DOCUMENT RESUME

ED 275 067 EA 018 917

AUTHOR	Roper, Susan Stavert; Hoffman, David E.
TITLE	Collegial Support for Professional Improvement: The Stanford Collegial Evaluation Program.
INSTITUTION	ERIC Clearinghouse on Educational Management, Eugene, Oreg.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Mar 86
CONTRACT	400-83-0013
NOTE	36p.; Prepared by the ERIC Clearinghouse on Educational Management, Eugene, Oregon.
AVAILABLE FROM	Publication Sales, Oregon School Study Council, University of Oregon, 1787 Agate Street, Eugene, OR 97403 ($4.00 prepaid; quantity discounts; $1.50 shipping and handling will be added to billed orders).
PUB TYPE	Collected Works - Serials (022) -- Reports - Descriptive (141)
JOURNAL CIT	OSSC Bulletin; v29 n7 Mar 1986
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Administrator Role; Elementary Secondary Education; *Formative Evaluation; *Participant Observation; Participation; *Peer Evaluation; Peer Relationship; Performance; Principals; *Professional Development; *Staff Development; Teacher Attitudes; *Teacher Evaluation; Teaching Conditions
IDENTIFIERS	*Collegial Teams; Isolation (Professional); Stanford Collegial Evaluation Program

ABSTRACT

To overcome teacher isolation and help make teaching a more attractive profession, educators might consider a peer support system based on the Stanford Collegial Evaluation Program. This paper describes the program's background, functioning, barriers to successful implementation, and how to neutralize them. The Stanford program uses a peer evaluation process and emphasizes formative, not summative, evaluations. Linked in reciprocal relationships, teachers collect information on their performance from collegial observations, student questionnaires, and self-assessment. They analyze strengths and weaknesses and prepare a plan for improvement. The program's effectiveness lies in use of first-hand observations and stress on factors that teachers consider important. The program has seven interdependent steps leading to an improvement plan. Major barriers to collegial evaluation such as teacher attitudes, lack of criteria validity and reliability, problems with administrator involvement, lack of teacher candor, and insufficient released time for observation may be overcome in a supportive school context by the right mixture of determination, good will, and humor. The key is wholehearted commitment to professional development. Included are 18 references and information on obtaining the implementation manual. (MLH)

Collegial Support for Professional Improvement

The Stanford Collegial Evaluation Program

Susan Stavert Roper
and David E. Hoffman

March 1986
Vol. 29, No. 7

Prepared by the ERIC Clearinghouse on Educational Management

OSSC BULLETIN

OSSC BULLETIN

Collegial Support for Professional Improvement

The Stanford Collegial Evaluation Program

Susan Stavert Roper
and David E. Hoffman

March 1986
Vol. 29, No. 7

Oregon School Study Council

ISSN 0095-6694

Individual Copy $4.00
Quantity Discounts:
10-24 copies - 15%
25-49 copies - 20%
50 + copies - 25%

Oregon School Study Council
1787 Agate Street
University of Oregon
Eugene, Oregon 97403
(503) 686-5045

Preface

At a time when attention is focused on improving the practice of teaching, it is encouraging to see increasing numbers of educators attracted to the benefits of teacher collegiality. Seven years ago, John Goodlad's Study of Schooling convincingly demonstrated that teachers' isolation from one another is a major impediment to their professional improvement.

Educators looking for a workable model of collegial support should carefully consider the Stanford Collegial Evaluation Program, refreshingly described in the following pages by two scholars involved with the program since its inception at the Stanford Center for Research and Development in Teaching in the mid-1970s.

Since 1981 Susan Roper has been director of the School of Education and Psychology and director of teacher education at Southern Oregon State College, Ashland. She first became interested in collegial evaluation when she coordinated the field test of the Stanford Collegial Evaluation Program. She presented workshops on the program in several districts and, when she was a school administrator, persuaded teachers in her own school to participate. Roper has taught at the University of British Columbia and Stanford University. She received her Ph.D. in sociology of education from Stanford in 1971.

David Hoffman is professor of secondary education and coordinator of graduate studies in the School of Education and Psychology at Southern Oregon State College. His concern with the development of a collegial/professional model for faculty evaluation began as a result of his experience with collective bargaining in an educational setting. Dissatisfied with the tone and direction of collective bargaining in school settings, he has urged faculties to develop bargaining goals based upon a professional model of their occupation of teaching rather than only focusing on the traditional economic goals of collective bargaining in an industrial setting. Hoffman has conducted trials of the program among members of his own college faculty and within public schools in the southern Oregon region. Hoffman received his doctorate in educational curriculum and instruction from Stanford University in 1973.

This Bulletin is published by the Council in cooperation with the ERIC Clearinghouse on Educational Management, which commissioned the authors. The material will also appear in a forthcoming Clearinghouse publication on instructional leadership.

Philip K. Piele
Executive Secretary, OSSC

About ERIC

The Educational Resources Information Center (ERIC) is a national information system operated by the U.S. Department of Education. ERIC serves educators by disseminating research results and other resource information that can be used in developing more effective educational programs.

The ERIC Clearinghouse on Educational Management, one of several such units in the system, was established at the University of Oregon in 1966. The Clearinghouse and its companion units process research results and journal articles for announcement in ERIC's index and abstract bulletins.

Besides processing documents and journal articles, the Clearinghouse prepares bibliographies, literature reviews, monographs, and other interpretive research studies on topics in its educational area.

This issue of the OSSC Bulletin was prepared by the ERIC Clearinghouse on Educational Management with funding from the Office of Educational Research and Improvement, U.S. Department of Education under contract number 400-83-0013. The opinions expressed in this report do not necessarily reflect the positions or policies of the Department of Education or the Clearinghouse.

Contents

Preface iii

About ERIC v

Foreword ix

The Program's Background 1

Evaluation Must Fit the Task 2
Workers Must Have Faith in The Evaluation System 3

How Does It Work? 5

Step 1. Choosing a Partner 6
Step 2. Selecting Criteria 6
Steps 3 and 4. Conducting Self- and Student Assessments 8
Step 5. Observing a Colleague 9
Step 6. Conferring with a Colleague 10
Step 7. Developing a Plan for Improvement 11

Barriers to Collegial Evaluation: How Can They Be Overcome? 13

Teacher Attitudes, Habits, and Experience 13
Validity and Reliability 14
Role of Administrators 15
Teacher Candor and Criticism 16
Time for Observation 17
School Context 18
What Is Needed: A Commitment to Professional Development 19

Conclusion: Collegial Relationships Are a Springboard to Professional Renewal 21

How to Obtain the Implementation Manual 23

References 24

Foreword

A recent survey of teachers by the Metropolitan Life Insurance Company reinforces what many observers have concluded: Teaching must be made a much more attractive profession, or the inability of education to attract capable new teachers and retain experienced ones will soon become a crisis. The survey found that 27 percent of current teachers are likely to leave the profession for another occupation within the next five years.

