

A Two-Year Follow-Up of a Staff Development Program Designed to Change Teacher Behavior

Eugene Schaffer

University of North Carolina at Charlotte

Samuel Stringfield

Johns Hopkins University

Roberta Devlin-Scherer

Drake University

ABSTRACT

Two years after participating in a replication of the Stallings Effective Use of Time (EUOT) Program, ten teachers were re-observed and interviewed to determine the extent to which they had maintained the measured changes in their behavior patterns. Subjects were selected for the follow-up from a 27 EUOT teacher sample based on having exhibited substantially more, or substantially less than average pre-to post-workshop desired behavioral change. On the behavioral variable, All Academic Statements, both relatively high and low implementers remained somewhat above their pre-workshop means. Neither group remained significantly improved on the aggregated variable, Organizing/Off-Task Behaviors. Interviews revealed that two years after training the high implementers recalled more specific, behavioral recommendations than did teachers who two years previously had been low implementers. The implications of these findings for research and practice are discussed.

INTRODUCTION

Several studies have indicated that teacher inservice education programs, which are based on prior process product research can produce significant changes in teachers' in-class behavior patterns. Concluding an otherwise positive review of nine such studies, Gage (1985) added two caveats.

First, we have no knowledge of how long the effects of these experimental training programs last; no follow-up studies have been made. Second, all of these nine training efforts were conducted in the context of research enterprises, or experiments. It is not known whether similar training would be effective in the context of regular, routine staff development in the schools (pp. 34-35).

Those two external validity issues need to be addressed for teacher effectiveness quasi-experiments to be of value to the great majority of teachers and administrators. The goal of inservice education is to have a positive, lasting impact on teachers. By presenting classroom observational data gathered in conjunction with a local staff development effort, and data gathered two years after the completion of that effort, the current paper address aspects of both issues.

Studies of the extended impact of teachers' inservice education are rare. Two types of follow-up efforts will be reviewed here. The first are studies in which programs were followed for at least one year after formal completion, and in which the dependent variables included measures of teachers' in-class behavior. The authors were able to identify three studies which met these criteria. Second, given that the particular program followed in the current research was a replication of the Stallings Effective Use of Time Program (EUOT), we will review the two documented efforts to follow the effects of the EUOT.

Johnson and Sloat (1980) reported the effects of a university extension course which was designed to alter teachers' use of behavior modification in the classroom (e.g., provide more academic praise, fewer nonverbal negatives). The course included lectures, films, observations of model activities, peer observations, and two weeks of guided practice in the specified techniques. Low-inference classroom observational data were gathered on the 13 participating teachers between weeks 4 and 16 of the course. Follow-up data were gathered 5 and 12 months after the course, and during a reinstated feedback period immediately following the 12 month follow-up.

On most of the variables measured, the group moved in the desired direction throughout the course. Some changes were dramatic. For example, the rate of "Academic Praise" rose 350%. However, at the 5 and 12-month follow-ups, and during the reinstated feedback period, teachers' rates of specified behaviors were not significantly above the first phase of the study.

The authors concluded that the demand characteristics of training situations may have influenced teacher behavior during the course. That is, the teachers knew what they were "supposed to do" and may have done those things in an effort to please, or appear competent to, the researchers.

Two studies have followed graduates of microteaching courses over 1-3 years. The more effective program, in terms of one-year effect (Trinchero, 1975) was also the one in which time was almost certainly confounded with additional training on the measured variables. The subjects were continuing their involvement in an Education Master's degree program. In a relatively unconfounded microteaching context, Borg (1972) found that behavioral gains on 7 of 13 variables were maintained for 39 months. Confidence in the generalizability of these studies is limited by the fairly obvious demand characteristics of the situations and by reservations (See e.g., Berliner, 1969) concerning the transfer of training from microteaching to regular classrooms.

