American Educational Research Association

Effects of a Minimal Intervention on Teacher Behavior and Student Achievement
Author(s): Theodore Coladarci and N. L. Gage
Source: *American Educational Research Journal*, Vol. 21, No. 3 (Autumn, 1984), pp. 539-555
Published by: American Educational Research Association
Stable URL: http://www.jstor.org/stable/1162914
Accessed: 11/03/2010 14:52

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=aera.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

American Educational Research Association is collaborating with JSTOR to digitize, preserve and extend access to *American Educational Research Journal.*

http://www.jstor.org

American Educational Research Journal
Fall 1984, Vol. 21, No. 3, Pp. 539–555

Effects of a Minimal Intervention on Teacher Behavior and Student Achievement

THEODORE COLADARCI
University of Maine at Orono
and
N. L. GAGE
Stanford University

Results from recent classroom-based experiments suggest that teacher behavior can be modified and student achievement improved through minimal interventions where (a) teacher training is carried out in correspondence course fashion and (b) comprehensive classroom observations are not conducted. This experiment represents such an intervention, carried out in the fourth, fifth, and sixth grade classrooms of 32 volunteer teachers. The experimental group teachers received through the mail a series of teacher training packets containing recommendations for teaching practices derived from the results of four large-scale correlational studies of teaching. Before and after training, classroom observations were conducted for 2 hours on two occasions. Analyses indicated that the intervention did not effect significant change in training-related teaching practices or end-of-year student achievement. Discussion of these results addresses factors that probably mediate treatment implementation and, consequently, effects on student achievement in research of this kind.

The research reported here was funded by the National Institute of Education (NIE) (Grant No. NIE-G-79-0014). The opinions expressed herein do not necessarily reflect the position or policy of the NIE, and no official endorsement by the NIE should be inferred. A version of this paper was presented at the 1981 meeting of the American Educational Research Association, Los Angeles.

The authors wish to thank Peggy Bander and Georgea Mohlman Sparks for their assistance during the study; the students, teachers, and administrators of the participating schools; Lyn Corno, Meredith D. Gall, and David Rogosa for critically reading earlier versions; and Lorraine Stone for typing several drafts of this paper.

Correlational research on teaching conducted during the past 15 years has resulted in the conceptualization of *direct instruction* (Rosenshine, 1976), a loosely defined set of teacher, classroom, and curriculum variables considered to be foremost in explaining growth in student achievement in the elementary grades (e.g., Berliner, 1979; Berliner & Rosenshine, 1977; Brophy, 1979; Good, 1979; Medley, 1982; Rosenshine, 1976, 1979; Rosenshine & Berliner, 1978). Powell (1978) offered perhaps the most succinct presentation of key components of the direct instruction concept:

> The coverage of content is extensive, is allocated to academic tasks, and the time is not broken by frequent interruptions or changes of task. Students spend a good portion of the time allocated to instruction actually engaged in instructional tasks, and the teacher monitors and encourages task engagement on the part of the students . . . . The atmosphere in the classroom is one in which academic work is both recognized to be important and performed. (p. 29)

As the results of this research were being synthesized, there was a call for experiments in research on teaching so that causal, rather than associational, hypotheses could be examined (e.g., Dunkin & Biddle, 1974; Gage, 1972, 1978; Nuthall & Church, 1973; Rosenshine & Furst, 1973; see Walker, 1980, for a different view). It further was argued that treatments employed in these experiments initially should be minimal with respect to their duration, expense, and intensity (Brophy, 1977). If effective, such minimal interventions would have the decided advantage of inexpensive and swift exportability to the classroom.

To date, three classroom-based experiments have been conducted in which the treatments (a) were based on previous correlational research concerning process-product relationships; (b) had a direct instruction emphasis; and (c) represented, to varying degrees, minimal interventions (Anderson, Evertson, & Brophy, 1979; Crawford et al., 1978; Good & Grouws, 1979). Each of these experiments will be described (also see Gage & Giaconia, 1981).

The Anderson et al. (1979) experiment was conducted in white, middle socioeconomic status (SES), first grade classes. Experimental group teachers received a manual presenting an instructional model, based largely on the correlational findings of Brophy and Evertson (1974):

> The treatment was minimal in cost and time. In October, the researchers met with teachers in the treatment schools and described the purpose of the study. The teachers who agreed to participate read the manual describing the instructional model and met again with the experimenters to discuss it. There was no further training, and no attempts were made during the year to boost the treatment. (p. 195)

Observations were conducted on 15 to 20 occasions in all control group classes and in 10 of the 17 experimental group classes. A series of between-class regression equations was employed to assess treatment effects and observation effects (i.e., observed vs. unobserved experimental group classes). There was a significant treatment effect, favoring the experimental group, on a reading composite score. Further, there was no observation effect: The presence of classroom observers did not mediate the treatment effect on student achievement.