Couple this finding with an estimate by the U.S. Department of Education that by 1992 the supply of new teacher graduates will be 34 percent less than the need, and the shape of a crisis looms large indeed.

In the *Executive Educator's* "Endpaper" (March 1986), John J. Creedon, president and CEO of Metropolitan Life, states, "The bitter truth is that, although we might call teaching a profession, we treat teachers as if they are low-level workers whose only hope for advancement comes with leaving."

At least a partial solution to the problem of making teaching more attractive is to make teaching truly a *profession.* Let teachers operate as professionals, exercising judgment, participating in decision-making, and taking responsibility for their own professional development.

In the opinion of Susan Roper and David Hoffman, an essential characteristic of any profession is collegial relationships. Yet in most schools teachers are typically isolated from one another and thus cannnot profit from their combined judgment and expertise. A major step toward professionalism will occur, therefore, when schools are restructured to allow the kind of collegial support system that Roper and Hoffman describe.

If the principles of collegiality are put to use in schools, not only should the quality of teaching improve, but more qualified people are likely to be attracted to teaching because they will be able to function as professionals.

Readers of last month's OSSC Bulletin (*It Is Time for Principals to Share the Responsibility for Instructional Leadership with Others*, by Keith A. Acheson with Stuart C. Smith) will note that

the ideas presented in these two Bulletins are eminently compatible. When teachers observe and give feedback to one another, they relieve principals of a responsibility that many of them have neither the time nor the expertise to do very well anyway.

This proposition in no way diminishes the importance and authority of the principal, who will still be responsible for summative evaluation and for coordinating teachers' collegial observations. An additional challenge for the principal is to lead in establishing the overall school context or culture that research has shown to be necessary for collegiality to thrive. If cooperation, experimentation, and community are to become the prevailing norms in a school, it is the principal who must lead the way.

Stuart C. Smith
Editor, OSSC and
Director of Publications,
ERIC/CEM

The Program's Background

This Bulletin is intended to familiarize school administrators and teachers with the staff evaluation system known as the Stanford Collegial Evaluation Program. A distinguishing feature of the system is that it utilizes teachers as the primary evaluators of their peers. The program links teachers in reciprocal team relationships; that is, each participant takes his or her turn in the roles of evaluator and evaluatee.

An equally important element of the program is an emphasis on *formative*, rather than *summative*, evaluation: In the Stanford program, the purpose of collegial evaluation is to improve instruction, not to gather evidence of improper performance for disciplinary action. Teachers collect information about their performance from collegial observations, student questionnaires, and self-assessment. Based on this information, they analyze the strengths and weaknesses in their teaching and prepare a plan for improvement.

Underlying these features is the assumption that teaching is, or ought to be, more profession than labor or craft. If the program did not seek professional development as its overriding goal--that is, if it did nothing more than substitute teachers for administrators as the agents of evaluation--it would be just another fashionable technique to deal with the perennial problem of staff evaluation.

The authors of the Collegial Evaluation Program view staff development as a more significant goal of evaluation than formal rewards and sanctions. They argue that the traditional bureaucratic factory model of staff organization, in which staff development and evaluation are primarily the responsibility of management, is not suitable to the nature of the work of the teacher and the organization of the school. To appreciate this point of view, it is necessary to understand the origin and theoretical basis of the program.

Evaluation Must Fit the Task

The Stanford Collegial Evaluation Program was developed by sociologist Sanford Dornbusch and his associates at the conclusion of a decade-long research program. Prior to their research, Dornbusch and W. Richard Scott (1975) developed a theory of evaluation for work within complex organizations. They then turned to study a variety of public and private, professional and nonprofessional organizations ranging from college football teams to hospitals, research institutions, churches, universities, and public schools.

Dornbusch and Scott were particularly interested in the manner in which each organization they studied evaluated its personnel. Among their conclusions was this one: To achieve an effective evaluation system within an organization, the type of tasks performed by those being evaluated has to be considered. If one were evaluating the quality of work at an assembly plant, it might be appropriate to pass over the actual behavior of workers on the assembly line and concentrate on measuring the degree to which the final product matches job specifications. Such an approach, however, would be inappropriate, in their view, for judging the quality of work of nurses and teachers, because patients and students are not material objects that react to treatment in a predictable or, necessarily, measurable way.

To determine whether a teacher practices a level of instruction and curricular design that is creative, analytical, or appropriate to the students, first-hand observation and evaluation of that practice are required. The evaluator must also be knowledgeable of and sympathetic to the task at hand.

J. Victor Baldridge and his colleagues (1976) proposed that bureaucratic evaluation systems (those designed for traditional line-and-staff organizations) "work best on *inert* tasks for which the resistance to successful completion of these tasks is predictable." In the case of teachers, writing a lesson plan according to school guidelines might be considered an example of an inert task. The evaluator need not be present at the time the plan is written to determine whether it matches stated criteria. In this case, the end product can be trusted as evidence of the quality of work.

In contrast to these inert tasks, the most essential tasks performed by professionals are *active* tasks. For the teacher, the actual teaching of the lesson--including inclass adjustments and invention based on student reactions--is an example of an active task. Relying on student test scores as singular evidence of whether the teacher has been creative and has exercised good judgment in teaching a given class is a very limited and sometimes misleading means of determining quality of teaching performance. There are simply too many other variables that affect test scores

to make the assumption that the teacher's performance, alone, determined them.

Active tasks are not routine. That is, they do not always respond to the "tried and proven" methods of solution. Although various strategies and operations may be available to accomplish the task, these may not always produce predictable results. The worker thus must deal with a variety of situations requiring analysis, invention, judgment, and use of a well-developed repertoire of strategies.

Because no one best method will guarantee the accomplishment of an active task, Dornbusch and Scott argue that the likelihood of success is increased by utilizing the judgment and expertise of the professional worker. The prospect of success is increased even further by bringing together the combined judgment and expertise of a group of professionals. This is the meaning of our use of the term *collegial.* For the professional worker, collegiality is most crucial when dealing with active tasks. The fact the occupation of teaching has seldom practiced this logic is largely a result of the organizational structure of most schools. That is, teachers are isolated from their peers at the most crucial point of work -- when they teach.

