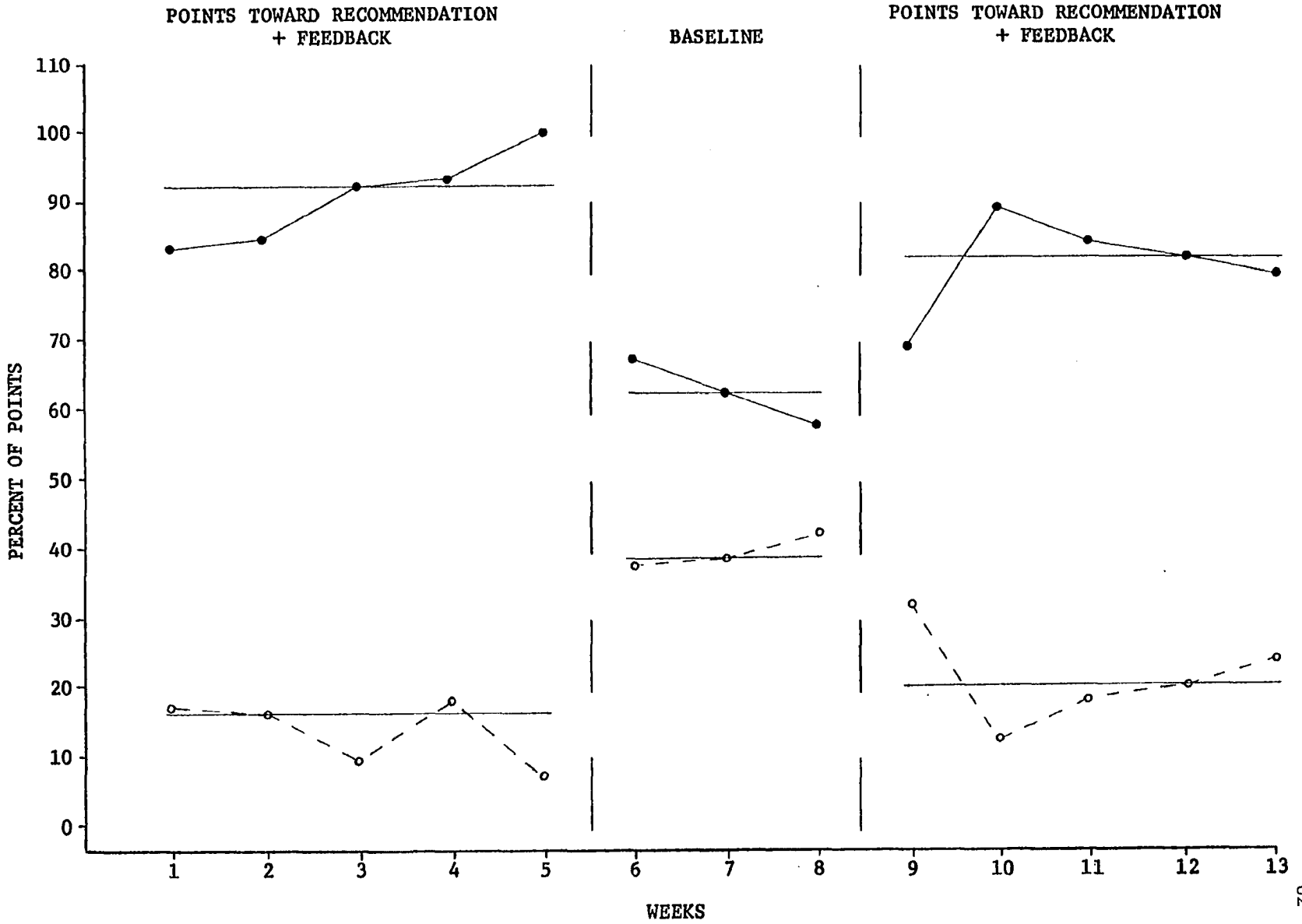
I sampled approximately 27% of the total possible occurrences of research-task completion in this study. Separate reliability percentages on the dependent variables are: review article -- 95%; data presentation -- 89%; individual meeting attendance -- 97%; log -- 95%; self-reported hours -- 92%; writing -- 95%; self-editing -- 95%.

Reliability percentages on the quantitative output of the MA students' work range from 92 to 100% with a mean of 95%. These are: review article words -- 98%; log words -- 94%; words written -- 100%; words self-edited -- 93% (these are all based on ± 5 words); self-reported hours -- 92% ($\pm \frac{1}{2}$ hour). An exact word and hour total (that which was totalled by the MA student on the document) was not used because one supervisor consistently rounded these numbers for his data sheet while the secondary observer recorded them directly.

Dependent variables

Figure 6 shows the percent of required points completed and not completed in each week of the study for all 15 subjects. The MA

Figure 6: Mean percent of required points completed (dots connected with a solid line) and not completed (circles connected with a dashed line) in each week of the study for the total group. There was a mean of 90 required points per week (range 86-96); N = 15.

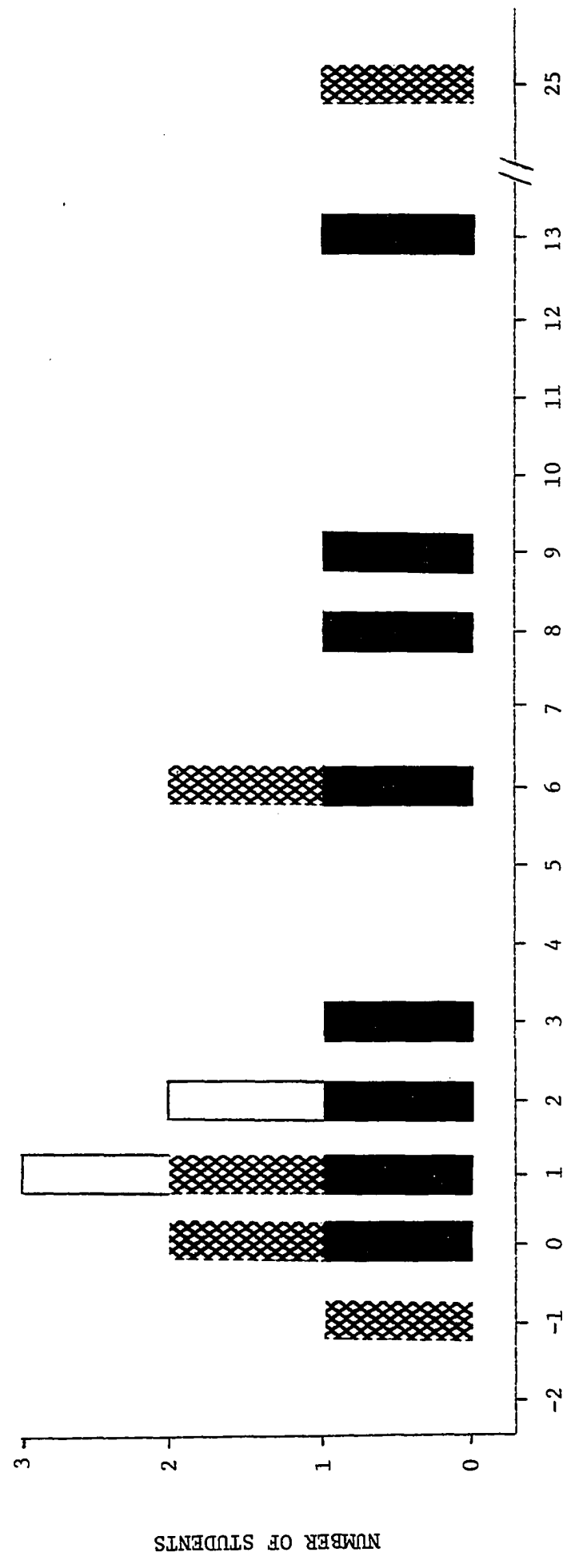
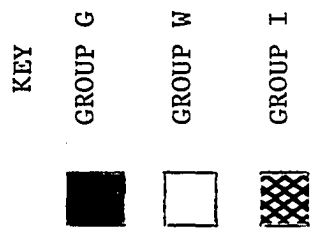


students showed a high rate of completion in the first experimental phase (median 92% completed), dropped abruptly in the baseline phase (62%), and recovered much of the original level of performance in the second experimental phase (82%). The percent of required points not completed exhibited a similar pattern: a median of 16% in the initial experimental phase, a rise to 38% in the baseline phase, and a near complete return to 18% in the second experimental phase. Tasks done extra inflated the percent of points completed in the first experimental phase only. (All other graphs in Study II (Figures 8-10) are also analyzed in terms of the percent completed curves.)