THE EUTP AND THE STALLINGS OBSERVATION SYSTEM

In "Implications of Research for Staff Development Programs," Griffin (1983) said of the EUOT, "The Stallings model is one of the few inclusive attempts to utilize actual teaching behavior as a basis for determining subsequent staff development activities" (p. 420). Detailed descriptions of the Stallings program have been published elsewhere (e.g., Stallings, Needels, & Stayrook, 1978; Stallings, 1980; Bush, 1985; Anderson, 1985). The three central features of the EUOT are direct, quantified, classroom observation; extended small group workshops which include behavioral feedback; and the training of local EUTP leaders.

THE STALLINGS OBSERVATION SYSTEM

Each teacher in the EUOT is observed for three one-hour periods before, and again after the workshop series. Data are gathered using the Stallings Observation System (SOS).

The SOS is a complex, low inference observation system designed to be sensitive to differing instructional methods interpersonal interactions and classroom environments. Using the SOS, an observer alternately gathers data on the teacher and his/her immediate environment, and on the entire class. In its present form, or with minor modifications, the SOS has been a major process data gathering instrument in several school and teacher effectiveness studies (e.g., Stallings & Kaskowitz, 1974; Stallings, 1980; Goodlad, 1983; Stringfield, Teddlie & Suarez, 1985; Stringfield & Teddlie, 1988).

Data from two sections of the SOS were analyzed in the current study. The first SOS subsystem is the Five Minute Interaction (FMI). The FMI consists of a series of frames in which each teacher behavior and/or interaction is coded into four categories: who (e.g., teacher), to whom (e.g., small group), what (e.g., asks a direct question), and how (e.g., with positive emotion). During a FMI, a minimum of one frame is recorded every six seconds, yielding 50 or more frames per five minute observation period. The analysis produces 19 separate FMI variables.

One FMI variable was analyzed for this paper: All Academic Statements (AAS). It is an aggregated variable and includes all academically related talking by teachers to students (individually or in groups) and students' academic talking to teachers. AAS has been found by Stallings (1980) to be both related to achievement and alterable.

The second SOS subsystem, labeled the Classroom Snapshot (CS), yields data on the activities of each adult and student in a classroom at a given moment. It contains codings for size of groups and types of materials being used. The CS is designed to gather data on 13 activities (e.g., reading silently), by eight types of materials (e.g., textbook), by four types of actors (e.g., teachers, individual students), by four recipient groups (e.g., large groups).

The teacher activity analysis of the CS yields 13 variables. Used for clinical training purposes with individual teachers, these categories are discussed separately with each teacher. However, for research purposes, the fact that the frequency of occurrence of some of the categories is very low produces instability in some analyses. In the present paper, note will be made of individual items, but statistical analyses of CS data will focus on two aggregated CS variables: Interactive Teaching (IT) and Organizing/Off-Task activities (O/OT).

The CS variables aggregated to product IT include reading aloud, instruction and explanation of new materials, review and discussion, and practice drill. Stallings (1980) found IT to be both a positive predictor of student achievement gain and an alterable variable. Passing out papers and lining up students for recess are examples of organizing activities. Off-task activities include social interactions and negative interactions (ex. discipline). Spending a large percentage of class time in O/OT has been found negatively correlated with achievement.

Several studies of the reliability of the SOS and its precursor have been conducted. Stallings and Kaskowitz (1974) conducted four reliability studies. They reported that on the great majority of the variables interrater agreements above .70 could be achieved.

They further concluded that data from the instrument tended to be reliable across days. Their observers gathered data in over 100 classes for two full days. The day one to day two reliability (Spearman-Brown Formula) ranged from .68 to .89 on individual variables. Sirotnik (1984) analyzed interrater agreements on

variables from a modified SOS, and found acceptable reliability quotients for all frequently occurring variables.

Neither Stallings and Kaskowitz nor Sirotnik examined the reliability of aggregated variables, though those could be reasonably assumed to be higher than the reliabilities of their parts. Stringfield, Teddlie and Suarez (1985) presented data relevant to the across time stability of one of the aggregated SOS variables: Interactive Teaching. Gathering six hours of CS data on 5 teachers in the fall and again in the spring of the 1984-1985 school year, the authors noted that the rank ordering of the five teachers did not change from fall to spring. The maintaining of rank ordering over time is one criterion suggested by Rogosa, Floden and Willett (1984) for assessing the consistency of individual differences over time.