Workers Must Have Faith in The Evaluation System

In traditional bureaucracies, such as factories, as well as in professional institutions, such as hospitals, staff evaluation has at least two closely related purposes: the control or modification of the worker's behavior and the differential distribution of reward and sanction. The logic of evaluation is that the behavior of workers--whether they be teachers, assembly line workers, or clergy--is expected to be influenced by the evaluations they receive. Whoever has supervisory authority is assumed to exercise control over the workers and their work.

This logic holds, however, only if workers believe that the evaluation system is sound. That is, workers must believe that what is performed or produced by their efforts is fairly and accurately observed and is measured against mutually acceptable standards. In actual practice, Dornbusch's team found many instances of organizations whose evaluation programs did not meet these conditions. This was particularly the case among school teachers. If one examines the conditions of work of the individual classroom teacher, much of what constitutes professional work activity (that is, teaching) takes place in isolation from direct observation by colleagues, supervisors, or administrators. One result of this situation is that the teacher may be evaluated primarily on inert and visible task performance. Accordingly, Dornbusch found that many teachers:

> report a belief that principals pay more attention to certain aspects of their performance that are easily visible, such as record keeping. And that the details of what goes on in the classroom are not known to their organizational superiors. . . .The same lack of knowledge of what goes on in individual teacher's classrooms extends to other important aspects of schools as well. . . .Principals did not know the materials used by teachers in the classrooms. Other teachers did not know the materials and grouping patterns that their fellow teachers used. (1976, p. 6)

The primary work of teachers is *teaching*. But in most schools this task takes place behind closed doors in isolation from one's peers and administrators. Any effective teacher evaluation system must respond to this condition of the teacher's work. A formative evaluation design assumes the opening up of classrooms to frequent and ongoing observation. It also projects a goal of ongoing dialogue among teachers about the instructional and curricular aspects of their work. During collegial evaluation teachers will establish a discourse on the topic of professional practice.

Before we describe the operational procedures of the Stanford Collegial Evaluation Program, it is instructive to note Gary Natriello and Sanford Dornbusch's findings about the characteristics of effective staff evaluation systems. The researchers found several features that were related to both teacher satisfaction with a system and the likelihood that the system could improve teaching performance. They found that teacher satisfaction was highly related to these conditions:

(1) greater frequency of classroom observations
(2) common criteria used to evaluate performance
(3) teacher involvement in the formulation of criteria
(4) greater frequency of feedback from supervisors or observers (Natriello and Dornbusch 1980-81)

Negative feedback, by itself, did not appear to dissatisfy teachers, but infrequent evaluation and imposed criteria did. If teachers are to be satisfied with an evaluation system and are expected to improve their performance accordingly, they need to feel that they have some control over the tasks being evaluated. The structure of the Stanford Collegial Evaluation Program, as we will describe, incorporates these elements of effective professional staff evaluation.

How Does It Work?

This chapter summarizes the basic steps of the Stanford Collegial Evaluation Program. We chose to present this program because it contains two elements that some other collegial programs do not: flexibility and reciprocity. Because teachers decide the criteria to use in evaluating one another, the program can be adapted by instructors at virtually every level of schooling. It has been field tested at the college level and with elementary and secondary teachers.

The program is also flexible in terms of the number of teachers that can participate. An entire faculty, a department or teaching team, or two interested teachers can participate. The heart of the program is the teacher pair. Although we believe that schools will benefit from more teachers participating, the program can work with only two interested teachers.

Most of the peer or collegial evaluation programs we examined in our review of the literature trained a selected group of teachers to evaluate their peers. Whether formally or informally, the trained teachers were often placed in a quasi-supervisory role over their untrained colleagues. In the Stanford program teachers evaluate one another. This reciprocal arrangement means that teachers are not thrust into a supervisory role. As colleagues observing and conferring with one another, their goal is mutual assistance.

The Stanford program has seven interdependent steps: (1) choosing a partner, (2) selecting criteria, (3) self-assessment, (4) evaluation by students, (5) observation, (6) conferences, and (7) planning a program of improvement. A weakness in any step will impair the others. For example, if the criteria are vague, the observations will not be focused and the conferences will probably not provide much useful information. If the conferences lack candor and specificity, the plans for improvement will suffer.

The entire sequence of self-assessment, student assessment, observations, and conferences requires ten to twelve hours spread over a month or two.

Step 1. Choosing a Partner

Should administrators assign teachers to pairs or should teachers choose their own partners? There are advantages and disadvantages with each of these options. One advantage of having administrators select teacher pairs is that they can try to match teachers whose skills complement one another. A disadvantage is that teachers may feel that one will be labeled "weak" and another "strong." Moreover, teachers increasingly resent administrative fiat in their own professional development activities.

Teachers who participated in the field test of the Stanford program said they preferred to choose their own partners. They disagreed, however, on the type of relationship the teaching pair must have to be successful. Some felt that they needed to work with a close friend. As one put it, "Working with a fellow teacher on this program required a lot of respect and trust. You've really got to like one another. It's almost like a marriage" (Roper, Deal, and Dornbusch 1976). Others said they would be uncomfortable in giving or receiving criticism with a friend. As one teacher said, "You don't have to love someone to work with them."

Teacher partnerships varied widely. Experienced teachers worked with inexperienced teachers, biology teachers with English teachers, sixth-grade teachers with second-grade teachers. All these combinations "worked." The main ingredient for good partnerships was mutual respect.

Step 2. Selecting Criteria

Selecting criteria to guide the observations is a five-step process in the program: (1) identifying a pool of possible criteria, (2) selecting criteria independently, (3) agreeing on criteria, (4) checking criteria for specificity and observability, and (5) listing criteria on the observation and self-assessment forms.

The pool of possible criteria is virtually limitless. One source is the teacher's own professional goals. A second is the district or school goals. Research on effective teaching in the past few years provides a number of specific and observable behaviors that teachers could use as criteria. These are especially useful to elementary school teachers focusing on basic skills. Educational researchers have generated a host of observational instruments that can be adapted by teachers. Good and Brophy's book *Looking in Classrooms* (1984) is a good source for instruments and reviews of studies.

Many districts have trained their administrators and teachers in Madeline Hunter's program Instructional Theory Into Practice

(ITIP). ITIP breaks down the teaching process into specific skills. The fact that teachers share a common lexicon from ITIP training makes it ideally suited to a collegial evaluation program.

Some participants in the collegial evaluation program found results of the student questionnaire helpful in generating criteria. Other teachers suggested that observations of outstanding teachers assisted them in identifying criteria.

