DOCUMENT RESUME

ED 131 030 SP 010 590

AUTHOR Carroll, J. Gregory
TITLE An Experiment on the Effectiveness of Training for Teaching Assistants.
PUB DATE 3 Sep 75
NOTE 25p.; Paper presented at the annual meeting of the American Psychological Association (84th, Washington, D.C., September 3, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS College Instruction; *College Teachers; *Educational Objectives; *Graduate Students; *Teacher Behavior; Teacher Education; *Teaching Assistants; Training Objectives; *Training Techniques

ABSTRACT

The results of this study show that an experimental training program for graduate teacher assistants (TAs) affected the observable teaching behavior of the participants in two respects--(1) participants made significantly greater use of instructional objectives and (2) they engaged in significantly more student-centered teaching. The subjects for this study were 19 novice TAs and 705 students enrolled in the introductory psychology course at Cornell University during the fall of 1975. The independent variable of this experiment was defined as TA participation in either an experimental or a control version of a seminar entitled, "The Teaching of Psychology." At random, ten TAs were assigned to the experimental group, and the other nine TAs were assigned to the control group. The experimental version of the seminar consisted of scheduled reading assignments on college teaching, at least two individual conferences for each TA with the course instructor, an individual video critique session for each TA, one unstructured group meeting with the course instructor, and five formal workshop sessions. The training program was designed by the investigator to encompass a broad range of teaching competencies. Several general skills related to overall course design were emphasized: (1) specifying instructional objectives; (2) utilizing objectives in the design of instruction and tests; and (3) making the cognitive levels of objectives congruent with the cognitive levels of instructional activities and tests. Specific skills were emphasized for facilitating student-centered discussions. The control version of the seminar consisted of fewer and unscheduled reading assignments, two unstructured group meetings with the course instructor, and the opportunity for each TA to view a videotape of his or her teaching, but alone and without any critique. (MM)

AN EXPERIMENT ON THE EFFECTIVENESS OF TRAINING FOR TEACHING ASSISTANTS[1]

A paper presented at the annual meeting
of the American Psychological Association,
Washington, D.C., September 3, 1976

J. Gregory Carroll[2]
Cornell University

[1] This paper is based on the author's doctoral dissertation (1976).

[2] Now at the Faculty Resource Center, University of Cincinnati.

As the adequacy of training for college teachers has been criticized more and more widely during the past fifteen or twenty years, countless pre-service orientation programs, in-service workshops, seminars, apprenticeship programs, intern programs, extern programs, and the like have been developed. During roughly the same period, the recruitment of teaching assistants (TAs) has been extended--for both pedagogical and budgetary reasons--to the undergraduate level itself, and the use of undergraduate TAs may have helped to heighten the perceived need for training as well as to increase the number and diversity of training programs pressed into service.

The recent development of TA training programs seems commendable, but it is surprising to note the lack of empirical research on the effects of such training. In the words of Barak Rosenshine, "We are instruction rich, and data poor" (1974, p.4). Although the literature contains dozens of superficial descriptions of innovative training programs and numerous catalogs of materials for teacher training, only a handful of empirical studies used student performance and satisfaction measures to

assess the effectiveness of TA training programs. Most published reports either failed to provide evaluative data or merely presented interview or questionnaire data from the TAs themselves.

Only four studies have been identified that (1) included student variables among the criteria of effectiveness, (2) reported reliable observational data to assess whether differences in training led to differences in specified teaching behaviors, and (3) utilized a true experimental design as defined by Campbell and Stanley (1963). All four involved training programs that focused on a few highly specific teaching competencies. Two of the four (Haber, 1973; Koffman, 1974) failed to show any more than chance differences in the observed teaching behaviors or in student outcomes. The remaining two studies involved training in interaction analysis and heuristic questioning techniques only, and both were conducted with mathematics TAs. Training in interaction analysis alone was not very effective (Daniels, 1970), but training in both interaction analysis and heuristic questioning was effective (Tubb, 1974).