EUOT WORKSHOPS

The second feature of the Stallings program is a series of five highly interactive, small group workshops. In an effort to maximize participation, group size is limited to six to ten teachers plus leaders.

In a series of five 2.5 hour meetings, teachers review their SOS profiles, aspects of the teacher effectiveness literature (e.g., classroom management, Academic Learning Time, use of praise, grouping), and their own efforts at behavioral change.

Following the fifth meeting, three additional hours of SOS data are gathered. A final meeting is held in which teachers receive a printout comparing their pre- and post-workshop behavior profiles.

TRAINING LOCAL EUOT LEADERS

The third feature of the EUOT is the training of local teacher trainer/leaders. In the first stage of EUOT, an external consultant trains local personnel in the use of the SOS, and in the running of workshops. Typically, the consultant runs a workshop series and either directly or through video and audio-tapes, observes workshops run by the trainees (for a more complete description of this process, see Stallings, Needels and Stayrook, 1979; Bush, 1984).

In the current study, 6 local leaders were trained by one consultant. Following Stallings, Needels and Stayrook (1979), the trainees were not allowed to observe classes until they reached the criterion of .80 or above on the SOS variables. They did not lead groups until they had observed the consultant leading a group, and they received feedback on their efforts after each workshop meeting.

EUTP FOLLOW-UP STUDIES

Two follow-up studies of the EUOT have been reported. One met the criterion of being behavior based, the other met the second criterion of being conducted at least a year after the original training. Neither met both criteria.

Stallings (1980) observed teachers at three points during one school year. Between the fall and mid-winter observations, the 25 experimental teachers attended EUOT workshops. Experimental teachers made significant fall to winter gains as measured on a precursor of the SOS, and maintained those gains at the 3 month follow-up. By contrast, the 19 control classes did not gain fall to winter, and became "more lax and less task oriented" (p. 15) at the spring observations.

Bethune (1983) conducted a one year follow-up of teachers in a replication of the EUOT. In a self-report survey, teachers reported continued use of workshop concepts and raised student achievement.

As happens in any program, some teachers in the Stallings research implemented suggested behaviors to a greater extent than did others. The

complacency of teachers (Stallings, 1983), value of the change as seen by teachers, and the teachers' prior familiarity with the practices (Sparks, 1983) all have been identified in prior research as contributing to differential inservice effectiveness.

This distinction between high and low implementers of staff development raises questions for staff development and adds nuances to the question of long term effect. How great are the differences between high and low implementers? Are the differences maintained over time? The current study was conducted to examine the two year residual effects of the EUOT on teachers who had been high or low implementers.

METHOD

Twenty-seven volunteer teachers from two small town and rural districts in North Carolina participated in four offerings of the EUOT during the winter and spring of 1982. Among the subjects, teaching experience ranged from four to over 30 years, grade level of instruction ranged from first grade to high school, and subjects taught by high school teachers included English, Science, and Mathematics.

As recommended by Stallings (1983), three one hour SOS observations were gathered on each teacher before and after participation in the workshop series. An analysis of pre-post change for the 21 teachers on whom complete data bases were available indicated that the group moved in the desired direction on 27 of 32 behavioral variables analyzed (sign test prob. $< .001$), and that the mean teacher behavior rate on two of three aggregated variables (AAS and O/OT) moved significantly in the desired direction. Change in IT tended to move in the desired direction, but the shift did not reach statistical significance. No between training group differences were significant (Stringfield, Schaffer, & Devlin-Scherer, 1986).

To obtain samples of high and low implementing teachers for the follow-up study, regressed gain scores were computed for each of the 21 teachers on the three aggregated behavioral variables (AAS, IT, and O/OT). The five teachers whose pre- to post-workshops gains were above the group average on all three variables were considered to be high implementers. Five teachers' gains were below the mean on all 3 variables, these teachers were considered to be low implementers. Those 10 teachers were asked to participate in the follow-up study, and all agreed to participate.