Despite having a multitude of sources from which to select criteria, teachers reported that this was the most difficult step of the program. The problem was that many of the criteria of most interest to teachers were very difficult to assess, whereas unimportant criteria were often more easily observed. For example, noise level in class was fairly easy to evaluate, but the more important aspect of student engagement was more difficult.

One suggestion was to identify goals and develop performance indicators for each goal; these performance indicators would then serve as the criteria. For example, for the goal "motivates students to participate in discussion," two performance indicators might be (1) the number of students called upon by the teacher and (2) the length of time teachers waited for an answer.

A good general rule is that teachers should agree on how they plan to measure a criterion before they include it on their list. Another rule is that this list should not include more than five criteria. Having too many can cause the observations to lack focus and can impair the quality of information teachers share with one another during conferences.

After the partners have selected their criteria independently, we recommend they spend a good share of their first meeting together agreeing on a common list. A joint list will enable them to exchange information about the same concerns. Teachers will find their observations and conferences to be more helpful if they share common criteria than if they examine totally different areas.

Although teachers in the field tests struggled with the task of identifying a few important and observable criteria, they did not see this as a waste of time. As one teacher put it, "Selecting criteria forced me to clarify my own educational philosophy. I had to decide what was really important to me and then try to operationalize my goals so they could be observed" (Roper, Deal, and Dornbusch 1976).

Steps 3 and 4. Conducting Self- and Student Assessments

Participants complete two forms of self-assessment. One parallels the evaluation by their colleague; the other parallels their evaluation by students. They rate themselves on the same shared criteria that their partners used to observe them. This allows them to check for areas of agreement and discrepancy between their own and their partners' evaluations. Many teachers found that discrepancies between their self-assessments and the assessments of their colleagues or students generated new criteria.

One point of caution in the use of self-assessments: Although they tended to agree on areas of strength and those needing improvement, teachers were usually harder on themselves than they were on one another.

On a positive note, having a colleague agree with a self-assessment helped motivate that teacher to modify behavior. For example, one teacher wrote on her self-assessment form, "In discussion I tend to rely on the same students who always have the answers, and I do not phrase open ended questions to include everyone." Her colleague noted, "Two boys spoke often, a few girls spoke occasionally, but no one else entered the discussion." After the second observation, her partner encouraged her by saying, "The discussion included more students and some who had not previously participated. You praised the newcomers --Good!" (Roper, Deal and Dornbusch 1976).

The other self-assessment questionnaire consists of items matched to those on the student questionnaire. Teachers can compare their own perceptions to those of students on such areas as classroom climate, the procedures used to evaluate students, students' interest in class, and so forth. For example, teachers would rate themselves on the item, "How clear to your students are your explanations of the class work?" Students would rate the teacher on the item, "How clear are your teacher's explanations of classwork?"

Again teachers found a great deal of similarity between their self-assessments and the responses of their students. However, by examining student responses, teachers found some interesting disagreements. Occasionally students felt that work was "too difficult" or "too easy." The teachers also learned that some students felt they could not get good grades "no matter how hard they tried." On the other hand, teachers gained a sense of consensus in such areas as how much students appreciated their sense of humor, patience, and knowledge of their subject area.

Teachers learned to look at the distribution of responses on the student questionnaire as well as the average. One young teacher was dismayed to see that she only received a mean of

"seldom" in response to the question, "How often do you talk to your students outside of class time?" In examining the distribution of responses, however, she found four students who said they talked with her "very frequently." In thinking it over, she realized that she spent most of her time outside of class with about four students.

A careful comparison of the self-assessment forms with student and colleague responses helped participants identify the main areas they wanted to include in their improvement plan. It also provided a means of reviewing their criteria. Some teachers produced revised criteria that drew upon their own input as well as that of their students and colleagues.

Step 5. Observing a Colleague

We began this report with the statement that teachers are seldom the subject of observations by their peers. Teachers spend most of their professional lives in isolation from one another. It should not be surprising that when presented with the prospect of being observed by a fellow teacher, they are often anxious. The tradition of the closed classroom is a difficult one to break. Yet, the most essential element of the Stanford program is *collegial* observation.

Participating in the Stanford Collegial Evaluation Program requires a minimum of two reciprocal observations. Teachers can allay their anxiety somewhat if they choose the time to be observed. The pair should decide at their first meeting if they are comfortable with their partner taking notes during the observation. Note taking usually helps improve the quality of information the partner can share, but for some teachers an observer in their classroom busily writing notes is so distracting as to be counterproductive. In those cases, the observer should plan on five or ten minutes immediately following the observation period to complete the observation form. Simply rating the teacher on the criteria, writing comments relevant to the criteria, and giving a few suggestions completes the observation form.

Although the minimal requirement of the program is two observations, we strongly encourage more. The number of observations depends on the amount of time teachers can be freed from their classroom responsibilities to observe. For secondary teachers, prep periods are used. For elementary teachers and those secondary teachers who do not have prep periods, a supportive administrator is essential. A principal or counselor who believes in the program can take over the class to free the participants for observations. Another alternative for some is to send their students to another teacher for a period.

In the reports from the field tests, teachers frequently said

that observing was every bit as helpful to them as being observed. Although they learned a great deal from their partner's assessment of their teaching, many said they learned as much from seeing their colleague in action. We were continually impressed with the benefits teachers found in observing one another. This finding adds to our belief in the benefits that can be gained by opening up the classroom to the supportive scrutiny of one's peers. It also confirms our earlier contention that collegiality is a dynamic process in which one is involved simultaneously as both teacher-observer and learner. It is a powerful--and inexpensive--staff development activity.

Step 6. Conferring with a Colleague

There are three conferences in the collegial evaluation program. The first conference is scheduled soon after the partners have completed their initial observations of each other. The purpose of this conference is to mutually report what they have witnessed in each other's classrooms, discuss any discrepancies between the assessments by self and colleague, and identify strengths and areas needing improvement.

The second conference follows a similar pattern but emphasizes suggestions for improvement. It is scheduled immediately after the second round of observations. Whereas the major purpose of the first conference is to gather data upon which to focus the next observation, this second conference should give the teacher some concrete ideas about how to improve. For both of these conferences, teachers should review their partner's observation forms beforehand. They should plan about an hour of uninterrupted time for each of the first two conferences.

Supporting our belief in the strength of the collegial process, teachers reported that they learned more from the conferences with their peers than with administrators. If the partner observed something that needed to be improved, the typical pattern was to suggest or to mutually determine an alternative strategy.

Teachers viewed this type of criticism as more helpful and less threatening than evaluations from administrators or supervisors. This testimony assured us that participants viewed the system as primarily a *formative* mechanism for improving instruction, rather than a means of reporting improper performance for disciplinary action.

