Making Sense of Distributed Leadership: The Case of Peer Assistance and Review
Author(s): Jennifer Goldstein
Reviewed work(s):
Source: *Educational Evaluation and Policy Analysis,* Vol. 26, No. 2 (Summer, 2004), pp. 173-197
Published by: American Educational Research Association
Stable URL: http://www.jstor.org/stable/3699559
Accessed: 11/09/2012 11:13

http://www.jstor.org

Making Sense of Distributed Leadership: The Case of Peer Assistance and Review

Jennifer Goldstein
Baruch College/City University of New York

This article explores a case of shifting leadership responsibility for teacher evaluation. Peer Assistance and Review (PAR) formally involves teachers in the summative evaluation of other teachers—although the boundaries of the involvement are often vague. Since teacher evaluation has traditionally been the domain of school principals, involving teachers in teacher evaluation raises questions about how those faced with the new role make sense of it and enact it. The article draws on theories of professions, organizations, and institutions to examine the implementation of PAR in one large urban school district. Findings suggest that, despite positive sentiments about the policy across stakeholder groups, those involved wanted principals to remain a central figure in the evaluation of teachers in PAR. Education's hierarchical norms, the difficulty of conducting evaluations, district leadership, and program ambiguity are identified as challenges to distributing leadership.

Keywords: *distributed leadership, organizational change, peer assistance and review, sensemaking, teacher evaluation*

TEACHER peer assistance and review (PAR) involves teachers in the summative evaluation of other teachers. While a variety of education policies increase teachers' leadership responsibility by placing them in such roles as mentors, curriculum developers, peer coaches, and researchers, only PAR increases teachers' formal authority by altering district organizational structures for teacher evaluation. As such, PAR challenges education's norms in particularly novel and potent ways.[1] This article examines PAR as a salient case of distributed leadership,

a growing theoretical perspective with few empirical studies.

I conducted this study of PAR during the first statewide implementation of the policy. In 1999, California Assembly Bill 1X marked the first time PAR had been instituted statewide, at a time when no major district had implemented the policy in over a decade. The legislation gave the state's roughly one thousand school districts a de facto mandate to have PAR programs in place to serve veteran teachers receiving unsatisfactory evaluations from their principals. Mentor funds

that districts were already receiving from the state would be phased out, and districts would only continue to receive that money by putting a PAR program in place.

Since its inception in K-12 public education over 20 years ago, little empirical research has been conducted on PAR. This study involves in-depth research over the course of a year and a half, and follow-up data gathered one year later. As such, it is a more extensive study of PAR policy than any undertaken to date. It explores the way those involved with implementing PAR in one urban district made sense of the policy, and how leadership responsibility for teacher evaluation was or was not redistributed. Questions about PAR's outcomes for the processes and quality of teacher support and evaluation, while critical, are topics for companion papers (see Goldstein 2002; 2003a). This article addresses the following three research questions:

1. How do educators within a school district make sense of the new role of teacher as evaluator of other teachers that comes with PAR?

2. How, if at all, is leadership responsibility for teacher evaluation redistributed as a result of the new role?

3. What are the implications for distributing leadership in organizations?

The introductory sections of the article explain PAR policy, discuss the conceptual underpinnings of the analysis, and describe the study's design and methods. The middle sections of the article present data from the study and address the research questions. The article ends with a discussion of study limitations and areas for future research.

Peer Assistance and Review

Background of PAR

PAR, historically referred to as "peer review" prior to the California legislation, experienced a very specific birth in Toledo, Ohio, in 1981. The teacher union president, frustrated with the caliber of new teachers hired by the school district, suggested that the district create an intern program. In 1981, after eight years of this proposal, a new negotiator for the district finally offered to support the new teacher intern program if the union would take responsibility for intervention with seriously ineffective tenured teachers. The union accepted, and the "Toledo Plan" of peer review came into existence (Kerchner & Koppich,

1993). It is still the most well known blueprint of the policy. Over the next two decades, Cincinnati and Columbus, Ohio; Poway and Mt. Diablo, California; Rochester, New York; Dade County, Florida; and Salt Lake City, Utah, were among the handful of school districts that undertook peer assistance and review.[2]

Most commonly, PAR involves "consulting teachers" or CTs—teachers identified for excellence and released from teaching duties full-time for two to three years—who provide mentoring to teachers new to the district or the profession, and intervention for identified veteran teachers experiencing difficulty. Although there are PAR programs without full-time release of consulting teachers, full-time release is the original Toledo model. The consulting teachers report to a district-wide joint teacher/administrator board, called the "PAR panel." The panel is typically co-chaired by the union president and the director of human resources (or some other high ranking district office administrator). Teachers in either the new or veteran category are called "participating teachers" or PTs[3]. A PT must meet specified quality standards or face removal from the classroom, as determined by the panel based on the recommendation of the consulting teacher, sometimes in concert with the principal. The panel's employment recommendation is then passed to the Superintendent, who makes a recommendation to the school board, the ultimate arbiter of personnel decisions.

Teacher Evaluation with PAR

As outlined, PAR addresses both support to and evaluation of teachers. Those involved with PAR often emphasize support as the most meaningful aspect of the policy. However, it is the evaluative aspect of PAR that is new to most educators, often controversial, and the focus of this study because of the particularly compelling way it challenges education's norms. Support is included in this discussion only as it relates to evaluation.

Those implementing PAR intend it to improve teacher evaluation several ways. Principals are increasingly described in literature as too overwhelmed or ill equipped to provide quality instructional leadership, including teacher evaluation (Copland, 2001; Cuban, 1988; Grubb, Flessa, Tredway, & Stern, 2003; Wise et al., 1984); PAR creates a full-time position dedicated to support and evaluation. Principals tend not to be well-trained to conduct educative evaluations, and traditional principal-led teacher evaluation usually

divorces assessment from support and professional development (Darling-Hammond, 1984; Wise et al., 1984); PAR links support and evaluation structurally as the same person conducts both, typically focused on rubrics for performance standards (Goldstein, 2003b). Principals often evaluate in isolation and can make capricious decisions; consulting teachers must defend their evaluative decisions to a panel of educators (Goldstein, 2003b). Some principals avoid the interpersonal conflict of giving negative evaluations to people they may know well, instead preferring to pass them to other schools (Bridges, 1986); PAR, as a district-based[4] rather than a site-based policy, draws from a pool of CTs and may avoid this dilemma. Principals have long complained of their inability to fire underperforming teachers due to the time and expense of union battles; PAR puts the teacher union at the table as a partner in teacher quality control.

Ultimately, with traditional teacher evaluation, nearly everyone is found to be competent (Koppich, 2000); notably, with PAR, this is not the case. Prior studies of PAR have found that 40 to 70% of teachers in intervention leave the classroom (Darling-Hammond, 1984; Hewitt, 2000; Kelly, 1998; Murray, 1999). Most leave voluntarily but dismissals do occur and are consistently seen to increase substantially with PAR. In addition, the remaining intervention cases successfully improve (Darling-Hammond, 1984; Hewitt, 2000; Murray, 1999), a finding in opposition to prior literature that has concluded that teachers cannot be "remediated" beyond their early years in the profession (Bridges, 1986). Similarly, PAR has proven more rigorous than traditional mechanisms for the gatekeeping of new teachers (Brown, 1993; Darling-Hammond, 1984; Goldstein, 2003b).

PAR in California and the Case District

Recall that the Toledo model of PAR involved both beginning teachers and veterans. Indeed, in most districts implementing PAR prior to California's legislation, the program began with beginning teachers as the less controversial part of the policy, and later expanded to veterans once the idea of teachers conducting teacher evaluations was established in a district. California's AB 1X, however, was borne from the desire to address the public's concern about ineffective veteran teachers; AB 1X's focus on veteran teachers reveals its intention as accountability policy rather than induction policy. However, while not

requiring it, the law allowed for the inclusion of beginning teachers in the program. As a result of this policy flexibility, and varying opinions about the wisdom of PAR among educators, school districts across the state created PAR programs that looked quite different from one another. Given the extremely small number of veteran teachers typically receiving unsatisfactory evaluations from principals, the relatively large amount of money attached to the policy through the mentor funds, and the enormous need to provide support to large numbers of new and often unprepared teachers, some districts took the opportunity to design PAR programs with CTs providing support to and conducting evaluations of *both* new and veteran teachers. The district presented in this study is one such district.

Conceptual Framework

Distributing Leadership for Teacher Evaluation

Formal teacher evaluation is a leadership function typically under the purview of principals. While administrator jurisdiction over teacher evaluation may now be taken for granted, the current hierarchical structure of the U.S. public education system was merely imported from the factory model of the industrial era a century ago (e.g., Tyack, 1974; Lortie, 1975). Education became routinized and standardized, with layers of management for supervision. The process that divided educational administration from teaching at that time was a strategy used by "upwardly mobile groups seeking to set themselves above their current peers" (Abbott, 1988, p. 106), a move that generated a largely male administration and female teaching force. Little (1988) claims that teachers, under this system, have traditionally viewed professional obligations to one another as intrusive at worst and loosely invitational at best, what Feiman-Nemser and Floden (1986) term a norm of noninterference. With this view, the responsibility for maintaining teacher quality and hence responsibility for evaluation resides hierarchically above teachers in the chain of command, in administration.

A countervailing vision for education, however, is one that would flatten the hierarchy and vest teachers with authority and responsibility for the quality of practice. One policy approach to improving educational quality over the past two decades has been to alter education's longstanding hierarchical authority structure, distributing lead-

ership responsibility beyond administrators to include teachers.[5] Research has suggested that increasing the leadership responsibility of teachers has positive outcomes for teacher quality and professionalism (e.g., Hart, 1987; 1995). However, the normative environment of public education has not been particularly conducive to this change, and such policies often fall short of realizing their full implementation or desired outcomes (Little, 1990).

Charging teachers with formal responsibility for the evaluation of other teachers in particular creates the potential for a struggle between teachers and administrators over occupational boundaries. Occupations hold jurisdiction (Abbott, 1988) over certain work tasks; over time this link becomes cognitively ingrained and taken-for-granted. However, occupations exist within an ecology of interdependence in which jurisdictional boundaries shift, jurisdictions are seized by one occupation from another, while the latter struggle to defend their territory (Abbott, 1988). PAR signals a potential jurisdictional shift between teachers and principals over the leadership function of teacher evaluation.[6] Such a shift creates a novel configuration of stretching responsibility for teacher evaluation across teacher leaders and principals (Spillane, Halverson, & Diamond, 2001).