Crawford et al. (1978) employed a sample of 33 middle SES third grade classes. There were three experimental conditions: observation only, minimal training plus observation, and maximal training plus observation. Minimally trained teachers simply received by mail at weekly intervals a series of five training packets containing recommended teaching practices. These recommendations were based on a detailed examination and synthesis of the results of four large-scale correlational studies (Brophy & Evertson, 1974; McDonald & Elias, 1976; Soar, 1973; Stallings & Kaskowitz, 1974). The maximally trained teachers, in addition to receiving the weekly packets, attended weekly 2-hour meetings in the 5-week period during which the training packets were delivered. These meetings were devoted to review, discussion, videotape viewing, and role playing. Classroom observations were conducted for approximately 16 full school days for each of the 33 teachers before, during, and after the training. Results indicated the minimal and maximal training conditions had equivalent effects on class achievement on a vocabulary posttest. Together, they were .69 of a standard deviation (*SD*) above the mean of the control group classes. There was no comparable effect on a reading comprehension posttest.[1]

Forty lower SES fourth grade classes participated in the experiment conducted by Good and Grouws (1979). After holding an introductory meeting for all teachers and their principals, the researchers described the instructional model to the 21 experimental group teachers for approximately 90 minutes. The instructional model, and the corresponding teacher's manual, were based on correlational findings summarized in Good and Grouws (1977). Two weeks after treatment began, an additional 90-minute meeting was held to answer questions about the program. Almost all of the teachers were observed on six occasions between October and

[1] Interestingly, minimally trained teachers were found to implement more of the training recommendations than the maximally trained teachers. This difference, however, may be partly illusory. The minimally trained teachers were initially higher than the maximally trained teachers on a measure of verbal fluency and a measure of structuredness, both of which correlated positively with implementation. However, a difference in implementation, albeit a small one, remained after adjusting for these initial differences.

the end of January. A class-level analysis of variance indicated a significant treatment effect, favoring the experimental group classes, on a mathematics test administered in mid-December. That a treatment effect was detected early in the school year is noteworthy and, indeed, encouraging. Good (1979) later reported that the experimental group classes still held an advantage on achievement at the end of the school year when the district carried out its regular testing—roughly 3 months after formal observations were completed.

The findings of Crawford et al. (1978) regarding the minimally versus maximally trained teachers, the absence of an observation effect reported by Anderson et al. (1979), and the results of Good and Grouws (1979) were cited in support of the minimal intervention in research on teaching:

> Although more research on implementation is needed, two tentative conclusions are warranted: (1) elaborate delivery systems may not be necessary for effectively training inservice teachers to perform specifically identified classroom behaviors, and (2) observation of teachers does not necessarily have to be a part of the inservice training. (Good, 1979, p. 57)

Although encouraging claims have been made concerning a minimal intervention (e.g., Good, 1979, 1980; Good & Grouws, 1979), such an intervention, however, had yet to be undertaken as defined above. That is, no intervention had been minimal with respect to both the delivery of training and the conduct of classroom observations. Thus, the implication of the minimal-maximal parity reported by Crawford et al. (1978) was clouded by the comprehensive observations conducted in the classrooms of all teachers. Similarly, the implication of the effects on trained-but-unobserved teachers reported by Anderson et al. (1979) was obscured by the initial meetings attended by all teachers. And neither of these factors was manipulated in the study conducted by Good and Grouws (1979): All teachers attended initial meetings, and observations were conducted in all classrooms.

The present study examined a minimal intervention. Unlike those of the three experiments discussed above, this intervention was minimal in that (a) the treatment consisted solely of mailing training materials to the experimental group teachers, and (b) only a few brief classroom observations were made.

METHOD

Subjects

The initial sample comprised 32 volunteer teachers and their fourth, fifth, and sixth grade students in a large, urban school district in the San Francisco Bay Area. Because of subsequent complications, the number of classes on which the achievement data analyses were based was reduced to 28, representing 631 students, most of whom are black and low in SES. (Insufficient achievement data were available for several classes, and there was treatment contamination involving an experimental group teacher and a control group teacher who taught in the same school.)

Instruments and Procedures

The teacher education packets. The teacher education packets (TEP) were those developed and used by Crawford et al. (1978). As already noted, the thousands of process-product correlations presented in the technical reports cited earlier were examined and considered as the basis for prescriptive statements (see Crawford et al., 1978, Vol. 1, pp. 25–31). The interpretations of 125 selected correlation coefficients provided the basis for three packets of teaching recommendations.