Positive reinforcement characterized many of the postobservation conferences. Teachers were able to see how their good teaching behaviors led to specific student responses. For several participants, this was the first time in their careers that they had evidence of their effectiveness as teachers--evidence they

valued because it came from a highly credible source.

During the field tests, topics in the conferences covered the total range of teacher and student behavior, including classroom climate, subject matter presentation, teaching style, management techniques, appropriateness of curriculum and materials, pacing, and student engagement.

Step 7. Developing a Plan for Improvement

Developing a plan for improvement is the most important step of the program.

First, teachers review all the information generated during the program. Second, they meet in the teaching pairs to agree on specific strategies for improvement. Third, they determine how to assess their progress in carrying out the improvement plan.

Prior to the final conference, each teacher should carefully review all the information that has been accumulated on his or her own teaching: the two types of self-assessment completed, the colleague's observation forms (minimum of two), and the student questionnaire data. Based on these sources, the teacher can identify a few strengths and weaknesses. We recommend that each teacher list no more than five weaknesses. He or she may wish to jot down a few ideas for improvement plans to share with the partner.

If a teacher's performance is of such quality that the identification of weaknesses becomes an exercise in tedium--scraping the barrel for insignificant mistakes--then the plan for improvement can focus on other areas of professional development. In such cases, we suggest that the partners use the opportunity to establish plans to develop and try out new techniques or materials, thus building skill in areas not normally a part of their regular repertoires.

The final or wrap-up conference between the collegial pair is mainly for the purpose of helping each other develop the improvement plan. Scheduled for two uninterrupted hours, this conference is usually held within two weeks after the second postobservation conference. The partners will discuss the major strengths and weaknesses in their teaching and will agree on specific strategies for improving their teaching in each area identified as a weakness. Once specific activities are listed on the Professional Development Plan, teachers will assist each other in determining how they can evaluate their progress.

In the field tests, improvement plans covered the whole range of teaching behaviors. Some required a major restructuring of the classroom, such as elimination of ability grouping. Others were

relatively minor, such as displaying student art work. When partners discussed assessing their progress in implementing their plans, many agreed to come in and observe each other. The improvement plan thus served as the criteria for the next round of observations. For plans related to student knowledge, teachers used tests and evaluation of classroom assignments to evaluate their progress. For plans related to student attitudes, they administered the student questionnaire again, using it as a posttest to assess their program.

When teachers work together to examine essays, math assignments, art work, shop projects, or video tapes of the performance of their students in music or in physical education classes, their cooperation takes on the characteristics of fellow professionals seeking to improve the quality of their craft. Improved diagnosis of the behavior of their students can emerge from such arrangements so that teachers may better understand the sources of performance disabilities and ultimately develop more effective strategies to deal with them.

Once the improvement plan was drafted, some teachers in the field tests chose to share it with their principal. Their purpose, besides providing evidence of their devotion to professional development, was to elicit the principal's support in their efforts to improve. Occasionally improvement plans required additional resources and the administrator needed to be convinced to provide them.

A large number of teachers were so gratified with this program--both its overall usefulness and the opportunity to work with their colleagues--that they planned to go through it again.

If the reader is growing somewhat skeptical at our optimism, we invite you to read on. In the next chapter we attempt to consider some of the major obstacles to collegial evaluation.

Barriers to Collegial Evaluation: How Can They Be Overcome?

Like all strategies for improving our schools, it is much easier to advocate them than to implement them. Because we believe it is easier to overcome obstacles that are foreseen, we have included this section in the hope it will alert potential participants in collegial evaluation to common problems that may occur.

Some of the barriers to the program are teachers' attitudes, habits, and experience. Others relate to the role of administrators. Still others concern problems endemic to teacher evaluation such as reliability, validity, and time for observation. After discussing these barriers, we will consider why an unsupportive school context is an especially difficult obstacle to overcome. Despite this formidable list of barriers, we are convinced that they can be overcome with the right mixture of determination, good will, and humor.

Teacher Attitudes, Habits, and Experience

The one attitude that is most destructive of any teacher improvement effort is "teachers are born, not made." If teachers are really convinced that personality is the only important variable in their profession, they will most certainly not be motivated to improve. We have found in our staff development work with teachers over the last several years that this attitude is often no more than an expression of anxiety about trying something new. Most teachers know better than anyone else that they can always sharpen their skills, learn something new, and improve their performance.

When the Stanford Collegial Evaluation Program is presented, teacher fears need to be discussed. Teachers need to be assured that the program is voluntary, confidential, and solely for their own professional growth. They need to be assured that everyone who has spent most of his or her professional life behind the closed doors of a classroom will naturally be anxious about being observed by a peer. They need to know that isolation is contrary to quality evaluation and staff improvement within any profession

where so much rides on the actual "practice" of one's craft.

Most teachers will quickly agree that measuring only the end product of teaching through the use of such devices as standardized achievement and aptitude tests or grade point averages is not a particularly valid way to determine the quality of what actually transpires in the classroom between teacher and pupil. As we have stated previously, the most appropriate occasion to evaluate the active task of teaching is at the time that it is in progress. If teachers can be convinced of this perspective, they will be open to assurances that teachers who have participated in collegial or peer support programs are very positive about their experience. They will also discover that it is much easier to open that classroom door again, after it has been opened the first time.

Even though isolation is a habit in our profession, habits can be changed. Teachers working in teams in open space schools, for example, expressed much less concern about being evaluated by their peers than did teachers in self-contained classrooms (J.W. Meyer and others 1971).

A related attitude is that teachers have no business evaluating one another because they lack training in supervisory skills. Usually, however, most teachers will recognize that administrators have pretty sparse training in this area themselves. In fact, teacher supervision and evaluation is a very small part of most administrative certification programs. Moreover, the goal of this collegial evaluation program is not to provide a scientifically objective picture of a teacher's performance, but to give the teacher some direction in improving that performance. The important thing to stress is that experienced teachers do have skill in teaching methodology and in curriculum development. It is primarily from this store of experience that they will operate in the collegial process. We have found that when teachers realized that they could provide their peers with good ideas for improvement, they began to worry less about their relative inexperience.

Validity and Reliability

Many of the limitations of collegial evaluation apply to teacher evaluation in general. The most serious of these are validity and reliability. As Charles McIntrye (1978) points out, there is no "measurable, independent, generally agreed upon outside criteria of teaching effectiveness." McIntye is quite correct in his statement that the "whole activity is severely embarrassed by the lack of criteria." Linda Darling-Hammond and her colleagues (1983) also list "observer bias, insufficient sampling of performance and poor measurement instruments" as additional threats to both reliability and validity.