The notion of distributing leadership across multiple actors in a school rather than solely with the site administrator has gained currency since the mid-1990s. Smylie, Conley, and Marks (2002) identify three models of distributing leadership in the recent literature: (a) leadership as the performance of tasks rather than the holding of roles (Heller & Firestone, 1995); (b) leadership as an organization-wide resource of power and influence, the interaction between individuals rather than the actions of individuals (Ogawa & Bossert, 1995; Leithwood & Jantzi, 2000); and (c) leadership as a social distribution that is "stretched over" two or more leaders in their interactions with followers in particular situations (Spillane et al., 2001). Since PAR does create a formal leadership role for teachers, it is perhaps a less esoteric model than that envisioned by some of the theoretical work on distributed leadership. Nonetheless, this study of PAR builds upon the literature on distributed leadership in several ways. First, the focus is the task of teacher evaluation, and the potential for separating the task from the administrative role of the principal. Second, the policy draws on talent across the district organization, teachers in addition to administrators. The study seeks to uncover

the resultant interaction of the stakeholder groups, and the influence the stakeholders have separately or collaboratively in deciding evaluative outcomes. Third, and in particular, the actors in the study were constantly engaged in negotiating the ways (research question #1) and degree (research question #2) that responsibility for teacher evaluation was stretched over the CTs, principals, and PAR panel. Few empirical studies exist of the third model of distributed leadership in particular (Smylie et al., 2002). As such, this article addresses a gap in an emerging area of scholarship.[7]

Making Sense of New Roles

While research on PAR is sparse, the literature identifies a shifting of teachers' roles to include the evaluation of other teachers (Darling-Hammond, 1984; Kelly, 1998; Kerchner, Koppich, & Weeres, 1997; Murray, 1998). Murray (1998) identifies various issues raised by PAR that require "changes in the beliefs and nature of authority relations" (Murray, 1998, p. 203), including redefining teachers' relationships with one another (Costa & Garmston, 2000; Darling-Hammond, 1997) and redefining teachers' relationships with administrators (also Kelly, 1998). Most educators at the K-12 level, both teachers and administrators alike, have little prior experience with teachers in the types of leadership roles created by PAR, and the shift is complex and challenging (Little, 1988).

Creating new institutionally legitimate roles is not only about creating new positions and structures—such as those established, however vaguely, by California's PAR legislation—but also fundamentally about creating new definitions and cognitive frames (e.g., Meyer & Rowan, 1977; Scott, 1995). If action leads to cognition (Weick, 1995), how are the cognitive frames around appropriate roles for teachers (and principals) shaped by the introduction of a PAR program? The taken for granted "way things are done" (Scott, 1995) in education has almost exclusively meant a conception of teachers and teaching in a classroom with a group of children (Little, 1988). Education has not established a cognitive frame through which to make sense (Weick, 1995) of teachers in other roles, despite the growing proliferation of teacher leadership policies. Feiman-Nemser (1998), for example, has identified how cognitive frames prevent most teachers in mentorship roles from viewing themselves as teacher educators. An adult providing assistance to a teacher, an adult manag-

ing a school, an adult developing policy—no cognitive link exists to connect these images of "experts" to the identity "teacher." This research focuses on the unlinked image of teachers as *evaluators,* exploring how teachers and those with whom they work come to make sense of this new role.

The literature on sensemaking highlights this phenomenon by which actors give structure to the unknown. Sensemaking is most evident at times when expectations break down, when patterns of behavior are interrupted, and hence new cognitive frames needed (Weick, 1995). Violations of social conventions like those enacted with PAR have the potential to transform an occupation, because the role innovation creates new ways of seeing and doing (Van Maanen & Barley, 1984). Spillane, Reiser, and Reimer (2002) note that the changed behavior required by new policy implementation involves complex cognitive processes in which implementing agents notice, frame, interpret, and construct meaning for policy messages. Sensemaking provides an appropriate frame for the study of the new teacher leadership roles brought by PAR because, surrounded by ambiguity, the consulting teachers and those with whom they work are not merely performing new roles but are also in the act of framing, interpreting, and constructing new roles, authoring as well as reading the new policy (Weick, 1995).

Tensions and Ambiguities of Teacher Leadership

While a multitude of teacher leadership programs and policies in recent years appear designed to challenge education's hierarchical norms, little change has occurred in the teaching profession (Grant & Murray, 1999; Tyack & Cuban, 1995) as teacher leadership policies have often been undermined. Little (1990) aptly notes that lead teachers "engage in a precarious form of improvisation . . . for an audience whose sympathy is far from certain" (Little, 1990, p. 341). In a review of the literature on teacher mentorship, she found a predictable pattern of conservative institutional responses to new teacher leadership policies, where "the institution makes moves to render the role harmless—and thus useless" (Little, 1990, p. 15, citing Bird, 1986). Little notes that even when program goals are accepted, there is an inability to overcome long-established norms and patterns of behavior. While the programs reviewed by Little do not specifically involve teacher evalu-

ation, the fact that the tendency has been to blunt the impact of even less controversial policies suggests that tension and jurisdictional struggle surround the implementation of PAR. Kerchner and Caufman (1995) argue that because PAR involves teachers making substantive judgments about teacher quality, it "places teachers in a social and intellectual position they have never before occupied" (Kerchner & Caufman, 1995, p. 115).

Role theory informs the tension and ambiguity that occurs with teacher leadership, and defines "role" as "particular behaviors and expectations tied to particular position labels" (Ickes & Knowles, 1982). Creating a new position label or status such as "consulting teacher" is distinct from defining the behavioral enactments that will accompany the role. Some roles involve a known and commonly shared social structure, so that the accompanying behavioral enactments are fairly defined (Stryker & Statham, 1985). Consulting teachers, as this article will elucidate, were faced with a new position for which the accompanying behaviors were largely undefined. Similarly, principals, faced with the new CT role, needed to generate new understandings about their own role vis-à-vis teacher evaluation. The ambiguity surrounding the authoring of these two roles naturally generated tension in the context of a normative environment not accustomed to teachers evaluating other teachers (Smylie & Denny, 1990). That tension informed the sensemaking process and redistribution of leadership or lack thereof.

Design and Methods

Rosemont Unified School District

The study employed an embedded single-case design (Yin, 1989) of one urban K-12 school district in California, the Rosemont Unified School District.[8] Rosemont has approximately 100 schools and 3,000 teachers, and is ethnically and economically diverse. Its history includes unpleasant relations between teachers and the district, due in part to district personnel policies that relied for many years on teachers to whom the district had not granted permanent contracts. The teachers could be fired at will in the spring and rehired in the fall, leaving them without benefits or job security over the summer. While this saved the district money, and gave principals the flexibility to fire many of their teachers very easily, these teachers eventually earned

contracts and became a teaching force with memories of mistreatment not easily forgotten. Animosity toward the district, and towards site administrators whom they had learned not to trust, ran high. As would be expected in any large district, there were a great variety of attitudes and relationships, and many teachers experienced positive relationships with site administrators. Nonetheless, it would be a fair generalization to say that many Rosemont teachers held unfavorable opinions of administrators. Consulting teachers as a group tended to view principals as well intentioned but over-burdened, and not to be relied upon for instructional leadership. Noted the lead CT: "If all principals were [instructional leaders] like we would like them to be, arguably you might not need PAR. But that's not going to happen." Finally, many administrators were suspicious of the teacher union, by law the district's partner in the PAR program, viewing it as an obstacle to educational improvement. Rosemont's teacher union president, like the teacher union presidents in prior PAR districts, was an advocate of the policy. In his eyes, PAR would serve two purposes: increase accountability for poorly performing teachers, and reduce principals' ability to fire new teachers at will.

Rosemont selected ten consulting teachers for the first year of implementation, who supported 88 beginning teachers and three veteran teachers across 28 schools. Due to both design intentions[9] and fiscal limitations, year one saw only a partial implementation of the policy in 28 schools. CTs had caseloads of nine to ten PTs, with the lead CT carrying half a caseload.[10] This was a reduction from the 12 to 15 PTs specified in the contract, in recognition of the large role CTs would play in program development in year one. Due to the very small number of veteran teachers in the program, this article primarily focuses on CTs' work with new teachers.

Design

The embedded structure of the study allowed me to examine the process of sensemaking occurring for both teachers and administrators, while the single-case design allowed for a fine-grained examination of one situated process of role shift and related sensemaking. I selected Rosemont after an extended pilot study in the district, as well as a pilot across a handful of other districts. The site was selected based on the degree of "interrup-

tion" (Weick, 1995) occurring, and hence the opportunity to witness new rules being written. In addition, the site was chosen because it had a prior experiment with PAR, perhaps increasing the potential for meaningful implementation in the first year of the program. In other words, while all districts needed to have a PAR program in place to continue receiving state mentor funding, many districts intended to limit the program to the minimum required by the new law—namely, something available to those teachers receiving unsatisfactory evaluations from principals, typically a very small number. Because of my interest in the possibility for jurisdictional shift, I sought out a site that was planning a more comprehensive program than that required by the law. Because of Rosemont's prior experience with PAR, key figures in the district saw the state legislation and attached funding as an opportunity to do what they previously could not afford. In addition, examining the initial development and implementation of the program matched this study's goal of witnessing the process of sensemaking. Finally, I was fortunate to be granted wide access in Rosemont.

The study was designed over a year and a half[11] using a role complement sample (Little, 2000), which allowed for a focus on the consulting teachers (CTs), while also looking across levels of the system based on which other actors were primarily connected to the CTs. The sample included the district's nine PAR panel members (teachers and administrators) and ten PAR CTs. In addition, three of the ten CTs were chosen for more in-depth data collection. This choice was influenced by their demography (years of experience, gender, and ethnicity) and degree of engagement in making sense of the reform. Mentees and principals were then included in the study based on their connection to the case study CTs. However, I also sought out interviews with principals and mentees not linked to the case study CTs who might represent divergent or unrepresented viewpoints (Miles & Huberman, 1994).[12]

The study's sample focused on the educators involved with PAR, as they were the ones engaged in making sense of the policy. In year one of PAR in Rosemont, because implementation was only partial, knowledge of the program was minimal beyond those directly involved. The purpose of the study was not to gauge the spread of knowledge about the policy, but rather to ex-

amine the ways those responsible for initially enacting the policy did so.

A note on the principals involved in the research is warranted. One major way the district chose where to place limited PAR resources in the first year of implementation was principals' interest in the program; PAR was placed in schools where the principal had signed up for it. The field of potential principal informants was narrowed further by the self-selection of those who were willing to participate in the research.

Methods

The study relied on observations, interviews, and surveys. Data collection began with the inception of the PAR panel in the spring prior to the year of implementation. It included the selection of the CTs by the panel, a two-day summer professional development retreat for panel members and CTs that brought them together as a group for the first time, and another week of CT meetings in August just prior to the start of school. In this way I was able to document the environment of messages about PAR into which the actors were entering and beginning their new roles. I then attended all panel meetings and hearings (approximately monthly) and almost every CT meeting (weekly) for one year, which provided rich opportunities to view sensemaking in action. Panel meetings were usually two hours in length, while panel hearings and CT meetings typically lasted a full day. These meetings were tape recorded and scripted. In total I observed approximately 311 hours of meetings.