The first packet, behavior management and classroom discipline, was based on the findings that classes characterized by a general unruliness and a poorly articulated system of rules are also characterized by frequent nonengagement in academic activities and student difficulty in attending to academic tasks. Teachers were informed of ways to manage their classes, largely in light of Kounin's work (1970; also see Brophy & Putnam, 1979). The packet cautioned teachers against disciplinary errors that prolong or compound the problem—specifically, the disciplinary errors regarding *timing* and *target.* Further, this packet encouraged teachers to develop a *system of rules*, which let students know—without always having to consult the teacher—what they can and should do during a given period. Finally, to curb misbehavior, as well as to identify and respond to students in need of assistance, teachers were encouraged to monitor activities when students are engaged in seatwork.

The second packet, instructional methods, highlighted the importance of large-group instruction, frequent use of question-and-answer sessions, and use of visual aids and phonics exercises in reading activities. In addition, with seatwork assignments, this packet informed teachers of the importance of assigning work of appropriate difficulty, using textbooks and workbooks (rather than games, toys, and machines), and minimizing the amount of time devoted to organizing and giving directions.

The third packet, questioning and feedback strategies, pertained to the manner in which the teacher selects students to respond to questions, the difficulty level of the questions asked, and the provision of feedback to the student's response. The summary listing of TEP recommendations is presented in the Appendix.

An introductory packet briefly discussed the TEP's general rationale and provided a classroom vignette illustrating a teacher whose practices largely conformed to the TEP recommendations. A final packet reviewed and summarized the preceding packets. A sixth packet presented an additional classroom vignette, illustrating teaching practices that were consistent with and inconsistent with the TEP recommendations. The six packets were mailed individually to experimental group teachers in December and January. Finally, the teachers received three summary sheets corresponding to the second, third, and fourth packets, respectively. These sheets were intended to provide a succinct and accessible review of the contents of the three packets and were mailed one per week, beginning in mid-February.

Teachers and their classes were randomly assigned to experimental conditions. Experimental group teachers were asked to become familiar with the TEP and, further, to follow the various recommendations in their teaching. Project staff did not meet with any teachers to discuss the training materials or to facilitate implementation.

Classroom observation instrument. Observers used this four-page instrument, adapted from a measure used by Crawford et al. (1978), to record their judgments and estimates on low-inference and high-inference variables. Each of the 26 items in the observation record reflected components of the TEP. The alternatives for each item in the observation schedule were scored so that a high value represented a high degree of conformity to a particular recommendation in the TEP. Thus, the observation records yielded rough estimates of the extent to which teacher behavior—of both experimental and control group teachers, before and after training—reflected the TEP recommendations.

Teachers were observed on four 2-hour occasions, twice before and twice after training. Each of the four observations was coded to obtain a total score representing a teacher's general conformity to recommendations (CTR) across all items. The two fall CTRs were averaged for each teacher, as were the two spring CTRs.

It was the CTR total, rather than the CTR item score, that was employed in the analyses of the observation data reported here. Analyses that employed the CTR total were considered more meaningful for two reasons. First, the sum of scores on n positively correlated items has greater reliability than the score on each item considered individually. Second, the most compelling and defensible analysis was one of the program as a whole (i.e., total CTR) simply because the discrete teaching recommendations were not manipulated independently. Although analyses that focus on the discrete teacher behaviors may have proved intriguing, the inevitable intercorrelation among these behaviors would have precluded any clear and meaningful interpretation (Crawford & Stallings, 1978).

The Comprehensive Tests of Basic Skills. The Comprehensive Tests of Basic Skills (CTBS), a battery chosen by the school district testing committee for regular use in the district, was employed as the measure of academic achievement. Test scores from the spring 1978 and 1979 administrations served as the pretest and posttest, respectively. Each student's test scores were combined to yield a reading total score and a mathematics total score which, in turn, were combined as a total score.

RESULTS

The statistical analyses focused on two main questions. First, did the intervention appreciably alter the training-related teaching practices of experimental group teachers? Second, did the intervention produce significant increments in academic achievement for the students in the experimental group classes?

Group Differences on CTR

Before treatment effects on teacher behavior were examined, CTR reliability was assessed. This reliability was independent substantively of the question of treatment effects on student achievement. Hence, reliability was estimated on the original sample of 32 teachers as well as on the final sample of 28 teachers.

For the full sample ($N = 32$), the two fall CTRs correlated .27, as did the two spring CTRs. The reliability of the sum of the two fall CTRs ("fall total CTR") and that of the two spring CTRs ("spring total CTR"), estimated by applying the Spearman-Brown formula, were both .43. In the present design, these estimates are equivalent to generalizability coefficients (Cronbach, Gleser, Nanda, & Rajaratnam, 1972). As such, they represent the ratio of between-teacher variance to the total observed-score variance, the latter comprising between-teacher variance and the nested combination of variance attributable to interactions involving teachers, occasions, and observers.[2] For the restricted sample ($N = 28$), the correlation between the two fall CTRs was reduced from .27 to .18, and that between the two spring CTRs was reduced from .27 to .21. With the Spearman-Brown formula applied, the estimated reliabilities of the fall total CTR and spring total CTR was .31 and .28, respectively. The internal consistencies (e.g., Cronbach, 1970, p. 160) of the observations on each of the four occasions ranged from .66 to .76.