Figure 7 shows that individual performance of 12 of the 15 subjects closely supported the average group performance. In the student evaluation at the end of the study a subject from Group G indicated she did not look at her weekly feedback form when it was handed out. She was one of the three that did not seem to be affected by the positive and negative points. One of the two Group I subjects who showed a negative score also had a negative value in Study I (refer to Figure 2). In the evaluation this student indicated she used her own self-management program on thesis work.

Figures 8 and 9 show the results on the separate research tasks. Review article, log, writing, self-editing and individual meeting attendance demonstrate clear behavioral effects -- in both experimental phases the median percents completed are at least 15% higher than the baseline percent completed median; review article is 15% higher, low 35% higher, writing 35% higher, self-editing 31% higher, individual meeting attendance 15% higher (based on an average of the

Figure 7: Frequency bar graph shows the relative performance of the 15 subjects for the entire study. Subjects who exhibited scores greater than zero earned more negative points in the baseline condition than they earned in the experimental condition. Subjects who exhibited scores less than zero earned more negative points in the experimental condition. "G" = member of the generating group; "W" = member of the writing group; "I" = member of the implementing group.



TOTAL NUMBER OF NEGATIVE POINTS EARNED IN
BASELINE CONDITION RELATIVE TO TOTAL EARNED IN EXPERIMENTAL CONDITION

- Figure 8: Percent of required points completed (dots connected with a solid line) and not completed (circles connected with a dashed line) in each week of the study for:
- a) Review Article -- mean of 10 required points per week (range 4-12); N = 10.
 - b) Log -- mean of 12 required points per week (range 7-13); N = 12. The log summaries were due in Weeks 5 and 13.
 - c) Writing -- mean of 14 required points per week (range 9-22); mean of seven subjects per week (range 5-9).
 - d) Self-Editing -- mean of 11 required points per week (range 9-15); mean of five subjects per week (range 4-7). The active-passive voice requirement started with Week 1; first-third person and content review became part of the requirement with Week 6.

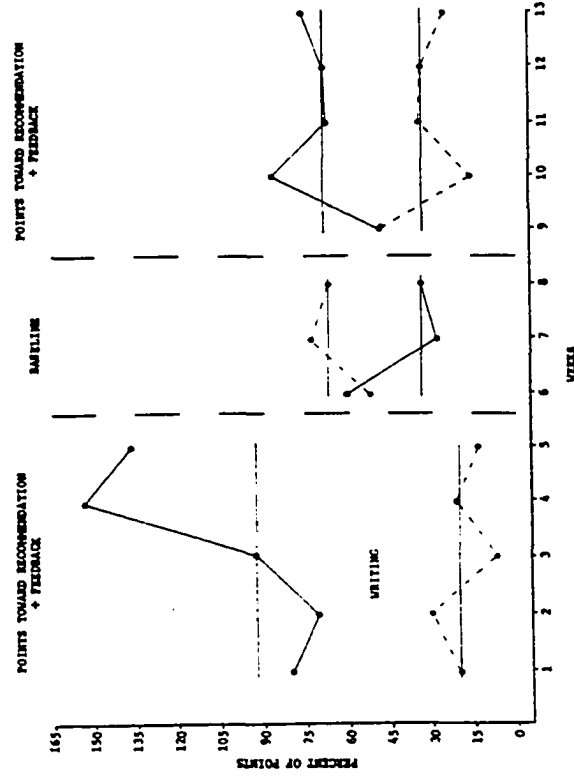
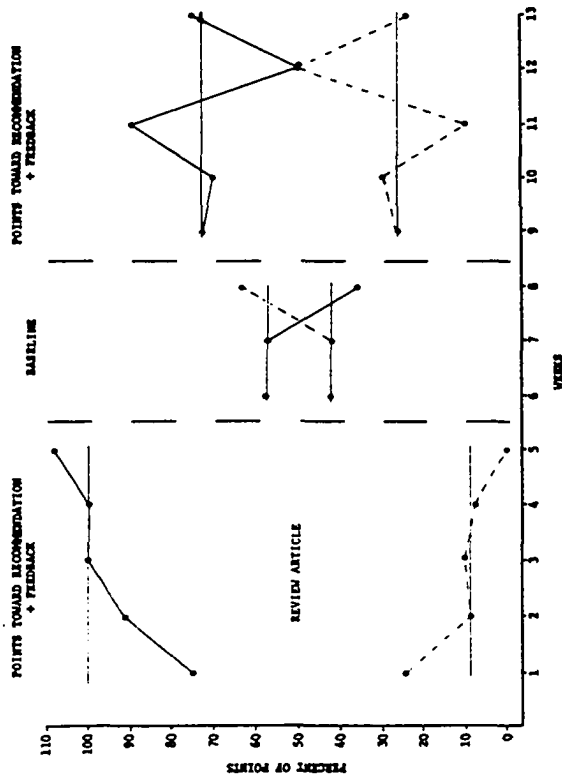
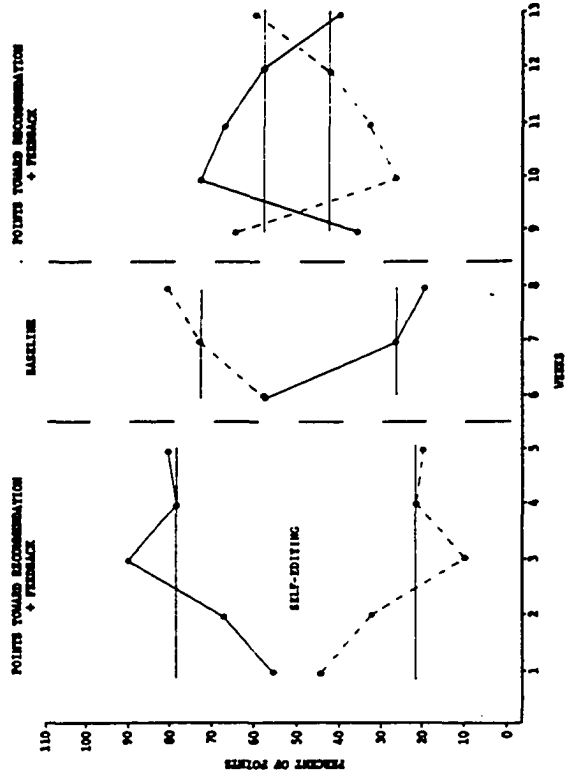
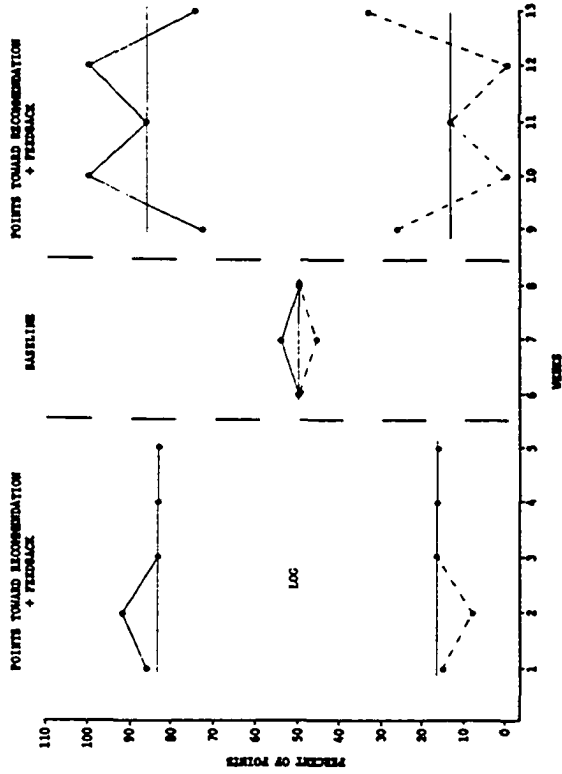
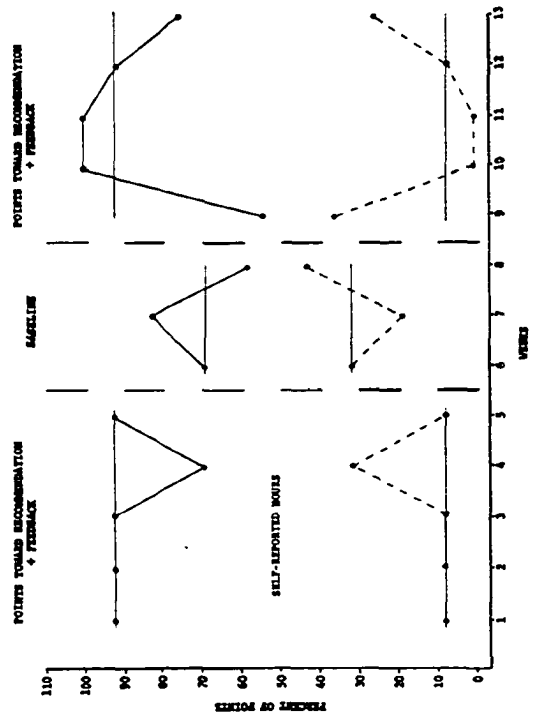
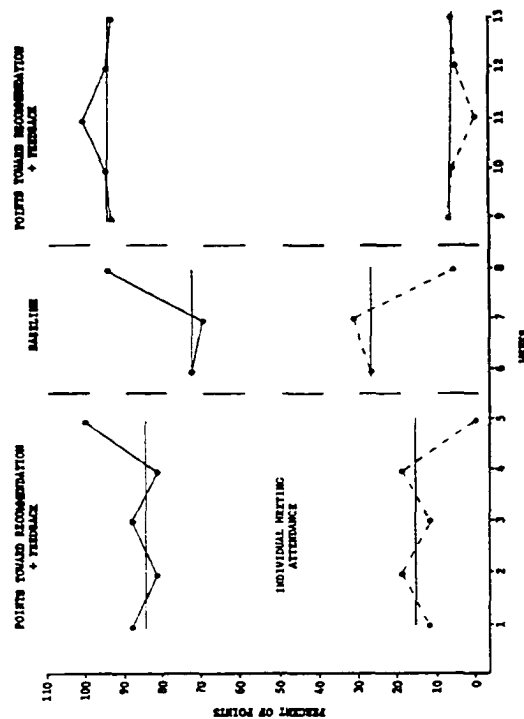
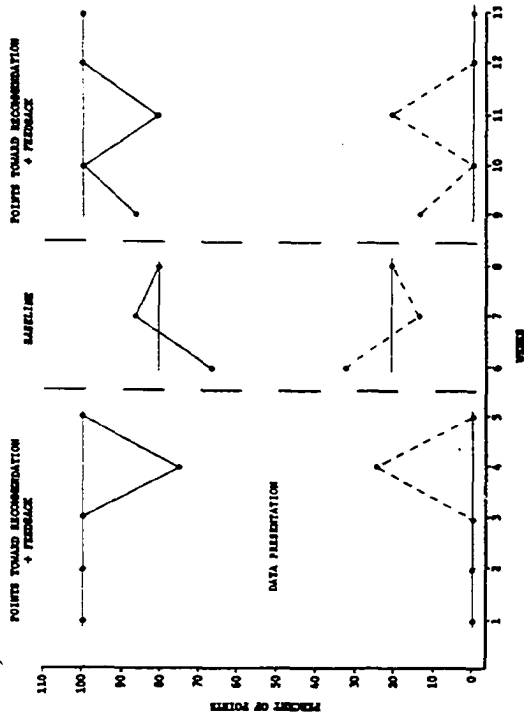