Collegial evaluation relies heavily on classroom observation. The fact that these observations are conducted by untrained observers who have limited time to observe does not recommend the program for its reliability or validity. It is important to keep in mind, however, that, while collegial evaluation may exacerbate these problems, ratings by trained and experienced supervisors have generally been found to lack interrator reliability and validity (Medley 1982, Peterson and Kauchack 1982, and Darling-Hammond and others 1983). Whether the principal or the colleague is the observer, these basic problems remain.

Participants should know that these problems exist -- that they are not going to be producing "hard data" for a research study. They will, however, learn something about their own performance and how to improve it. We have found that teachers are seldom bothered by subjectivity. In fact, they are often skeptical about so-called "objective" measures of competence of the active task of teaching. Their reactions to various attempts to quantify their behavior echo the sentiment of Elliot Eisner (1976) when he admits to feelings of "uneasiness" regarding objective approaches. He concedes that "a high degree of precision is possible when one counts but is diminished when one is required to judge." Yet he laments that these more precise approaches "somehow fail to tell the whole story."

Role of Administrators

Establishment of a collegial system of evaluation demands that administrators clearly articulate the purpose and structure of the process to their faculties. Beyond the walls of the school, administrators must also use all their communication and political skills to break down the "management-employee" view that many lay boards have about what constitutes the appropriate organizational structure of their schools. It will take fancy communication footwork to educate and convince lay boards that the word *professionalism* is indeed defined by the existence of an evaluation system that is both collegial and formative.

Clearly, administrators do have a role in collegial evaluation, but the lack of reliability and validity makes us question the wisdom of administrators using this program as part of the formal system for evaluation. For one thing, since teachers choose the criteria, there is no way to get an assessment of overall competence. Teachers will focus their observations on specific aspects of their teaching.

Also, we fear that including the program as part of the formal evaluation system would subvert its intent. For example, one teacher could be rated much lower than another because he or she selected criteria that focused the observation on a weakness. Another teacher, not so brave, chose criteria that were neutral or

highlighted strengths. To protect themselves, teachers might select criteria that make them look good, fail to provide constructive criticism of one another, and complete the self-assessment forms superficially. The program would degenerate into pointless busy work that has little hope of leading to instructional improvement. For these reasons, teachers have been unanimously opposed to using this program as part or instead of a summative evaluation.

We are not the only ones who have run into teacher opposition regarding administrator participation in a collegial approach. Tom Bird and Judith Warren-Little (1983) had a similar reaction from teachers in a staff development program for improving teacher skills called "coaching." In considering the role of administrators in the coaching process, teachers asked two questions. First, how could administrators who participated in coaching keep "coaching" and "evaluation" separate? They concluded, "They can't." Second, "Would administrators have access to what teacher-coaches learned in the course of coaching? The decision was, no."

Teachers' reluctance to share their vulnerabilities with their supervisors is understandable. What is less understandable, however, is their reluctance to be candid with one another.

The dilemma here is obvious: a profession takes responsibility for evaluation and improvement of professional practice among its members. This means, at a minimum, there must be opportunity for open observation of one another while at work and for an exchange of frank evaluations of that work. If members of the profession fear, however, that disclosure will threaten their security, they will be strongly motivated to avoid both the observation and the evaluation that follows.

Teacher Candor and Criticism

In the field tests of the Stanford program and similar collegial observation projects, there is a definite bias toward positive assessment. Reporting on this phenomenon, Bird and Warren-Little say, "The impulse was somehow to avoid ever giving any indication that the teacher observed had done anything less than a sterling job."

This problem is not confined to elementary and secondary schools but is alive and well in our universities. McIntyre cites evidence suggesting that colleagues rate one another's teaching "more leniently" than students do and explains that "perhaps this means when those results have been found, that faculty solve the collegial problem simply by being generally nice to each other. They do unto others as they hope others will do unto them, but this yields rather uniform ratings which aren't much use in making

discriminations".

It was clear in talking to teachers who participated in the field tests of the Stanford program that one of the primary benefits was praise from their colleagues. In listening to their reports, we were struck by the terrible loneliness of a profession that isolates adults from one another. Teachers are starved for some kind words. The fact that these words came from people whom they respected and who faced similar situations made them all the more welcome. No doubt praise will be an important part of teacher conferences in the program, but if teachers are to use collegial interaction to improve, they must go beyond praise.

We do not know of any shortcut solution to this problem. Our experience with teachers who have been through this program more than once gives us hope that the more teachers participate, the more candid they become. It takes time to break the norm of never criticizing (constructively or otherwise) one's colleagues. Even when invited to do so, this norm interferes. We did see a difference between teacher interns and experienced teachers. Interns were much more likely to be critical than were experienced teachers. Of course, the interns are expected to need considerable assistance as novices. This difference argues for making the program a part of induction experiences for teachers beginning their initial full-time positions. These beginning teachers are most in need of collegial support and likely to be most open to constructive criticism.

We do not want to overemphasize the problem of lack of candor. Experienced teachers did give each other helpful criticism, mostly in the form of suggestions for improvement. This is perfectly acceptable and perhaps even the best form of criticism. Nevertheless, our concern is that some serious problems in a teacher's performance were occasionally ignored by the colleague. In introducing this program, we need to pass on the words and wisdom of Tom Bird and Judith Warren-Little:

> If a powerful analysis of teaching is to be shared, persons' teaching practices cannot be regarded as private or personal, but must be regarded as tools of a profession which are open to evaluation. Such a situation poses risks which require that the participants meet as equal professionals sharing both their confusion and their success. (1983)

Time for Observation

The problem of freeing teachers for observation seems almost too obvious to mention. It is amazing, however, that a whole program can fail to get off the ground because teachers cannot be freed for two to three hours. We believe that this is less a problem of time than it is a problem of priorities, because we

have seen many occasions where districts manage to release teachers to attend workshops that are consistent with district goals. (By the way, we applaud these efforts.) We would, however, like to convince administrators and school boards that what teachers have to share with one another is as valuable as what they will learn from an outside expert -- and potentially more lasting.

Finding time for collegial evaluation, then, is not the main issue. Convincing the powers that be that teachers are professionals who learn best from one another is the central issue. In the financial crunch facing so many school districts it would seem easier to "sell" a program that does not require high priced consultants, expensive materials, and disruption of classes than the more typical inservice experience that often requires all three. Strange as it seems, districts will often pay the price for the legitimacy of the expensive "expert" rather than put those resources into using their own staff as experts. Lack of time is a symptom, not a cause, for the more basic problem of lack of support for collegiality.