Because classes were randomly assigned to the experimental conditions, the difference in fall CTR between the control group and experimental group was expected to be practically negligible. Given the design of the present study, the suitability of the analysis of covariance (ANCOVA) was

[2] In the present design, teachers and occasions were crossed. Each observer, however, did not observe all teachers; hence, observers were nested within teachers. Further, because not all observers observed on each of the four occasions, observers similarly were nested within occasions.

initially entertained. The utility of ANCOVA in such a design would lie more in the consequent reduction of error variance than in the posttest adjustment for initial random differences on the pretest (e.g., Linn & Slinde, 1977). Reducing the error term, of course, results in a more sensitive statistical test; but, because this reduction is proportional to the magnitude of the pretest-posttest correlation, ANCOVA is of little use where this correlation is less than approximately .30 (Elashoff, 1969). Such a correlation was not expected (and, ultimately, not obtained) between the fall and spring CTR and, consequently, the use of ANCOVA was considered unnecessary. Rather, *t* ratios were computed for the difference between the spring CTR means.

Table I presents the means and standard deviations of the CTR totals by experimental condition. Although favoring the experimental group, the spring difference in CTR between the two groups was not statistically significant. These data indicate that, as a whole, treatment implementation was poor: Training-related teaching practices of the experimental group teachers were not altered appreciably.

The mean differences presented in Table I can, however, be examined at a descriptive level. For the full sample, the experimental group pretreatment CTR mean fell below the corresponding control group mean −1.67 raw score points, or −.23 *SD* (pooled). After treatment, in contrast, the experimental group means was slightly above the mean of the control group (1.37 raw score points, or .20 *SD*). These small differences were more pronounced in the restricted sample, where the standardized mean differences were −.46 *SD* and .23 *SD*, respectively. Table I also shows that

TABLE I

Conformity-to-Recommendations (CTR): Within-Group Means and Standard Deviations for the Fall and Spring Observations for Full and Restricted Sample

CTR	Control ($N = 16$)		Experimental ($N = 16$)		t[a]
	M	*SD*	*M*	*SD*	
Full sample					
Fall	72.38	7.14	70.63	7.92	−.66
Spring	69.19	5.85	70.56	7.70	.57
	Control ($N = 13$)		Experimental ($N = 15$)		
Restricted sample					
Fall	74.15	5.39	70.97	8.07	−1.20
Spring	68.92	3.95	70.40	7.94	.61

[a] No *t* reported here is statistically significant ($\alpha = .05$).

CTR in each group declined from fall to spring. This decline was considerably more marked for the control group, however: Raw change in CTR from fall to spring for control group teachers was −5.23, whereas the corresponding figure for experimental group teachers was −.57. Statistical significance notwithstanding, the mean differences in total CTR suggest that the teacher training may have retarded a decline from fall to spring in the incidence of training-related teaching practices among experimental group teachers.

Group Differences on Achievement

The effect of the TEP on student achievement was assessed through an alternative to conventional ANCOVA (see Rogosa, 1980). As in ANCOVA, the vertical distance between the two sample within-group regressions was evaluated at the point on the covariate corresponding to the weighted average of the two group means. However, the test statistic for this measure is independent of the degree of homogeneity of the population within-group regressions. This analysis "may be thought of as a safer ANCOVA in that the procedure retains its statistical properties when the [homogeneity] assumption is violated" (Rogosa, 1980, p. 312).

Between-class, stepwise regression analyses were performed on achievement with the Statistical Package for the Social Sciences (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). Treatment (T)—a dummy variable coded 0 (control) or 1 (experimental)—was entered on the first step. The CTBS pretest was treated as a covariate (X) and entered on the second step. Treatment and covariate were multiplied and entered on the third step, representing the TX interaction. The unstandardized regression coefficient for T at the second step is equivalent to the conventional ANCOVA treatment effect: the adjusted mean difference on the dependent measure. The square of this value is the numerator of the test statistic (see Rogosa, 1980, p. 312, Equation 12). The estimated variance-covariance matrix of the regression coefficients T and TX, needed to calculate the denominator of the test statistic, was obtained from the SAS SYSREG procedure (SAS Institute, 1979).

Analyses reported here were conducted with grade levels combined. This was accomplished through a within-grade, student-level linear transformation ($M = 50$, $SD = 10$) of CTBS raw scores which, subsequently, were aggregated at the class level. Treatment effects corresponding to within-grade analyses did not differ significantly from those presented here.

Relevant descriptive statistics are presented in Tables II and III. The former provides CTBS pretest and posttest means and standard deviations. Table III presents the within-group (i.e., within-treatment) regression equations for the dependent measures.

