Using Collaborative Performance Appraisal to Enhance Teachers' Professional Growth: A Review and Test of What We Know

J. BRADLEY COUSINS
Associate Professor, Faculty of Education, University of Ottawa, 145 Jean-Jacques Lussier, Ottawa, Ontario Canada K1N 6N5

Current efforts to restructure education focus on the professionalization of teaching (e.g., Murphy, 1991), which presupposes that teachers assume responsibility for their professional growth and development. Yet unrelenting public demands for accountability have forced school districts to implement and maintain performance appraisal policies that acknowledge the supervisors' responsibility for the enhancement of teachers' professional growth. Many such policies advocate a collegial or collaborative appraisal model that expect teachers and supervisors to work as partners in stimulating teachers' professional development. This model is in significant conflict with traditional "inspectorial" approaches to performance appraisal in schools. Research on collaborative performance appraisal is in its infancy, but sufficient studies have now been conducted to warrant an assessment of what is known. Such a review is provided in the present article, but, in addition, a second objective is pursued, that of assessing current supervisors' views about what constitutes effective collaborative appraisal practice. An exploratory survey study based on knowledge generated by the review of literature is reported to accomplish this goal.

Many current school district performance appraisal policies and practices are badly out of sync with contemporary educational restructuring initiatives. For example, Marczely (1992) found that most school districts in Ohio are still using 'trait' and 'rating' methods to evaluate professional teaching staff while at the same time claiming to value professional growth and instruction as legitimate purposes for teacher evaluation. Teacher evaluation of this sort is virtually antithetical to teacher growth and development (McGreal, 1990). Heafele (1992) refers to the still widely practiced "tell and sell" and "tell and listen" models of appraisal. Both presuppose a central role for supervisors in the identification of teachers' performance weaknesses and the development of remedies framed by specified performance objectives. In both models teachers have little or no control of the process: "The hierarchy of power is affirmed, and the principal the dispenser of rewards and punishment, possesses all of it" (Haefele, 1992, p. 337). Research on the impact of performance appraisal systems has clearly indicated the need for reform. For example, Lawton, Hickcox, Leithwood,

and Musella (1986) in a large-scale survey of Ontario teachers and administrators found that the most frequently sited purpose for doing appraisals was "to comply with policy." An abundance of evidence suggests that appraisals are often viewed by administrators and teachers as benign, ritualistic exercises that are time consuming, unfulfilling, perfunctory, and superficial (Darling-Hammond, Wise, & Pease, 1983; Good & Mulryan, 1990; Haefele, 1992; Housego, 1989; Lawton et al., 1986; Murphy, 1987; Rothberg & Buchanan, 1981; Sandell & Sullivan, 1992; Sergiovanni, 1992; Stiggins & Bridgeford, 1985).

In response to pressure to reform, many school systems have recently developed appraisal policies that directly embrace the function of stimulating teachers' professional growth. In an increasing number of districts dual track systems have been mounted that distinguish between teacher evaluation for the purposes of supporting summative personnel decisions (e.g., promotion, tenure, placement, dismissal) and accountability requirements, on the one hand, and professional development, formative evaluation, supervision for growth, or general professional improvement activities, on the other. Duke and Stiggins (1990) and Popham (1998a, 1988b) call for the complete separation of teacher evaluation and growth enhancement processes because of the inherent conflict of accountability and growth objectives. Among the more salient justifications for this assertion is that distinctions made by supervisors between these two fundamental functions may not always be sharp in either conception or in practice. Resulting uncertainties act to inhibit teachers' inclination to innovate, experiment, or take risks. The focus in the present study is on the growth-enhancement functions, or tracks, of performance appraisal policies.

Conceptual Framework

Figure 1 shows a framework adapted and expanded from that developed by Lawton et al. (1986) for conceptualizing the performance appraisal process. The framework portrays in a temporal sequence the major components of the growth-oriented appraisal process as well as factors influencing it, on the one hand, and consequences of it, on the other. These components are briefly described below.

Factors Influencing the Process

A variety of organizational and individual factors and conditions will shape the growth-oriented appraisal process and, indirectly, the impact attributable to it. These factors include attributes associated with the supervisor (e.g., leadership style, training, time available), the teacher (e.g., desire for constructive feedback, growth objectives, experience, knowledge of self) and the organization (e.g., administrative support, policy history, culture, policy specifications). Two demographic characteristics—school level and gender—are of particular interest in the present study. These are discussed in more detail later in the article.

Growth-Oriented Appraisal Process

The process is best conceived as nonlinear and cyclical. At a macro level, preparatory activities are followed by information collection and observation and, in turn, feedback and follow-up. In practice, however, a teacher and supervisor may, and indeed ought to, work their way through this cycle several times throughout the course of the process. The subcomponents of the cycle are described as follows.

Preparation.   Variables associated with preparation activities include criteria for evaluation, standards of performance, and the provision and conduct of preconferences or meetings. Key to collaborative supervision is the process of negotiation that ensures that teachers have significant and meaningful input into the establishment of growth objectives.

Data Collection.   Preparatory activities determine data collection parameters, which include types of data to be collected (e.g., observation notes and scripts, interview notes, videotape), sources of data (e.g., supervisor, teacher, student), methods used (e.g., observations of practice, self-assessment), and their frequency of use.

Feedback and Follow-up.   Feedback and follow-up activities might take the form of face-to-face discussion between supervisor and teacher, written summaries of observations, collaborative review of videotapes, or some combination of verbal and written feedback. In many cases, a post-observation interview (hereafter post-conference) is held with the teacher in order to digest and deliberate the findings of the process.

Impact

In the case of growth-oriented appraisal processes, impact will be manifest in the teacher's conceptual development or learning about her or his own performance, or, more importantly, in terms of changes in performance. Impact will also be evident in motivational and attitudinal dimensions. Conceptual and affective consequences may be positive or negative, intended or unintended.

Prior Research

Duke and Stiggins (1990) noted that empirical research on the use of teacher evaluation systems for the purposes of enhancing professional growth is badly lacking. Nonetheless, empirical research and reviews of practice concerning the nature and impact of performance appraisal systems has developed sufficiently in recent years to offer an increasingly clearer picture of what exemplary practices look like. Presently this literature is reviewed within the structure presented in Figure 1.

Figure 1. Conceptual framework.
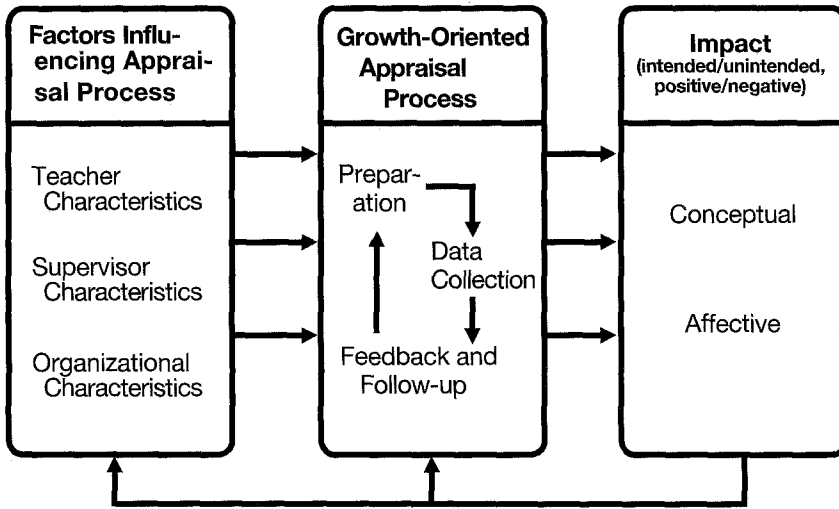
Factors Influencing the Process