Figure 9: Percent of required points completed (dots connected with a solid line) and not completed (circles connected with a dashed line) in each week of the study for:

- a) Individual Meeting Attendance -- mean of 15 required points per week (range 10-15); N = 15.
- b) Data Presentation -- mean of six required points per week (range 2-10); N = 6.
- c) Self-Reported Hours -- mean of 12 required points per week (range 7-13); N = 12.



two experimental phase medians).

Data presentation and self-reported hours show equivocal results -- their experimental condition medians are 20% and 23%, respectively, over their baseline medians -- but they both have more than the minimum number of overlapping data points.

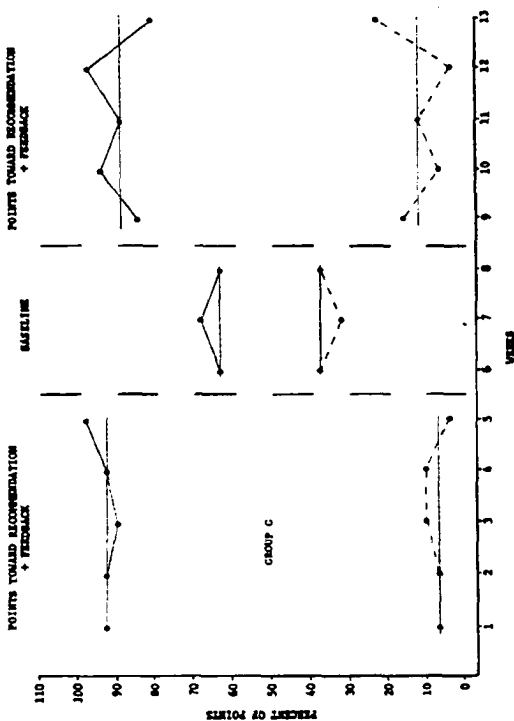
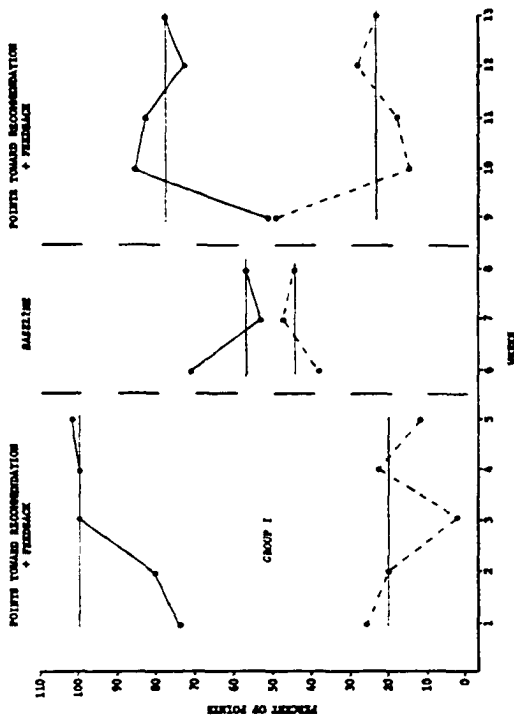
I have excluded the group meeting attendance task from the results of Study II. I left the response definition for this task the same as Study I but two of the doctoral supervisors could not arrive at the group meeting until 5:00 P.M. because of a conflicting class. The supervisors did not check attendance until the start of the small group meeting which was now one and one-half hours after the stated deadline in the response requirements. So the percent of points completed in this study was incorrectly raised (median percents completed for the three phases were 96%, 100% and 100%).

Figure 10 represents the performance of the MA students at two levels of progress of thesis research. (I excluded Group W from this second study because there were only two members and the low number of required points per week -- mean of five points -- produced too much variability in the percent completed curves. I still included these points in the total group average and the separate research task compilations.) Group G and Group W show a clear effect -- in both experimental phases for Group G the median percents completed are a minimum 26% higher than the baseline median; Group W is 21% higher.

The data recorded by the doctoral supervisors on the students' quantitative output on research tasks generally matched the effects

Figure 10: Percent of required points completed (dots connected with a solid line) and not completed (circles connected with a dashed line) in each week of the study for:

- a) Group G -- mean of 43 required points per week (range 34-47); N = 8.
- b) Group I -- mean of 42 required points per week (range 36-47); N = 5.



observed for the other dependent variables. Separate word totals on review article, log and hours' total on self-reported hours (students who were required 12 hours thesis work per week), were above the minimum requirements in the first experimental phase, equal to or below the minimum requirements in the baseline phase, and then above the minimum requirements in the second experimental phase. In this second study the word totals for writing and self-editing were far above the minimum in the first experimental phase, far below in baseline, and higher than the baseline results but still lower than the minimum requirements in the second experimental phase.

