Teacher evaluation and the demoralization of teachers

Chris Bradford [a], [*], Melissa Braaten [b], [**]

[a] *Department of Curriculum & Instruction, University of Wisconsin-Madison, 225 N. Mills St., Madison, WI, 52706, USA*
[b] *School of Education, University of Colorado Boulder, 249 UCB, Boulder, CO, 80302, USA*

HIGHLIGHTS

- Teacher evaluation engages tensions between accountability and improvement.
- Teacher evaluation activities may contribute to teacher demoralization.
- Management logics may amplify performativity not authentic professional growth.
- Professional competence and identity is shaped by teacher evaluation activities.

ARTICLE INFO

ABSTRACT

Drawing from two years of ethnographic data, this case study details one middle school science and mathematics teacher's experience of new statewide teacher evaluation processes. Initially, these processes held potential as professional learning opportunities. However, limited opportunities for sense-making about what counts as "good" teaching foreclosed on teacher learning contributing to teacher demoralization. Evaluation processes eroded teacher professional identity by reshaping notions of professional competence in ways that did not make sense to teachers. Insights from this study inform debates about teacher evaluation and contribute to scholarship questioning the relationship between accountability policies in education and teachers' professional growth.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Even when working on Focus Lessons and Close Readings,[1] I was sent the message that I *wasn't* doing these things, and I am not a good teacher … Even though Vice Principal Fry said my [teacher evaluation] scores were good, I was left feeling like I would rather do less rather than more because my effort is not recognized anyway. (Jenny, middle school science and mathematics teacher).

Jenny was not always this downtrodden. Hearing "I would rather do less" came as a surprise after our two-year study. Many teachers had lost patience with the new teacher evaluation process, but Jenny's remarks did not sound like the Jenny we had come to know. Jenny was confident as a teacher and passionate about wanting to do more, not less. She had a strong vision for science teaching, a willingness to try new ideas, and a desire to investigate problems associated with teaching and learning. So, it came as a surprise that this fourteen-year veteran teacher was now seriously questioning whether her efforts were worthwhile.

How did Jenny's experience with teacher evaluation—something she initially viewed as a learning opportunity and a chance for administrators to get to know her teaching better—ultimately call into question Jenny's identity as teacher? Jenny's experience with new teacher evaluation systems draws attention to long-standing dilemmas about educational accountability and improvement (e.g., Stiggins, 1986), the pitfalls of superficial performativity rather

* Corresponding author.
** Corresponding author.
E-mail addresses: chris.bradford@wisc.edu (C. Bradford), melissa.braaten@colorado.edu (M. Braaten).

[1] Focus Lessons and Close Readings are *structured teaching* instructional routines anchoring Glacier City School District's Strategic Initiative (Fisher & Frey, 2013). Teachers were expected to enact instructional routines without modification. This was enforced through "walkthrough" observations by administrators. Close Reading is described in more detail below.

than authentic professional growth (e.g., Ball, 2003; Edgington, 2013; O'Leary, 2014; Rennert-Ariev, 2008), and the importance of teachers' professional competence and identities (e.g., Carlone, Johnson, & Scott, 2015; Greeno & Gresalfi, 2008; Sfard & Prusak, 2005).

In this case study, we analyze Jenny's experiences with a teacher evaluation system introduced as one component of data-centric accountability and improvement initiatives—reform efforts prioritizing certain data and decisions to meet improvement goals linked to accountability benchmarks for schools, districts, and states. This study addresses the following questions, from a teacher's perspective:

a) How do teacher evaluation activities construct notions of quality teaching?
b) In what ways are teacher evaluation activities generative for a teacher's learning, identity, and professional practice and in what ways are they constraining?

Rather than fostering the type of professional learning that teachers and instructional leaders crave, this case presents a narrative of how performing *for* a teacher evaluation system came at a cost—briefly stripping Jenny of qualities that defined her as a teacher—contributing to teacher demoralization (Santoro, 2011, 2018). We hope that this cautionary tale extends our understanding of how teachers experience certain reform efforts.

2. Framing teacher evaluation

Sparked by priorities set in the Race to the Top (RTTT) initiative, states and districts revamped teacher evaluation with two goals: 1) increasing accountability measures gauging the quality of teachers and teaching, and 2) creating professional learning opportunities to develop a shared sense of what "counts" as quality teaching (Kane, McCaffrey, Miller, & Staiger, 2012). Early optimism derived from a series of projects developing new measures of effective teaching including the Classroom Assessment Scoring System (CLASS, Hamre & Pianta, 2007; Pianta & Hamre, 2009), the Danielson Framework (Danielson, 2013), and Protocol for Language Arts Teaching Observation (PLATO, Grossman, Loeb, Cohen, & Wycoff, 2013).