INSTRUMENTS

Two instruments were used in the follow up study. To determine teachers' continuing level of EUOT implementation, two separate one hour SOS observations were made. No teacher was observed twice on one day. Only regular academic periods were observed.

All follow-up data were gathered by observers involved in pre- and post-EUOT SOS data gathering. All observers were required to demonstrate reliability of the SOS before each observation cycle. Efforts were made to maximize the similarity of conditions of observation across the three data gathering cycles (e.g., observing comparable instructional groups engaged in comparable instructional content).

After all classroom observations were completed, an open-ended survey was constructed to determine teachers' recollections of the EUOT, and the extent to which they believed it continued to affect their instructional practices. Teachers were asked to recall the major precepts of EUOT. Some teachers needed prompting in order to recall themes. The interviewees were then given the titles

Table 1

Means and Standard Deviations for Three SOS Aggregated Variables for High and Low EUOT Implementers at Three Points in Time, and Contrasting Total Sample Data

Implementation	(n)	Pre-EUOT		Post-EUOT		Follow-Up	
		X	(SD)	X	(SD)	X	(SD)
All Academic Statements							
High	(5)	81.9	(17.0)	95.8	(3.1)	93.1	(3.2)
Low	(5)	74.7	(15.2)	87.1	(4.3)	83.4	(7.7)
Total	(21)	82.0	(12.5)	92.2	(4.5)		
Organizing/Off-Task							
High	(5)	14.6	(14.4)	1.3	(2.9)	20.0	(20.0)
Low	(5)	13.3	(13.3)	13.5	(4.8)	18.0	
Total	(21)	15.1	(13.5)	8.8	(8.3)		

of the five workshops and asked to remember specific concepts or activities they had learned.

After the content of the workshops was re-established, the teachers were asked to recall their use of EUOT. Probes were designed to elicit specific examples. Teachers were asked to describe changes they had observed in student behavior or performance as a result of the workshops.

ANALYSIS

Following Borg (1972), follow-up analyses were conducted only on variables which had shown significant pre- to post- workshop gain; in this case, AAS, O/OT and a sign test for the 32 variable set.

On AAS and O/OT, two way analyses of variance (high vs. low implementing groups, and pre-workshop vs. follow-up observations) were computed. Separate pre- to followup sign tests were computed for the 32 variable change for the two groups. Content analyses of the open ended questions were undertaken.

Table 2

Analysis of Variance: Implementation Level (High vs. Low) by Time (Pre-workshop vs. Follow-Up) for Two SOS Aggregated Variables

Variable	Source	df	SS	MS	F	p
All Academic Statements	Between Ss					
	Group (G)	1	360.25	360.15	1.96	
	S (G)	8	1466.37	183.30		
	Within Ss					
	Time (T)	1	496.11	496.11	4.42	.069
	G X T	1	8.03	8.03	0.07	
	T X S (G)	8	897.01	112.13		
Organizing/Off-Task	Between Ss					
	Group (G)	1	13.61	13.61	0.03	
	S (G)	8	3814.34	476.79		
	Within Ss					
	Time (T)	1	127.51	127.51	0.56	
	G X T	1	0.61	0.61	0.00	
	T X S (G)	8	1808.34	226.04		

RESULTS

An analysis of individual SOS items indicated that the high implementers moved in the desired direction from pre to post on 29 or 32 items (sign test $p < .0001$). From pre- to follow-up, however, it dropped to 18 of 32 (n.s.). Similarly, the lowest implementers gained pre- to post- on 24 of 32 variables ($p < .01$), but by follow-up the number of desired-direction changes had fallen to 17 of 32 (n.s.). Table 1 presents descriptive data on the high and low implementer teachers pre, post-and follow-up on the two aggregated behavioral variables, AAS and O/OT.

Table 1 shows an almost parallel movement between the groups on the aggregated All Academic Statements. Both groups began near or at acceptable (80%+) levels. Both moved to extremely high levels of AAS and then dropped slightly at the two-year follow-up. As can be seen in Table 2, the main effect for Time (pre- to follow-up) is marginally significant ($p < .067$), and the interaction and group effects are not significant.