Table IV presents the estimated treatment effect and the corresponding 95% confidence interval for each measure. These estimates of the treatment effects consistently spanned zero, comprising both positive and negative values. From these data, it appears teacher training was ineffective in improving student achievement.

One could plausibly assume that, irrespective of experimental condition, there would be variability in CTR. To be sure, not all experimental group teachers would be expected to demonstrate the same degree of CTR; teacher attitudes, beliefs, motivations, and so on, doubtless were operating (Doyle & Ponder, 1977). And the assumption could not be made that, by virtue of their group assignment, control group teachers would demonstrate no CTR whatever; one also would expect natural variability in CTR here.

TABLE II

CTBS Pretest and Posttest Means and Standard Deviations

Test	Control ($N = 13$)		Experimental ($N = 15$)		Pooled[a] ($N = 28$)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Pretest						
Reading total	50.40	5.71	49.68	4.82	50.00	5.24
Mathematics total	50.15	5.23	49.88	4.56	50.00	4.87
Total score	50.29	5.78	49.76	4.87	50.00	5.30
Posttest						
Reading total	50.98	5.32	49.21	4.83	50.00	5.13
Mathematics total	50.37	4.54	49.70	4.07	50.00	4.29
Total score	50.70	4.99	49.43	4.63	50.00	4.83

Note. $N = 28$. Grades were pooled through a within-grade *T*-score transformation ($M = 50$, $SD = 10$) of student level scores.

[a]Experimental conditions pooled.

TABLE III

Within-Group Regressions of CTBS Posttest on Pretest

Posttest	Control ($N = 13$)		Experimental ($N = 15$)	
	Constant	*b* for pretest	Constant	*b* for pretest
Reading total	12.290	.768	3.995	.910
Mathematics total	19.915	.607	11.983	.756
Total score	16.463	.681	7.069	.851

Note. $N = 28$. Grades were pooled through a within-grade *T*-score transformation ($M = 50$, $SD = 10$) of student-level scores.

TABLE IV

Estimates of Class-Level Treatment Effects

Test	Treatment effect	F (1, 23)[a]	Lower and upper end points of a 95% confidence interval
Reading total	−1.171	1.234	−3.333, .991
Mathematics total	−.481	<1	−2.895, 1.993
Total score	−.866	<1	−2.999, 1.267

Note. $N = 28$. Grades were pooled through a within-grade T-score transformation ($M = 50$, $SD = 10$) of student-level scores.
[a]No F reported here is statistically significant ($\alpha = .05$).

A supplemental analysis was conducted to assess the degree to which the teachers' CTR—whether naturally occurring or attributable to the training—correlated with student achievement. Such an analysis is informative in the present context in that it yields additional evidence concerning the relevance of the teacher training to student achievement. In this sense, the effects of a program or treatment can be evaluated by examining *all* teachers, regardless of the experimental condition to which they initially had been assigned.

A full-year measure of total CTR was obtained by averaging total CTR across the four occasions; this served as the *process* measure. *Product* was a residual score based on the CTBS posttest total, obtained by regressing the CTBS posttest total on the pretest at the student level. These residuals were then aggregated at the class level and, in turn, correlated with the process measure. The resulting correlation ($N = 28$) between total CTR and residual achievement, grades pooled, was $r = .29$ ($p > .10$). This correlation increased considerably, however, with the removal of one discrepant case ($r = .40$, $p < .05$). (The three within-grade process-product correlations were positive, although not statistically significant.) Thus the general content of the training, derived from previous process-product research (Brophy & Evertson, 1974; McDonald & Elias, 1976; Soar, 1973; Stallings & Kaskowitz, 1974), ostensibly had pedagogical value in the present context as well.

DISCUSSION

As an experiment, this study failed to corroborate the positive results obtained previously in similar classroom-based experiments (Anderson et al., 1979; Crawford et al., 1978; Good & Grouws, 1979). Toward the end of the school year, the experimental group teachers did not evidence markedly greater conformity to the training recommendations than that exhibited by the control group teachers. Further, as would be predicted from this first finding, the classes of these two groups of teachers were not appreciably different in end-of-year academic achievement.

There was a priori reason to expect the desired change in teaching practices among experimental group teachers. After all, the intervention in the present study was the same as the minimal training condition found effective in the study by Crawford et al. (1978). Further, the experiments conducted by Anderson et al. (1979) and Good and Grouws (1979) similarly did not involve a comprehensive delivery system, yet both studies resulted in positive change in training-related teaching practices among experimental group teachers. The poor treatment implementation in the present study probably was due in large part to the major methodological difference between this study and these three previous experiments: *Both* training and classroom observations were held to a minimum.