The low number in the range of required points per week for review article (Figure 8 Caption) resulted when the chief faculty sponsor allowed the students to forego a review article in order to concentrate on preparing weekly data for the small group meeting. This occurred in the last three weeks of the study -- in Figure 8 the percent completed data points for Weeks 12 and 13 contain only four and eight required points (which, in this case, equals four and eight subjects), respectively.

Six of the eight members of Group G were absent and contributed no data points for the Tuesday tasks of Week 11 and the Thursday tasks of Week 12 (Figure 10). These absences caused the low values in the range of required points per week in Figure Captions 8-10.

Table V shows the rate of tasks done extra across phases for Study I and Study II. There was a marked decrease in the second study in the number of extra tasks (down 75%) and the number of students who completed extra tasks (down 45%).

Table V: Record of tasks that earned extra positive points, and number of students contributing for each phase of Study I and II.

TABLE V

RECORD OF TASKS THAT EARNED EXTRA POSITIVE POINTS,
AND NUMBER OF STUDENTS CONTRIBUTING FOR EACH PHASE
OF STUDY I AND II

	Number of tasks done extra and number of contributing students	
	<u>Study I</u>	<u>Study II</u>
First Experimental Phase	4 subjects 9 tasks	3 subjects 3 tasks
Baseline Phase	3 subjects 7 tasks	2 subjects 2 tasks
Second Experimental Phase	4 subjects 7 tasks	1 subject 1 task
Mean for Entire Study	4 subjects 8 tasks	2 subjects 2 tasks

As a member of Group I, the experimenter's data are included in the results of Study I and Study II. The group data for both studies accurately represent the individual performance of the experimenter. I compared the median percents completed for Group I (Figures 5 and 10) and the experimenter's percent of points completed in each of the six individual phases of the two studies -- four of the six percent of points completed for the experimenter are with $\pm 10\%$ of the corresponding median percents completed for the group.

The doctoral supervisors and MA students filled out the questionnaire on the dependent variables and general procedures of the supervising system before the start of the study. Fourteen subjects scored a mean of 89% (range 75-100%) -- one student failed to take the questionnaire; the three supervisors scored a mean of 95% correct (range 90-100%).

Eleven of the 15 subjects filled out the evaluation at the end of the study. The results of the student evaluation were similar to those of the first study. The individual meeting with the doctoral supervisor again received the most favorable rating. The students indicated that this supervising system was well-organized; the behavioral contingency relationships of the two conditions, the phase changes, and the feedback forms were clear, while the points toward recommendation again received a less clear rating. Only one student reported ever finding an error on the feedback form; few students indicated they compared their group's performance to their own performance. Poor ratings still occurred for the large group meeting (remained too long and unstructured); and now, the students

indicated the small group meeting need time allotments so that each member could satisfactorily discuss their data and/or review article. Poor ratings occurred for the log summary and the self-editing procedures -- they were rated as ineffective and needless work; the weekly log earned mixed reviews again. The other research tasks received the subjects' approval. The subjects indicated very few acts of "hoarding" (that is, when the student completed a research task in the week but he or she did not turn it in to the supervisor until some later week). Few of the subjects completed any of the optional Graduate College procedures.

Nine of the 11 respondents (82%) said they preferred this type of supervising system -- with its weekly meetings, specification of research tasks, deadlines for completion of the tasks, and extrinsic consequences contingent on completion -- over a less structured approach. One subject (9%) disliked the aversive features of this type of supervision; one subject did not complete this question on the evaluation.

Discussion

The results from Study II again demonstrate that points toward recommendation combined with weekly feedback on the subject's performance were effective in controlling research activities. This systematic replication (Sidman, 1960) provides increased reliability of the experimental results. Twelve of the 15 subjects earned proportionately less negative points in the experimental condition than in the baseline condition. Subject performance on the separate research tasks better matched the contingency requirements present in the experimental and baseline conditions. There were clear effects for five of the seven tasks -- review article, log, writing, self-editing, individual meeting attendance; and only two equivocal effects -- data presentation, self-reported hours -- due to variability of the data. Both Group G and Group I showed a clear behavioral effect in this second study. The BAB design better demonstrated the potency of the treatment package -- 10 of the 13 weeks (77%) in Study II were in the experimental condition; yet, for six of the 10 graphs (seven research tasks and three groups) included in both studies the median percents completed in the first experimental phase were higher in Study II than in Study I. (In only two of the 10 graphs for Study II were the second experimental phase medians higher than for those of Study I, but the length of the Study II experimental phase was nearly twice as long.)

A couple of tentative explanations can be given for the better performance exhibited in Study II. First, the new point scale, that

increased the value of completion (and increased the aversive consequence for non-completion) of the writing and self-editing tasks (Figure 8), may have been responsible for the greatly improved control over these tasks. Group I's performance (Figure 10) that better matched the experimental and baseline condition requirements may have been similarly affected.

Second, this supervising system generated an average of eight to 14 hours per week of research behavior for 13 graduate students over seven months of time (the other four subjects participated for one study only). It is possible that the subjects' self-management techniques increased in effectiveness over this period of time.

Third, progress in the completion of a long term task is probably mildly rewarding. Those 13 subjects who participated in both studies were farther along towards final completion. (When asked to indicate the best feature of the supervising system on the final evaluation during Study II, one student wrote, "My thesis is near completion", and another, "I did some work on my thesis".) Successful completion of the research tasks in Study II earned this extra reward for each of these students.

And fourth, the weekly feedback form in Study II included a norm referenced component. However, in the evaluation only two of the students indicated that they consistently compared their performance with their group's performance across the semester. There were many more instances of 100% task completion for individual subjects in a week than for the group in which he or she was a member. The negative or corrective feedback value of the form probably decreased for

members when the group's percentage of points completed was below their own. A high probability may exist that low group performance serves to justify an individual's own low performance rather than acting as a guide for improved work. A possible solution would be to make the group feedback component cumulative -- the greater number of good performers would keep the average high.

The results of the data presentation task for this second study (Figure 9) show that the very high rate of percent completion did not continue across all conditions (see Figure 4). In Study II there were four students who prepared data for the entire 13 weeks, five students with five weeks of implementation, and two students with irregularly occurring tasks (each presented data three times across the semester). When data presentation became a weekly task with regularity for a majority of the subjects the point and feedback contingencies began to show an effect. If all subjects were required to do this task for all weeks of the study, the behavior might show a reversal in the baseline condition.

The combined results from the two studies provide strong evidence that the treatment package -- points toward recommendation and weekly feedback on performance -- can control the research behavior of masters' students. Sixteen of the 24 graphs (67%) from Figures 1 through 10 have shown clear effects; four of the 24 (17%) have shown equivocal effects; the remaining four have shown no effects present (17%). All of these separate parts -- total group, individual performance, the research tasks, the three groups based on thesis progress -- indicate the Behavioral Research-Supervising System can

lead to sustained progress on the preparatory tasks that are involved in doing a master's thesis.

Equally important as the empirical results is the participants own evaluation of the system. The defining features of the Behavioral Research-Supervising System are -- specification of the separate research tasks, placement of weekly deadlines for completion on the research tasks, presentation of extrinsic consequences following completion of the task, a weekly meeting schedule in which participants discuss research issues. Based on the two end of the semester evaluations, 78% of the students preferred this type of supervising system; 9% of the students considered certain aspects, such as deadlines and negative points, as too aversive to warrant continued use; 13% indicated neither a strong preference nor disapproval.

STUDY III

Studies I and II showed that the Behavioral Research-Supervising System was effective in controlling MA students' thesis research behavior. Study II compared the students' performance under the Behavioral Research-Supervising System with other graduate students from the Psychology Department under other supervising systems (collectively grouped in the Standard Supervising System). I developed a questionnaire that measured reported occurrence of research activities in the supervising system that the student worked in and the amount of supervision time provided by these systems. Although the subjects were not randomly selected, the distribution of a questionnaire sampled a large number of students, and these students represented a range of supervising systems within the Department. The MA students from the BRS System also filled out this same questionnaire.