The EUOT criterion for aggregated Organizing/Off Task is study 20% or with less. an Both acceptable high and level low of organizing implementers and began the study with an acceptable level of organizing and off-task time. The high implementers dropped to near zero, then rose to 20% at follow-up. The Low implementers did not drop at post-workshop, and rose slightly at follow-up. Both groups' criterion means range, and pre-, post- and follow-up were within the criterion range, and changes were not significant.

The open ended survey of high and low implementers revealed that all ten participants valued the sessions and most considered the workshops helpful. As one teacher stated, "None of us begrudged the time, because we could see progress and how our job was made easier." Yet there were distinctions between the groups.

Regarding the feedback data they received on their classroom behavior, the high implementers reported that the workshops supplied them with present and future assistance. One teacher stated, "I could take the strategies right back to the classroom and use them the next day." Another found, "Much I could use immediately...other ideas I stored away till the beginning of the next year." Some stated that they gained an awareness of the process of teaching, "I could see what went on every minute. I was conscious of my every movement; my every decision!" One high implementer stated, "I'm more sensitive to the impact of my behavior."

In clear contrast, the low implementers tended to report that their pre-workshop behavior patterns had been reinforced by the workshops and that they had no need for major changes. As one stated, "The workshops confirmed what I knew." Another observed, "It is good to have a review and a chance to talk to friends." The social theme was echoed by a third teacher, "I like getting together with friends that I rarely talk to during the day."

When teachers were asked to name specific things they had learned and implemented based on the workshops, the low implementers noted the general sharing of ideas and the workable solutions generated in the workshops. The high implementers named many more *specific* skills and behaviors they had learned or sharpened. The increased use of reading aloud, varying the number of activities in a class period, distributing opportunities for responses among students and increased behavioral awareness of students were all named as specific changes the high-implementer teachers had tried to establish and maintain.

DISCUSSION

This two-year follow-up of a research based, behaviorally oriented staff development program was designed to address two questions. First, did the EUOT, which had produced clear pre- to post-workshop gains on most measures in the larger sample, produce significant long term effect on teacher behavior? Second, were between-teacher differences in implementation levels maintained over two years?

Regarding the maintenance of change across time, the results were mixed. Classroom observational data gathered using the SOS tended to indicate that an enhanced rate of AAS remained after two years. The teachers' mean level of O/OT behavior, by contrast, had returned to pre-EUOT (though not unacceptable) level. Sign tests for directional change across the full 32 variable profile indicated a highly significant pre- to post-EUOT effect; however, this gain was not maintained two years after the workshops.

The ten teachers began the study with fairly high levels of academic interaction, and low levels of organizing/off task activities. Most of the teachers in the original study, and many of those in the follow up study, obtained pre-EUOT scores on AAS and O/OT which were within the ranges suggested by Stallings (1980) for effective teaching. Post EUOT scores, especially for the high-implementer group, were extremely high. It is possible that the moderate regressions were in part results of a ceiling effect.

Further, the EUOT does not assume that the highest possible implementation scores are optimal in all contexts. Soar (1968, 1977) and Soar and Soar (1979) have noted that there may be an optimal level of academic work that varies from grade to grade and class to class. Stallings (1989) has noted that context must be considered in determining an optimal score in any class. An advanced placement algebra class might strive to maintain the extremely high levels of AAS noted in the high implementer, post-EUOT data. Maintaining such levels of academically focused activity may be neither practical nor desirable in a mixed ability second grade reading group. Given that the current study included teachers from two districts and a variety of grades and content areas, a strict interpretation of the quantitative data would indicate that the teachers maintained reasonably high levels of AAS in the follow-up. In the judgment of observers who gathered data pre-, post-, and follow-up; however, several classes' quantitative drops represented regressions in the quality of instruction.