Classes in the present study were observed for a maximum of 8 hours throughout the entire school year: two 2-hour periods in fall and again in spring. Crawford et al. (1978), in contrast, obtained classroom observations for approximately 16 full days throughout the school year—before, during, and after treatment. While the manifest function of classroom observations is to obtain information concerning classroom characteristics and events, the latent function of such observations may be to facilitate treatment implementation. Minimally trained teachers in the Crawford et al. study unwittingly may have come to regard the relatively frequent and lengthy classroom observations as a kind of supervision or monitoring. If so, the conduct of classroom observations probably would have enhanced the compliance of these experimental group teachers with the training recommendations. The failure of experimental group teachers in the present study to implement the training recommendations, then, may have resulted from the relatively infrequent and brief classroom observations.

The plausibility of this conjecture must be evaluated in view of the results, reported by Anderson et al. (1979), which suggest that the absence of observers in the classes of experimental group teachers did not reduce treatment implementation. On the basis of their finding, one might argue that the comparatively low amount of classroom observation in the present study cannot be held responsible for the ineffectiveness of its training in bringing about the desired changes in teaching practices among experimental group teachers. There remains, however, a possibly fundamental difference between the present study and the one conducted by Anderson et al. (1979). In the present study, teachers never met with project staff for discussion, questions and answers, and the like; the TEP simply were mailed to the experimental group teachers. Anderson et al. (1979), in contrast, met twice with all experimental group teachers (i.e., including those trained but unobserved)—once to describe the purpose of the study and distribute the training material, and the second time to discuss the instructional model presented in the training material. These meetings, like those conducted by Good and Grouws (1979), probably fostered treatment implementation. First, the meetings doubtless were informative, facilitating understanding of the instructional model and its applicability. Second, by holding these meetings, the project staff were in a position to communicate enthusiasm for the training and personal concern for the teachers. The teachers' perception of the enthusiasm and concern could favorably dispose the teachers to the overall project and, in turn, enhance subsequent implementation. In short, these initial meetings with teachers may have served to address the general factors thought to influence the implementation of change proposals (e.g., Charters & Jones, 1973; Doyle & Ponder, 1977; Fullan & Pomfret, 1977; also see Mohlman, Coladarci, & Gage, 1982). The meetings in the Anderson et al. (1979) study, then, may have offset the absence of classroom observers for the trained but unobserved teachers.

It appears that for an intervention to be successful, the project staff must be engaged with participating teachers in some fashion. As Fullan and Pomfret (1977) stated within the larger context of curriculum and instruction implementation, "There is no substitute for the primacy of personal contact" (p. 391). Such contact might well be a necessary condition for successful implementation and, consequently, associated change in targeted outcomes. However, the dynamics of this contact and the corresponding relationships with implementation of change proposals remain to be clarified through systematic research (e.g., Mohlman, 1982).

Good and Grouws (1979) argued that their findings, along with those of Anderson et al. (1979) and Crawford et al. (1978), indicated that classroom-based experiments

> are capable of yielding improvement in student learning that are practically as well as statistically significant. Such data are an important contradiction to the frequently expressed attitudes that . . . brief, inexpensive treatments cannot hope to bring about significant results. (p. 361)

The results of this study should temper such optimism concerning the promise of the minimal intervention in research on teaching.

APPENDIX

Summary of Teaching Recommendations

Behavior Management and Classroom Discipline

1. Teachers should have a system of rules that allows pupils to attend to their personal and procedural needs without having to check with the teacher.

2. Teachers should prevent misbehaviors from continuing long enough to increase in severity or spread to and affect other children.

3. Teachers should attempt to direct disciplinary action accurately—that is, at the child who is the primary cause of a disruption.

4. Teachers should keep overreactions to a minimum (even though overreactions are probably effective in stopping the misbehavior).

5. Teachers (and aides, if present) should move around the room a lot, monitor pupils' seatwork, and communicate to pupils an awareness of their behavior, while also attending to their academic needs.

Instructional Methods

6. When pupils work independently, teachers should ensure that the assignments are interesting and worthwhile and still easy enough to be completed by each pupil working without teacher direction.

7. Teachers should keep to a minimum such activities as giving directions and organizing the class for instruction. They can do this by writing the daily schedule on the board, insuring that pupils know where to go and what to do, and so on.

8. Teachers should spend at least one-third to one-half of their time teaching larger groups of pupils (more than eight children). When they do teach smaller groups or individuals, they should take steps to make sure that the other pupils in the class have work to which they can attend.

9. Teachers should make abundant use of textbooks, workbooks, and other pencil-and-paper activities. These have been found to be associated with higher pupil achievement. But the use of games, toys, and machines has not been found to be associated with higher pupil achievement.

10. Teachers should provide visual demonstrations and phonics exercises in conjunction with reading activities.

11. Teachers should frequently conduct public (i.e., addressed to a larger group or the whole class) question-and-answer sessions concerned with the academic subject matter at hand. With less academically oriented pupils, teachers may find it helpful to initiate some brief private discussions concerning personal matters.