Method

Subjects and setting

Seventy-two graduate students in the Department of Psychology at Western Michigan University responded to this questionnaire -- 57% were in the Applied Behavior Analysis program, 14% in the Experimental Behavior Analysis program, 18% in the Clinical program, 8% in the School Psychology program, 3% in the Industrial program. Approximately 85 students received a questionnaire.

There were two qualifications for filling out the questionnaire: 1) the student currently had to have graduate admission to an MA degree granting program in the Department of Psychology, Western Michigan University, or 2) if the student had already completed his or her thesis research, the work had to have been done in the Psychology Department at Western Michigan University. Roughly 20% of the respondents had already completed their thesis research (that is, the final thesis copy had been accepted by the Graduate College); all of these students were from the Standard Supervising System. Twenty-four percent of the 72 respondents were from the Behavioral Research-Supervising System.

I distributed the questionnaire in seven separate graduate classes, three weeks from the end of the Winter 1977 school semester. These classes averaged 10 to 15 students.

This approach did not provide a random sampling of all graduate students available from all supervising systems present in the

department, but a good cross-section of students (and therefore, supervising systems) was collected. Twenty-one systems were included in this study. This total represented 15 of 21 (71%) current faculty members and six faculty members no longer at Western Michigan University. The median number of graduate students from each supervising system was three. (Of the six faculty members no longer in the department, five of the six students had already completed their thesis research.)

General procedures

The professors teaching these classes allowed the students to complete the questionnaire at the start of the class period or at the end. No faculty member refused my request to hand out the form in their classes.

Prior to distributing the forms in the classroom, I explained the two possible qualifications for a respondent and the three categories on the informed consent form -- the data collected were being used for research purposes, a student could have his or her data withdrawn from the study at any time upon written request, and all data collected would remain anonymous since the results would be pooled.

The design of the questionnaire was as follows: the student filled in his or her name, date of entrance to the graduate program, under which faculty sponsor, and to which program. The student then answered how much individual, group, and other sources of research supervision he or she received in the sponsor's thesis supervising system. The student then indicated whether the system had specified

deadlines and consequences for completion and non-completion of the research tasks. The student gave the starting date of the literature review research, a pilot study or the formal implementation of the thesis design, and the write-up of the thesis document. Additionally, I asked if the student kept a log and tally of hours worked on thesis activities, prepared weekly data, had somebody else edit his or her writing; how many drafts were written, the dates of the orals meeting, final acceptance of the thesis by the Graduate College, and if the student took a full or part time class load during the thesis preparation. Dates were recorded to the month and year; the questionnaire took seven to 10 minutes to complete.

Results and Discussion

Advising system

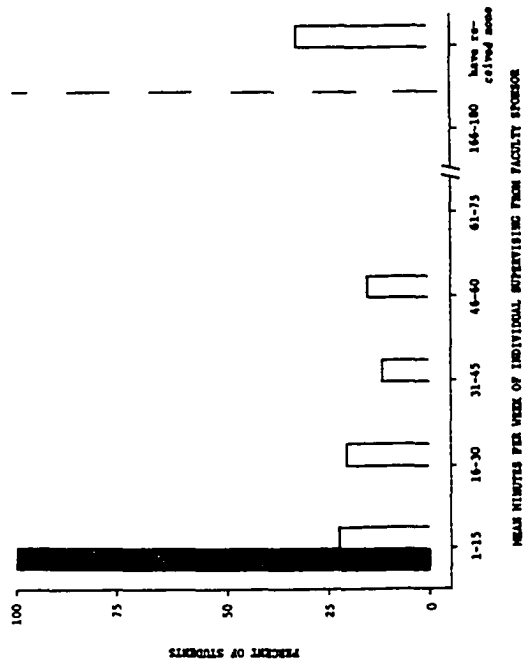
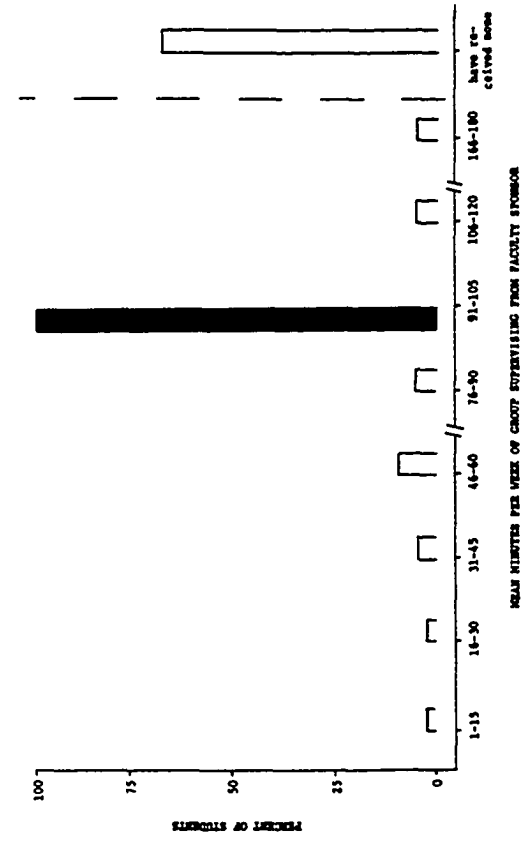
Based on the reported data of the questionnaire, 71% of the Behavioral Research-Supervising System students and 79% of the Standard System students were full time graduate students during the time of thesis preparation. In Figure 11, 68% of the Standard System students received some individual supervision from their faculty sponsor on thesis research; 100% of the BRS System students received this form of supervision. (Thirty-two percent of the Standard System students had received no research supervision to date -- see Table VI for statistical analysis of these students.) Thirty-two percent of the Standard Supervising System students received supervision in a group format; 100% of the BRS System students have received over 90 minutes per week of group supervision from their faculty sponsor.

Only 15% of the Standard System students received any other source of direct supervision -- Ph.D. student, Thesis Committee members, friends, research group in setting where design implemented -- whereas, 100% of the students in the BRS System had doctoral candidates as research supervisors. In the Standard System 25% of the students had deadlines set for research tasks and 3% had extrinsic consequences contingent on the deadlines; 100% of the BRS System students had weekly deadlines and consequences contingent on the deadlines.

Research tasks

Figure 12 shows the occurrence of specific research tasks as a

Figure 11: Mean minutes per week of supervision from faculty sponsor for Behavioral Research-Supervising System students and Standard Supervising System students in two different formats -- individual and group. There are 17 BRS System students and 54 Standard System students in each graph.



KEY

■ RESEARCH-SUPERVISING SYSTEM STUDENTS

□ STANDARD SUPERVISING SYSTEM STUDENTS

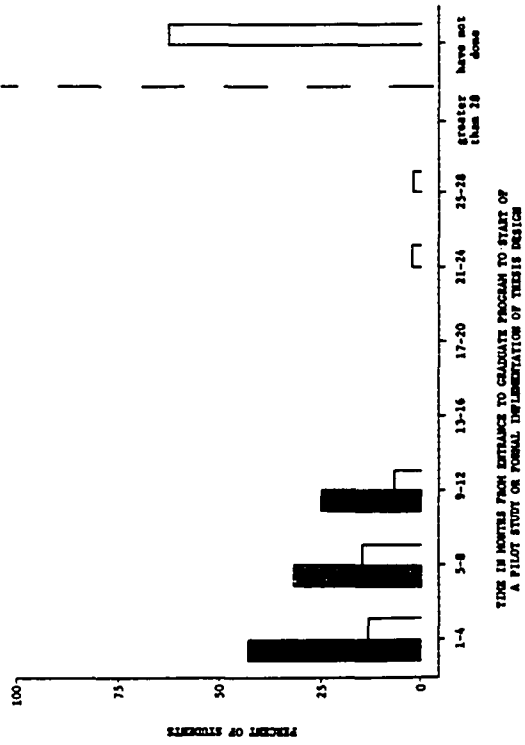
Table VI: An analysis of students from the standard supervising system who have not received any research supervision.