In addition to observing meetings, I conducted semistructured interviews with panel members and CTs in the fall and spring, and interviewed the three case study CTs in the winter. I interviewed 11 principals and 15 PTs, as well as a few key district level informants such as the Superintendent. Interviews lasted between thirty minutes and three hours; all but three were tape recorded and transcribed.

A multiwave survey approach complemented the continuous fieldwork (Miles & Huberman, 1994) by (a) drawing attention to themes to look for in the fieldwork; (b) serving as a method of triangulation for findings in the fieldwork; and (c) broadening the sample of principals and mentees beyond those interviewed. All panel members and CTs completed a survey. Sixteen out of 28 principals returned surveys (57%); to-gether with interview data, 20 principals were included in the study (77%). Fifty-seven out of 91 PTs returned surveys (63%); together with interview data, 61 mentees were included in the study (67%) (for a breakdown of observation hours, as well as interview and survey data collection, see Appendix).

I interwove data collection and analysis from the outset of the study. Ongoing analysis with a variety of tools informed the data collection process: summaries of fieldnotes; analytic memoing; and coding (descriptive and interpretive early on, moving towards patterns later in the study) (Miles & Huberman, 1994). Earlier stages of the process emphasized exploration, moving towards theory development and confirmation as the study progressed. I used the qualitative software QSR NUD*IST 4 for data management.[13]

Making Sense of PAR
in an Ambiguous Policy Context

Introducing PAR in Rosemont required panel members, consulting teachers, principals, and participating teachers to begin to make sense of it. Spillane et al. (2002) argued that sensemaking is the interaction of three elements—prior knowledge and experience, policy signals, and social situation, including beliefs about those experiences, signals, and situations. They note: "The fundamental nature of cognition is that new information is always interpreted in light of what is already understood and believed" (Spillane et al., 2002, p. 394). As argued in the conceptual framework, most educators' prior knowledge about teacher evaluation is that it is the purview of principals. Critical to understanding the interpretation of PAR, however, are whether Rosemont's educators *supported* or *believed in* the traditional method of principal-led teacher evaluation, what policy signals they received about the new alternative to that method, and how the alternative supported or challenged the existing social situation. Using Spillane et al.'s (2002) framework as a guide, the article now turns to the first research question—"how do educators within a school district make sense of the new role of teacher as evaluator of other teachers that comes with PAR?" This section explores beliefs about teachers conducting evaluations and the policy signals given about PAR. Social situation is addressed later in the article.

Program Ambiguity

A certain degree of program ambiguity is to be expected with the implementation of most new policies (March, 1988). While this ambiguity is not novel, it is highly relevant to understanding how individuals make sense of—and subsequently make decisions about—their roles and relationships within new programs, and is therefore worth exploring here.

While there was widespread confusion about an array of PAR program details, the most significant for our purposes involved the summative evaluations of teachers in PAR—who would be conducting them and how? Only a few key players, notably the union president and district co-chair, designed the foundation of Rosemont's PAR program. They selected the other panel members, who had varying degrees of prior knowledge about PAR. The panel members as a group then hired the CTs. This contributed to a tiered array of understandings about what the program would look like.

In addition, the contract language was left sufficiently vague to leave the control of evaluation unclear. A distinction was made between "review" and "evaluation," such that the PAR program was responsible for the review of classroom performance for teachers in the program, while the principal was responsible for the evaluation of performance outside the classroom for the same group of teachers. However, even within the definition of "review," the contract required the PAR Panel to "examine documented interactions between the teacher, Consulting Teacher, and *principal*," and discuss the recommendations with the Consulting Teacher and *principal*" (emphasis added). Hence, the principal was also a participant in the review process.[14]

Finally, while the contract may have provided a jumping-off point for the initial implementation, it was only one of many messages from the environment about the policy's meaning (Coburn, 2001). In addition to the contract language, panel members interpreted the program for CTs, CTs interpreted the program for principals and PTs, and CTs interpreted the program for one another every week in their all day meeting. Principals received policy signals from the district panel co-chair, an Associate Superintendent who sat on the panel, and the superintendent. Significant among the early messages was the repeated directive to CTs from the panel to "be diplomats"

and "sell" the program to principals. The district co-chair of the program also suggested, as the program began in August, that its continued survival was in jeopardy because the support of the Superintendent was unclear.

A change in Superintendent had occurred, such that the Superintendent who had approved the PAR contract was no longer superintendent when the program began. The new Superintendent, CTs and principals were told, wanted principals to be "instructional leaders." As he commented in his interview, "It's a program I want to keep, although there are some things about it I would want us to revisit and to talk about, and that would be how to include principals in this process in a way that they're really involved in an ongoing way in the teaching and learning and supervision process . . . Great program if it included this very important part for me." He goes on to express his commitment to principals as the ultimate instructional leader of their buildings, and impatience for the "principal overwhelm" complaint:

> I believe that the final person to sign off on an evaluation needs to be the principal. And, if the principal is outside of that process, how can they sign off on something that they haven't been involved in? . . . I always say to principals who say to me [that they don't have time to do all their evaluations well], I just sort of raise my eyebrows and think, hmm, what's that about? It is your job.

Many principals who signed up for PAR in fact either did not understand its evaluative aspect, confusing it with other support programs, or did not understand where the locus of control for evaluation would be, confusing it with the standards-based teacher evaluation system (STES) for principals that was being piloted by the district at the same time. As one CT summarized:

> [Principals] had no idea whatsoever [how the program was meant to operate]. They had no clue. Any clue that I gave them was along the way. They were always surprised and thrilled and like, wow, you're doing three formal observations, you have so much work to do! But they had absolutely no idea what it was about. They were confused about the difference between PAR and STES. They didn't understand the component of evaluation that I was responsible for. One would turn in an evaluation on the PT [in addition to the one] I did. I think a lot of them got the mentoring part, but they were

surprised to see me so much, like oh, you're here again!?

As a result of the ambiguity surrounding responsibility for teacher evaluation with PAR, there were two potentially (though not inherently) conflicting messages about PAR evaluation: one that handed responsibility for the evaluation of classroom teaching to the CTs, and another that told CTs to work closely with principals who were supposed to be the instructional leaders of school buildings.

Differing PAR Orientation

Within this context of varied messages about PAR in the environment was the reality that not everyone involved supported the idea of a teacher evaluating other teachers. Table 1 displays responses to two survey questions addressing respondents' orientation to the PAR program. Respondents are in three groups: administrators (principals and administrators on the panel),

teacher leaders (CTs and teachers on the panel), and PTs. Item 1 asked respondents to indicate who should be responsible for assessing whether teaching standards are being met, while Item 2 asked who should be responsible for removing teachers not meeting those standards from the classroom. There was agreement across all respondents that it is site administrators' role to assess standards, although there were significant differences between all three groups of respondents, with site administrators agreeing with the statement most strongly. Yet there was similar agreement across respondents that assessing standards is also the role of teachers, with no significant between-group differences. However, the groups reported uncertainty in Item 2 about whether teachers should be responsible for removing other teachers from the classroom when standards are not met, with no significant between-group differences. Instead, respondents agreed with the statement that it is site administrators' role to enforce the standards that in Item 1 it

TABLE 1
Attitudes about Responsibility for Teacher Assessment and Removal, ANOVA by Job Type

PAR Orientation Item[i]	Full Sample	a: Adm.[ii] (*n* = 18)	b: TLs[iii] (*n* = 15)	c: PTs (*n* = 57)	Group *F* Value
1. Whose role should it be to assess whether teaching standards are being met?					
District administrators	3.15	3.38	3.73	2.91	$F(2,81) = 2.57$
Site administrators	4.33	4.94[bc]	4.53[ac]	4.07[ab]	$F(2,88) = 10.59$***
Teachers	4.50	4.38	4.60	4.51	$F(2,87) = 0.81$
Teachers' union	3.51	3.18[b]	4.29[ab]	3.40[b]	$F(2,83) = 4.75$*
Community	3.19	2.78	3.00	3.38	$F(2,85) = 2.07$
Universities	2.82	2.38	2.79	2.96	$F(2,82) = 1.49$
The state	2.72	2.53	3.21	2.64	$F(2,81) = 1.48$
2. Whose role should it be to remove those teachers not meeting standards from the classroom?					
District administrators	3.90	4.73[c]	4.29	3.57[a]	$F(2,83) = 6.95$**
Site administrators	4.40	4.72[c]	4.54	4.26[a]	$F(2,88) = 2.19$
Teachers	3.61	3.53	4.08	3.52	$F(2,82) = 2.55$
Teachers' union	3.89	4.31[c]	4.57[c]	3.59[ab]	$F(2,84) = 7.54$***
Community	2.90	2.56	2.69	3.07	$F(2,83) = 1.12$
The state	2.67	2.63	3.00	2.60	$F(2,81) = 0.45$

Note. *$p \leq .05$, **$p \leq .01$, ***$p \leq .001$. Bonferroni multiple comparisons were used to test all pairs of group differences. If mean values for one column show a superscript for another column, those two groups are different at the $p \leq .05$ familywise error rate.
[i]Questions were asked on a Likert scale where 1 = *Strongly disagree*, 2 = *Disagree*, 3 = *Uncertain*, 4 = *Agree*, and 5 = *Strongly agree*.
[ii]Adm. = Administrators (principals and PAR panel administrators).
[iii]TLs = Teacher Leaders (consulting teachers and PAR panel teachers).

had been alright for teachers to participate in assessing. Administrators, in fact, reported more agreement for site administrator responsibility than teachers on both items, suggesting that despite some administrator acceptance of teacher involvement in these leadership functions, administrators nonetheless still believed these functions to be more appropriately the domain of principals. Recall from the study design that PAR in Rosemont was only in schools where the principal had signed on; as a sample, the principals responding to this survey are likely those most positively inclined towards the program. Administrators' attachment to administrator responsibility for enforcing standards therefore seems particularly salient, as it could likely be stronger for the universe of possible respondents.