Several authors advocate a sharp distinction between *decision-oriented* (summative) and *growth-oriented* (formative) appraisal systems to the point of insisting that organizational members with different roles have responsibility for the different functions (Duke, 1990; Duke & Stiggins, 1990; Greene, 1992; Popham, 1988a, 1988b; Stiggins & Duke, 1988). Others recognized that while traditional inspectorial models of appraisal are ineffective and should be abolished, it is not practical or sensible to divorce the two functions of teacher evaluation (Hunter, 1988; McLaughlin & Pfeifer, 1988a, 1988b). Most agree, however, that the intended purposes of appraisal ought to be made explicit (Glasman, 1974; Joint Committee, 1988; Popham, 1988a; Stiggins & Duke, 1988) and that methods for data collection ought to match the stated purpose (Stiggins & Bridgeford, 1985). One promising way to change the traditional 'mind-set' about performance appraisal and to put much greater emphasis on professional growth is to involve staff directly in the development of performance appraisal systems, a strategy that continues to receive strong conceptual and empirical support (Duke, 1990; Duke & Stiggins, 1990; Glasman & Paulin, 1982; Greene, 1992; McLaughlin, 1990; Natriello, 1990; Natriello & Dornbusch, 1980/81; Ryan and Hickcox, 1980; Stiggins & Bridgeford, 1985).

It is important that school system administrators actively promote and support growth-oriented appraisal policies. They can do this in at least three ways. First, training should be provided for both supervisors and teachers in order to enhance knowledge and skills and to promote "shared language" (Duke, 1990; McGreal, 1982, 1990; McLaughlin, 1984, 1990; McLaughlin & Pfeifer, 1988a; Ryan & Hickcox,

1980). Second, resources should be provided in the form of supplies and equipment, and time for classroom-based observation and supervisor-teacher interaction (Glasman, 1974; McLaughlin, 1984; Ryan & Hickcox, 1980). Finally, it is important to be clear about procedures to the point of specifying them in official policy documents (Joint Committee, 1988; McGreal, 1982; McLaughlin, 1984; Natriello & Dornbusch, 1980/81). The potential for confusion about growth enhancement and summative evaluation activities is high (Duke & Stiggins, 1990; Greene, 1992) and needs to be kept in check by explicit, clear, and accessible documentation.

Beyond organizational factors, attributes of both supervisors and teachers are crucial to the success of the growth enhancement process. Supervisors must be willing and able to separate evaluation and growth enhancement responsibilities both conceptually and in practice. A need for a collegial orientation to the process, one that invites and respects participation by the individual being supervised, is paramount (Duke, 1990; Duke & Stiggins, 1991; and Haefele, 1992). Cousins (1988), for example, found that participatory factors and motivation to grow were powerful determinants of principals' use of appraisal data for improving their own performance. In the absence of authentic participation the process is unlikely to move beyond ritualistic superficiality and mediocrity.

Growth-Oriented Appraisal Process

Preparation.   A variety of criteria or explicit dimensions of performances should to be made available for teachers to consider in advance of the process (Ryan & Hickcox, 1980). Such criteria ought to be research-based (Glasman & Paulin, 1982; Greene, 1992; Haefele, 1992; Housego, 1989; Medley, Coker & Soar, 1984) and a small, managable set should be selected for focus at any given time (McGreal, 1982). Scriven (1987) argues that research-based criteria are not valid indicators of performance, but his argument is limited to decision-oriented, summative teacher evaluation. Research-based indicators of performance provide valuable parameters for discussion, deliberation, and negotiation in a professional growth enhancement context. Several researchers strongly support the case to individualize appraisal for growth (Duke, 1990; Duke & Stiggins, 1990; Hickcox, 1990; Green, 1992; Huberman, 1988; Leithwood, 1991; McGreal, 1982; K.D. Peterson, 1990; P.L. Peterson & Comeaux, 1990; Sandell & Sullivan, 1992), which is achieved through the process of collaborative goal or performance objective setting that is respectful of the teacher's agenda for personal professional development (Alfonso & Goldberry, 1982; Duke, 1990; Erffmeyer & Martray, 1990; Glasman & Paulin, 1982; McGreal, 1982; McLaughlin, 1984; McLaughlin & Pfeifer, 1988a; Stiggins & Duke, 1988). McLaughlin (1984, 1990; McLaughlin & Pfeifer, 1988a), Iwanicki (1990), Stiggins and Bridgeford (1985), and Murphy (1987), among others, underscore the importance of integrating individual professional goals with those of the organization. While teachers are likely to negotiate on behalf of their own professional self-interests, it is incumbent on supervisors to negotiate on behalf of the organization, the probable consequences of which are the

strengthening of school improvement initiatives and the enhancement of the meaning of growth objectives to individual teachers working within the organizational culture.

Data Collection.   Historically, observation of classroom performance has been the mainstay of data collection regarding teacher performance appraisal (Medley et al., 1984; Murphy, 1987; Stiggins & Duke, 1988). Stodlosky cautiously extolled the virtues of using observation for development purposes.

> In formative evaluation, direct observation may be very appropriate if too much is not made of any given observation. Direct observation can provide useful occasions are what teaching is all about and may provide a very appropriate focus for discussing improvement. (1984, p. 17)

But several researchers advocate the use of multiple and innovative (e.g., videotape) data collection methods and sources (Darling-Hammond et al., 1983; Duke & Stiggins, 1990; Good & Mulryan, 1990; McGreal, 1982; Murphy, 1987; P.L. Peterson & Comeaux, 1990; Stiggins & Duke, 1988).

Self-assessment (Darling-Hammond et al., 1983; Stark & Lowther, 1984; Stiggins & Bridgeford, 1985), peer assessment (Kauchak, K. Peterson, & Driscoll, 1985; Rothberg & Buchanan, 1981; Stiggins & Bridgeford, 1985), and assessment of student performance (Hickcox, 1990; Murphy, 1987; Popham, 1988a; Ryan & Hickcox, 1980; Stiggins, 1989) if suited to growth goals and agreed upon have the potential to enrich the availability of information for reflection. Good and Mulryan (1990) advocate the conservative use of rating scales but only when supplemented with other data. McGreal (1990) is less generous. "The literature does not offer much support for the use of ratings in formative or summative teacher evaluation" (p. 49). Several recent studies (Kauchak et al., 1985; K.D. Peterson, 1987; Riner, 1987) strengthen the argument for assessing student progress as a source of feedback for teacher development, but as Brauchle, McLarty and Parker (1989) note, teachers need to develop their skills in compiling satisfactory evidence of student growth and evaluators need to better differentiate high-quality compilations of such evidence from indications of student growth.

Natriello (1984, 1990; Natriello & Dornbusch, 1980/81) found that teachers prefer to have fairly frequent observations of their performance even if feedback is negative, as was the case in some instance. According to Rothberg and Buchanan (1981), teachers are open to longer, more frequent observation that is less threatening, more informal, and guided by known objectives. There is no need to limit involvement in collaborative appraisal to those with supervision responsibility (Duke & Stiggins, 1988; Greene, 1992; Haefele, 1992). In fact, recent evidence suggests that teachers prefer to obtain feedback from compatible peers (Sandell & Sullivan, 1992). Finally, some researchers adamantly support training in data collection and analysis techniques and instructional supervision for those in supervisory capacities (Duke & Stiggins; Thies-Sprinthall & Sprinthall, 1987) and the Joint Committee on Standards for Educational Evaluation (1988) supports the use of data control mechanisms in order to protect privacy.