TABLE VI

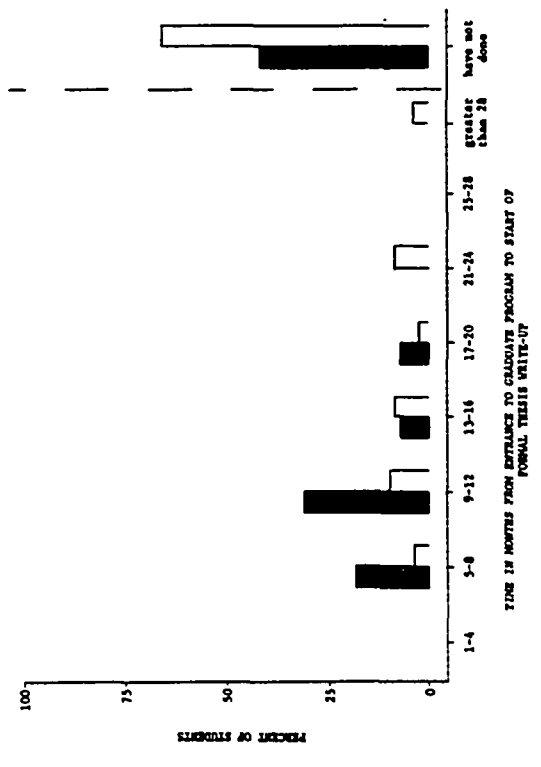
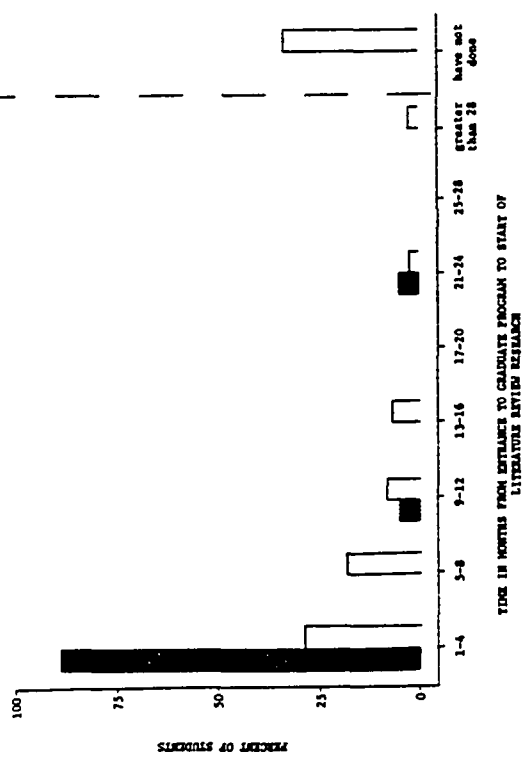
AN ANALYSIS OF STUDENTS FROM THE STANDARD SUPERVISING SYSTEM
WHO HAVE NOT RECEIVED ANY RESEARCH SUPERVISION

	<u>Percent</u>	<u>Months since entrance to graduate program</u>	
		<u>Median</u>	<u>Range</u>
Students with no individual supervision from faculty sponsor	32%	7 mos.	3-51 mos.
Students with no group supervision from faculty sponsor	67%	7 mos.	3-51 mos.

- Figure 12: Time in months from entrance to graduate program to start of three research tasks for Behavioral Research-Supervising System students and Standard Supervising System students.
- a) Literature Review Research -- there are 17 students from the BRS System, 51 students from the Standard System.
 - b) Pilot Study or Formal Implementation -- there are 16 students from the BRS System, 43 students from the Standard System.
 - c) Formal Thesis Write-Up -- there are 17 students from the BRS System, 53 students from the Standard System.



KEY
 ■ RESEARCH-SUPERVISING SYSTEM STUDENTS
 □ STANDARD SUPERVISING SYSTEM STUDENTS



function of the amount of time the student has been in the graduate program for the two supervising systems. Eighty-eight percent of the BRS System students started their literature review within the first semester (1-4 months) after admittance; 28% of the Standard System students did this within the first semester. Not shown on this graph is the fact that 41% of the BRS System students (compared to 14% of the Standard System students) started this task in the first month after admittance. (See Table VII for the statistical analysis of the students who have not started these research tasks.)¹

Seventy-five percent of the BRS System students started a pilot study or their formal thesis implementation within the first two semesters (1-8 months) after admittance to the graduate program; 26% of the Standard System students started this task within the first two semesters.

Forty-eight percent of the BRS System students started their formal write-up within one year (1-12 months) after admittance; 13% of the Standard System students commenced the formal write-up in this time span. In the BRS System 94% of the students kept a weekly research log and 100% kept a weekly tally of hours worked on thesis activities; the totals for the Standard System students were 26% and 11%, respectively.

¹One student from the Standard Supervising System started his literature review research and implementation of a pilot study four months prior to his entrance into the graduate program.

Table VII: An analysis of students from the Behavioral Research-Supervising System and Standard Supervising System who have not started the research tasks.

TABLE VII

AN ANALYSIS OF STUDENTS FROM THE BEHAVIORAL RESEARCH-SUPERVISING SYSTEM AND STANDARD SUPERVISING SYSTEM WHO HAVE NOT STARTED THE RESEARCH TASKS

	<u>Percent</u>	<u>Months since entrance to graduate program</u>	
		<u>Median</u>	<u>Range</u>
Students who have not started literature review research			
BRS:	0%	—	—
Standard:	33%	7 mos.	3-19 mos.
Students who have not started a pilot study or implementation of thesis			
BRS:	0%	—	—
Standard:	62%	7 mos.	3-21 mos.
Students who have not started formal thesis write-up			
BRS:	41%	7 mos.	3-15 mos.
Standard:	66%	7 mos.	3-21 mos.

Conclusions

The results for Study III show that MA students in the Behavioral Research-Supervising System generally receive more direct research supervision -- more absolute meeting time and from more sources -- than the Standard Supervising System students. Measured from entrance to the graduate program, the students in the BRS System generally start to work on the early component tasks that lead to a completed Master's Thesis within a shorter time interval than the Standard Supervising System students. Also, the BRS System students continue working on these preparatory tasks so that steady, weekly progress toward completion was maintained (these results were shown in Studies I and II). It is probable that the weekly output of the Standard Supervising System students is not nearly as regular across semesters, and steady progress is not maintained.

GENERAL CONCLUSIONS

The defining features of the Behavioral Research-Supervising System are -- specification of the separate research tasks, placement of weekly deadlines for completion on the research tasks, presentation of extrinsic consequences following completion of the tasks, a weekly meeting schedule in which the participants discuss research issues.

The results from Studies I and II show that points earned toward recommendation letters combined with weekly feedback on the subjects' performance are effective in controlling research behavior. All levels of analysis in the system -- total group average, individual performance, the eight research tasks, the three groups based on thesis progress -- support this interpretation. Over 75% of the subjects approved this type of research supervising system. The results from Study III show that students in the Behavioral Research-Supervising System receive more direct supervision time, more sources of supervision, and start work on research activities within a shorter time interval from admittance to the program than other comparable graduate students in the department.

From the results of the three studies it would seem that the Behavioral Research-Supervising System could produce these outcomes.

- 1) Better quality masters' theses -- for example, there would be a more thorough literature review, a clearer writing style, and better functional analyses of the independent and dependent variables; the thesis would make a contribution to the field of behavior analysis

and it would have a higher probability of publication and presentation. 2) The final product would be completed in a shorter period of time. These two issues are assertions -- they have not been proven -- but they may serve to guide future research and technological development in this system.

RECOMMENDATIONS

I present the following as suggestions for further research and technological development of the Behavioral Research-Supervising System:

1) The first research proposal for the new MA student should address this issue -- how do we know we have a problem? -- the student would gather descriptive data to show the setting warranted a master's thesis in order to change the problem. The first proposal would outline this prepilot investigation.

2) If the results indicated a problem, the student would prepare the formal research proposal. After editing by the doctoral supervisor, the student would send a copy to his or her Orals Committee and would arrange to have a group meeting or individual meetings to discuss the proposal.

3) Based on student feedback, a second proposal due in four weeks is unnecessary; it is better to keep revising the initial one.