The survey also asked respondents to identify their ideal balance of teacher/principal involvement in and responsibility for teacher evaluation from a choice of four statements

(see Table 2 below). Respondents were then asked to talk about the choices they made on their surveys in follow-up interviews. Table 2 shows the responses to these four survey items. The results show the variation in beliefs among the core group of policy implementers. For Item 3, "expert teachers should only be involved in the formative assessment of other teachers, not their summative personnel evaluations," teacher leaders reported significantly less agreement than both PTs and administrators, suggesting that as a group they were less willing to limit teachers' roles in teacher evaluation to formative assessment. This is unsurprising—yet confirmatory—given the self-selection of this group into participation in PAR. In Item 5, "expert teachers should have primary responsibility for the summative personnel evaluations of teachers in PAR, with limited principal involvement," administrators reported significantly less agreement than teacher leaders and PTs. However, there was universal disagreement

TABLE 2

Attitudes about Principal and Teacher Leader Responsibility for Evaluation of Teachers in PAR, ANOVA by Job Type

PAR Orientation Item[i]	Full Sample	a: Adm.[ii] (n = 18)	b: TLs[iii] (n = 15)	c: PTs (n = 57)	Group F Value
3. Expert teachers should only be involved in the formative assessment of other teachers, not their summative personnel evaluations.	2.80	3.06[b]	1.87[ac]	2.95[b]	$F(2,90) = 4.89**$
4. Expert teachers should participate in the summative personnel evaluations of teachers in PAR, in collaboration with principals.	3.97	3.81	4.27	3.95	$F(2,90) = 0.90$
5. Expert teachers should have primary responsibility for the summative personnel evaluations of teachers in PAR, with limited principal involvement.	3.02	2.06[bc]	3.47[a]	3.21[a]	$F(2,90) = 7.55***$
6. Expert teachers should have sole responsibility for the summative personnel evaluations of teachers in PAR, with no principal involvement.	1.96	1.42[c]	1.57	2.24[a]	$F(2,90) = 5.96**$

Note. $*p ≤ .05 **p ≤ .01, ***p ≤ .001$. Bonferroni multiple comparisons were used to test all pairs of group differences. If mean values for one column show a superscript for another column, those two groups are different at the $p ≤ .05$ familywise error rate.
[i]Questions were asked on a Likert scale where 1 = *Strongly disagree*, 2 = *Disagree*, 3 = *Uncertain*, 4 = *Agree*, and 5 = *Strongly agree*.
[ii]Adm. = administrators (principals and PAR panel administrators).
[iii]TLs = teacher leaders (consulting teachers and PAR panel teachers).

for Item 6, "expert teachers should have sole responsibility for the summative personnel evaluations of teachers in PAR, with no principal involvement," although PTs—the group least likely to be established in the norms of the district—nonetheless reported significantly less disagreement than administrators.

At the same time, despite the variation in orientation between stakeholder groups to the PAR model, the data show a convergence toward the idea of teacher leaders and principals collaborating on evaluations. Responses to Item 4, "expert teachers should participate in the summative personnel evaluations of teachers in PAR, in collaboration with principals," produced both the highest overall mean for the four items, as well as the highest mean for each group with no significant between groups differences.

Policy Signals: Ambiguity as a Response to Differing PAR Orientation

According to those leading the program, the panel co-chairs and lead CT, the ambiguity of how the responsibility for evaluation would be divided was key to the success of the program. The teacher union president, who preferenced survey item number four above, explained:

> [Survey item number] four is a political statement as much as anything else. "Limited" [in number five] is a charged word. It implies attack upon administrators, upon somehow challenging their qualifications to do the job or something like that. Four, this is the statement that I have to make and I have to keep making and I have to keep making again. Four is the only thing that you can say, because it is vague. It just says collaboration. Then it means it's open, it's embracing.

The district panel co-chair's ideal balance of responsibility stretched across answers four *and* five, saying individual CT-principal pairs would decide for themselves how to enact the policy. Asked by CTs in a meeting at the beginning of the year, "who is ultimately responsible for teacher evaluations?" he responded:

> I appreciate your desire to get that clarity, but I don't know that I can give it up front. A lot of it will depend on the consulting teacher and the principal. For example, the CT must present a formal presentation to the Panel, but how that gets formed is dependent on a trio of people—

the consulting teacher, principal, and PT. That principal's involvement may look very different in [different schools]. What I don't want to do is start with hard and fast rules that we have decided for principals. That's the way to kill the program.

Finally, the lead CT echoed the sentiments of this administrator and the union president, viewing ambiguity itself as a good thing for program success.

One panel member noted in frustration, "There has been no conversation, or at least it seems to me that I haven't been involved in any conversation to ask, in this program, what should be in place? I don't think those conversations in terms of what's the mission statement or the goal or the what are we trying to do [have happened]." Looking back on the year, one CT commented: "It was sort of like you were always feeling you weren't getting a straight answer. I don't think it was that [the panel was] trying to be evasive, I think it was that things hadn't been thought out sometimes, and sometimes they might have been thought out but the situation was ambiguous. But it got frustrating. It was sort of like just tell me, you know, I just want you to tell me this is the way it is."

What the three program leaders articulated exemplifies Baier, March, and Saetren's (1988) argument that intentionally increasing the ambiguity of a policy is a common method for gaining support for it across stakeholders with differing goals and beliefs. As a result, "official policy is likely to be vague, contradictory, or adopted without generally shared expectations about its meaning or implementation" (Baier et al., 1988, p. 59). Noted one teacher on the panel, "Collaboration. Well, what's the difference between collaborating with them and limited principal involvement? Collaborating—meaning we sit down together and write [the evaluation]? We sit together and go over the evidence? There could be some strings to that. Limited? Limited is the way it is currently, although you could make the argument that it could be collaborative the way it is currently." Rather than view this ambiguity as a problem of implementation, we can understand it as a quite pragmatic approach to gaining the support necessary for the policy to exist, given the disparate orientations toward teacher responsibility for teacher evaluation. The dilemma is that while a policy may then *exist*, the ambiguous context

allows for infinite interpretation and enactment possibilities.

Who is Responsible for Teacher Evaluation?

In response to the ambiguous context of PAR, Rosemont's stakeholders interpreted responsibility for teacher evaluation in disparate ways. Table 3 shows the variety of responses by different stakeholders to the interview question: "who is ultimately responsible for the summative evaluation of a PT in PAR?" The disparity of panel member responses to this question perhaps highlights more than any other piece of data the program's lack of clarity, given the program's design and the fact that the panel made ultimate recommendations to the Superintendent regarding continued PT employment.

In addition, the CTs looked to the panel as the policy making arm of the program, as defined by Rosemont's contract language. Panel members' lack of a unified opinion about responsibility for evaluation, therefore, made clarity among other stakeholders difficult. One panel teacher expressed this sentiment:

> That part [evaluation] I'm not so sure about. I guess because we're in such a transitional phase, is it still the principal? Is it the principal and the new system under PAR with the consulting teachers? Is it the consulting teacher? Is it the panel who actually listens to the cases and makes the decisions or ultimately makes a ruling one way or the other? So I think that part is unclear, and this is the part that is somewhat frustrating to me. I'm a panel member and I'm unclear. I have no idea really. I don't know what the arrangement has been.

Most surprising was that two panel members—one teacher and one administrator—identified the principal as still ultimately responsible for the evaluations, demonstrating the strength of established practices and beliefs. Conversely, no principals identified themselves as the ultimate evaluator. Also interesting is that while two out of three panel administrators identified the CT as the ultimate evaluator, and none identified the panel, the same number of panel teachers identified the panel as the ultimate evaluator, and none identified the CT. Panel teachers expressed concern about too much weight falling on CTs.

Finally, PTs—who appeared to receive rather clear policy signals from CTs—generally identified their CT as their evaluator, both in interviews as shown in Table 3 and on the survey. One PT in interviews, and 15 on the survey, demonstrated the perception of a joint effort between CTs and principals for evaluations. Nineteen out of 57 PTs chose responses on the survey that included the panel—although all PT responses were combinations of evaluators that also included the CT. No PT chose "Principal and Panel" as responsible for his or her evaluation.

How Did Rosemont's Educators Make Sense of PAR?

Recall that sensemaking is comprised of three constructs: prior knowledge, policy signals, and social situation, and beliefs about all three of these (Spillane et al., 2002). While those needing to make sense of PAR mostly began with the same prior knowledge about teacher evaluation models, namely that evaluation is the purview of princi-

TABLE 3
Interview Responses to the Question, "Who is Ultimately Responsible for the Summative Evaluation of a PT in PAR?" *

Respondent	Response						
	CT	Principal	Panel	CT/principal	CT/principal/panel	Uncertain	Total
Consulting teacher	2		2		1	2	7
Principal	5		1		1	3	10
Panel administrator	2	1					3
Panel teacher		1	2			2	5
Participating teacher	11			1		1	13
Total	20	2	5	1	2	8	38

*Not all study participants are displayed. Due to some initial interviews occurring prior to the question being added to the protocol, and the semi-structured interview format, not all respondents were asked this question.

pals, the data presented in this section have shown that: (a) stakeholders differed in their belief systems about the appropriate division of responsibility for evaluation between principals and teachers, with administrators less positive about teacher authority for evaluation than teachers, and all groups most positive about the idea of collaboration between teachers and principals; (b) stakeholders received different policy signals about PAR that left responsibility for teacher evaluation unclear, due in large part to program ambiguity; and (c) these varying beliefs and policy signals resulted in disparate interpretations of the import of the policy, most notably that even those charged with running the program held varied opinions about ultimate responsibility for teacher evaluation.

With the varied interpretations engendered by the sensemaking process in the background, the article now turns to the second research question by examining PAR as enacted: "how, if at all, is leadership responsibility for teacher evaluation redistributed as a result of the new role?"

Enacting PAR in an Ambiguous Policy Context

I posited in the conceptual framework that ambiguous spaces, like that in Rosemont described thus far, provide the fertile ground for the possibility of jurisdictional shift (March, 1994; Weick, 1995). In fact, despite both the variety of beliefs about giving evaluative responsibility to teachers, and the stated preference for collaboration, when asked *what* that collaboration looked like in practice, the descriptions uniformly assigned more involvement to the CT. For example, one CT commented, "Now, I didn't go to [principals] and say, 'Oh, let's do this summative together.' [But I gave it to them and said] 'Here's their summative. Look it over and I want to make sure that we're on the same page.' So we talked about it. That's what I consider having collaboration." Similarly, a principal defined collaboration this way:

> The CT should have the prime responsibility, but the principal still needs to collaborate. The writing should be limited principal's writing. Collaborating would mean reviewing the [written evaluations] with the CT. [The principal] doing at least one observation, not a formal one, but at least one full informal one with note-taking [per year]. It would still be a lot less than the CT involvement with that PT.

Other responses included such language as a CT using the principal as "another pair of eyes," or the CT being "in the classroom much more than the principal."

After an initial period of building trust, most principals were content to let CTs work independently without too much communication between the two. Several CTs noted how principals wanted to talk to them more at the beginning of the year, but once they had essentially proven themselves, principals were happy to let them go about their business. Noted one CT, "Once they had confidence in me it was like 'Hi,' small talk, 'Bye.' [I'd ask,] 'Do you want to meet about this stuff or do you just want me to put it in your box?' [They'd respond,] 'Just put it in my box. You're doing a great job.' " The exception was cases of PTs who were not meeting standards, about whom there tended to be a fair amount of communication between CTs and principals.