Feedback and Follow-up.   Enhanced communication skills and good quality discussion and dialogue are constructive aspects of the development process (Duke, 1990; Joint Committee, 1988; P.L. Peterson & Comeaux, 1990; Stiggins & Duke, 1988). Such feedback is motivating, and has stimulated teachers to reflect on their performance (McLaughlin & Pfeifer, 1988a; Stiggins & Duke, 1988). Duke and Stiggins (1990) raised the issue of whether communication should be simply descriptive and nonjudgmental (as in reciprocal peer coaching) or whether evaluative feedback should be provided in growth-oriented appraisal processes. The assumption underlying the provision of nonjudgmental feedback is that the teacher is her or his own best judge (or worst critic, as the case may be). But providing feedback that is nonevaluative is extraordinarily difficult even when there is serious intent (Kilbourn, 1990). Constructive feedback might be conceptually represented on a continuum. At one extreme vivid description of behaviors would be the focus with the colleague observer operating as a mirror or videotape for the teacher, as in the case of reciprocal peer coaching. At the other, as in the case of mentoring, for example, the intention would be to review observations and provide evaluative comments and suggestions for improvement. Providing nonjudgmental feedback is no easy task (Kilbourn, 1990). Some would argue that any questions or comments, by virtue of their capacity to focus the observation, are inherently evaluative (J.A. Ross, personnel communication, April 1993). The extent to which feedback should be judgment-free is a complex issue and is likely mediated by such variables as teacher experience. In any case, supervisors should develop good listening skills (Stiggins & Duke, 1988) and attempt ongoing and frequent communication (Natriello & Dornbusch, 1980/81). McLaughlin and Pfeifer (1988a) emphasized that feedback should be timely, specific, and credible.

The destination of written reports is an important issue for the growth-oriented process. For teachers to seriously embrace risky, or even just novel, instructional activities they must be free from sanction and know it (Duke, 1990). Reports destined for central office personnel files are likely to give the wrong signals to teachers. Only people with a legitimate need to use the data should be given access to reports (Joint Committee, 1988). Finally, it is important to develop and implement follow-up activities. A plan of action should be negotiated and agreed upon during the post-conference or at some point following the period of observation (Joint Committee, 1988; Ryan & Hickcox, 1980). The participation of the teacher in the construction of this plan is essential to a meaningful process.

Impact

The intended impact of growth enhancement activities is generally acknowledged as influence on behavior or the stimulation of improvements in performance (Lawton et al., 1986; Stiggins & Duke, 1988), but impact defined in this manner has been observed to be quite low (Darling-Hammond et al., 1983; Lawton et al., 1986; Murphy, 1987 P.L. Peterson & Comeaux, 1990). Others argue that impact is likely to be more apparent if defined in terms of cognitive or conceptual development, which may take

the form of new knowledge, adjusted focus, new skills, and changed classroom practice (Kilbourn, 1990). On the other hand, impact may be defined in affective terms and might include effects on morale, commitment, and job satisfaction. Some positive outcomes may be indirect or unanticipated. For example, the supervisory relationship may change to become more trusting and collegial, teachers may take the growth-oriented appraisal process more seriously, or they may develop an attitude of inquiry about their own development (Kilbourn, 1990). Certain negative effects may also accrue. The process may be anxiety producing, may be negatively perceived if identified problems cannot be corrected (Natriello, 1990), or may lead to defensiveness, frustration, wasted time, work overload, or superficiality (Kilbourn, 1990). Finally, impact has been described in terms of both individual and organizational constructs (Alfonso & Goldberry, 1982; Natriello, 1990). Whereas conceptual development and job satisfaction correspond to the former, mobilization of resources, careful and conscientious curriculum implementation, and organizational commitment are benefits associated with the latter.

Current Supervisors' Views and Opinions

Prior research on the nature and impact of performance appraisal systems suffers from at least the following deficiencies. First, most of the research is conducted within the context of comprehensive appraisal policies that address both accountability and performance improvement demands (e.g., Kauchak et al., 1985; Lawton et al., 1986; McLaughlin & Pfeiffer, 1988a); there is little research evidence exclusively focused on the systematic enhancement of teacher growth using prespecified performance appraisal policies as a backdrop. Second, research that has focused more directly on the dynamics and consequences of collaborative performance appraisal has been limited to qualitative research methods (e.g., Greene, 1992; Stiggins & Bridgeford, 1985). Methods that enable detected trends and themes to be generalized over school systems are needed. Third, much of the empirical literature has focused on teachers' attitudes toward evaluation and performance appraisal processes (e.g., Kauchak et al., 1985; Natriello, 1984; Rothberg & Buchanan, 1981; Sandell & Sullivan, 1992; Stiggins & Bridgeford, 1985). Little has been written about supervisors' perceptions of the appraisal process. Finally, very little research in this area has been guided by a prespecified conceptual framework. While interpretive designs have been of high utility in exploratory modes of research, research-based knowledge in the area has developed sufficiently to identify key variables and relationships in advance. The present study seeks to overcome these breeches in our knowledge. Specifically, a questionnaire was developed within the parameters of the foregoing conceptual framework and literature review and administered to individuals with responsibility for teacher supervision. Two sets of research questions guided the investigation:

1. To what extent do supervisors' self-reported opinions about using collaborative performance appraisal to enhance teacher growth correspond to exemplary practice

as identified in the research literature? What are the most salient gaps in current practice?

2. Can supervisors' opinions about enhancing teachers' growth through collaborative appraisal be differentiated on the basis of demographic and personal attributes (i.e., gender, school level)?

Although little has been written on the impact of gender on successful supervision (Shakeshaft, 1989), recent reviews of research have revealed gender differences in leadership and supervision styles. In a metaanalysis of research Eagle, Karau, and Johnson (1992) found that more recent studies have produced higher gender stereo-typic findings than have older ones. In their words women who occupy the principal role are more likely than men to treat teachers and other organizational subordinates as colleagues and equals and to invite their participation in decision making. Men evidently adopt a less collaborative style and are relatively more dominating and directive than are women (p. 91). Also, Shakeshaft (1989, p. 329) purported that women are more likely as supervisors to focus on instructional issues and matters concerning the child whereas men are more drawn to administrative problems. Based on these assertions, it is reasonable to posit that women supervisors would be more inclined toward the collaborative perspective on growth-oriented appraisal.

School level (elementary, secondary) was also selected for study because prior research suggests it is likely to be predictive of supervisors' beliefs and actions. Previous research has shown, for example, that teachers have greater classroom autonomy in secondary schools and that elementary schools are more likely to be comparatively tightly coupled, "rational bureaucratic" organizations with narrower sets of goals, and a higher degree of integration (Firestone & Herriot, 1982; Hallinger & Murphy, 1987; Herriot & Firestone, 1984). Given smaller staff size and tendencies toward integration, one might expect supervisors in elementary schools to be more receptive toward collaboration than their secondary school counterparts.

Method

Sample.   The sample was a convenience sample of elementary and secondary school personnel with responsibility for teacher supervision. It was drawn from six separate sources in a large geographic region in central Ontario. These were 1) 53 principals, vice-principals, and department heads from four east-central rural school districts enrolled in a multisession in-service program on enhancing teacher growth; 2) 17 secondary school department heads from a single suburban school district engaged in a multisession in-service program designed to produce multidimensional profiles of teacher growth (Cousins, 1991); 3) 11 secondary school vice-principals from an additional east-central school district attending a single session workshop on using profiles of growth to enhance teachers' professional development; 4) 32 aspirant school administrators from several Ontario school districts attending a multisession preservice program leading to Ontario provincial principal certification; 5) 23 princi-

pals and vice-principals from four boards in west-central Ontario attending a multises-sion refresher in-service program on school improvement; and 6) 13 secondary school department heads from an east-central school district attending a multisession leader-ship training program. The total achieved sample was 152. Response rates were 100% for sources 1, 2, 3, but dropped to 45%, 60%, and 65% for sources 4, 5, and 6, respec-tively. For these latter sources participation was voluntary. Of those providing background information 64 (46%) were women, 66 (48%) were in the elementary panel, and 70 (48%) were in supervisory roles other than principal or vice-principal (i.e., department head, division chair, lead teacher, consultant, coordinator).