4) The chief problem with the 'implementing' tasks is that their completion is not related closely enough to the 'writing' tasks. As now designed, the output documents for review article, data presentation, log, and self-reported hours are not optimally useful when the student is writing the manuscript. These preparatory tasks must be completed in such a way that each effectively controls the researcher's behavior when preparing the thesis write-up. For example, the Literature Review form (Appendix C) is important for analyzing the merits of each article read. But a sheaf of these

completed forms probably does not most effectively help the student write the Introduction section of a thesis. I propose another form (Appendix E -- Research Article List) that contains a section for the complete citation of the article (this easily becomes the thesis Reference page) and a section with specific categories for indicating the possible relevance of the article read to the student's own thesis.

5) Several subjects in the evaluation commented that they sometimes read marginally useful or relevant articles during the semester just to avoid the negative points. I think the first requirement of the semester for the review article task should be the completion of 14 articles, to be read one article each week in that semester. The supervisor would approve the articles on the list at the first small group meeting (or delete some and suggest others).

6) Data presentation is the most closely related preparatory task to the final thesis write-up. It will determine the scope and particular emphasis of analysis in the write-up. In the present system though, I think formal data analysis remains unspecified. A preliminary step in this direction would have the MA student list every graph he or she is recording data on and, at the end of the semester, the student would state the results in terms of the independent variable(s), the experimental design, and implications for future research.

7) I also think that a presentation of the results with graphs to the thesis group should be a semesterly requirement of the data

presentation task.

8) Students were critical of the short time for the small group meeting on Thursday. This meeting should not be lengthened but an agenda should be followed with time allotments provided for each student in the group, because students criticized the discussion drift into non-research areas.

9) All of the Graduate College procedures that are generally applicable to all the MA students should have group-wide due dates for completion. These would be tasks like submitting the Candidacy form, taking the English Qualifying test and a meeting with each of the faculty members on the student's committee (once a semester).

10) The students criticized the weekly log as unnecessary and rarely relevant to the thesis preparation. The log summary, as a review technique, was called redundant and of no value. Again, I believe the problem is that the proper form has not been designed for this research task. All the log entries made during implementation must be in proper format for the writing stage. (Similarly, a log form should be designed to make the output from the generating stage valuable for the researcher when implementing.) The categories of the weekly log form could be the sections normally found in a thesis write-up -- introduction, subjects, reliability, discussion, etc. -- the MA student records an entry under, say, five of the eight selected categories with a 125 word minimum total. Then for the log summary (due twice a semester), the student transfers the most relevant entries to a master sheet that then becomes a thesis outline for preparing each thesis section.

11) Also included with the log should be an implementation sheet; this shows the four months of the current semester and on this calendar the student circles the exact date of important events related to the thesis implementation -- for example, condition changes, when specific handouts were distributed, when subjects added or withdrew from experiment).

12) In the student evaluation for self-reported hours, the students neither praised nor criticized this task. I believe that a prospective behavioral scientist should try to get his or her professional productive behavior under the cue control of a graph and other quantitative measures. This is a good step in the development of effective self-management techniques (Foster, 1974). The self-reported hours form (Appendix F) will show the time spent per day and per week on all research activities from this supervising system.

13) The weekly totals could be graphed across the 15 weeks of the semester or on a yearly (52 weeks) basis as part of the response requirement.

14) Generating a thesis design leads to implementing, and implementing leads to writing. The write-up is the most important document because it is the summary of all the other stages. By this time preparation of the final write-up should already be assured -- the log summaries function as a thesis outline; the Introduction section comes from the Research Article list; the Method is written during the semester the student implements; Results and Discussion come from the summary and the data presentation and data analyses tasks. I believe that the weekly word total for writing could now rise to

1500 words, so that the student could complete the write-up in the semester following implementation. I would not want to increase the point value of this task any more because the single point tasks would then start to lose their significance as points completed on the feedback form. The cue control of research behavior could then be diminished.

15) The self-editing task, as currently designed, is of questionable utility. The MA students rated the requirements as confusing, and the self-edited words as colorful but with no editing rules or intuitive behavior developed as a result. I think the most important areas of emphasis should be active voice, topic sentences for paragraph organization, and short, cumulative sentences. These seem to be the most clearly evident in "readable" writing.

16) A couple of students each semester still had trouble completing all of their required research tasks. These are two possible changes in the supervising system. First, the weekly feedback form could also show cumulative group performance. Because the personal performance of these students was consistently below that of their group, this would presumably provide them with more effective negative or corrective feedback. Second, the present weekly deadlines for task completion could be broken down into even smaller time intervals. For example, if the student typically does not complete the writing and editing tasks (which are due on Tuesday of the thesis week), then 250 words written and edited could be due at the Thursday small group meeting and the remainder due at the normal time of the individual meeting. The next level could be 150 words written and edited due on

each weekday of the thesis week (thus providing five deadlines).

For further research approaches I would like to see the research behavior of the generating students (Group G) tracked as they enter their second year under the system and compared to the performance of the current Group I members. Similarly, one could compare the data of new Group G members to the current Group G's performance. Of particular importance to the system would be the research 'performance' of the recently completed students who have left the system for other graduate programs or jobs. This canvassing could be in the form of a questionnaire to BRS participants and to other graduate students from the department. Would the subjects from this supervising system be more likely to set up similar systems in order to ensure completion of long term projects they are now facing? What is the research record of the participants when not under the system -- would they tend to do more or less than other graduate students?

These questions have important implications for the training of scientists. I believe that the quickest approach to developing scientific behavior is the establishment of rule-governed behavior (Skinner, 1969). This behavior will become maximally effective when shaped by the contingencies of actually doing research. However, the development of contingency-shaped behavior takes time. It is doubtful that the intrinsic rewards produced by the research activities (the most important probably being the learned reward of control (Malott et al., in press)) will maintain the rate of completion over this long period of time. Research behavior is developed and maintained primarily by extrinsic consequences. The graduate students have

grades, stipends, and attention and approval from faculty members. Research behavior on the faculty level is maintained by salaries, subordinate staff members, audiences at conventions, in classrooms and from journal readership when publishing. The control of scientific behavior by intrinsic rewards is not an unimportant goal because effective consequences no longer have to be arranged by the community. But control by extrinsic consequences is foremost.

Therefore, in order to support maintenance of the subject's research behavior, the supervising system should do this: the graduate students should be encouraged to present their research in front of an audience. Presentation is a professional activity that is rewarded by the scientific community; it can become a powerful reward for research activities that extend over many months. The weekly, small group, and large group meetings in the supervising system already provide some social rewards for discussing research. The next stage in the development of presentation skills would be the formal presentation of the results of the thesis implementation (with graphs, tables, handouts) in the large group meeting. This situation tends to be very "supportive" -- that is, a lot of positive feedback and reward statements. Next would come a presentation at a departmental colloquia, and then to regional and national conventions. Similarly, graduate students who have completed their thesis work could be asked to return for a presentation of this research to the large group in the Thursday meeting (or of research they are currently engaged in). This approach may be the most expedient way to program maintenance of the effects of the supervising system.

Of interest to the system designer is a component analysis of the two independent variables. If points toward recommendation was effective by itself then the weekly task of preparing written feedback could be eliminated. What effect would feedback alone have or are both variables needed in concert to effectively control research behavior? What would the addition of cumulative group performance, as a category on the feedback form, have on the performance of the group members?

Of major theoretical importance to the field is the precise contribution of rule control in this system.