One principal demonstrated the role of district paperwork in demarcating jurisdiction. She commented, "[CTs] turned in all the paperwork and gave most of the information to the panel. The principals could add things if they wanted to. But the consulting teachers are going to be turning in the paperwork this year, not me. The district sends us scanned paperwork we have to do for the other teachers, but I didn't get any for [those teachers in PAR]." Halverson, Brown, and Zoltners (2001) note that artifacts such as this evaluation paperwork are "externalized representations of ideas and intentions used by practitioners in their practice." The lack of evaluation paperwork given to principals displayed the idea and intention that they were not conducting evaluations on PAR PTs. Indeed, here a principal made sense of PAR and her role in it by noting simply that she did not have evaluation paperwork for those PTs that were in PAR. Another principal, making the shape of a box with his fingers, highlighted with chagrin the "tombstone-like" space given to him for comments on the official PAR evaluation paperwork. Again, the artifact suggested to this principal his minimal role in PTs' evaluations.

One could argue that these characterizations of the enactment of PAR demonstrate a *limiting* of CT agency when compared to the traditional Toledo model of PAR that gives sole responsibility for evaluation to the CT. Nonetheless, these characterizations all give primary agency to the CT in the evaluation process, not the principal.

Panel Hearings and Employment Decisions

PAR panel hearings provided perhaps the most informative lens through which to view the enactment of PAR policy, and the question of whether or not any redistribution of responsibility for teacher evaluation took place in Rosemont. At the PAR panel hearings, held in November, January, and April, CTs presented reports about their PTs to the panel and usually a few other district level administrators in attendance. The reports included how the PTs were performing and what the CT had done to support them. The panel in turn made suggestions regarding support. At the spring hearing, and in some cases even earlier, CTs presented their renewal or nonrenewal recommendations and the panel voted on a decision. In the fall, the panel felt that principals should attend the first hearing. They believed that if principals saw PAR in action, they would like what they saw and support the program; at that time, the purpose of principal attendance at the hearings was principals' *gathering of knowledge* about PAR. By the second hearing, however, evidence of a drift toward "collaboration," it was expected that principals would attend hearings; indeed, for the remainder of the year, principals who did not attend were seen as unsupportive of PAR and not upholding their share of the bargain (ultimately only one of 28 principals never attended). Panel members came to view principal *input* about PTs as necessary, without acknowledging this shift.

The hearings therefore provided rich data on the distribution of leadership because the CT and principal literally sat side by side at a table, presenting to the panel. The dynamics of these presentations were, not surprisingly, quite different from one CT-principal pair to the next. Typically, the CT did most of the presenting, with the panel then asking the principal what he or she had to add. But many principals offered information throughout the presentations, and in a few cases the pair truly seemed to work as colleagues, finishing one another's sentences. In most cases, the CT and principal were in agreement that a PT was either meeting standards or in need of dismissal. It was those cases where the CT and principal were not in agreement or the outcome was unclear, however, that provide fodder for analysis and demonstrate what was ultimately the authority of CTs' voices. Two such examples highlight the emergent CT jurisdiction for teacher evaluation.[15]

In one, a PT named Timothy was not renewed for employment, though by all accounts he was doing a passing job by what had been Rosemont's traditional teacher quality standards. While everyone involved agreed he was smart and creative, his philosophy of education and his interests outside of teaching led them to the impression of a "winging it" approach, and his reluctance to change made them skeptical of his likelihood for improvement. Because no flagrant crimes or otherwise egregious errors were being made, the consulting teacher (Caroline), principal, and panel members all described him as someone who would have "slipped through" without PAR and "slid by" for the duration of his teaching career. For that reason, the union president described Timothy's nonrenewal as "historic." The principal—who believed the right decision had been made—indicated that she likely would not have acted on her belief that Timothy was below standard had she been solely responsible for the evaluation, without PAR. Rather, she had been persuaded by Caroline, who clearly felt the weight of the decision most heavily. Caroline agonized over making the decision, but felt confident that she had provided Timothy with extensive support and opportunities to improve.[16]

What the data from this example show—directly and indirectly through the use of language—is the degree to which those involved viewed Caroline as responsible for the ultimate decision, and therefore the gatekeeper for teacher quality. Although the panel and principal clearly supported Caroline in the decision, and the panel retained in name decision-making power with respect to the recommendation to the Superintendent, the language used by the various parties belied the degree of agency Caroline actually had in determining the outcome of the story. She told the panel, "I finally made the decision to nonrenew," while the principal reported to the panel that Caroline "has already made the decision." Caroline described the panel hearing as a "rubber stamp" on her decision, expressing the desire that the panel be "a little tougher on us."

In the case of a second challenging PT, the principal hired an uncredentialed teacher named Kim one week prior to the start of school, but quickly concluded that Kim was not meeting standards. While Eva, the CT, was initially skeptical of Kim's chances for success, she was persuaded by the progress Kim was able to make, and defended

Kim's renewed employment in the district. Eva became a buffer or mediator, communicating the principal's concerns to Kim and providing Kim with focused support to address those concerns. Part of doing this involved translating the principal's broad concerns into concrete specifics on which Kim might improve. Ultimately, Eva diffused the principal's criticism of Kim at the panel hearing by demonstrating Kim's growth on the teaching standards.[17] The principal's complaints about Kim seemed vague and unsupported by comparison. Kim was renewed for employment in the district and placed at another school.

PAR put these two CTs in the new role of teacher as evaluator of other teachers. For Caroline, this meant making the decision to fire a teacher, and for Eva, it meant in essence defending a teacher from her principal. Caroline led the principal to the nonrenewal decision they made "together," while Eva spoke out against the principal's opinion. While both principals still had a voice in the evaluations, the voices of the consulting teachers appeared to be preferenced in both examples. In the first, the consulting teacher voice was preferenced by the principal herself, as the principal was persuaded by Caroline's assessment. In the second, the consulting teacher voice was preferenced by the panel, which gave Eva's assessment more weight than the principal's. The preferencing of these CTs' voices over those of the principals suggests a shift in jurisdictional control, and hence leadership responsibility, for teacher evaluation.

The consulting teachers were vested with this authority for PTs' renewal recommendations in large part because of the perceived quality of the evaluations they had conducted. For Caroline, this authority rested on the amount of support she had given, so that Timothy's lack of growth had meaning (Goldstein, 2003b). It was argued earlier that principals are not able to give thorough support since their jobs have so many other requirements (Copland, 2001; Grubb et al., 2003). For Eva, her authority rested on the skill with standards-based evaluation that the panel perceived her to have, which allowed her to concretely demonstrate growth on a performance rubric (Goldstein, 2003b). Principals in Rosemont had not for the most part been trained in standards-based evaluation and instead were still evaluating teachers with the "forty-five minute observation and a checklist" model.

As the various actors made sense of PAR, they were in the act of creating new roles for themselves and rules for teacher evaluation, defining appropriate behavioral enactments to attach to their new position labels. In doing so, they challenged assumptions about teacher evaluation and authority relations in education. With Caroline, the data show a consulting teacher collaborating as a peer with an administrator, and ultimately making the decision that has historically been the administrator's to make. Notably, the decision was to fire a mediocre teacher; in this example, CT jurisdiction challenged assumptions about acceptable teacher performance. Indeed, 11 of the 88 new teacher PTs were fired, and all three veterans left the classroom (Goldstein, 2003a). With Eva, the data show a principal unable to solely decide the fate of an emergency credentialed teacher, as she is forced to contend with a second and perhaps more powerful voice than her own. In this example, CT jurisdiction subtly challenged the assumed authority of principals to determine a new teacher's employment outcome.

Negotiating New Roles
for Teachers and Principals

Thus far the article has explored how stakeholders made sense of an ambiguous policy context, and presented examples of how the policy was enacted. This final data section looks at how PAR, as enacted, challenged stakeholders' existing authority relations and teacher evaluation practices—the social situation into which PAR was placed. How did stakeholders respond to the jurisdictional shift just described? Stated differently, how did they make sense of the enactment of PAR? As such, this section cuts across the first two research questions 1 and 2.

The difficulty of naming teacher leadership

While the data show CTs holding significant authority for teacher evaluation, CTs largely avoided *attributing* jurisdiction for teacher evaluation to themselves. For example, one CT noted in October, "I would hope that I'm the eyes of everybody that's making the decision . . . I don't want to say it's me [evaluating], I don't want to say it's the principal, I don't want to say it's the Panel. I think it's a collaborative decision." By February, however, her language attached a sense of authority for evaluation to the role she played as eyes for the group, suggesting a growing comfort with the

role of evaluator: "I'm still clarifying [to principals] that I'm the one that's going to do the evaluation. You're going to have input. We bring in both [to the panel]. They're still not clear. I still have principals saying, 'Now who do I have to evaluate? Do I have to go in there and observe?' I say, 'No, I'll do that. I'll look, then we'll sit down and look at the information together.' " Despite articulating this sense of authority, she deferred ultimately to the panel: "I really feel that the evaluation, the part that says you stay . . . or you go, really needs to be the panel." Another CT, looking back on the year, said she had collaborated with principals yet gave this description: "Principals gave me very limited feedback, but they gave me an impression of the person. And they were a sounding board for what I said about the person." Caroline, the CT who made the decision to non-renew Timothy, explained why the panel was ultimately responsible for his nonrenewal:

> Clearly the panel was responsible for the final evaluation, because . . . well, that's that whole ambiguity. What does the word evaluate mean? I mean I'm evaluating all the pieces and I'm making a recommendation, but I am not really making the evaluation. It really is the panel that's doing it. And that's what I do tell [my PTs]. I say, "I make a recommendation. They make the final decision." I don't use the word evaluation.

The response signaled her discomfort with responsibility for Timothy's employment decision.

While the enactment of PAR largely demonstrated more responsibility for evaluation by CTs than principals, naming this as limiting principal power proved too radical a response, both on the survey and in interviews. Literature suggests that, due to the norms against it, educators do not often recognize leadership by teachers even where it exists (Bascia, 1998; Johnson, 1984; Wasley, 1991). The favored term "collaboration" seemed to be used as a euphemism for the reality of limited principal involvement, a more acceptable way for teachers to participate in evaluation. By and large respondents felt that relationships between CTs and principals had been collaborative—although both CTs and principals wanted it to be more so—and were slow to acknowledge that CTs had made evaluative decisions. Different CTs defined their role vis-à-vis evaluation differently, but most seemed happy to name the "final decision" as someone else's responsibility.