Specific Methods for Asking Questions and Providing Feedback

12. In selecting pupils to respond to questions, teachers should use the technique of calling on a child by name before asking the question, as a means of insuring that all pupils are given an equal number of opportunities to answer questions.

13. Teachers should avoid calling on volunteers more than 10 or 15% of the time during question-and-answer sessions. It is also advisable to discourage pupil "call outs" to questions asked of other children (except possibly from less academically oriented children who may benefit from this type of activity).

14. In the interest of promoting smooth, task-oriented discussions, teachers should not encourage large numbers of pupil-initiated questions and comments. It is also important for teachers to listen carefully to pupils' opinions and, if a disagreement is called for, to express such disagreement to the child.

15. With less academically oriented pupils, teachers should ask easier questions—questions that can almost always be answered correctly. When questioning more academically oriented pupils, teachers should ask more difficult questions—questions that are answered incorrectly about one-fourth of the time.

16. Teachers should give praise only for really outstanding work; also, praise is likely to be more effective with less academically oriented pupils. Mild criticism is effective in communicating higher expectations ("you can do better") to more academically oriented pupils.

17. With less academically oriented pupils, teachers should always aim at getting the child to give some kind of response to a question. Rephrasing, giving clues, or asking a new question can be useful techniques for bringing forth some answer from a previously silent pupil or one who says "I don't know" or answers incorrectly.

18. With more academically oriented pupils who generally become actively involved in discussions, teachers should concentrate on getting the correct response. Therefore they should redirect questions to other pupils if the more academically oriented pupil answers incorrectly.

19. Teachers should give the answer (to both more and less academically oriented pupils) if the response is at least partly correct. Teachers should not simply repeat the same questions if any pupil (either more or less academically oriented) answers incorrectly, says "I don't know," or remains silent.

20. With more academically oriented pupils, teachers should give brief feedback extensively (80% or more of the time) during private, one-to-one discussions. When dealing with less academically oriented pupils, teachers should use approximately equal amounts of brief and longer feedback, tailoring the duration of their reactions to the needs of the child in each situation.

21. During reading group instruction, teachers should give a maximal amount of brief feedback, and provide fast-paced activities of the drill type.

22. During public question-and-answer sessions, teachers should occasionally give a detailed "why" explanation in answer to a question.

REFERENCES

ANDERSON, L., EVERTSON, C., & BROPHY, J. (1979). An experimental study of effective teaching in first-grade reading groups. *Elementary School Journal, 79,* 193–223.

BERLINER, D. (1979). Tempus educare. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings, and implications* (pp. 120–135). Berkeley, CA: McCutchan.

BERLINER, D., & ROSENSHINE, B. (1977). The acquisition of knowledge in the classroom. In R. Anderson & W. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 375–396). Hillsdale, NJ: Erlbaum.

BROPHY, J. (1977). *Training teachers in experiments: Considerations relating to nonlinearity and context effects.* Paper presented at the annual meeting of the American Educational Research Association, New York City.

BROPHY, J. (1979). Teacher behavior and its effects. *Journal of Educational Psychology, 71,* 733–750.

BROPHY, J., & EVERTSON, C. (1974). *Process-product correlations in the Texas Teacher Effectiveness Study: Final report* (Research Report No. 74-4). Austin: Research and Development Center for Teacher Education, University of Texas.

BROPHY, J., & PUTNAM, J. (1979). Classroom management in the elementary grades. In D. Duke (Ed.), *Classroom management* (pp. 182–216). The 78th yearbook of the National Society for the Study of Education, Pt. 2. Chicago: University of Chicago Press.

CHARTERS, W., & JONES, J. (1973). On the risk of appraising non-events in program evaluation. *Educational Research, 2*(11), 5–7.

CRAWFORD, J., GAGE, N. L., CORNO, L., STAYROOK, N., MITMAN, A., SCHUNK, D., STALLINGS, J., BASKIN, E., HARVEY, P., AUSTIN, D., CRONIN, D., & NEWMAN, R. (1978). *An experiment on teacher effectiveness and parent assisted instruction in the third grade* (3 vols.). Stanford, CA: Center for Educational Research at Stanford.

CRAWFORD, J., & STALLINGS, J. (1978). *Experimental effects of in-service teacher training derived from process-product correlations in the primary grades.* Paper presented at the annual meeting of the American Educational Research Association, Toronto.

CRONBACH, L. (1970). *Essentials of psychological testing* (3rd ed.). New York: Harper & Row.

CRONBACH, L., GLESER, G., NANDA, H., & RAJARATNAM, N. (1972). *The dependability of behavioral measures: Theory of generalizability for scores and profiles.* New York: Wiley.

DOYLE, W., & PONDER, G. (1977). The practicality ethic and teacher decision-making. *Interchange, 8*(3), 1–12.