The results of four quasi-experimental studies (Costin, 1968; Hockett, 1972; Lewis & Orvis, 1973; Yaghlian, 1972) generally indicated that training programs designed for a wider scope of teaching competencies yielded significant effects on student achievement and student ratings of instruction. The training programs of these studies tended to span a range of topics such as course design, communication, classroom climate, questioning, testing, influence styles, student participation, and the use of student feedback. The validity of these quasi-experimental studies, however, was susceptible to the influence of history, maturation, and interaction effects, as set forth in Campbell and Stanley (1963).

The present study with psychology TAs utilized a true experimental design to test whether a training program that was highly general in scope (1) affected the participants' teaching behavior, and (2) improved their student ratings of instruction. It was not feasible to measure the effectiveness of training in terms of student achievement. Since the

TAs of this sample each taught seminars on different topics within introductory psychology, there was no basis for comparing the performance of students from different discussion sections.

A third objective of the study, however, was to determine how the TA training program affected the relationship between student ratings of instruction and student performance in the seminars.

The independent variable of this experiment was a dichotomous variable defined as TA participation in either an experimental or a control version of a seminar entitled "The Teaching of Psychology." Participation in the experimental version was mandatory for the ten TAs selected at random for the experimental group, and the control version was mandatory for the nine TAs randomly assigned to the control group.

The experimental version consisted of scheduled reading assignments on college teaching, at least two individual conferences for each TA, an individual video critique session for each TA, one unstructured group meeting, and five formal

workshop sessions. This training program was designed by the investigator to encompass a broad range of teaching competencies. Several skills related to overall course design were emphasized: specifying instructional objectives, utilizing objectives in the design of instruction and tests, and making the cognitive levels of objectives congruent with the cognitive levels of instructional activities and tests (Bloom, 1956). Analyzing the structure of knowledge which underlies psychological research documents (Gowin, 1975) was a subject matter competency that the training sought to enhance. Specific skills for facilitating student-centered discussions (McKeachie, 1969) included being an active listener, establishing a conducive classroom atmosphere, being aware of nonverbal communication, obtaining feedback from group members, using varied and appropriate questioning techniques, and providing closure. Detailed documentation of the training procedures is given in Carroll (1976).

The control version of the seminar consisted of fewer and unscheduled reading assignments on college teaching, two

unstructured group meetings with the course instructor, and the opportunity for each TA to view a videotape of his or her teaching alone and without any critique. During both meetings with the control group TAs, the course instructor had the participants share their own insights, but he purposely added few substantive comments of his own.

Seven dependent variables were defined and measured as follows.

1. Use of objectives: the degree to which the TA specified clearly formulated objectives and maintained agreement among objectives, instruction, and tests.

2. Cognitive levels of questions: six mutually exclusive categories of questions as defined in Bloom (1956). The six levels, from lower order to higher order, are: memory, comprehension, application, analysis, synthesis, and evaluation. This variable was measured by trained, independent raters who analyzed classroom questions and seminar quiz questions.

3. Student-centered versus instructor-centered teaching: teaching characterized by accepting feelings, giving praise,

using students' ideas, and asking questions as opposed to lecturing, giving directions, and criticizing. This variable was measured by four trained, independent raters using the Indirect/Direct ratio (I/D) from Flanders Interaction Analysis (1970).

4. Student talk ratio (S/T): the ratio of minutes of student talk in class to minutes of student talk and TA talk in class, as measured by the number of tallies in the respective Flanders categories.

5. Student evaluation of TA effectiveness: the mean of student ratings on three questionnaire items related to teaching ability.

6. Student evaluation of seminar effectiveness: the mean of student ratings on four questionnaire items related to outcomes of the seminars.

7. Student performance: numerical grade achieved by the student for overall seminar performance. The numerical grade was determined by a seminar quiz, a term paper, and class participation.

Hypotheses

Seven hypotheses were formulated to predict differences between the experimental group and the control group. On the basis of previous quasi-experimental research, the hypotheses predict the expected direction of difference. The first four hypotheses predict outcomes related to the first major objective of the experiment, i.e., to determine whether the experimental TA training program affected the teaching behavior of the participants.

1. The experimental group TAs will be rated higher than the control group TAs on the use of objectives.

2. There will be greater congruity among the cognitive levels of classroom questions and quiz questions for the experimental group than for the control group. That is, the mean of the correlations between the frequency of each TA's seminar questions at the various cognitive levels and the frequency of that TA's quiz questions at the various cognitive levels will be higher for the experimental group as a whole than it will be for the control group.

