References

Aronson, E., Turner, J., & Carlsmith, J. Communicator credibility and communication discrepancy as determinants of opinion change. *Journal of Abnormal and Social Psychology*, 1963, *67*, 31-36.

Carroll, J. Assessing the effectiveness of a training program for the university teaching assistant. *Teaching of Psychology*, 1977, *4*, 135-138.

Carsrud, A. Undergraduate tutors: Are they useful? *Teaching of Psychology*, 1979, *6*, 46-49.

Eagly, A. Comprehensibility of persuasive arguments as a determinant of opinion change. *Journal of Personality and Social Psychology*, 1974, *29*, 758-773.

Fremouw, W., Millard, W., & Donahoe, J. Learning-through-teaching: Knowledge changes in undergraduate teaching assistants. *Teaching of Psychology*, 1979, *6*, 30-32.

Janis, I., & Field, P. A behavioral assessment of persuasibility: Consistency of individual differences. *Sociometry*, 1956, *19*, 241-259.

Katz, A. Inexpensive animal learning exercises for huge introductory laboratory classes. *Teaching of Psychology*, 1978, *5*, 91-93.

Murray, H. Predicting student ratings of college teaching from peer ratings of personality types. *Teaching of Psychology*, 1975, *2*, 66-69.

Murray, H. Classroom behaviors of social science lecturers receiving low, medium and high teacher ratings. *Ontario Universities Program in Instructional Development Newsletter*, 1977, *Number 14*.

White, K., & Kolber, R. Undergraduate and graduate students as discussion section leaders. *Teaching of Psychology*, 1978, *5*, 6-9.

Note

Address requests for reprints to Dr. Albert N. Katz, Department of Psychology, University of Western Ontario, London, Ontario N6A 5C2, Canada.

A Behavioral System for Supervising Undergraduate Research

Gary D. Gant, Michael J. Dillon, and Richard W. Malott
Western Michigan University

Five of eight students in a systematic behavioral control course showed 20% better performance during the intervention phase.

Most of us, including students, often have trouble completing long-term research projects because rewards may be weak or infrequent, deadlines absent, and other more immediately pressing activities intrude. To overcome such difficulties, we designed a behavioral research supervisory system for undergraduate students based in part on a supervisory system for students conducting master's thesis research. (Dillon, Kent, & Malott, in press). This supervisory system for undergraduate research had four basic features:

1. Written descriptions of required tasks, criteria for task completion, and systems procedures. The students received these written descriptions at the beginning of this study. Such written descriptions have improved desired performance in other settings (Pommer & Streedbeck, 1974; Meyers, Artz & Craighead, 1976).

2. Deadlines for completion of required tasks. We specified weekly deadlines for most tasks with less frequent deadlines for the remaining tasks. Others have suggested the general value of deadlines (Van Houten, Hill & Parsons, 1975; Miller, Weaver & Semb, 1974; Malott, 1971).

3. Rewards and aversives[2] for completing and not completing required tasks. On a weekly basis, the students received performance records on their number of complete and incomplete tasks for each particular week and their cumulative complete and incomplete percentages across the semester. Such performance records have proven useful in other settings (Panyan, Boozer & Morris, 1970; Parsonson, Baer & Baer, 1974).

At the beginning of this study, we also announced that any requested letters of recommendation would include these cumulative performance records of all students. So, these records may have functioned in a manner similar to quiz scores in an academic course—with favorable records acting as rewards and unfavorable records acting as aversives, due, in part, to the value-augmenting effects of the announcement of their future use in letters of recommendation. Performance records combined with other delayed uses have improved the completion of a variety of tasks (Pierce & Risley, 1974; Pommer & Streedbeck, 1974).

4. Weekly meetings between each undergraduate student and a graduate supervisor. During weekly meetings, the supervisors monitored tasks, marked task-completion graphs, and reviewed requirements for the next week. Similarly, Panyan, et al. (1970) improved the percentage of tasks completed by attendants at an institution for retarded children with a weekly review of performance records and their public posting.

The goal of this supervisory system was to ensure that undergraduate students steadily performed all activities needed for completing research projects. This study also experimentally examined effects on the percentages of specified research tasks completed and not completed by the students after announcements that letters of recommendation would either include or not include performance records.

Method. We conducted this study in the Psychology Department of Western Michigan University. The setting within the department was the Student Centered Educational Program (SCEP), an accelerated program for 100 first and second year undergraduate students taught by one faculty advisor and a staff of six graduate and 29 undergraduate students.