School Context

To discover those conditions in the schools that foster or impede collegial evaluation, we draw on the experiences of others who have studied similar programs. We hope that school administrators will review this information and make a more informed decision than we could make concerning the appropriateness of the collegial evaluation model for their school or district.

In a peer clinical supervision project, Shirley McFaul and James M. Cooper (1984) trained twelve teachers to conduct evaluations using clinical supervision. They reasoned that the well-respected program of clinical supervision, with its emphasis on teacher involvement, was an ideal mechanism for collegial evaluation. The "key" research question was this: "Is the form and spirit of the model congruent with teachers' attitudes and abilities as well as with the environment?"

They were disappointed with their findings. Although all but one of the teachers went through the motions of clinical supervision, they did not do so thoroughly. Conferences were missed or shortened, data were insufficient, indepth analysis was lacking, and there was a propensity to come up with only one alternative or a simplistic solution. Their experience made them question both the "willingness" and "ability" of teachers to engage in this program. Their explanation: "The underlying assumptions of the peer clinical supervision model were incongruent with the school context."

McFaul and Cooper identified four contextual patterns that, in

their opinion, undermined clinical supervision: "isolation and fragmentation, stratification, standardization and reactionism." In this school teachers worked alone and rarely interacted with one another on professional issues. They were fragmented into several groups. The principal contributed to this fragmentation by favoring one group of teachers over another. Their autonomy was severely curtailed by pressure to teach the same curriculum in the same way. The principal's unpredictable style, coupled with frequent changes from the district office, left teachers reacting to their environment rather than attempting to control it.

In short, the context of this school was hostile to the values of professionalism, individualization, and collegiality. These values are at the heart of a peer supervision program.

Bird and Warren-Little also identified school context as crucial in the success of their peer-coaching program. They describe schools in which peer coaching was successful as "adaptable." The adaptable schools cultivated the norms of "experimentation, evaluation and collegiality--the shared expectation that a faculty is always getting better together."

Referring to these norms, Gary D. Fenstermacher and David Berliner (1985) give us a description of a "well maintained" school:

> These norms permit, indeed encourage, teachers to talk easily to one another about what they are doing and how it is working; they engender a shared set of words and concepts for describing classroom events; and they encourage trying out new ideas and openly reporting the results. They are schools where funds, facilities, incentives, time, and personal support are provided commensurate to the tasks to be performed and the goals to be attained.

These research studies are useful for two reasons. First, the similarity of the findings is convincing evidence that school context is a variable crucial to the success of such programs. Second, the latter two studies, which describe similar programs that succeeded, provide guidelines for identifying environments conducive to collegial evaluation.

What Is Needed: A Commitment to Professional Development

In 1975 Elizabeth Dillon wrote a perceptive and amusing newsletter on the topic, "How to Kill a Staff Development Program." She said that, contrary to popular opinion, it is not easy to destroy good programs--a team effort is essential.

First, she advised the district's central office staff. They should avoid showing any interest in the project, fail to assign responsibility for its implementation, and make sure that the objectives of the project have no relation to district goals.

Her counsel to building administrators wishing to kill staff development activities was even more concrete. Do not express a positive attitude toward the activity, she admonishes. In fact, convey to teachers that you, too, are being forced to participate against your better judgment. Above all, do not stay for the entire training. Frequent interruptions by the principal effectively convey the message that the whole endeavor is a waste of time.

These strategies by administrators have proved quite effective, but, in the last analysis, teachers are the ones who can sound the death knell for a professional development program. Some of the more subtle strategies are to mutter frequently, "We tried that and it didn't work" or "We're already doing that." Smiling with "amused tolerance at your colleagues who are enthusiastic about staff development activities" is another good one. Of course, there are also the tried and true techniques: refusing to participate; focusing attention on the really crucial topics of coffee breaks, room temperature, and parking; and sharing your conviction that society is darn lucky teachers will even work with "today's impossible kids," let alone expecting them to sacrifice time in staff development activities as well.

The point is that no matter how worthwhile or well-organized a professional development activity is on paper, it can be subverted by the lack of commitment of school administrators and teachers.

Reading Dillon's advice, it is easy to conclude that negative voices, if they are loud enough, and negative behaviors can successfully undermine professional growth efforts. Our hope was lifted, however, when we learned from Dale Mann's 1978 study of innovative federal projects that it is not necessary to have one hundred percent of a faculty or even a majority of the staff supporting a project. Mann found that backing from a "critical mass" of about one quarter of the staff was sufficient.

Applied to the Stanford Collegial Evaluation Program, Mann's finding means that two to four pairs of teachers in most elementary schools and six to twelve pairs in secondary schools would constitute a "critical mass." With administrative support and a good introduction, we think the Stanford Collegial Evaluation Program will be attractive enough to win commitment from this size of a cadre of teachers.

Conclusion: Collegiality Is a Springboard to Professional Renewal

The guiding principles of the Stanford Collegial Evaluation Program are legitimacy, clarity, and visibility. Selecting a respected colleague as one's partner in the program provides the legitimacy. Choosing important, specific, and observable criteria help to focus observations that lead to clear and useful suggestions in conferences. Teachers are visible to one another through the observation periods.

In implementing this program in the public schools, we discovered a few other principles relevant to collegial evaluation. The first is that teachers are anxious to improve and have a pretty good idea of where they most need improvement. The second is that they have a wealth of information to share with one another. A third is that they will listen to one another and their students. This combination of attitudes and behaviors made the Stanford Collegial Evaluation Program a success for most participants.

Because of this successful record of implementation, we feel frustrated when we consider the barriers to collegial evaluation presented in the previous chapter. We feel frustrated for this reason: Those schools that face formidable barriers may be the very ones most in need of collegial evaluation. Teachers who feel the most anxiety about observing one another may be the very teachers who will profit the most. Administrators who are not generally supportive of professional development could benefit the most from helping teachers to implement improvement plans. Those schools, in short, that are not "well-maintained" or "adaptable" are the ones where collegial evaluation could make the most impact.

We would like to convince educators that it is through programs such as the Stanford program for collegial evaluation that the norms of collegiality, experimentation, and evaluation can be built. Although we are not unmindful of the difficulties involved in implementing collegial evaluation in a school context adverse to innovation, we also want to do more than preach to the already converted. Yes, collegial evaluation will be more likely

to succeed in those schools that do not have the barriers we have described. No, we cannot wait to begin such programs until these schools have overcome these barriers or we will wait forever.