DUNKIN, M., & BIDDLE, B. (1974). *The study of teaching.* New York: Holt, Rinehart and Winston.

ELASHOFF, J. (1969). Analysis of covariance: A delicate instrument. *American Educational Research Journal, 6,* 383–402.

FULLAN, M., & POMFRET, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research, 47,* 335–397.

GAGE, N. L. (1972). *Teacher effectiveness and teacher education: The search for a scientific basis.* Palo Alto, CA: Pacific Books.

GAGE, N. L. (1978). *The scientific basis of the art of teaching.* New York: Teachers College Press.

GAGE, N. L., & GIACONIA, R. (1981). Teaching practices and student achievement: Causal connections. *New York University Education Quarterly, 12*(3), 2–9.

GOOD, T. (1979). Teacher effectiveness in the elementary school. *Journal of Teacher Education, 30,* 52–64.

GOOD, T. (1980). Research on teaching. In G. E. Hall, S. M. Hord, & G. Brown (Eds.), *Exploring issues in teacher education: Questions for future research* (pp. 51–73). Austin: Research and Development Center for Teacher Education, University of Texas.

GOOD, T., & GROUWS, D. (1977). Teacher effects: A process-product study in fourth-grade mathematics classrooms. *Journal of Teacher Education, 28,* 49–54.

GOOD, T., & GROUWS, D. (1979). The Missouri Mathematics Effectiveness Project: An experimental study in fourth-grade classrooms. *Journal of Educational Psychology,* 71, 335–362.

KOUNIN, J. (1970). *Discipline and group management in classrooms.* New York: Holt, Rinehart and Winston.

LINN, R., & SLINDE, J. (1977). The determination of the significance of change between pre- and posttesting periods. *Review of Educational Research, 47,* 121–150.

MCDONALD, F., & ELIAS, P. (1976). *The effects of teaching performance on pupil learning. Beginning Teacher Evaluation Study: Phase II, Final report* (Vol. 1). Princeton, NJ: Educational Testing Service.

MEDLEY, D. M. (1982). Teacher effectiveness. In H. E. Mitzel, J. H. Best, & W. Rabinowitz (Eds.), *Encyclopedia of educational research* (pp. 1894–1903.). New York: Free Press.

MOHLMAN, G. (1982). *Assessing the impact of three inservice teacher training* *models.* Paper presented at the annual meeting of the American Educational Research Association, New York.

MOHLMAN, G., COLADARCI, T., & GAGE, N. L. (1982). Comprehension and attitude as predictors of implementation of teacher training. *Journal of Teacher Education, 33,* 31–36.

NIE, N., HULL, C., JENKINS, J., STEINBRENNER, K., & BENT, D. (1975). *Statistical package for the social sciences* (2nd ed.). New York: McGraw-Hill.

NUTHALL, G., & CHURCH, J. (1973). Experimental studies of teaching behavior. In G. Chanan (Ed.), *Towards a science of teaching* (pp. 9–25). London: National Foundation for Educational Research.

POWELL, M. (1978). Research on teaching. *The Educational Forum, 4,* 27–37.

ROGOSA, D. (1980). Comparing nonparallel regression lines. *Psychological Bulletin, 87,* 307–321.

ROSENSHINE, B. (1976). Classroom instruction. In N. L. Gage (Ed.), *The psychology of teaching methods* (pp. 335–371). The 75th yearbook of the National Society for the Study of Education, Pt. 1. Chicago: University of Chicago Press.

ROSENSHINE, B. (1979). Content, time, and direct instruction. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings, and implications* (pp. 28–56). Berkeley, CA: McCutchan.

ROSENSHINE, B., & BERLINER, D. (1978). Academic engaged time. *British Journal of Teacher Education, 4,* 3–16.

ROSENSHINE, B., & FURST, N. (1973). The use of observation to study teaching. In R. Travers (Ed.), *Second handbook of research on teaching* (pp. 122–183). Chicago: Rand McNally.

SAS INSTITUTE, INC. (1979). *SAS user's guide.* Raleigh, NC: Author.

SOAR, R. (1973). *Final report: Follow through classroom process measurement and pupil growth (1970–1971).* Gainesville: Institute for the Development of Human Resources, University of Florida.

STALLINGS, J., & KASKOWITZ, D. (1974). *Follow through classroom observation evaluation (1972–1973).* Menlo Park, CA: SRI International.

WALKER, D. F. (1980). *Practical inquiry and impractical experimentation.* Paper presented at the annual meeting of the American Educational Research Association, Boston.

AUTHORS

THEODORE COLADARCI, Assistant Professor, College of Education, University of Maine at Orono, Orono, ME 04469. *Specialization:* Educational psychology.

N. L. GAGE, Margaret Jacks Professor of Education and, by courtesy, Professor of Psychology, School of Education, Stanford University, Stanford, CA 94305. *Specializations:* Educational psychology, research on teaching.