Education's norm of "being nice"

Given all that is known about principal reticence to give negative evaluations, it should not be surprising that the CTs—especially in their first year in the role—would be slow to embrace the title of evaluator. Most CTs defined their role in interviews as one of supporter of new teachers, and tended to mention evaluation almost as an aside or afterthought. During the hiring process, the panel asked prospective consulting teachers whether they were prepared to make a decision that a participating teacher be dismissed; all those hired said yes, although some laughed later in the year over the ease with which they had given that response. Responsibility for evaluation clearly comes with a price. Caroline described losing sleep over the decision to not renew Timothy, and having bad dreams about it once the decision was made: "I felt like I was committing violence in a way, even though I tried to keep in mind that I'm doing this for the sake of kids and I felt in my own mind that I had weighed it very carefully." While providing support is often perceived as nurturing, evaluation—typically defined as separate from support—is often conceived to mean rejecting the female norm of "being nice" held by many teachers (Murray, 1998; Moir, 1999). One CT poignantly described her own process in recommending renewal for a teacher she actually deemed not meeting standards. She highlighted the mediating role played by the principal (principal #1), compared to what might have occurred had the same PT been at a different school (with principal #2):

> If [the PT] was at [a different school with principal #2], he wouldn't have made it through. So that's how the collaboration [between the CT and principal] kind of infests itself. Principal #1 and I both work with compassion. Maybe this person needs more time, blah, blah, blah. Principal #2, you're either making it or you're not and she has extremely high standards of her teachers. If [the PT] was a first-year teacher [at principal #2's school] and that stuff was going on in his classroom, [principal #2] would have said, "he's not coming back here," and I would have been affected by that. Then I would have been like, "I agree, he's out" . . . I needed support in the final opinion. It's like somebody needed to shake me up a little bit to get to [a nonrenewal]. I needed to get the principal in the classroom a lot more. I need them to document evidence so that it is not just [me]. Why

188

do I need that? I don't know. I think I just do right now.

This particularly reflective CT acknowledged here that as she constructed her CT role, she made sense of the jurisdiction granted to her by modifying it and redefining it as stretched over the CT-principal unit.

The positive response from principals

Given the currently overwhelming nature of principals' responsibilities, most principals viewed PAR as a welcome relief from a small portion of their administrative load. One CT described principals' reactions to her as, "Thank God someone is doing the evaluation here. It's one less thing I have to do. And, you know, this is the way it should be." One principal, asked whether he wanted to continue with the program, said, "Yes, I'd take [the CT] because she's helping me [*laughter*], because she's doing it. I mean, she's doing it!" Principals were very positive in their year-end program evaluations about CTs and the support they received from PAR. While they may have had mixed feelings about their decreased involvement in teacher evaluation, principals were overwhelmingly positive about PAR after seeing it in action. All principals hoped to have PAR available to them the following year. For the most part, principals' need to reduce their workload was greater than concerns about a reduction in a piece of their authority.

Given that administrators were initially less favorable than CTs to the idea of teacher authority for teacher evaluation (see Tables 1 and 2), it is perhaps surprising that principals were more willing than CTs to name CTs as responsible for evaluation (Table 3). In general, Rosemont's stakeholders believed that the quality of the evaluations with PAR were quite strong. Specifically, all stakeholder groups reported improvements for teacher evaluation and accountability (Goldstein, 2003a). It may well be that action leads to cognition (Weick, 1995), and principals were impressed by the CTs and warmed to the PAR model.

The call for instructional leadership

While principals seemed willing to grant jurisdictional control of evaluation to CTs, they still experienced tension around the reduction of their own role. While quick to offer that they do not have the time to conduct evaluations well, some

principals remained uncomfortable with evaluations being conducted by someone else. Principals were conflicted because they wanted to be, and the superintendent expected them to be, instructional leaders. Yet they recognized that CTs were doing a better job than they could. One noted:

> [PAR is] helpful, it gets a job done and a job done well, no question about it. But to me it's a little sad. Someone is coming along and doing [teacher evaluations] and that's great because the job needs to be done, but I always see that as my job, a principal's job. It's a concession to reality, really. Essentially I have all these ideas about being a principal, getting things done, but that's just not the reality. So this is filling a role, but I don't think that's the way it should be. It's this administrator's compromise.

One principal—the one who never made it to a panel meeting—dealt with this tension by continuing to conduct evaluations. He explained:

> What I've chosen to do here is to do informal evaluations, because technically you can't evaluate somebody twice and I don't want to interfere with the relationship [between the PT and CT]. But I also feel an obligation to get into the classrooms and to see people and so what I've done is set up a series of four evaluations, informal evaluations, observations, and then I'll do an informal summary at the end. I've told the teachers that they can share with their PAR CT if they want or they can, you know, throw it in the garbage can or whatever they want, but that I feel that I need to have that connection with them. So with that I've been happy letting the PAR CTs do what they want.

This comment may reflect an internal resistance to relinquishing a piece of control. At the same time, this principal's commitment to observing the teachers in his building is both understandable and laudable—highlighting the conflict or "compromise" PAR raised for some site leaders.

The shift toward collaboration

Despite positive sentiments across stakeholders about the program as enacted, a decided shift towards formal collaboration between CTs and principals was seen. For example, the summative evaluation paperwork moved in three years of PAR implementation from including a box for principals to comment on out-of-classroom

performance, to adding a line for principals to officially sign-off, to formally including principal comment and signature alongside that of CTs' for each of the six California Teaching Standards.[18] This shift must be understood in the context of the perceived technical improvement to evaluation resulting from CT jurisdiction that was highlighted above. The fact that stakeholders were positive about the program and believed improvements in evaluation were occurring makes any shift away from CT jurisdiction particularly salient.

In a recent study of three schools instituting distributed leadership in various ways, Sebring, Hallman, and Smylie (2003) found that two of the three schools "emasculated" their programs after two to three years of good progress. None of these schools were engaged in distributing leadership for teacher evaluation, which this article has argued pushes on education's hierarchical norms in a particularly powerful way. This may explain the shorter duration of PAR in Rosemont as originally implemented, compared to these programs.

How was Leadership Responsibility for Teacher Evaluation Redistributed?

This study of PAR demonstrates that teachers can evaluate teachers. In some circumstances and organizational structures, teacher evaluators are accepted and even highly valued by both administrators and other teachers. The data highlight a variety of expanded roles for the CTs: gatekeepers, who chose to recommend nonrenewal of some PTs; buffers, who recommended retaining some PTs despite principals who wanted them dismissed; and colleagues of principals, who collaborated on decision-making. Conducting evaluations was difficult for CTs, and seeing themselves as evaluators also proved difficult. The transition to being one's brother's keeper is not easy (Wasley, 1991; Kerchner and Koppich, 1993; Kerchner et al., 1997). Nonetheless, CTs—in varying degrees of collaboration with principals—conducted summative evaluations and made decisions about the continued employment of other teachers. They reported those decisions to the panel, which was responsible for the "final decision" to be recommended to the school board. However, out of 91 PTs, there was no case of the panel disagreeing with a CT.[19] Yet, despite this reality of teachers evaluating their peers, there was the desire by most involved for more collaboration between CTs and principals, although descriptions of collaboration often looked like CT jurisdiction. Over time, the program shifted towards more codified collaboration.

Implications for Distributing Leadership in Organizations

The remainder of the article turns from data and description to the third research question and the issue of implications. What lessons about distributed leadership might organizations learn from Rosemont's initial experience with PAR? In this section, the data presented above are placed against an existent theoretical model for distributed leadership. The data do not quite fit, and I attempt to build upon the model.

While I have argued thus far that the desire for collaboration among stakeholders was a sign of resistance to CT jurisdiction for evaluation, it is nonetheless understandable that PAR principals and CTs wanted to work together—especially given principal isolation and CTs conducting evaluations for the first time. The new rules (March, 1994) for teacher evaluation being generated in Rosemont were about collaboration: the consulting teachers and principals were expected to collaborate, with a variety of practical reasons provided. The collaborative approach to evaluation suggests a model of distributing leadership that is less about dividing tasks than sharing responsibility for them.

Stretched Over: Two Models of Distributing Leadership

The data from Rosemont suggest two models of conceiving distributed leadership. In one, tasks are divided between actors; CTs held jurisdiction for teacher evaluation with limited involvement by principals, while principals focused their energy on other leadership responsibilities. In the second, tasks are shared among actors; CTs and principals were expected to work together on teacher evaluation, in essence a joint leadership project. The former describes what was mostly seen in Rosemont in the first year of the program, while the latter describes what most stakeholders reported wanting—although recall that their descriptions of "collaboration" actually suggested primary responsibility for evaluation by CTs. Said differently, their descriptions of task sharing looked like task division.

Spillane, Diamond, and Jita (2000) propose a similar distinction. They term the sharing of tasks

"collective leading," where multiple actors coenact leadership tasks. Collective leading depends on reciprocal interdependency, where the practice of one leader requires input from the others. Since task sharing or collaboration was desired but not widely enacted between CTs and principals in Rosemont in year one, the data presented in this article cannot elucidate that schema.

The flip side in Spillane et al.'s framework is "interdependent leading," where various leaders hold responsibility for tasks in a chronology or assembly line. In the case of a veteran teacher in PAR, for example, interdependent leading would be the principal giving an unsatisfactory evaluation, the CT subsequently making a renewal or nonrenewal recommendation, and the panel making the employment decision. The data showed, however, that the panel essentially accepted CTs' recommendations,[20] questioning the degree of interdependence present. New teachers, the bulk of those in the program, enter PAR immediately without an evaluation from the principal to place them there. So in the case of PAR in Rosemont, the division of leadership tasks may have been too loosely coupled (Weick, 1995) to qualify as interdependent. For this reason I am using the term task division.[21]

Factors Enabling Task Division and Task Sharing in Rosemont

Hierarchical norms

As demonstrated above, principals are overwhelmed by their jobs and welcomed PAR as a relief from a piece of their responsibilities. Yet the norms against teachers holding authority for personnel evaluations challenged movement in this direction. The transition from authoritarian to participative leadership is a difficult one for principals (Kerchner and Koppich, 1993), who are expected to be instructional leaders but are nonetheless asked to move over for teacher leadership (Little, 1988); the data showed several principals, despite their support for PAR, conflicted about their disengagement from the process. Central office administrators often oppose the type of organizational changes brought about by teacher leadership and peer review (Kerchner & Koppich, 1993); the Rosemont data showed the superintendent and an assistant superintendent *on the panel* as not fully supportive of PAR's key concept of teacher responsibility for teacher evaluation, instead viewing principals as ultimately responsible. Despite

the pragmatic attraction of lightening principals' responsibilities, the norms against doing so contributed to the shift toward task sharing.