As we have conceptualized it, collegial evaluation is primarily a staff development program. Even in schools lacking the conditions outlined by researchers as essential for successful innovations, there are still some staff development activities going on. Over two billion dollars are spent on staff development every year. This comes down to somewhere between $1,000 and $1,700 per teacher per year (Fenstermacher and Berliner 1985). It is obvious that resources are available. Our plea is to begin using those resources in ways that can improve the school context as well as the individual performances of teachers. We believe that the Stanford Program of Collegial Evaluation is one option for enhancing both the organization and the individuals within the organization.

Collegiality is not just a process; it is an attitude about how one relates to the responsibilities of one's profession. What it says to practicing teachers is that the quality of their work is dependent upon community as well as individual effort. We do not see collegial evaluation as merely a specific and promising process of evaluating teachers in the short run. Our ambitions are more substantial. We see collegial evaluation as both a mechanism and a goal. While this paper is devoted to defining the operation of the system, our broader goal is the establishment of collegial relationships between teachers as an essential and abiding characteristic of the profession. We believe this characteristic holds the most promise as a means of obtaining effective and self-generating professional renewal.

How to Obtain the Implementation Manual

Readers of this Bulletin who wish to implement the Stanford Collegial Evaluation Program may purchase the manual from the Oregon School Study Council.

The Field Test Edition (fall 1976) of *The Collegial Evaluation Program: A Manual for the Professional Development of Teachers* was written by Sanford M. Dornbusch, Terrence E. Deal, Deborah Plumley, and Susan S. Roper and was published by the Stanford Center for Research and Development in Teaching, Stanford, University.

The manual contains all the directions and forms needed to implement the program. It also explains the rationale for each step, incorporates examples of successful practices from other teachers, and offers suggestions to help teachers receive the greatest benefit from this experience.

Copies, at $10.00 each, may be obtained from the Oregon School Study Council, 1787 Agate Street, University of Oregon, Eugene, OR 97403. A $2.00 handling charge is added to billed orders. Quantity discounts are available.

References

Baldridge, J. Victor; Dornbusch, Sanford; Garbarino, Joseph W.; and Jacobs, Richard M. "Faculty Evaluation: Tool for Eliciting Desired Behaviors." *Journal of Dental Education*, 40,8 (1976), pp. 528-540.

Bird, Tom, and Warren-Little, Judith. *Finding and Founding Peer Coaching: An Interim Report of the Application of Research on Faculty Relations to the Implementation of Two School Improvement Experiments.* Boulder, Colorado: Center for Action Research, Inc., March 1983. ED 238 141.

Darling-Hammond, Linda; Wise, Arthur E.; and Pease, Sara R. "Teacher Evaluation in the Organizational Context: A Review of the Literature." *Review of Educational Research*, 53,3 (Fall 1983), pp. 285-328.

Dillon, Elizabeth A. "How to Kill a Staff Development Program." *Staff Development Newsletter*, 2, 2 (September 1975). Professional Development Associates, P.O. Box 4303, Austin, Texas 78765.

Dornbusch, Sanford. "A Theory of Evaluation Applied to the Schools." Paper presented at the Sociological Research Association, New York, 1976.

Dornbusch, Sanford; Deal, Terrence E.; Plumley, Deborah; and Roper, Susan. *The Collegial Evaluation Program: A Manual for the Professional Development of Teachers.* Field Test Edition. (Copies are available from the Oregon School Study Council, 1787 Agate Street, University of Oregon, Eugene, OR 97403, for $10.00.)

Dornbusch, Sanford, and Scott, W. Richard. *Evaluation and the Exercise of Authority.* San Francisco: Jossey-Bass, 1975.

Eisner, Elliot W. "The Use of Qualitative Forms of Evaluation for Improving Educational Practice." *Educational Evaluation and Policy Analysis*, 1, 6 (November-December 1976).

Fenstermacher, Gary D., and Berliner, David. "Determining the Value of Staff Development." *The Elementary School Journal*. 85, 3 (January 1985), pp. 281-314.

Good, Thomas, and Brophy, Jere D. *Looking in the Classrooms*. Cambridge, Massachusetts: Harper & Row, 1984.

Mann, Dale. "The Politics of Training Teachers in Schools." In Making Change Happen, edited by Dale Mann. New York: Teachers College Press, 1978.

McFaul, Shirley, and Cooper, James M. "Peer Clinical Supervision: Theory vs. Reality." *Educational Leadership*, (April 1984), pp. 4-9.

McIntyre, Charles J. "Peer Evaluation of Teaching." Paper presented at the American Psychological Association Convention, Toronto, Canada, August 1978. ED 180 295.

Medley, D.M. *Teacher Competency Testing and the Teacher Educator*. Charlottesville, North Carolina: Association of Teacher Educators; and the Bureau of Educational Research, University of Virginia, 1982.

Meyer, J.W.; Cohen, E.G.; Brunetti, F.A.; Molnar, S.R.; and Salmon, E.L. *The Impact of the Open Space School upon Teacher Influence and Autonomy: The Effects of an Organizational Innovation*. Stanford, California: Stanford Center for Research and Development in Teaching, Stanford University, *R & D Technical Report*, No. 21, October 1971.

Natriello, Gary, and Dornbusch, Sanford. "Pitfalls in the Evaluation of Teachers by Principals." *Administrator's Notebook*, 1980-81, EJ 261 536.

Peterson, K., and Kauchak, D. *Teacher Evaluation: Perspectives, Practices and Promises*. Salt Lake City, Utah: Salt Lake City Center for Educational Practices, University of Utah, 1982.

Roper, Susan Stavert; Deal, Terrence E.; and Dornbusch, Sanford. "Collegial Evaluation of Classroom Teaching: Does it Work?" *Educational Research Quarterly*, 1, 1 (Spring 1976), pp. 56-66.

OSSC GOVERNING BOARD

Jean Fagnan, School Board Member, Reedsport SD 105
Charles Wiltse, Superintendent, Baker SD 5
George Benson, Superintendent, Centennial SD
Lee Christiansen, Washington ESD
Philip K. Piele, Executive Secretary, Oregon School Study Council
Ralph C. Rands, Associate Professor, Division of Educational Policy and Management, University of Oregon
Milt Baum, Associate Superintendent of School Management Service Division, State Department of Education
Betty Knight, School Board Member, Oregon School Boards Association

OSSC STAFF

Philip K. Piele, Executive Secretary
Stuart C. Smith, Editor
Mary Cihak Jensen, Patrick J. O'Connor, Graduate Research Assistants