We believe that a partial explanation of this moderate regression from the extremely high post-EUOT scores lies in the continuing culture and climate of schools. Three recent studies have nested measures of classroom process variables within school effectiveness designs (Barr & Dreeban, 1983; Mortimore et al., 1988; Teddlie, Kirby & Stringfield, 1989). In each case clear mean classroom process differences among schools were found. These studies indicate that a cultural and behavioral set of norms exists within schools. Though large scale studies of efforts to change behavioral norms within schools are lacking, it is reasonable to assume that such norms, if unchanged, will reassert themselves on any small group of teachers. Such an assumption would predict a moderate regression to the local mean after any intervention.

Anecdotal evidence supports the role of school culture in the current study. At one school, for example, teachers attending the EUOT informed their principal that indiscriminate use of the school intercom system reduced students time on task. The well-intended principal's response was to continue using the intercom at his convenience, but to begin each interruption with an apology. This served to lengthen the interruptions. After several weeks, the teachers chose to discontinue complaints about the interruptions, accepted this dysfunctional behavior as an unfortunate part of a generally well-regarded principal.

There are aspects of a class over which individual teachers have marginal control (e.g., the intercom, choice of spelling texts). Changes in those areas are likely to require huge efforts, efforts which teachers are likely to eventually view as being cost-ineffective. Such a determination on the parts of individual professionals can be expected to result in the moderate regressions noted in the current study.

There are other areas over which an individual professional may exert considerable control (e.g., calling on all students, accepting no call-out responses, not grading students' papers during class time). An individual's focus on those areas can result in permanent change. We believe that the long-term gains in rates of teacher-student academic interactions found in the current study are evidence of such change.

The second issue addressed in this study concerned the sustaining of differences in levels of implementation by various teachers. The highly significant behavioral differences in pre- to post-EUOT gains between high and low implementers disappeared on all measures at follow-up.

Questionnaire and interview data from the teachers revealed a unanimous enthusiasm for the program. In contrast to the classroom observational data, clear differences remained between teachers who were originally high vs. relatively low implementers. High implementers' remembrances of the workshops contained many more specific, behavioral suggestions for change. Lower implementers remembered the workshops as reinforcing their pre-workshop teaching patterns.

Two years after the EUOT was conducted in the two district, originally high implementers remembered more of the information from the program. Yet except for the area of All Academic Statements, they appeared to be acting on their knowledge no more than the original relatively low implementing teachers. It is possible that the study's sample was not sufficiently large to demonstrate significant effects.

We believe that a more plausible explanation involves the continuing culture and climate of the schools, discussed above. Regardless of differences in initial levels of enthusiasm and implementation, teachers returned to schools with ongoing cultures and behavior patterns. We believe that those norms reasserted their influence on teachers' inclass behaviors after the innovation was completed. If confirmed by additional analyses, such a finding would argue for a focus on teachers nested within schools in future change-maintenance studies. Our data suggest that while individual teachers can, in isolation from the larger school community, develop and maintain knowledge gains; in the absence of long-term support, teacher's *behavior* patterns tend to revert to school and system level norms.

Three implications may be drawn from the current research. First, more long-term impact studies are needed. Many successful teacher level quasi-experimental change studies have been reported. The current paper is only the second to have reported an extended follow-up.

Second, until additional follow-up studies report more maintenance of effect than was found in the present analyses, caution should be exercised in claiming substantial impact as a result of the application of teacher effectiveness research. Change which does not stabilize in the field is of limited value.

Third, research is needed on the variables determining the level of maintenance of teacher change. Pullan and Pomfret (1977) have offered a plausible set of opening hypotheses in this area. Pullan, Bennett, and Rolheiser-Bennett (1989) continue to find two dimensions of teacher-change limitation. The first concern, that most innovations are "abstract, vague, complex and unproven" has been addressed in the design of and prior research on the EUOT. The second, that insufficient attention has been paid to follow-through

or supporting implementation, is addressed during EUOT implementation, but is typically not followed-up. We believe that further follow-up studies should examine the role of the school and district in sustaining innovation. Research in this area could add considerable value to both the school improvement and teacher effectiveness literatures. ■