3. The teaching of the experimental group TAs will be more student-centered than that of the control group TAs.

4. The seminars taught by the experimental group TAs will have a higher student talk ratio than the seminars taught by the control group TAs.

Hypotheses five and six predict outcomes related to the second major objective of the experiment, to determine whether the experimental training program improved the effectiveness of instruction as judged by the participants' students.

5. The experimental group will have higher student evaluations of TA effectiveness than the control group.

6. The experimental group will have higher student evaluations of seminar effectiveness than the control group.

Hypothesis seven addresses the third objective--to investigate whether the training program affected the correlation between student ratings and student performance in the participants' seminars.

7. For the students taught by the experimental group TAs, there will be a significant positive correlation between

student evaluations of TA effectiveness and student performance in seminar.

The correlation coefficient for the control group could not be predicted with any confidence. That coefficient could reasonably be expected to fall anywhere within the wide range already reported in the literature.

Subjects

The subjects for this experiment were 19 novice TAs and 705 students participating in the introductory psychology course at Cornell University during the fall semester of 1975. Four of the TAs were graduate students and fifteen were undergraduates. Six were male and thirteen were female. All TAs conducted weekly seminars in their own area of expertise as a supplement to the lectures given by the course instructor. The graduate TAs each met with four sections per week; the undergraduates each met with two sections per week.

The TAs were matched on verbal aptitude, and ten were randomly assigned to the experimental group, nine to the

control group. There were no significant differences between the two groups of TAs on sex, grade level, verbal aptitude, cumulative average, or class size ($p > .10$).

An analysis of student characteristics indicated that there were no significant differences between the experimental group and control group students on sex, college, grade level, approximate cumulative average, major, or primary reason for taking the course ($p > .05$).

Procedure

The experimental design was the posttest-only control group design (Campbell & Stanley, 1963). Since the novice TAs were assigned randomly to the experimental and control groups and since there were no significant differences in the TA characteristics and student characteristics reported earlier, it was assumed that there was no initial bias favoring one group of TAs over the other. The TAs were aware that their teaching was being studied for the purpose of group comparisons, but when questioned at the end of the semester, they could not identify specific variables being measured.

Observational data were obtained by videotaping one class session of each TA during eriod near the end of the semester. The recordi[illegible] classes was counterbalanced so that approximately the same number of experimental and control group TAs were recorded each week and so that approximately the same number of each group were teaching for the first, second, third, or fourth time that week.

Four trained raters independently analyzed each videotape using Flanders Interaction Analysis. Data on the cognitive levels of classroom questions and quiz questions were obtained from four other trained, independent raters using a modified version of the Teacher-Pupil Question Inventory (Davis & Tinsley, 1968). Table 1 contains the coefficients of inter-rater reliability.

Table 1

Inter-Rater Reliability

Raters	Coefficient of Agreement
FIA	$\bar{\pi} = .84$
TPQI, classroom	$\bar{r} = .73$
TPQI, quiz	$r = .74$

use of objectives, t (17) = 1.60, p < .07. Thus, hypothesis one was confirmed.

Table 2

t-Test of the Use of Objectives

Group	Mean	s.d.	df	t	p
E	4.02	.44	17	1.60	.07
C	3.70	.43			

Hypothesis two stated that the experimental group TAs would achieve greater congruity among the cognitive levels of classroom questions and quiz questions than the control group TAs. A Pearson correlation coefficient was computed for each TA by pairing the number of classroom questions at each cognitive level with the number of quiz questions at that level. The mean coefficient for each group of TAs was then computed using the Fisher Z transformation, and a t-test was performed. As shown in Table 3, the difference between groups was not significant, t (17) = .90, p > .18. Thus, hypothesis two was rejected.

Table 3

Congruity

Group	Mean	s.d.	df	t	p
E	.4				
			17	.90	.18
C	.13	.70			

The Flanders data indicated that the teaching of the experimental group TAs was more student-centered than that of the control group TAs. As shown in Table 4, the I/D ratios of the experimental group were significantly higher than those of the control group, $t\ (17) = 1.69$, $p < .06$. Thus, the third hypothesis was confirmed.