The difficulty of evaluation

The data showed that principals and panel members developed confidence in CTs and their recommendations, suggesting the possibility of a successful task division model. Nonetheless, if firing teachers were easy, principals would do it more often. Instead, they avoid the conflict of negative evaluations by passing teachers around to other schools (Bridges, 1986). CTs, however, recommended nonrenewal of PTs at unprecedented rates, although they were reluctant to be held singularly responsible for the decisions that they had in effect made. Hence Caroline decides to nonrenew Timothy, comments on the degree to which she felt the hearing was a "rubber stamp" situation, but reports in the same interview that the panel was ultimately responsible. This desire not to be the one blamed for a nonrenewal leads away from task division to task sharing.

District leadership

As we saw, Rosemont had a new superintendent in year one of the PAR program. After establishing positive district/union relations and signing PAR into contract with the prior superintendent, the union president now had to start fresh. The new superintendent had a firm conviction that principals needed to be instructional leaders. She accepted PAR because her predecessor had signed it into the contract, but she did so warily. The result was mixed signals about not only the details of the program but the program's prospects for continuation at all. These mixed signals went to CTs as well as principals already worried about their jobs (many of whom were dismissed later that year), and contributed heavily to a shift from task division to task sharing.

Ambiguity

As demonstrated, program ambiguity was an intentional strategy by those spearheading the PAR program in order to secure its support from potentially negative administrators. While increased program clarity and coherence would certainly have pleased many of those involved, a more coherent program might have created a dead program (Baier et al., 1988). The lesson, while not new, is that a shift towards greater pro-

gram coherence can come over time, as support is secured. As the data showed, administrators were very positive about PAR once they had been exposed to it. The dilemma is that there is no guarantee that by the time Rosemont is ready to give PAR greater coherence it will not have already become harmless, or at least more harmless than originally conceived, like so many teacher leadership policies before it. Indeed, the shift from task division to task sharing may foreshadow such a transformation. In the vacuum created by the lack of a unified definition of PAR, people regressed to that which was familiar—namely principal involvement in or control over teacher evaluation. The task sharing model is therefore a potential confirmation of Little's (1990) finding that districts move quickly to blunt the effects of new teacher leadership policies.

The push for collaboration in Rosemont, however understandable, can be viewed in the light of the body of literature that shows little change resulting from teacher leadership policies over time. Institutional theory sheds light on this phenomenon, as "the elements of rationalized formal structure are deeply ingrained in, and reflect, widespread understandings of social reality" (Meyer and Rowan, 1977, p. 533). As this article has emphasized, education's social reality has not previously included space for teachers in the leadership role they assumed in PAR, and the deeply ingrained notions of appropriate roles for teachers was challenged by the policy. Yet PAR's ambiguity made it a weak challenge. The undefined responsibility for teacher evaluation gave power to the attractive idea of sharing responsibility. While collaboration is a legitimate approach to leadership, the term itself is pregnant with ambiguity, and allowed for a drift away from teacher jurisdiction for teacher evaluation. However attractive the shared or collective model may be, institutional theory and prior research on teacher leadership policies suggest that the shared model may be just a stop on the way back to principal jurisdiction for teacher evaluation in Rosemont. This possibility highlights the ongoing challenge to distributing leadership in public education.

Study Limitations and Areas for Future Research

Cognitive shifts such as those required by PAR happen slowly. The study was intentionally de-

signed during the first year of PAR implementation in order to witness the interruption to institutionalized ways of conducting teacher evaluations, and the cognitive dissonance that would accompany that interruption. The study duration of a year and a half—even the two and a half year duration with the follow-up data—does not allow for the longitudinal view that might see more change in the way things are done over time. Specifically, a longer period of data collection might yield more information about the development of new cognitive frames for teachers evaluating teachers. Given the unanswered questions about the shift from task division to task sharing and what the latter model might look like, longitudinal data would certainly add to knowledge about PAR and processes of distributing leadership.

This study intentionally focused on beginning teachers in PAR rather than veterans, not only because beginning teachers made up the lion's share of participating teachers in Rosemont, but for pragmatic reasons. In year one of the program, as CTs were nervous about their new jobs and intervention cases were facing new accountability measures, I made the choice that interviewing and/or observing intervention teachers, had they even agreed to it, was too risky. I relied instead on secondary data sources for information about intervention cases, namely CTs' interview descriptions of their work with veterans, and descriptions of veterans' cases in CT meetings and panel hearings. Future research should certainly attempt to gather data from intervention cases who have important stories to tell.

At the outset of this study, the Rosemont PAR panel requested that the identity of the district remain confidential. I granted this freely, trusting it would make those involved more comfortable with my presence and able to speak more openly in interviews. In many ways, however, there is a loss of contextual information that would make this work more meaningful as a policy study. Many details have been omitted or changed in order to make the setting more generic.

Given Rosemont's approach of placing PAR in schools where the principal signed up for it, and the reality of voluntary participation in research, there was clearly a selection bias from the whole universe of principals in the case district to those who were included in the study. Principal enthusiasm for the program must be viewed

in this context, while principal resistance to the program can be seen as particularly salient. Future research would ideally be conducted in a PAR site or sites that have placed the program in schools in a less biasing manner.

An expanded and more specific list of response options on the survey for ideal CT and principal roles in evaluation (questions three to six) would have generated more meaningful data. Recall that all groups of respondents preferenced CT and principal collaboration. Given the ambiguity and positive connotation of the word "collaboration," and the negative reaction to the alternate "limited" principal involvement, future surveys would reap much more meaningful results by specifying concrete examples of CT and principal involvement (e.g., Little, 1982) and avoiding language characterizing that involvement.

This study did not attend to the particular qualities and qualifications of principals. Recent research suggests that principals' personal qualities play a large role in their ability to accept leadership by teachers (Sebring et al., 2003), and that more expert principals are actually more successful at recognizing expertise in mentors, stepping back, and facilitating the mentoring process from afar (Carver, 2002; Youngs, 2003). Smylie et al. (2002) note that it "is a paradox of teacher leadership that it requires administrative leadership to be effective" (Smylie et al., 2002, p. 182). Future research on PAR and distributed leadership should examine this important factor.

Notes

[1] For example, the September 1995 edition of *The elementary school journal* was devoted to the topic of teacher leadership. In it, the leadership function of teacher evaluation is mentioned in only one of seven articles.

[2] Peer review case studies have been conducted of Toledo (Darling-Hammond, 1984; Gallagher, Lanier, & Kerchner, 1993), Poway (Gallagher, Lanier, & Kerchner, 1993), and Rochester (Koppich & Kerchner, 1999; Murray, 1999; Grant & Murray, 1999).

[3] Districts use a variety of terms. "Consulting Teacher," "PAR Panel," and "Participating Teacher," while common, are certainly not used by all districts employing PAR programs.

[4] Some small districts have even formed consortia to share a PAR program.

[5] All appearing in 1986 and promoting teacher leadership were: *Tomorrow's Teachers* by The Holmes Group; *A Nation Prepared: Teachers for the 21st Century* by the Carnegie Task Force on Teaching as a Profession; *What's Next? More Leverage for Teachers* by the Education Commission of the States; and *Time for Results* by the National Governor's Association.

[6] See Grant and Murray (1999) for a defense of positioning teachers and principals as separate occupations.

[7] Another empirical study conducted simultaneously is Carver (2002), which examined four principal-mentor teacher pairs and their support of new teachers.

[8] Rosemont is a pseudonym.

[9] CTs serve in their positions for three years. Program leaders believed that gradual implementation would tier the CTs and maintain program expertise and memory by preventing the return of all CTs to the classroom in the same year.

[10] The lead CT, as is common with PAR, had a reduced case load of PTs in exchange for handling the administrative needs of the program and serving as mentor to other CTs.

[11] Follow-up data were gathered one year after the study's completion, but addressed issues that are not a focus in the analysis presented here.

[12] I did not interview the three intervention cases due to the sensitivity of their situations and the fragility of the new program.

[13] While such software can sometimes force researchers into analytic schema too early, I did not create my NUD*IST coding schema until all data had been collected. In addition, I created the schema from my own progressive coding patterns, unassisted by the software.

[14] It is worth noting that Lawrence (2000), the former union president of Toledo who initiated peer review policy there 20 years ago, has argued vehemently that principals should not be involved in the peer review process for legal reasons. His argument is that there needs to be one clear evaluator, otherwise there is a possibility for disagreement which can cause a loss to an unsatisfactory teacher in arbitration.

[15] See Goldstein (2002) for the full narrative cases from which these examples are drawn.

[16] CTs worked with PTs an average of an hour per week, often more in challenging cases, and were available to PTs for ongoing support by email and district-provided cell phones.

[17] PAR support and evaluation was aligned to the California Standards for the Teaching Profession. CTs spent significant professional development time becoming well versed in and calibrated on the standards, including supporting their assessments with evidence.

[18] Standards 1 through 5 pertain to in-classroom performance, while only standard 6 pertains to out-of-classroom performance. Recall that the PAR contract language formally gave principals responsibility for out-of-classroom performance, while leaving their role vis-à-vis in-class performance ambiguous.

[19] The panel was most likely to fulfill its checks and balances role by questioning the amount of support provided to the PT or the quality of the evidence

presented. In one case, a CT was asked to gather more evidence and present again a few weeks later.

[20] While this was evident in Rosemont in year one of implementation, more oversight by PAR panels has certainly been seen in established programs (see Gallagher et al., 1993).

[21] Task division might be thought of as "independent leading," but this ignores the presence of other active stakeholders. Traditional teacher evaluation by a principal would be independent leading. The key feature of the CT role in PAR that sets it apart from independent leading is the accountability to the PAR panel. In a future article focused on the role of the panel in PAR's model of teacher evaluation, I will build on this distributed leadership framework and propose a model of accountable independence, where CTs are engaged in independent leading with accountability, a professional model.

Appendix

TABLE A1
Observations—Number of Days (Total Hours)

	Spring 2000	Fall 2000	Winter 2001	Spring 2001	Total
Panel meetings	3 (8.5)	3 (6)	1 (2)	2 (4)	9 (20.5)
Panel hearings		3 (28.5)	2 (19)	3 (21)	8 (68.5)
Consulting teacher meetings		17 (108)	7 (42)	8 (48)	32 (198)
Consulting teacher professional development*		3 (24)			3 (24)
Total	3 (8.5)	26 (166.5)	10 (63)	13 (73)	52 (311)

Note. *Included here is only that professional development which was relevant to sensemaking about the PAR program and CTs' roles in it.

TABLE A2
Interviews Conducted

	Fall 2000	Winter 2001	Spring 2001	Total Number of Interviewees	Total Number of Interviews
Panel member	9		9	9	18
Consulting teacher	10	3	8	10	21
Principals		6	5	11	11
Participating teacher			15	15	15
Additional district office informants	1	1	1	3	3
Total	20	10	38		68*

Note. *67 interviews were actually conducted because one principal was also a member of the panel and is counted twice in the spring tally.