REFERENCES

- Anderson, S. (1985). Teacher training techniques from four observational perspectives. *Journal of Classroom Interaction*, 20(1), 16-28.
- Barr, R., & Dreeben, R. (1983). *How schools work*. Chicago: University of Chicago Press.
- Berliner, D. (1969). *Microteaching and the technical skills approach to teacher training*. Stanford, California: Technical Report No. 8, Stanford Center for Research and Development in Teaching.
- Bethune, G. (1983, April). *Academic learning time: The current status of the Stallings training*. Paper presented at the American Educational Research Association, Montreal.
- Borg, W. (1972). The minicourse as a vehicle for changing teacher behavior: A three-year follow-up. *Journal of Educational Psychology*, 63, 572-579.
- Bush, M. (1984). The complexity of institutionalizing a program: Acquisition of training, observing, and computing capability. *Journal of Classroom Interaction*, 20(1), 6-15.
- Fullan, M., Bennett, B., & Rolheiser-Bennett, C. (1989). *Linking classroom and school improvement*. Paper presented at the American Educational Research Association, San Francisco.
- Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research*, 47(1), 335-397.
- Gage, N. (1985). *Hard gains in the soft sciences: The case of pedagogy*. Bloomington, IN: Phi Delta Kappa.
- Griffin, G. (1983). Implications of research for staff development programs. *Elementary School Journal*, 83, 414-425.
- Goodlad, J. (1984). *A place called school*. New York: McGraw-Hill.
- Huberman, A., & Miles, M. (1984). *Innovation up close*. New York: Plenum.
- Johnson, J., & Sloat, K. (1980). Teacher training effects: Real or illusory. *Psychology in the Schools*, 17(17), 109-115.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters*. Somerset, England: Open Books.
- Sirotnik, K. (1984). An inter-observer reliability study of a modified SRI observation system. *Journal of Classroom Interaction*, 19(1), 28-38.
- Soar, R. (1968). Optimum teacher-pupil interaction for pupil growth. *Educational Leadership*, 26, 275-280.
- Soar, R. (1977). An integration of findings from four studies of teacher effectiveness. In G. Borich & K. Fenton (Eds.), *The appraisal of teaching: Concepts and process*. Reading, MA: Addison-Wesley.
- Soar, R., & Soar, R. (1979). Emotional climate and management. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings and implications*. Berkeley, CA: McCutchan.
- Sparks, G. (1984). Synthesis of research on staff development for effective teaching. *Educational Leadership*, 41(3), 65-72.
- Stallings, J. (1989). Personal communication.
- Stallings, J. (1983). *Effective Use of Time program*. Nashville, TN: Vanderbilt University.
- Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Researcher*, 9(12), 11-16.
- Stallings, J., & Kaskowitz, D. (1974). *Follow through classroom observation evaluation (1972-1973)*. Menlo Park, CA: SRI International.
- Stallings, J., Needels, M., & Stayrook, N. (1979). *How to change the process of teaching basic reading skills in secondary schools: Phase II and Phase III*. Menlo Park, CA: SRI International.
- Stringfield, S., & Teddlie, C. (1988). A time to summarize: Six years and three phases of the Louisiana School Effectiveness Study. *Educational Leadership*, 46(1), 43-49.
- Stringfield, S., Teddlie, C., & Suarez, S. (1985). Classroom interaction in effective and ineffective schools: Preliminary results from Phase III of the Louisiana School Effectiveness Study. *Journal of Classroom Interaction*, 20(2), 31-37.
- Stringfield, S., Schaffer, E., & Devlin-Scherer, R. (1986). A study of the generalizability of teacher change quasi-experiments. *Journal of Classroom Interaction*, 21(2), 9-15.
- Teddlie, C., Kirby, P., & Stringfield, S. (1989). Effective versus ineffective schools: Observable differences in the classroom. *American Journal of Education*, 97(3), 221-236.
- Trincherio, R. L., Jr. (1975). *Three technical skills of teaching: Their acquisition, stability and effects on student performance*. Unpublished doctoral dissertation. Stanford University.