Instrument.   Data were collected using the Teacher Supervision Survey (TSS), a seven-page questionnaire developed and pilot tested by the author. The instrument was based on the conceptual framework (Figure 1) and a review of prior research. The final version of the instrument could be completed in 20–30 minutes and was prefaced with the following definition of supervision.

> For the purposes of this survey, supervision is defined as performance appraisal activities that have as an expressed purpose: the *growth* or *development* of teachers. Such activities are variously referred to as developmental evluation, supervision for growth, or formative evaluation. Supervision is distinct from teacher evaluation activities designed to support personnel decisions such as promotion, tenure, or dismissal. Supervision occurs with competent teachers who are normally on permanent contract.

The questionnaire contained two major sections corresponding to current practice and supervisors' views and opinions. In the present study only the opinion data are reported. All questionnaire items were on a four-point Likert-type scale ranging from strongly disagree (1) to strongly agree (4). "Not applicable" was provided as an option.

Procedure.   In most cases, the TSS was administered and collected at in-service and preservice sessions. In cases where programs were designed to improve supervisors' skills in promoting teachers' professional growth (Sources, 1, 2, 3, and 6), the questionnaire was administered prior to the initial session (pretest) a warm-up for participation. At the principal refresher in-service program (5) participants were given at the session the TSS and a stamped self-addressed envelope in which to return it. Data were entered into the computer and analyzed using SPSSpc and a gender (female, male) be panel (elementary, secondary) 2×2 between group multivariate analysis of variance was conducted with variables associated with subcomponents of the concep-tual framework serving as clusters of dependent variables. Since the analysis calls for equal N within cells, cases were randomly eliminated leaving 26 cases per cell for a total of 104. It must be noted that while the procedure rendered gender and panel completely orthogonal, both were confounded with supervisors' role. Principals and vice-principals were more likely work in the elementary panel, X^2 (1, $N=104$)=17.28, $p<.001$, and to be male, X^2 (1, $N=104$)=7.57, $p<.01$. Accordingly, results of the multivariate analyses must be interpreted with caution.

Results

Table 1 shows descriptive statistics for each of the 30 attitudinal items, first for the total sample and then broken down by gender and panel category. Results of the 2 × 2 multivariate analysis of variance are reported in Table 2. Several of the questionnaire items reflect factors that influence the appraisal process (that is, teacher, supervisor, organization, and appraisal policy characteristics). Results concerning each of these categories are considered in turn. As seen in Table 1, respondents tended to accept teachers as being aware of their strengths and motivated to improve, but there was disagreement, on average, over whether teachers should be the ultimate judge of their own performance, or whether they are aware of their shortcomings. Also, it was revealed that perceptions about heavy workload intrude upon the growth-oriented appraisal process. A statistically significant multivariate main effect for panel was observed (see Table 2). Inspection of the univariate F's revealed that respondents from the secondary panel were more likely to believe that teacher workload was an obstacle than their elementary colleagues. There was also a suggestion that secondary supervisors were ambivalent regarding their beliefs about teacher's awareness of shortcomings, whereas elementary respondents tended to disagree. This latter finding approached, but did not achieve statistical significance. No other statistical main effects or interactions concerning teacher characteristics were found.

Supervisors, on the whole, agreed that their role is fundamental to an effective collaborative process. They acknowledged in the strongest terms that trust between the appraiser and the teacher is crucial and that supervisor training is important. Multivariate analyses revealed no differences in views about supervisors characteristics attributable to supervisor gender or panel.

Two items in the organization characteristics category met with respondents' disagreement. Clearly the availability of supervisors' time for the growth-oriented appraisal process and the size of their supervisory units were viewed as obstacles to the process by respondents (see Table 1). Supervisors were of the view, however, that their central administration was supportive of growth-oriented teacher appraisal activity. Elementary school supervisors were more likely than their secondary school counterparts to agree about district office support. No other statistically significant differences regarding organization characteristics were found.

The final category of potentially influential factors was characteristics of the appraisal policy itself addressed by five items in the questionnaire. As shown in Table 1, each of the items met with supervisors' moderate agreement. They tended to agree most strongly that supervisory roles are clearly defined and that supervision should be collegial. Table 2 shows that differences in opinion regarding policy issues were held be women versus men and elementary respondents compared with secondary. Women were less likely to agree that the supervisor-teacher relationship should *not* be expert-novice and that collegial supervision should *not* be used for accountability purposes. Women were also less likely than men to report that policy documents in their boards are adequate. Differences due to panel showed that respondents from elementary schools had a better sense of supervisory role responsibilities and were slightly more

Table 1. Descriptive Characteristics of Supervisors' Views and Opinions by Gender and Panel (Scale = 1 "Strongly disagree" to 4 "Strongly agree").

Abbreviated Item Stem	Total			Gender (N=104)				Panel (N=104)			
				Female		Male		Elementary		Secondary	
	Mean	SD	N	Mean	SD	Mean	SD	Mean	SD	Mean	SD
FACTORS INFLUENCING APPRAISAL PROCESS											
Teacher Characteristics											
1. Teachers are well aware of their strengths	2.93	0.63	149	2.98	0.57	2.92	0.55	2.89	0.66	2.81	0.88
2. Teachers are generally motivated to improve	3.09	0.54	150	3.06	0.56	3.11	0.51	3.04	0.57	2.88	0.55
3. Teacher is ultimate judge of own performance.	2.43	0.72	145	2.37	0.80	2.39	0.57	2.38	0.67	2.42	0.71
4. Teachers do not perceive heavy workload as prohibitive to process. [a]	2.27	0.80	149	2.30	0.80	2.17	0.86	2.48	0.71	1.94	0.83
5. Teachers are well aware of their shortcomings. [a]	2.40	0.67	149	2.42	0.69	2.45	0.57	2.33	0.68	2.50	0.54
Supervisor Characteristics											
6. Supervisor's subject area expertise is nonessential.	3.03	0.74	148	2.98	0.84	3.02	0.75	3.00	0.57	3.17	0.51
7. Supervisors need to be experts in instruction.	2.83	0.81	150	2.72	0.76	2.83	0.78	2.71	0.76	2.88	0.70
8. Trust between supervisor and teacher is extremely important.	3.79	0.42	150	3.81	0.44	3.75	0.43	3.83	0.39	3.77	0.47
9. Supervisors should be trained to carry out the role.	3.72	0.47	149	3.66	0.52	3.77	0.42	3.70	0.46	3.75	0.48
Organization Characteristics											
10. District administration is supportive of process.	3.23	0.69	143	3.18	0.65	3.41	0.61	3.46	0.54	3.12	0.70
11. Not difficult to find time for supervision. [a]	1.78	0.79	145	1.77	0.81	1.79	0.79	1.96	0.83	1.61	0.72
12. Number of teachers to supervise is manageable.	2.75	0.78	119	2.59	0.81	2.94	0.72	2.73	0.73	2.86	0.81
14. Unit not too large to supervise all teachers. [a]	2.32	1.04	129	2.16	0.93	2.32	1.05	2.30	1.00	2.21	0.99

Policy Characteristics

14. It is clear who is to supervise teachers in district.[a]	3.28	0.84	149	3.09	0.81	3.32	0.85	3.45	0.67	2.94	0.92
15. Support documents for supervision are adequate.	2.60	0.82	144	2.47	0.90	2.82	0.65	2.77	0.74	2.47	0.84
16. Supervision should not be expert-novice relationship.[a]	3.27	0.67	147	3.15	0.72	3.53	0.54	3.27	0.72	3.44	0.58
17. Supervision should be separate from teacher evaluation.	2.94	0.91	142	2.83	0.98	3.01	0.91	2.57	0.95	3.21	0.83
16. Supervision should not be used for accountability.[a]	2.81	0.84	144	2.70	0.86	2.94	0.79	2.57	0.81	3.02	0.81