Hill and Grossman (2013) note that optimism about teacher evaluation should be tempered by attention to contexts of schools and schooling. Specifically, they caution that new reforms may fail if they: 1) require more than regulatory changes, 2) fundamentally change core instruction, or 3) are introduced on top of other routines, practices, and reforms. Because new teacher evaluation efforts meet all three cautionary conditions, Hill and Grossman (2013) anticipate teacher evaluation activities struggling to function as professional learning opportunities for teachers. Their concerns highlight two important perspectives necessary for research on teacher evaluation: 1) a measurement and management perspective, and 2) a practice-based sense-making perspective.

2.1. Teacher evaluation as measurement and management

Teacher evaluation research typically aims for objective judgment—an unbiased measurement of teach*ing* and teach*ers* (e.g., Goldhaber, 2015). Goldhaber (2015) defines teacher evaluation as "an objective measure that does not rely on human interpretation of teacher practices, and by design, it is a system in which teachers are evaluated relative to one another rather than relative to an absolute standard" (p. 88). Measurement is typically conducted by an authority figure (i.e., administrator), but can include input from

coaches, colleagues, students, and self-evaluations. Evaluations focus primarily on classroom teaching practices, but can include professional practices like "maintaining accurate records" or "communicating with families" (Danielson, 2013). Information is gathered using assessment tools and delivered to the teacher in writing and sometimes discussion.

A measurement framing foregrounds a *logic of management* whereby teachers improve instruction in response to data in order to earn rewards and avoid sanctions (Firestone, 2014; Holloway-Libell, 2014). Management logic defines quality teaching as faithful replication and performance of instructional practices or curricular efforts (Firestone, 2014; Goldhaber, 2015; Hill, Charalambous, & Kraft, 2012). Teachers' adaptive expertise—fine-tuning instructional activities in light of dynamic local situations—may not register as high-quality teaching when teacher evaluation systems scan for fidelity of implementation.

The logic of management inherently depends on trust in the measurement tools themselves to provide all of the necessary information for improvement. The implication of this is that other sources of information, such as peer feedback or new perspectives gained after deliberation about a dilemma of practice, might be subject to bias or other untrustworthy influences (Horn, Kane, & Wilson, 2015). Horn et al. (2015) offer an important distinction between using performance metrics for instructional *management* versus *improvement*; information meaningful for school communities with an instructional improvement orientation was viewed as unnecessary and even suspicious in school communities oriented towards instructional management.

The management logic underpinning teacher evaluation can lead to performativity—inauthentic performances that are not genuine indicators of growth—which may be short-lived and ultimately demoralizing for teachers (e.g., Edgington, 2013; Santoro, 2011, 2018). Demoralization refers to processes that contribute to "discouragement and despair" experienced as teachers' vision of good teaching and sense of integrity as professionals comes into conflict with job expectations. As such, demoralization is not an emotion, discrete event, or individual idiosyncrasy; instead it is "borne out of ongoing value conflicts with pedagogical policies, reform mandates, and school practices" (Santoro, 2018, p. 3). One potential route for staving off teacher demoralization is to better understand how teachers make sense of new policies and practices. To do this requires a different approach to studying teacher evaluation activities that amplifies teachers' perspectives and experiences.

2.2. Teacher evaluation as practice and sense-making

We draw from research on teachers' workplace activities and re-examine how teacher evaluation operates as a *practice*. By *practice*, we mean, "coordinated, patterned, and meaningful interactions of people at work; the meaning of and the medium for these interactions is derived from an "activity" or "social" system that spans time and space" (Spillane, 2012, p. 114). In the case of teacher evaluation, this means engagement in activities, conversations, reflections, and other forms of participation *to make sense of what constitutes quality teaching*.

Notions of what constitutes quality teaching are embedded, negotiated, and represented through tools, routines, and norms used during teacher evaluation activities. Thus, each participant's understanding of a practice depends on his/her role and point of view (Feldman & Pentland, 2003). The active work of determining what constitutes quality teaching through sense-making in a community of practice is centrally about developing a shared sense of professional competence. Competence is constructed in practice through the interweaving of narratives that we tell ourselves (i.e.,