The subjects were eight undergraduate junior and senior psychology majors—six SCEP staff members and two students who had no involvement in SCEP other than their research. Five of the students received college credit for their research projects, two received one-year research grants from the University, and one received both the credit and the grant. All students conducted educational technology projects in SCEP. The students gave their informed consent to participate at the beginning of the study.

The research supervisors were the six graduate students of the SCEP staff—two PhD candidates and four MA candidates in the Psychology Department. The supervisors received practicum credit for their supervisory activities. They also gave consent to participate.

In written guidelines[3] provided to the students, we specified recurring tasks, which occurred nearly every week, and periodic tasks, which occurred less frequently; during the supervisory meetings, the supervisors also assigned a variety of nonrecurring tasks which were particular to each research project. The following paragraphs list the three groups of tasks, and Table 1 relates them to four general classes of research behavior, each important for training scientists.

Drawing presentation-quality graphs; (8) Making a public presentation of the results of the research project to an audience of psychology students and professors; (9) Writing a second preliminary proposal (for a research project the next semester); and (10) Completing a research system evaluation.

Examples of nonrecurring tasks specified by the supervisors included meeting with professors and preparing handouts for the students' subjects.

Table 1 categorizes required recurring, periodic, and nonrecurring tasks under four classes of research behavior: generating a research design, implementing the design, writing a report of the results, and presenting the results to an audience.

This study examined the effects on task-completion data of announcements that the Program Coordinator for SCEP (a PhD candidate) would include or not include all students' task-completion data in letters of recommendation. The first and third phases were intervention phases during which all task-completion data were included in letters of recommendation; the second phase was a baseline phase during which the task-completion data did not contribute to the letters of recommendation. The consent form described the independent variable and the two conditions, stating that at some point in the study the independent variable would be removed. The first phase lasted for eight weeks, the second phase for four weeks, and the third phase for three weeks. At the start of the second phase (baseline), we announced that task-completion data over the next few weeks would not be included in the letters of recommendation, although we

Table 1
Research Tasks Categorized Under Four Classes of Research Behavior

Generate Design	Implement Project	Write Report	Present Results
-preliminary proposal	-implementation schedule	-writing schedule	-committee review
-full proposal	-data reporting	-literature reviews	-presentation-quality graphs
-review by committee	-log reporting	-report writing	-public presentation
-literature review	-nonrecurring tasks: consent forms, writing quizzes, questionnaires, phase changes, reliability	-editing	-nonrecurring tasks: posters, transparencies, speech outlines
-nonrecurring tasks: alternate designs, meetings, discussions		-presentation-quality graphs	
		-nonrecurring tasks: flowcharts, rough drafts, statistical tests	

The eight weekly recurring tasks were the following: (1) Attending a 30-minute supervisory meeting; (2) Reporting the total hours worked (four hours minimum required); (3) Writing a literature review (100 words minimum); (4) Writing 200 words on a final report; (5) Editing paragraphs of the final report for topic sentences and supporting sentences; (6) Recording project events in a log (75 words minimum); (7) Displaying data; and (8) Marking the task-completion graph.

The ten periodic tasks were the following: (1) Taking a quiz on written descriptions of the research tasks and procedures; (2) Presenting a research proposal to a research review committee; (3) Writing a preliminary research proposal (70 words minimum); (4) Writing a final proposal (150 words minimum); (5) Developing an implementation schedule (with deadlines for each event); (6) Developing a writing schedule (with deadlines for each section); (7) expected the students to continue to complete their required weekly research tasks.

The supervisors also served as the primary observers. Their training consisted of reading the written descriptions of tasks and procedures for the supervisors and the students and taking a quiz over these descriptions. The first two authors served as secondary observers. To conduct reliability checks, we attended meetings between supervisors and students on an unannounced schedule and inspected the required tasks. To compute reliability percentages, we divided the number of agreements by the sum of agreements and disagreements and multiplied the quotient by 100.

The mean overall reliability between observations of the primary and secondary observers was 88% with a range of 50-100% over the 18 individual checks. Seven reliability checks occurred in the first phase, none in the second, and

11 in the third. We observed 18% of the total number of required research tasks in this study (that is, 136 of 739 tasks).