Table 4

I/D Ratio

Group	Mean	s.d.	df	t	p
E	.44	.21			
			17	1.69	.06
C	.31	.11			

Hypothesis four, which predicted higher student talk ratios for the experimental group than for the control group,

was rejected. Table 5 shows that there was no significant difference between groups on the S/T ratio, t (17) = .36, p > .36.

Table 5

S/T Ratio

Group	Mean	s.d.	df	t	p
E	.55	.20			
			17	.36	.36
C	.52	.20			

Hypotheses five and six stated that the student evaluations of the experimental group would exceed those of the control group for TA effectiveness and seminar effectiveness, respectively. As demonstrated in Table 6, both hypotheses were confirmed.

Table 6

Group	Mean	s.d.	df	t	p
TA effectiveness					
E	4.19	.38			
			17	1.34	.10
C	3.98	.30			
Seminar effectiveness					
E	3.53	.58			
			17	1.36	.10
C	3.18	.52			

Students rated the effectiveness of the experimental group TAs significantly higher than that of the control group TAs, $t(17) = 1.34$ $p < .10$. Students also rated the seminars taught by the experimental group TAs as being significantly more effective than the seminars taught by controls, $t(17) = 1.36$, $p < .10$.

In addition, as shown in Table 7, the overall Pearson correlation between the I/D ratios and the mean ratings of TA effectiveness was positive, $r = .53$, $p < .02$, two-tailed. Likewise, the correlation between I/D ratios and the mean ratings of seminar effectiveness was positive, $r = .57$, $p < .02$, two-tailed. Therefore it was concluded that the use of indirect teaching was positively associated with student ratings of the effectiveness of instruction.

Table 7

Variables	r	p
I/D & TA effectiveness	.53	.02[a]
I/D & seminar effectiveness	.57	.02[a]

[a]The significance level is two-tailed.

As indicated in Table 8, student ratings of TA effectiveness showed no significant difference between the nine TAs with the greatest congruity and the nine with the least congruity among cognitive levels of classroom and quiz questions, $t\ (16) = .58$, $p > .57$, two-tailed. Similarly, there was no significant difference between the high- and low-congruity groups on student ratings of seminar effectiveness, $t\ (16) = .54$, $p > .59$, two-tailed. Thus, there was no evidence of an association between the congruity variable and student ratings of instruction.

Table 8

High- vs. Low-Congruity Groups

Group	Mean	s.d.	df	t	p
	TA effectiveness				
High	4.14	.34			
			16	.58	.57[a]
Low	4.04	.40			
	Seminar effectiveness				
High	3.42	.62			
			16	.54	.59[a]
Low	3.27	.54			

[a]The significance level is two-tailed.

Finally, hypothesis seven predicted that for the experimental group there would be a significant positive correlation between student ratings of TA effectiveness and student performance in seminar. This correlation was computed by pairing the mean seminar grade of each seminar section with the mean rating of TA effectiveness from that section. The results are presented in Table 9. There is no evidence that the training program significantly increased the correlation between student performance and student ratings of either TA effectiveness or seminar effectiveness.

Table 9

Correlations with Student Performance

Group	r	p
TA effectiveness		
E	.21	.15
C	.13	.59[a]
Seminar effectiveness		
E	.18	.37[a]
C	.15	.53[a]

[a]The significance level is two-tailed.

For the nine high-congruity TAs, however, there was a significant correlation between student performance in seminar and student ratings of TA effectiveness, as shown in Table 10 ($r = .38$, $p < .08$, two-tailed). This correlation coefficient for the high-congruity TAs was significantly greater than the corresponding coefficient ($r = -.19$) for the nine low-congruity TAs ($p < .08$, two-tailed). As for the correlation between student performance and student ratings of seminar effectiveness, the coefficients were not significant for either the high-congruity group ($r = .24$, $p > .25$, two-tailed) or the low-congruity group ($r = -.12$, $p > .61$, two-tailed), and the two coefficients were not significantly different from each other ($p > .25$, two-tailed).