GROWTH-ORIENTED APPRAISAL PROCESS

Preparation

19. Have enough information to set expectations for teacher.[a]	3.09	0.62	140	3.04	0.71	3.14	0.60	3.11	0.70	3.04	0.61
20. Number of growth objectives should be limited.	3.38	0.54	150	3.46	0.50	3.28	0.53	3.37	0.49	3.38	0.57
21. Growth objectives should overlap with school/district priorities.[a]	2.05	0.79	148	1.98	0.78	2.17	0.80	1.76	0.65	2.34	0.80
22. Growth objectives should not be set just by supervisor.[a]	3.24	0.63	147	3.30	0.61	3.19	0.63	3.32	0.65	3.18	0.56

Data Collection

23. Student assignments and work should be used as feedback.	3.13	0.62	147	3.17	0.61	3.12	0.59	3.19	0.45	3.06	0.70
24. Many sources of information should be used in the process.	3.48	0.56	150	3.52	0.64	3.38	0.48	3.45	0.64	3.42	0.50

Feedback and Follow-up

25. Supervisors should be good listeners.[a]	3.71	0.76	149	3.61	0.88	3.75	0.68	3.73	0.70	3.62	0.89
26. Records should be kept on file only at teacher's request.	2.38	0.89	149	2.36	0.85	2.35	0.94	2.03	0.82	2.69	0.84

General

27. Supervisor need not be directly involved in process	2.01	0.85	147	2.23	0.90	1.88	0.78	1.92	0.87	2.13	0.79
28. Supervision process should be collegial.	3.34	0.65	149	3.32	0.75	3.30	0.61	3.22	0.71	3.40	0.66
29. Supervision should stimulate self-reflection in teachers.	3.73	0.45	150	3.72	0.45	3.72	0.45	3.71	0.46	3.71	0.46
30. Others should be involved in the supervision process.	3.19	0.62	144	3.26	0.63	3.19	0.52	3.20	0.68	3.24	0.48

[a] Item reversed from phrasing in questionnaire.

Table 2. Sources of Variation for 2 × 2 Multivariate Analysis of Variance ($N = 104$).

Source	Univariate df	Hotelling's Approx. F	Multivariate df	Probability
FACTORS INFLUENCING THE APPRAISAL PROCESS				
Teacher Characteristics				
Panel (P)	1	3.46	5, 87	$p < .01$
Gender (G)	1	0.13	5, 87	NS
P × G	1	0.71	5, 87	NS
Error	91			
Total	94			
Organization Characteristics				
Panel (P)	1	2.80	4, 59	$p < .05$
Gender (G)	1	1.15	4, 69	NS
P × G	1	0.56	4, 69	NS
Error	72			
Total	75			
Policy Characteristics				
Panel (P)	1	5.75	5, 81	$p < .001$
Gender (G)	1	4.58	5, 81	$p < .001$
P × G	1	1.80	5, 81	NS
Error	85			
Total	88			
GROWTH-ORIENTED APPRAISAL PROCESS				
Preparation				
Panel (P)	1	3.45	4, 86	$p < .01$
Gender (G)	1	1.13	4, 86	*NS*
P × G	1	0.39	4, 86	NS
Error	89			
Total	92			
Feedback & Follow-up				
Panel (P)	1	8.32	2, 97	$p < .05$
Gender (G)	1	0.57	2, 97	NS
P × G	1	0.23	2, 97	NS
Error	98			
Total	101			

accepting of local support documents than their secondary school colleagues. However, there was clear evidence to suggest that members of the secondary panel favored a more collegial performance appraisal model. This group was more likely to agree that supervision should be separate from teacher evaluation and that supervision should not be used to support accountability needs. Elementary school respondents, on the whole, were ambivalent about this distinction. Within groups, opinions were mixed as reflected by relatively high standard deviations.

The remainder of the items correspond to opinions about components of the actual appraisal process. These opinions were divided into preparation, data collection, feedback, and general categories. Respondents' opinions favored the statements describing effective collegial supervision practice in the preparation category with the exception of one item. Respondents were not persuaded that teachers' growth objectives should overlap with school and district priorities. The multivariate analysis yielded only a single main effect for panel. Inspection of univariate F's revealed that secondary personnel were more likely to favor overlap of growth objectives and local priorities than their elementary school counterparts, although on average, they remained in disagreement with the concept.

Respondents were generally supportive of the two items listed in the data collection category, although they clearly favored the use of multiple sources of data over the use student achievement information. No statistically significant differences were attributable to gender or panel in this category. The two items associated with the feedback and follow-up category revealed distinct views. Supervisors were highly accepting of the suggestion that they ought to be good listeners, but took exception to the proposed disposition of records on the process. On average, respondents disagreed that appraisal records should be kept in personnel files only at the teacher's request. However, as shown in Table 2, elementary and secondary supervisors were of a different mind on this issue with secondary respondents being generally in agreement. There were no effects attributable to respondent gender within this category.

Finally, general views about the appraisal process revealed that supervisors were strongly supportive of a growth-oriented process that stimulates self-reflection in teachers. They also clearly agreed that the process should be collegial and that it ought to involve colleagues other then the supervisor and teacher in question. However, supervisors dismissed the notion that they need not be directly involved in the process. Not shown in Table 2 is a multivariate interaction effect that was approaching but did not achieve statistical significance. This weak multivariate tendency was due to the univariate interaction between panel and gender regarding the statement that supervision should stimulate self-reflection, $F(1,90) = 6.79$, $p < .01$. Females in the elementary panel were slightly more disposed toward this outcome than their male counterparts, but the reverse was true in the secondary panel.

General Discussion

Collaborative Performance Appraisal: Is There Middle Ground?

The present study sought to define exemplary practice in the implementation of growth-oriented appraisal policies by conducting a thorough review of a burgeoning literature. The review was guided by a conceptual framework that lays out personal and organization conditions that foster collaborative appraisal and that describes the process itself as a cyclic operation involving preparatory, information generation and feedback activities. The framework also specifies process impact in conceptual and

affective terms. The review of prior research yielded a rich base of data that informs an image of effective practice. The participation of teachers in the process, taking an active role in negotiating growth objectives, deciding on the nature and frequency of information collection, and engaging in transactional, constructive feedback is clearly the most salient feature of the process. This observation is in keeping with findings in other domains that underscore the advantages of participation, for example, participatory program evaluation (e.g., Cousins & Earl, 1992) and participatory decision making (e.g., Duke & Gansneder, 1990; Smylie, 1992). It is also in keeping with current trends in educational restructuring, which emphasize the professionalization of teaching and advocate the dismantling of traditional power relationships in favor of a more collegial and participatory organizational structure (Leithwood & Jantzi, 1990; Murphy, 1991; Sarason, 1991). Within the domain of supervision, however, collaborative performance appraisal as defined here appears to fall somewhere between what is known, on the one hand, as neoprogressive approaches and, on the other, as neotraditional (Darling-Hammond & Sclan, 1992; Garman, Glickman, Hunter, & Haggerson, 1987; Tracy & McNaughton, 1989).

The neoprogressive orientation to supervision posits that the primary focus for growth is the teacher's agenda and that the process is agreement-oriented and collegial. Criteria are determined internally and are based exclusively on the needs of the individual within her or his particular context. According to Garman (Garman et al., 1987), the emphasis in supervision ought to be on personal empowerment and helping teachers to become reflective and inquiry minded. The neoprogressive view, some would argue, is consistent with the original spirit and intent of clinical supervision as espoused by its founding architects (Cogan, 1973; Goldhammer, 1969). Others would disagree. Theorists' whose views are aligned with a neotraditional posture are committed to the improvement of instruction but rely heavily on current knowledge about effective teaching to determine appropriate criteria for supervision. A strong advocate of this approach, Hunter (see Garman et al., 1987, and Tracy & McNaughton, 1989), implores that teachers need to see cause-and-effect relationships in their classroom practice and it is the supervisor's job to show them. This view of clinical supervision, which has its roots in behavioral psychology, purports not to be concerned about the standardization of teaching, but rather, the standardization of the invariance in the relationships between teaching and learning.