REFERENCES

- Baer, D. M., Wolf, M. M. and Risley, T. R. "Some current dimensions of applied behavior analysis." Journal of Applied Behavior Analysis, I (Spring 1968), 91-97.
- Bijou, S. W., Peterson, R. F. and Ault, M. H. "A method to integrate descriptive and experimental field studies at the level of data and empirical concepts." Journal of Applied Behavior Analysis, I (Summer 1968), 175-191.
- Childers, P. R. and Haas, V. J. "Effect of detailed guidance on the writing efficiency of college freshmen." Journal of Experimental Education, XXXIX (Fall 1970), 20-23.
- Foster, C. Developing Self-Control. Kalamazoo, Michigan: Behaviordelia, Inc., 1974, pp. viii + 136.
- Kazdin, A. E. "Methodological and assessment considerations in evaluating reinforcement programs in applied settings." Journal of Applied Behavior Analysis, VI (Fall 1973), 517-531.
- Keller, F. S. "Goodbye teacher..." Journal of Applied Behavior Analysis, I (Spring 1968), 79-89.
- Knezevich, S. J. Program Budgeting (PPBS). Berkeley, California: McCutchan Publishing Corporation, 1973, pp. xi + 335.
- Lloyd, K. E. and Knutzen, N. J. "A self-paced programmed undergraduate course in the experimental analysis of behavior." Journal of Applied Behavior Analysis, II (Summer 1969), 125-133.
- Malott, R. W. A behavioral -systems approach to the design of human services. In D. Harshbarger and R. Maley (Eds.), Behavior Analysis and Systems Analysis: An Integrative Approach to Mental Health Programs. Kalamazoo, Michigan: Behaviordelia, Inc., 1974, pp. vii + 403.
- Malott, R. W. and Whaley, D. L. Psychology. New York: Harper's College Press, 1976, pp. xviii + 684.
- Meyers, A. W., Artz, L. M. and Craighead, W. E. "The effects of instructions, incentive and feedback on a community problem: dormitory noise." Journal of Applied Behavior Analysis, IX (Winter 1976), 445-457.
- Miller, L. K., Weaver, F. H. and Semb, G. "A procedure for maintaining student progress in a personalized university course." Journal of Applied Behavior Analysis, VII (Spring 1974), 87-91.

Panyon, M., Boozer, H. and Morris, N. "Feedback to attendants as a reinforcer for applying operant techniques." Journal of Applied Behavior Analysis, III (Spring 1970), 1-4.

Sidman, M. Tactics of Scientific Research. New York: Basic Books, Inc., 1960, pp. x + 428.

Skinner, B. F. Contingencies of Reinforcement. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1969, pp. xv + 319.

REFERENCE NOTES

Malott, R. W., Tillema, M. and Glenn, S. An Introduction to the Analysis and Modification of Behavior. (In press)

Tillema, M. "Validation of instructional materials over free modifiers and cumulative sentences, and the effects of contingencies on using free modifiers and cumulative sentences." Master's thesis, Western Michigan University. (In preparation)

APPENDICES

.....

..

APPENDIX A

Weekly Supervisor Form

September

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20
21	22	23	24
25	26	27	28
29	30		

November

31	1	2	3	4	5	6	WK 6
7	8	9	10	11	12	13	WK 7
14	15	16	17	18	19	20	WK 8
21	22	23	24	25	26	27	WK 9
28	29	30					

October

	1	2	WK 1
3	4	5	6
7	8	9	WK 2
10	11	12	13
14	15	16	WK 3
17	18	19	20
21	22	23	WK 4
24	25	26	27
28	29	30	WK 5

December

	1	2	3	4	WK 10
5	6	7	8	9	10
11	12	13	14	15	16
17	18	19	20	21	22
23	24	25	26	27	28
29	30	31			

Supervisor _____

Thursday

Student	At-tend	Review Art.	Date
1)			
		wds	
2)			
		wds	
3)			
		wds	
4)			
		wds	
5)			
		wds	

Tuesday

Attend	Log Hrs.	750 Wds.	Self-Edit	Pos.	Neg.
wds	hrs	wds	wds		
wds	hrs	wds	wds		
wds	hrs	wds	wds		
wds	hrs	wds	wds		

APPENDIX B

Individual Feedback Form

<u>September</u>											<u>November</u>								
	1	2	3	4	5	6	7	8	9	10	11	31	1	2	3	4	5	6	WK 6
5	6	7	8	9	10	11	7	8	9	10	11	12	13	WK 7					
12	13	14	15	16	17	18	14	15	16	17	18	19	20	WK 8					
19	20	21	22	23	24	25	21	22	23	24	25	26	27	WK 9					
26	27	28	29	30	28	29	30												
<u>October</u>											<u>December</u>								
							1	2	WK 1	1	2	3	4	WK 10					
3	4	5	6	7	8	9	WK 2	5	6	7	8	9	10	11	WK 11				
10	11	12	13	14	15	16	WK 3	12	13	14	15	16	17	18					
17	18	19	20	21	22	23	WK 4	19	20	21	22	23	24	25					
24	25	26	27	28	29	30	WK 5	26	27	28	29	30	31						

Student _____ Supervisor _____

Week N	At-tend	Review Art.	Data	Attend	Log	Hrs.	750 Wds.	Self - Edit	Pos.	Neg.

Cumulative										
earned/possible +'s										-
earned -'s										-

- +1....response met criterion for that week (or behavior that was done extra)
- 1....response did not meet criterion for that week
-behavior not applicable to student

You are in the experimental or 'B' treatment condition at this time: positive and negative points are contingent on your performance; you will receive this feedback sheet weekly on your thesis work. Only under this condition of the experiment will data be supplied to Dr. Malott in determining recommendations. Data from the 'A' condition (which must be collected) will never see the light of day associated with your name.

APPENDIX C

Literature Review Form

TITLE: _____

AUTHOR(S): _____

JOURNAL: _____ YEAR _____ Vol. _____ PP. _____

AREA OR PROBLEM:

RESEARCH QUESTION:

SUBJECTS & SETTING:

DEPENDENT VARIABLE (specification, observation, reliability):

INDEPENDENT VARIABLE (specification, observation, reliability):

EXPERIMENTAL DESIGN:

RESULTS:

CONCLUSIONS:

RELATED SIGNIFICANCE:

POSSIBLE REFERENCES:

APPENDIX D

Reliability Sheet

<u>September</u>											<u>November</u>							
	1	2	3	4							31	1	2	3	4	5	6	WK 6
5	6	7	8	9	10	11					7	8	9	10	11	12	13	WK 7
12	13	14	15	16	17	18					14	15	16	17	18	19	20	WK 8
19	20	21	22	23	24	25					21	22	23	24	25	26	27	WK 9
26	27	28	29	30							28	29	30					
<u>October</u>											<u>December</u>							
					1	2	WK 1						1	2	3	4	WK 10	
3	4	5	6	7	8	9	WK 2				5	6	7	8	9	10	11	WK 11
10	11	12	13	14	15	16	WK 3				12	13	14	15	16	17	18	
17	18	19	20	21	22	23	WK 4				19	20	21	22	23	24	25	
24	25	26	27	28	29	30	WK 5				26	27	28	29	30	31		

Supervisor _____

MA Student _____

Secondary Observer _____

Condition _____

	<u>Thursday</u>			<u>Tuesday</u>						
	At-	Review	Data	Attend	Log	Hrs.	750	Self-	Pos.	Neg.
Secondary Observer:	tend	Art.					Wds.	Edit		
		wds		wds	hrs		wds	wds		
Primary Observer:										
		wds		wds	hrs		wds	wds		
Agreements:										
Disagreements:										

APPENDIX E

Research Article List

<u>ARTICLE</u>	<u>RELEVANCE</u>
1) Title:	DV(s) -
	IV(s)/Setting -
Authors:	Subjects -
	Results -
Journal:	Other Issues -
2) Title:	DV(s) -
	IV(s)/Setting -
Authors:	Subjects -
	Results -
Journal:	Other Issues -
3) Title:	DV(s) -
	IV(s)/Setting -
Authors:	Subjects -
	Results -
Journal:	Other Issues -
4) Title:	DV(s) -
	IV(s)/Setting -
Authors:	Subjects -
	Results -
Journal:	Other Issues -

APPENDIX F

Hours Form

	Wednesday	Thursday	Friday	Saturday	Sunday	Monday	Tuesday	
Meetings								<u>hrs</u>
<u>wds</u> Log								<u>hrs</u>
<u>wds</u> 750 wds								<u>hrs</u>
<u>wds</u> Res Prop								<u>hrs</u>
<u>wds</u> Self-edit								<u>hrs</u>
Imple- menting								<u>hrs</u>
<u>wds</u> Lit Rev								<u>hrs</u>
Data Pres								<u>hrs</u>
Grad Col Procds								<u>hrs</u>
<u>Tot wds</u>								<u>Tot hrs</u>

Tot hrs