Table 10

Correlations with Student Performance

(High- and Low-Congruity Groups)

Group	r	p
TA effectiveness		
High	.38	.08[a]
Low	-.19	.42[a]
Seminar effectiveness		
High	.24	.25[a]
Low	-.12	.61[a]

[a]The significance level is two-tailed.

The significant difference between the high- and low-congruity groups on the correlation between student performance and student ratings of TA effectiveness was interpreted as suggesting that the extreme inconsistency among published reports of such correlations (reviewed in Kulik & Kulik, 1974) may be attributable to inappropriate measures of student preformance. These data suggest that the variability in reported correlation coefficients may be distorted by the variability from one study to another in the degree to which

the measures of performance were congruent with course objectives and instructional tasks.

In sum, four general conclusions can be stated.

1. The experimental TA training program affected the observable teaching behavior of the participants in two respects. The participants made significantly greater use of objectives ($p < .07$), and they engaged in significantly more student-centered teaching ($p < .06$).

2. Student ratings of the effectiveness of instruction were significantly higher for the participants than for the controls ($p < .10$).

3. Regardless of training, the use of indirect teaching skills was positively correlated with student ratings of instructional effectiveness ($p < .02$, two-tailed).

4. Regardless of training, for the nine TAs with the greatest congruity among the cognitive levels of classroom and quiz questions, as compared to the nine with the least congruity, there was a significantly higher correlation between student ratings of TA effectiveness and student performance in seminar ($p < .08$, two-tailed).

References

Bloom, B. S. (Ed.) Taxonomy of educational objectives: Cognitive domain. New York: Longmans, Green and Co., 1956

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand-McNally, 1963.

Carroll, J. G. Effects of a teaching assistant training program on teaching behavior and student ratings of instruction. Doctoral dissertation, Cornell University, 1976.

Costin, F. A graduate course in the teaching of psychology: Description and evaluation. Journal of Teacher Education, 1968, 19, 425-432.

Daniels, J. W. Effects of interaction analysis upon teaching assistants and student achievement in introductory college mathematics. Doctoral dissertation, Indiana University, 1970.

Davis, O. L., & Tinsley, D. C. Cognitive objectives revealed by classroom questions asked by social studies student teachers. In Hyman, R. T. Teaching: Vantage points for study. Philadelphia: Lippincott, 1968

Flanders, N. A. Analyzing teaching behavior. Reading, Mass.: Addison-Wesley, 1970.

Gowin, D. B. The domain of education. Unpublished manuscript, Cornell University, Department of Education, 1975.

Haber, F. B. The effect of instruction in questioning strategies on teaching assistants' classroom performance. Doctoral dissertation, Arizona State University, 1973.

Hockett, J. C. An examination of changes in teacher behaviors and learner perceptions associated with a program for teaching assistants at the Florida State University: A pilot study. Doctoral dissertation, Florida State University, 1972.

Koffman, M. L. A study of the effects of instructional analysis and feedback on the classroom behavior and student achievement of university teaching assistants. Doctoral dissertation, University of Massachusetts, 1974.

Kulik, J. A., & Kulik, C. C. Student ratings of instruction. Teaching of Psychology, 1974, 1, 51-57.

Lewis, D. R., & Orvis, C. C. A training system for graduate student instructors of introductory economics at the University of Minnesota. Journal of Economic Education, 1973, 5, 38-46.

Maas, J. B. Cornell inventory for student appraisal of teaching and courses. Ithaca: Graphic Arts Press, 1972.

McKeachie, W. J. Teaching tips: A guidebook for the beginning college teacher (6th ed.). Lexington, Mass.: D. C. Heath, 1969.

Rosenshine, B. Teacher competency research. Unpublished manuscript, University of Illinois, Urbana, Bureau of Educational Research, 1974.

Scott, W. A. Reliability of content analysis: The case of nominal coding. Public Opinion Quarterly, 1955, 19, 321-325.

Tubb, G. W. Heuristic questioning and problem solving strategies in mathematics graduate teaching assistants and their students. Doctoral dissertation, Texas A & M University, 1974.

Yaghlian, N. University teaching: The impact of an in-service program for teaching fellows in chemistry. Doctoral dissertation, University of Michigan, 1972.