Is there middle ground? Tracy and McNaughton (1989) think not. They argue that the two perspectives have fundamentally different starting points and assumptions and that reconciliation is not possible. They suggest, however, that supervisors ought to employ situational leadership in determining the preferred model of supervision. Novice teachers, for example, might benefit more significantly from the guidance emanating from a neotraditional, hierarchical, expert-novice orientation, whereas their more experienced counterparts are more likely to profit from a self-directed process. But what of collaborative performance appraisal as defined in the present study? Clearly, neoprogressive principles are inherent in effective growth-oriented appraisal. Adherence to teachers' expressed needs and desires in the collaborative construction of growth objectives, agreeing information collection types and schedules, encouraging

peers and colleagues to be part of the process, providing nonjudgmental, constructive feedback, and leaving the disposition of records to the discretion of the teacher, are but to name a few. But some elements remain outside of the neoprogressive realm. Negotiations that include efforts on behalf of the supervisor to ensure that growth objectives at least partly overlap with school or system priorities imply organization responsibility and commitment. The mere establishment of performance objectives or standards regardless of the collaborative nature of their determination, for some, contradicts the neoprogressive, self-motivated, self-reflective view. It may be concluded, then, that collaborative performance appraisal as defined by the findings of recent research appears to occupy middle ground. If we accept this assertion we need to carefully study implementation in order to assess the gap between current and effective practice, to identify obstacles that are likely to intrude upon the development of effective practice, and to devise strategies to overcome these potential barriers. The exploratory survey study reported here is a preliminary attempt to accomplish this goal.

Current Supervisors' Views and Opinions

The present data are limited in three ways. First, the obtained sample was based on convenience sampling strategies, and although data were collected from a wide variety of school supervisors from a broad sample of school districts, limits on the generalizability of findings remain. Second, the data collection instrument was based on the literature review and pilot tested, but rigorous checks on reliability and validity remain to be done. Third, analyses of differences attributable to supervisor gender and school level are confounded with supervisory role. Elementary school respondents and males were more likely to be principals or vice-principals. The observed findings need to be replicated under more stringent methodological controls, but they serve as an adequate starting point. The findings are discussed with these caveats in mind.

What are the most salient gaps between supervisors' views and opinions and effective practice? Supervisors' views about the *conditions* supporting collaborative performance appraisal were generally favorable. They acknowledged that trust between supervisors and teachers was vitally important and that teachers are motivated to improve. Support from central district administration and supervisor training were seen as important factors. Supervisors also noted that subject-area expertise is not a fundamental requirement. But several gaps emerged concerning the respondents' views about conditions supporting collaborative appraisal. Three obstacles were particularly notable. The availability of time for supervisors to carry out the role, the prohibitive size of the supervision unit, and teachers' heavy workload were identified as significant barriers to effective practice. These obstacles are likely to be relatively easily remedied by strategies devised to involve others in the process and otherwise share responsibilities. Such strategies are in keeping with current restructuring initiatives and capitalize on principles associated with division of labor. Two additional

obstacles are likely to be more problematic to remedy. Respondents did not necessarily concur that teachers are the best judge of their own performance or that they were well aware of their shortcomings. Adherence to this view lodges the process very directly into a hierarchical model and will ultimately work to erode trust. Active participation and ownership of the process by teachers is fundamental to the stimulation of growth and reflective practice. Serious strategies designed to change the mindset of supervisors (McLaughlin, 1990) are likely to be powerful predictors of effective practice.

Respondents held several opinions about the appraisal *process* that were consistent with effective collaborative practice. They underscored the importance of multiple sources of data, acknowledged that the process should be collegial, *not* expert-novice, and that it should stimulate teacher reflection. They also supported the involvement of others in the process and the use of student data as a way of providing feedback to teachers. However, three significant gaps emerged in their views about effective practice. First, supervisors were not persuaded that growth objectives should overlap with school or district priorities. While teachers' voice and ownership in the process are key to growth, teachers working in isolation from peers and in directions not aligned with current initiatives are likely to rob teachers from the benefits of teachers' joint work (Cousins, Ross & Maynes, 1994) and to erode coherence in school improvement initiatives that both build upon and draw from cooperation and collaboration among staff and administration (Iwanicki, 1990; Leithwood & Jantzi, 1990; Murphy, 1987; Stiggins & Bridgeford, 1985). Supervisors need to maintain a delicate balance between their responsibility to negotiate on behalf of the organization, on the one hand, and the importance of the teachers' agenda in the process, on the other. The second emergent obstacle concerns the disposition of records. Supervisors did not concur that the disposition of records should be left to the teachers' discretion. Supervisors need to recognize that the stimulation of growth requires of teachers a certain propensity to take risks in the classroom and that teachers will be unlikely to take risks under threat of sanction (Duke, 1990). While central administration need to know for accountability purposes that genuine growth-enhancement activities are taking place, personnel files need not provide a compulsory home for detailed records about the achievement of growth objectives. Policies that require nothing more than a brief activity statement to appear in the board office file are more likely to engender a sense of trust in the process among teachers while simultaneously satisfying accountability demands. Finally, supervisors were of the view that they need to be directly involved in the appraisal for growth process. Given their concerns about unit size and the lack of time available for appraisal, this practice is highly impractical and will serve to render the process superficial and meaningless. As noted above, strategies designed to involve peers in mentoring or peer coaching relationships, for example, will supply the supervisor with the knowledge that legitimate growth enhancement is taking place while at the same time not place excessive demands on her or his time. Teachers need to take responsibility for their own growth and supervisors need to make work conditions enabling to this end. This does not imply direct involvement as a necessary requirement.

What differences are attributable to supervisor gender and school level? School level turned out to be more predictive of differing opinions than supervisor gender. Super-visors from the secondary panel were more likely to view teacher workload as an obstacle to participation in the growth-oriented process and to raise concerns about supervisory role clarity and support from district administration. However, they were considerably more likely to advocate a clear separation between teacher evaluation and supervision for growth and to acknowledge that the process ought not to be used for accountability purposes. These findings are possibly attributable to the fact that more principals were located in elementary schools in the present sample. In secondary schools, clear separation of the process may be more possible given the current depart-mentalized structure common to most high schools. Department heads are theoretically able to carry out the appraisal for growth role while teacher evaluation for personnel administration can be left to the school administrators. Similar structural conditions are not present in most elementary schools where principals assume responsibility for both functions. The present findings imply that principals in the elementary panel are unable to draw a sharp distinction between growth-oriented and accountability appraisal functions either conceptually or in practice. The need to involve others in the process is particularly important under such conditions. One possibility might be to involve teacher leaders, a common structural feature of many restructuring initiatives, in the growth enhancement role. However, evidence suggests that lead teachers do not covet supervision as an added responsibility (Smylie & Brownlea-Conyers, 1992; Smylie & Denny, 1990) since they do not wish to jeopardize their membership in the collegium. According to Smylie and Brownlea-Conyers (1992, pp. 164–165), teacher leaders acknowledge "principals prerogatives to perform leadership tasks, such as personnel evaluation, that might place teacher leaders in untenable positions with their teaching colleagues." Clearly, strategies designed to change the mindset of teachers and administrators about the purposes of appraisal for growth are likely to be essential to overcoming this obstacle.

Differences of opinion attributable to supervisor gender were limited to policy considerations. Male supervisors (who were more likely to be principals or vice-princi-pals) tended to find support documents adequate and to underscore the need for training. They were also less likely to view the role from an expert-novice perspective. Prior research suggests that women in leadership roles are more likely to adopt a collegial stance with staff (Tabin, 1991) and score higher than men on a variety of leadership and managerial dimensions (Boleman & Deal, 1992). The present data are not consistent with these prior findings, although further, better controlled research is called for.