Results. The students completed an average of 12% more research tasks during the intervention phases, that is, when we announced that their letters of recommendation would include their task-completion data for those phases, and they failed to complete an average of 10% less tasks during those intervention phases (see Figure 1). (Note that the students could complete more than 100% of their required tasks in a week by doing extra tasks, but extra tasks were

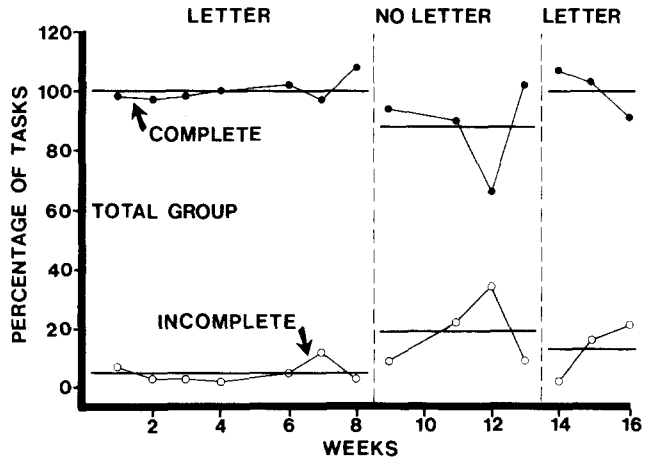

Figure 1. Mean percentage of complete and incomplete tasks for the total group across the 16 weeks of the study. Horizontal lines represent the mean for the phase. No data points appear on the fifth week, due to a blizzard, and on the tenth week, due to the University's Spring Break.

not considered when computing the percentage of incomplete tasks.) We performed a Friedman's two-way analysis of variance by ranks test. For the percentage of tasks complete, it showed $\chi_r^2 = 8.06$ at $p = .018$ level; for the percentage of tasks incomplete, $\chi_r^2 = 5.81$ at the $p = .055$ level.

A more detailed analysis also supports these general findings. Data for five of the eight students (63%) showed experimental control. On the average for those five students, the means of the intervention phases differed from the means of the baseline phases by 20%.

Although the data for three students failed to show experimental control, all three performed well. One student completed all of the tasks in the baseline phase and failed to complete only 4% in intervention; the second student failed to complete only 6% of the tasks in the baseline phase and 3% in intervention; the third student failed to complete 12% in the baseline phase and 11% in the intervention.

Six of the eight recurring tasks (75%) showed experimental control with the means of the intervention phases differing from the means of the baseline phase by 31%, while the data for meeting attendance and recording in a project log failed to show an improvement from the baseline phase to the following intervention phase. The announcements also showed control for the nonrecurring tasks with an average difference among the phase means of 23%, although for the percentage of incomplete tasks, the difference between the second and third phases was only 1%. The announcements, however, did not control the periodic tasks.

Discussion. The supervisory system did ensure that undergraduate students steadily completed activities in the four classes of research behaviors; they designed projects, implemented them, wrote reports on them, and presented them to professional audiences. Additionally, in an anonymous evaluation of the system, the students and the supervisors gave good ratings to most features, preferring this type of system with its specification of tasks, use of deadlines and weekly meetings, and presentation of rewards and aversives to no system at all.

Even though performance levels in intervention were higher than in the baseline phase, some remaining rewards or aversives may have contributed to maintaining considerable research activity during that phase. First, the students' control of the projects remained as a possible learned reward for completing research tasks (Malott, 1973; Malott, Tillema & Glenn, 1978). Evidence of control might have included predicted changes in the project data, desired results discovered in similar studies in the research literature, or their own high performance records. Although, as described in the introduction, the reward or aversive value of the performance records may have partly depended on their announced use in letters of recommendation, the records may have retained some value as a particularly distinctive evidence of control. Second, because the students were manipulating educational variables, any deterioration in their project procedures may have produced aversive reactions from the students' subjects or other staff members of SCEP. Nonetheless, the results of this study indicate that these two possible sources of rewards or aversives alone may be insufficient to maintain high levels of research activity.

An evaluative cost analysis (Krapfl, 1974) indicates that the educational benefits of this system probably outweigh the costs. As a rough approximation, we can compute the benefits in terms of hours of educational activity on the part of the undergraduate students and on the part of the graduate supervisors. According to their reports, the combined output of the eight undergraduate students averaged 64 hours per week, and the combined output of the six graduate supervisors averaged 12 hours per week; so the total output was 76 hours of fairly high-level activity per week for 14 advanced students. For purposes of comparison, we can view this combined output as what we would expect from 8.4 students in a three credit-hour course, assuming that each student would work about nine hours per week ($8.4 \times 9 = 76$). We can estimate the major part of the direct costs by looking at the average time of the graduate student who served as the research coordinator; that time was six hours per week. So in six hours per week, one graduate assistant can manage the equivalent of a high-level research or practicum course for 8.4 students. This seems like a reasonable investment of assistantship dollars and time.

As a refinement of the present study, one might assign point values to each requirement, weighing hard-to-control activities more heavily.

Others might replicate the four basic features of this study (written descriptions, deadlines, added rewards and aversives, and weekly monitoring) with different types of long-range projects (such as grant or program proposals or financial reports), different populations (behavioral health workers, nurses, or administrators), or different settings (behavioral health institutions or community-based settings).