TABLE A3
Surveys Returned

	Total Number of Surveys Disseminated	Total Number of Surveys Returned
Panel member	9	9
Consulting teacher	10	10
Principal	28	16
Participating teacher	91	57
Total	138	92

References

Abbott, A. (1988). *The system of professions: An essay on the division of expert labor.* Chicago, IL: The University of Chicago Press.

Baier, V. E., March, J. G., & Saetren, H. (1988). Implementation and ambiguity. In J. G. March (Ed.), *Decisions and organizations* (pp. 150–164). New York: Basil Blackwell Inc.

Bascia, N. (1998). Teacher unions and educational reform. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *The international handbook of educational change* (pp. 895–915). The Netherlands: Kluwer Academic Publishers.

Bird, T. (1986). *The mentors' dilemma.* San Francisco, CA: Far West Laboratory for Educational Research and Development.

Bridges, E. M. (1986). *The incompetent teacher: The challenge and the response.* Philadelphia, PA: The Falmer Press.

Brown, D. R. (1993). *An evaluation of the effects of a peer tutoring and assessment program on new teacher induction.* Unpublished doctoral dissertation. University of California, Santa Barbara.

Carnegie Task Force on Teaching as a Profession. (1986). *A nation prepared: Teachers for the 21st century.* New York: Carnegie Corporation.

Carver, C. L. (2002). *Principals' supporting role in new teacher induction.* Unpublished doctoral dissertation in teacher education. East Lansing, MI: Michigan State University.

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis 23*(2), 145–170.

Copland, M. A. (2001). The myth of the superprincipal. *Phi Delta Kappan 82*(7), 528–533.

Costa, A. L. & Garmston, R. J. (2000). Peer assistance and review: Potentials and pitfalls. In G. Bloom and J. Goldstein (Eds.), *The peer assistance and review reader* (pp. 63–79). Santa Cruz, CA: New Teacher Center at University of California, Santa Cruz.

Cuban, L. (1988). *The managerial imperative and the practice of leadership in schools.* Albany, NY: State University of New York Press.

Darling-Hammond, L. (1984). The Toledo (Ohio) public school intern and intervention programs. In A. E. Wise, L. Darling-Hammond, M. W. McLaughlin, & H. T. Bernstein (Eds.), *Case studies for teacher evaluation: A study of effective practices* (pp. 119–166). Santa Monica: RAND.

Darling-Hammond, L. (1997). *The right to learn.* San Francisco, CA: Jossey-Bass.

Feiman-Nemser, S. (1998). Teachers as teacher educators. *European journal of teacher education 21*(1), 63–74.

Feiman-Nemser, S., & Floden, R. E. (1986). The cultures of teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching, third edition.* New York: Macmillan Publishing Company.

Gallagher, J. J., Lanier, P., & Kerchner, C. T. (1993). Toledo and Poway: Practicing peer review. In C. T. Kerchner and J. E. Koppich (Eds.), *A union of professionals: Labor relations and educational reform* (pp. 158–176). New York: Teachers College Press.

Goldstein, J. (2002, April). *Results from the implementation of teacher peer assistance and review in California: A case study of one urban school district.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Goldstein, J. (2003a, April). *Complexity and courage? Examining the quality of teacher support and evaluation provided by peer assistance and review.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Goldstein, J. (2003b). *Teachers at the professional threshold: Distributing leadership responsibility for teacher evaluation.* Unpublished doctoral dissertation in Administration and Policy Analysis, Stanford University.

Grant, G., & Murray, C. E. (1999). *Teaching in America: The slow revolution.* Cambridge, MA: Harvard University Press.

Grubb, W. N., Flessa, J., Tredway, L., & Stern, J. (2003, April). A job too big for one: Multiple principals and other approaches to school leadership. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Halverson, R., Brown, F., & Zoltners, J. (2001). Distribution across artifacts: How designed artifacts illustrate school leadership practice. Unpublished manuscript.

Hart, A. W. (1987). A career ladder's effect on teacher career and work attitudes. *American Educational Research Journal, 24*(4), 479–503.

Hart, A. W. (1995). Reconceiving school leadership: Emergent views. *The Elementary School Journal, 96*(1), 9–28.

Heller, M. F., & Firestone, W. A. (1995). Who's in charge here? Sources of leadership for change in eight schools. *The Elementary School Journal, 96*(1), 65–86.

Hewitt, D. (2000). The Cincinnati plan. In G. Bloom and J. Goldstein (Eds.), *The peer assistance and review reader* (pp. 115–121). Santa Cruz, CA: New Teacher Center University of California, Santa Cruz.

Ickes, W., & Knowles, E. S. (1982). *Personality, role, and social behavior.* New York: Springer-Verlag.

Johnson, S. M. (1984). *Teacher unions in schools.* Philadelphia, PA: Temple University Press.

Kelly, P. P. (1998). *Teacher unionism and professionalism: An institutional analysis of peer review programs and the competing criteria for legitimacy.*

Unpublished doctoral dissertation in teacher education. East Lansing, MI: Michigan State University.

Kerchner, C. T., & Caufman, K. D. (1995). "Lurching toward professionalism: The saga of teacher unionism." *The Elementary School Journal, 96*(1), 107–122.

Kerchner, C. T., & Koppich, J. E. (1993). *A union of professionals: Labor relations and education reform.* New York: Teachers College Press.

Kerchner, C. T., Koppich, J. E., & Weeres, J. G. (1997). *United mind workers: Unions and teaching in the knowledge society.* San Francisco, CA: Jossey-Bass.

Koppich, J. E. (2000). Peer assistance and review: Enhancing what teachers know and can do. In G. Bloom and J. Goldstein (Eds.), *The peer assistance and review reader* (pp. 19–31). Santa Cruz, CA: New Teacher Center, University of California, Santa Cruz.

Koppich, J. E., & Kerchner, C. T. (1999). Developing careers, building a profession: The Rochester Career In Teaching plan. Unpublished manuscript.

Lawrence, D. (2000). California's opportunity. In G. Bloom and J. Goldstein (Eds.), *The peer assistance and review reader.* Santa Cruz, New Teacher Center, University of California, Santa Cruz.

Leithwood, K., & Jantzi, D. (2000). The effect of different sources of leadership on student engagement in school. In K. A. Riley and K. S. Louis (Eds.), *Leadership for change and school reform: International perspectives* (pp. 50–66). New York: Routledge/Falmer.

Little, J. W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. *American Educational Research Journal, 19*(3), 325–340.

Little, J. W. (1988). Assessing the prospects for teacher leadership. In A. Lieberman (Ed.), *Building a professional culture in schools* (pp. 78–106). New York: Teachers College Press.

Little, J. W. (1990). The mentor phenomenon and the social organization of teaching. *Review of Research in Education 16,* 297–351.

Little, J. W. (2000). Personal communication.

Lortie, D. (1975). *Schoolteacher.* Chicago, IL: The University of Chicago Press.

March, J. G. (1984). How we talk and how we act: Administrative theory and administrative life. In Sergiovanni, T. J., and Corbally, J. E. (Eds.), *Leadership and organizational culture: New perspectives on administrative theory and practice* (pp. 18–35). Urbana, IL: University of Illinois Press.

March, J. G. (1988). *Decisions and organizations.* New York: Basil Blackwell Inc.

March, J. G. (1994). *A primer on decision making: How decisions happen.* New York: The Free Press.

Meyer, J. W., & Rowan, B. (1977). Institutional organizations: Formal structure as myth and ceremony. In W. W. Powell and P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 41–62). Chicago, IL: The University of Chicago Press.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis.* Thousand Oaks, CA: Sage Publications.

Moir, E. (November, 1999). Presentation to the Teacher Union Reform Network. Santa Cruz, CA.

Murray, C. E. (1998). Teacher peer review: Possibility or pipedream? *Contemporary Education 69*(4), 202–204.

Murray, C. E. (1999). Rochester teachers struggle to take charge of their practice. In B. Peterson and M. Charney (Eds.), *Transforming teacher unions* (pp. 46–49). Milwaukee, WI: Rethinking Schools.

Ogawa, R. T., & Bossert, S. T. (1995). Leadership as an organizational quality. *Educational Administration Quarterly, 31*(2), 224–243.

Scott, W. R. (1995). *Institutions and organizations.* Thousand Oaks, CA: Sage Publications.

Sebring, P. B., Hallman, S., & Smylie, M. (2003, April). *When distributed leadership is called back.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Smylie, M., Conley, S., & Marks, H. M. (2002). Exploring new approaches to teacher leadership for school improvement. In J. Murphy (Ed.), *The educational leadership challenge: Redefining leadership for the 21st century. (101st yearbook of the National Society for the Study of Education, Part I.)* (pp. 162–188). Chicago, IL: National Society for the Study of Education.

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

Spillane, J. P., Diamond, J. B., & Jita, L. (2000, April). Leading instruction: The distribution of leadership for instruction. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Spillane, J. P., Halverson, R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher 30*(3), 23–28.

Spillane, J. P., Reiser, B., & Reimer, T. (2002). "Policy implementation and cognition: Reframing and refocusing implementation research." *Review of Educational Research Quarterly, 72*(3), 387–431.

Stryker, S., & Statham, A. (1985). Symbolic interactionism and role theory. In G. Lindzey and E. Aronson (Eds.), *The handbook of social psychology* (pp. 311–378). New York: Random House.

Tyack, D. B. (1974). *The one best system.* Cambridge, MA: Harvard University Press.

Tyack, D. & Cuban, L. (1995). *Tinkering toward utopia.* Cambridge, MA: Harvard University Press.

Van Maanen, J., & Barley, S. R. (1984). Occupational communities: Culture and control in organizations. *Research in Organizational Behavior 6,* 287–365.

Wasley, P. A. (1991). The practical work of teacher leaders: Assumptions, attitudes, and acrophobia. In A. Lieberman and L. Miller (Eds.), *Staff development for education in the '90s: New demands, new realities, new perspectives* (pp. 158–183). New York: Teachers College Press.

Weick, K. E. (1995). *Sensemaking in organizations.* Thousand Oaks, CA: Sage Publications.

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). *Teacher evaluation: A study of effective practices.* Santa Monica, CA: RAND.

Yin, R. K. (1989). *Case study research: Design and methods.* Newbury Park, CA: Sage Publications.

Youngs, P. (2003, April). *Responding to changes in state teacher policy: How principal leadership shapes mentors' practices.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Author

JENNIFER GOLDSTEIN is Assistant Professor, School of Public Affairs, Bernard Baruch College, City University of New York, One Bernard Baruch Way, Box C-305, New York, NY 10010; Jennifer_Goldstein@baruch.cuny.edu. Her areas of specialization are educational policy and leadership.