Implications for Policy and Practice

The data reported in the present study are exploratory and warrant further investigation, but they suggest some clear alternatives for policy direction and practice. First, the distinction between teacher evaluation and supervision to stimulate teacher growth

needs to be made more explicit and district staff (supervisors and teachers alike) would benefit from training programs designed to foster this distinction. McLaughlin (1990) asserted that changing the organizational culture concerning teacher appraisal is among the most formidable tasks confronting both policy makers and implementors. The present data do not diverge from this view. Second, the involvement of nonadministrative personnel in supervision for teacher growth needs to be reexamined closely. Teachers may not be attracted to this responsibility. Strong arguments for enhancing teachers' participation in school-level decision making (Murphy, 1991; Smylie, 1992) have been countered by sobering data concerning teachers' needs and the impact of such activities on relationships within the school (Duke, Showers & Imber, 1982; Duke & Gansneder, 1990; Weiss, Cambone, & Wyeth, 1992). Policy-makers need to support leadership development programs that make direct links to the supervisor's role in fostering teacher growth in collegial and nonhierarchical ways. Finally, efforts should be made to ensure reasonable overlap between teacher growth objectives and school and district priorities. Likely consequences of such efforts would be enhanced organizational efficiency and amplification of meaning of individual growth objectives. Supervisors need to attend, however, to the critical importance of teacher input into the appraisal process.

Implications for Research

Finally, the present data help to inform an agenda for research. First, supervisory practices were differentiated on the basis of characteristics of supervisors. These findings need to be verified and expanded using more tightly controlled research methods. What other characteristics (e.g., organizational role, age, tenure in the role) are likely to predict supervisors' beliefs and opinions? The results of such research will be useful in shaping policy initiatives and foci for training.

Second, self-reported attitudinal data need to be validated by other methods and sources of evidence including direct observation of supervision processes. Participant observation designs are likely to be extraordinarily illuminating of the dynamics of effective practice. Third, attention needs to be focused on the conceptual and affective impact of the growth-enhancement process. Which aspects of the process can be most effective? In what ways are factors affecting collaborative performance appraisal predictive of impact? Fourth, training issues need to be studied more closely. Which dimensions of supervisory practice and beliefs are likely to be improved through in-service and preservice opportunities? What sorts of professional development experiences will be beneficial to supervisors? Finally, understanding about discrepancies between beliefs and practice needs to be clarified. To what extent do district policies force supervisors to compromise their beliefs? What other organizational facets have similar effects?

Acknowledgments

References

Alfonso, R.J., & Goldsberry, L. (1982). Colleagueship in supervision. In T.J. Sergiovanni (Ed.)., *Supervision of teaching* (pp. 90–104). Virginia: Association for Supervision & Curriculum Development.

Blase, J.J. (1990). Some negative effects of principals' control-oriented and protective political behaviour. *American Educational Research Journal*, 27(4), 727–743.

Bolman, L.G., & Deal, T.E. (1992). Leading and managing: Effects of context, culture, and gender. *Educational Administration Quarterly*, 28(3), 314–329.

Brauchle, P.E., McLarty, J.R., & Parker, J. (1989). A portfolio approach to using student performance data to measure teacher effectiveness. *Journal of Personnel Evaluation in Education*, 3(1), 17–30.

Cogan, M.L. (1973). *Clinical supervision*. Boston: Houghton Mifflin.

Cousins, J.B. (1988) *Factors influencing knowledge utilization: Principals' use of data concerning their own performance*. Unpublished doctoral dissertation. Toronto: University of Toronto.

Cousins, J.B., (1991). Using profiles of teacher growth for teacher supervision. *Canadian Educational Leadership Network News*, 2(2), 4–8.

Cousins, J.B., & Earl, L. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, 14(4), 397–418.

Cousins, J.B., Ross, J.A., & Maynes, F.J. (1994). The reported nature and consequences of teachers' joint work in three exemplary schools. *Elementary School Journal*. 94(4), 441–465.

Darling-Hammond, L., & Sclan, E. (1992). Policy and supervision. In C.D. Glickman (Ed.). *Supervision in transition*, 1992 Yearbook of the Association for Supervision and Curriculum Development, (pp. 7–29). Alexandria, VA: ASCD.

Darling-Hammond, L., Wise, A.E., & Pease, S.R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research*, 53(3), 285–328.

Duke, D.L. (1990). Developing teacher evaluation systems that promote professional growth. *Journal of Personnel Evaluation in Education*, 4(2), 131–144.

Duke, D.L., & Gansneder, B. (1990). Teacher empowerment: The view from the classroom. *Educational Policy*, 4(2), 145–160.

Duke, D.L., Showers, B.K., & Imber, M. (1982). Teachers and shared decision making: The costs and benefits of involvement. *Educational Administration Quarterly*, 16(1), 93–106.

Duke, D.L., & Stiggins, R.J., (1990). Beyond minimum competence: Evaluation for professional development. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers* (pp. 116–132). Newbury Park: Sage.

Eagly, A.H., Karau, S.J., & Johnson, B.T. (1992). Gender leadership style among school principals: A meta-analysis. *Educational Administration Quarterly*, 28(1), 76–102.

Erffmeyer, E.S., & Martray, C.R. (1990). A quantified approach to the evaluation of teacher professional growth and development and professional leadership through the goal-setting process. *Journal of Personnel Evaluation in Education*, 3(3), 275–300.

Firestone, W.A., & Herriott, R.E. (1982). Prescriptions for effective elementary schools don't fit secondary schools. *Educational Leadership*, 40(3), 51–53.

Garman, N.B., Glickman, C.D., Hunter, M., & Haggerson, N.L. (1987). Conflicting conceptions of clinical supervision and the enhancement of professional growth and renewal: Point and counterpoint. *Journal of Curriculum and Supervision*, 2(2), 152–177.

Glasman, N.S., (1974). Personnel evaluation research: Implications for instruction improvement. *The Canadian Administrator*, 13(6), 29–33.

Glasman, N.S., & Paulin, P.J. (1982). Possible determinants for teacher receptivity to evaluation. *Journal of Educational Administration*, 20(2), 148–171.

Goldhammer, R. (1969). *Clinical supervision: Special methods for the supervision of teachers*. New York: Holt, Rinehart and Winston.

Good, T.L., & Mulryan, C. (1990). Teacher ratings: A call for teacher control and self-evaluation. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers.* (pp. 191–215). Newbury Park: CA: Sage.

Greene, M.L. (1992). Teacher supervision as professional development: Does it work? *Journal of Curriculum and Supervision*, 7(2), 131–148.

Haefele, D.L. (1992). Evaluating teachers: An alternative model. *Journal of Personnel Evaluation in Education*, 5, 335–345.

Hallinger, P., & Murphy, J. (1987). Instructional leadership in the school context. In W. Greenfield (Ed.), *Instructional leadership: Concepts, issues, and controversies.* (pp. 179–203). Boston: Allyn & Bacon.

Herriott, R.E., & Firestone, W.A. (1984). Two images of schools as organizations: A refinement and elaboration. *Educational Administration Quarterly*, 20(4), 41–57.

Hickcox, E.S. (1990). Ideas for performance appraisal: Breaking the mould. *Canadian School Executive*, 10(3), 11–13.

Housego, I.E.. (1989, September). Principals' evaluation of teachers in a British Columbia school district: Are the reports professional or bureaucratic documents? *The Alberta Journal of Educational Research*, 35(3), 196–216.

Huberman, M. (1988). Teacher careers and school improvement. *Journal of Curriculum Studies*, 20(2), 119–132.

Hunter, M. (1988). Effecting a reconciliation between supervision and evaluation: A reply to Popham. *Journal of Personnel Evaluation in Education*, 1(3), 275–279.