We might improve the quality of students' output in all four areas of research (generating and implementing a design, writing and presenting the results) by requiring the appropriate observable activities: For example, quality writing might include frequent use of active voice (see APA, 1974) and the use of "free modifiers" (words set off from the main clauses of sentences by commas, dashes, or parentheses) and "cumulative sentences" (free modifiers appearing at the end of sentences) (see Tillema, 1977; Christensen, 1968; Note 4).

Additional areas of research might include studying the concurrent effects of this supervisory system on students' behavior (for example, do students increase their total activity level or drop other activities when they conduct research under a system like this) and studying future effects (do students use similar systems to resolve problems of procrastination in other areas for themselves, friends, or co-workers).

References

American Psychological Association. *Publication Manual* (2nd ed.). Washington, DC: Author, 1974.

Christensen, F. The problem of defining a mature style. *English Journal,* 1968, *57,* 572-579.

Dillon, M. J., Kent, H. M., & Malott, R. W. A supervisory system for accomplishing long-range projects: An application to master's thesis research. *Journal of Organizational Behavior Management,* (in press).

Krapfl, J. E. Accountability through cost-benefit analysis. In D. Harshbarger & R. F. Maley (Eds.), *Behavior analysis and systems analysis: An integrative approach to mental health programs*. Kalamazoo, MI: Behaviordelia, Inc., 1974.

Malott, R. W. *Contingency management in education* (2nd ed.). Kalamazoo, MI: Behaviordelia, Inc., 1971

Malott, R. W. *Humanistic behaviorism and social psychology.* Kalamazoo, MI: Behaviordelia, Inc., 1973.

Malott, R., Tillema, M., & Glenn, S. *Behavior analysis and behavior modification: An introduction.* Kalamazoo, MI: Behaviordelia, Inc., 1978.

Miller, L. K., Weaver, F. H., & Semb, G. A procedure for maintaining student progress in a personalized university course. *Journal of Applied Behavior Analysis,* 1974, *7,* 87-91.

Meyers, A. W., Artz, L. M., & Craighead, W. E. The effects of instructions, incentive, and feedback on a community problem: Dormitory noise. *Journal of Applied Behavior Analysis,* 1976, *9,* 445-447.

Panyan, M., Boozer, H., & Morris, N. Feedback to attendants as a reinforcer for applying operant techniques. *Journal of Applied Behavior Analysis,* 1970, *3,* 1-4.

Parsonson, B. S., Baer, A. M., & Baer, D. M. The application of generalized correct social contingencies: An evaluation of a training program. *Journal of Applied Behavior Analysis,* 1974, *7,* 427-437.

Pierce, C. H., & Risley, T. R. Improving job performance of neighborhood youth corp aides in an urban recreation program. *Journal of Applied Behavior Analysis,* 1974, *7,* 207-215.

Pommer, D. A. & Streedbeck, D. Motivating staff performance in an operant learning program for children. *Journal of Applied Behavior Analysis,* 1974, *7,* 217-221.

Tillema, M. *Bringing about and maintaining a mature sentence style in undergraduate students.* Unpublished master's thesis, Western Michigan University, 1977.

Van Houten, R., Hill, S., & Parsons, M. An analysis of a performance feedback system: The effects of timing and feedback, public posting, and praise upon academic performance and peer interaction. *Journal of Applied Behavior Analysis,* 1975, *8,* 449-457.

Notes

1. This report is based on a Master's thesis submitted by the first author to the Psychology Department, Western Michigan University. We thank Brad Huitema for his assistance in determining the statistical analysis. The thesis includes a detailed description and analysis of the system described in this report. Portions of the research were presented at the 1978 Midwestern Association of Behavior Analysis convention. Interested readers can inquire about a microfilm of the thesis by writing the D. B. Waldo Library, Western Michigan University, Kalamazoo, MI, 49008. Readers may obtain reprints of this report from Dr. Richard W. Malott, Department of Psychology, Western Michigan University, Kalamazoo, MI 49008.
2. We tend to maximize contact with rewards and minimize contact with aversives (Malott, Tillema, & Glenn, 1978, p. 6). Readers should see Malott, et al., 1978, for a discussion of the value of rewards and aversives as technical, behavioral terms roughly equivalent to positive and negative reinforcers.
3. Readers can obtain a complete copy of the current researchers' and supervisors' guidelines and forms from Dr. R. W. Malott.
4. A tested, programmed text, *Structuring mature sentences: a manual teaching the use of free modifiers and cumulative sentences* (Tillema, 1977), is available from Dr. R. W. Malott.