Iwanicki, E.F. (1990). Teacher evaluation for school improvement. In J. Millman & L. Darling-Hammond (Eds.)., *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers.* (pp. 158–171). Newbury Park, CA: Saga.

Joint Committee on Standards for Educational Evaluation. (1988). *The personnel evaluation standards: How to assess systems for evaluating educators.* Newbury Park: Sage.

Kauchak, D., Peterson, K., & Driscoll, A. (1985). An interview study of teachers' attitudes toward teacher evaluation practices. *Journal of Research and Development in Education*, 19(1), 32–37.

Kilbourn, B. (1990). *Constructive feedback: Learning the art.* Toronto: OISE Press.

Lawton, S.B., Hickcox, E.S., Leithwood, K.A., & Musella, D.F. (1986). *Development and use of performance appraisal of certificate education staff in Ontario school boards. Vol. I: Technical Report.* Toronto: The Queen's Printer for Ontario.

Leithwood, K.A. (1991). The principal's role in teacher development. In B. Joyce (Ed.), *Changing school culture through staff development.* (pp. 71–107). Alexandria VA: Association for Supervision and Curriculum Development.

Leithwood, K.A., & Jantzi, D. (1990). Transformational leadership: How principals can help reform school cultures. *School Effectiveness and School Improvement*, 1(4), 249–280.

Marczely, B. (1992). Teacher evaluation: Research versus practice. *Journal of Personnel Evaluation in Education*, 5, 279–290.

McGreal, T.L. (1990). The use of rating scales in teacher evaluation: Concerns and recommendations. *Journal of Personnel Evaluation in Education*, 4, 41–58.

McGreal, T.L. (1982). Effective teacher evaluation. *Educational Leadership*, 39(4), 303–305.

McLaughlin, M.W. (1984). Teacher evaluation and school improvement. *Teachers College Record*, 86(1), 193–207.

McLaughlin, M.W. (1990). Embracing contraries: Implementing and sustaining teacher evaluation. In J. Millman & L. Darling-Hammond (Eds.) *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers.* (pp. 403–415). Newbury Park: CA: Sage.

McLaughlin, M.W., & Pfeifer, R.S. (1988a). *Teacher evaluation: Improvement, accountability, and effective learning.* New York: Teachers College Press.

McLaughlin, M.W., & Pfeifer, R.S. (1988b). Teacher evaluation: Organizational change, accountability, and improvement. In E.S. Hickcox, S.B. Lawton, K.A. Leithwood, & D.F. Musella (Eds.), *Making a difference through performance appraisal.* (pp. 113–140). Toronto: Ontario Institute for Studies in Education Press.

Medley, D.M., Coker, H., & Soar, R.S. (1984). *Measurement-based evaluation of teacher performance: An empirical approach.* New York: Longman.

Murphy, J. (1987). Teacher evaluation: A comprehensive framework for supervisors. *Journal of Personnel Evaluation in Education,* 1(2), 157–180.

Murphy, J. (1991). *Restructuring schools: Capturing and assessing the phenomena.* New York: Teachers College Press.

Natriello, G. (1984). Teachers' perceptions of the frequency of evaluation and assessments of their effort and effectiveness. *American Educational Research Journal,* 21(3), 579–595.

Natriello, G. (1990). Intended and unintended consequences: Purposes and effects of teacher evaluation. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers.* (pp. 35–45). Newbury Park: Sage.

Natriello, G., & Dornbusch, S, (1980/81). Pitfalls in the evaluation of teachers by principals. *Administrator's Notebook,* 29(6), The University of Chicago.

Peterson, K.D. (1987). Teacher evaluation with multiple and variable lines of evidence. *American Educational Research Journal,* 24(2), 311–317.

Peterson, K.D. (1990). Assistance and assessment for beginning teachers. In J. Millman, & L. Darling-Hammond (Eds.). *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers.* (pp. 104–115). Newbury Park: Sage.

Peterson, P.L., & Comeaux, M.A. (1990). Evaluating the systems: Teachers' perspectives on teacher evaluation. *Educational Evaluation and Policy Analysis,* 12(1), 3–24.

Popham, W.J. (1988a). Judgment-based teacher evaluation. In S.J. Stanley & W.J. Popham, *Teacher evaluation: Six prescriptions for success.* (pp. 56–77). Alexandria, VA: ASCD.

Popham W.J. (1988b). The dysfunctional marriage of formative and summative teacher evaluation. *Journal of Personnel Evaluation in Education,* 1(3), 269–273.

Riner, P.S. (1992). A comparison of the criterion validity of principals' judgments and teachers' self-ratings on a high-inference rating scale. *Journal of Curriculum and Supervision,* 7(2), 149–169.

Rothberg, R.A., & Buchanan, L.L. (1981). Teacher perceptions of teacher assessment. *Phi Delta Kappan,* 62(7), 527.

Ryan, D.W., & Hickcox, E.S. (1980). *Redefining teacher evaluation: An analysis of practices, policies, and teacher attitudes.* Toronto: Ontario Institute for Studies in Education Press.

Sandell, M., Sullivan, K. (1992, June). Teacher disillusionment and supervision as a part of professional development. *The Alberta Journal of Educational Research,* 38(2), 133–140.

Sarason, S.B. (1990). *The predictable failure of educational reform: Can we change course before it's too late?* San Francisco: Jossey-Bass.

Scriven, M. (1987). Validity in personnel evaluation. *Journal of Personnel Evaluation in Education,* 1(1), 9–23.

Sergiovanni, T.J. (1992). Moral authority and the regeneration of supervision. In C.D. Glickman (Ed.), *Supervision in transition,* 1992 Yearbook of the Association for Supervision and Curriculum Development. (pp. 203–214). Alexandria VA: Association for Supervision and Curriculum Development.

Shakeshaft, C. (1989). The gender gap in research in educational administration. *Educational Administration Quarterly,* 25(4), 324–337.

Smylie, M.A. (1992) Teacher participation in school decision making: Assessing willingness to participate. *Educational Evaluation and Policy Analysis,* 14(1), 53–67.

Smylie, M.A., & Brownlea-Conyers, J. (1992). Teacher leaders and their principals: Exploring the development of new working relationships. *Educational Administration Quarterly,* 28(2), 150–184.

Smylie, M.A., & Denny, J.W. (1990). Teacher leadership: Tensions and ambiguities in organizational perspective. *Educational Administrative Quarterly*, 26(3), 235–259.

Stark, J.S., & Lowther, M.A. (1984). Predictors of teachers' preferences concerning their evaluation. *Educational Administration Quarterly*, 20(4), 76–106.

Stiggins, R.J. (1989). A commentary on the role of student achievement data in the evaluation of teachers. *Journal of Personnel Evaluation in Education*, 3(1), 7–15.

Stiggins, R.J., & Bridgeford, N.J. (1985). Performance assessment for teacher development. *Educational Evaluation and Policy Analysis*, 7(1), 85–97.

Stiggins, R.J., & Duke, D. (1988). *The case for commitment to teacher growth*. Albany: SUNY Press.

Stodlosky, S.S. (1984). Teacher evaluation: The limits of looking. *Educational Researcher*, 13(9), 11–18.

Tabin, Y. (1991). Joining the old boy's club? Women's careers as school principals in British Columbia. *CELN News*, 3(1), 10–13.

Thies-Sprinthall, L., & Sprinthall, N.A. (1987). Experienced teachers: Agents for revitalization and renewal as mentors and teacher educators. *Journal of Educational*, 169(1), 65–79.

Tracy, S.J., & MacNaughton, R.H. (1989). Clinical supervision and the emerging conflict between the neo-traditionalists and the neo-progressives. *Journal of Curriculum and Supervision*, 4(3), 245–256.

Weiss, C.H., Cambone, J., & Wyeth, A. (1992). Trouble in paradise: Teacher conflicts in shared decision making. *Educational Administration Quarterly*, 28(3), 350–367.