

## Principal evaluation – linking individual and building-level progress: making the connections and embracing the tensions

Sally J. Zepeda, Philip D. Lanoue, Noris F. Price & Albert M. Jimenez

To cite this article: Sally J. Zepeda, Philip D. Lanoue, Noris F. Price & Albert M. Jimenez (2014) Principal evaluation – linking individual and building-level progress: making the connections and embracing the tensions, *School Leadership & Management*, 34:4, 324-351, DOI: [10.1080/13632434.2014.928681](https://doi.org/10.1080/13632434.2014.928681)

To link to this article: <http://dx.doi.org/10.1080/13632434.2014.928681>



Published online: 10 Oct 2014.



Submit your article to this journal [↗](#)



Article views: 228



View related articles [↗](#)



View Crossmark data [↗](#)

## Principal evaluation – linking individual and building-level progress: making the connections and embracing the tensions

Sally J. Zepeda<sup>a\*</sup>, Philip D. Lanoue<sup>b</sup>, Noris F. Price<sup>b</sup> and Albert M. Jimenez<sup>c</sup>

<sup>a</sup>*Department of Lifelong Education, Administration and Policy, University of Georgia, Athens, GA, USA;* <sup>b</sup>*Clarke County School District, Athens, GA, USA;* <sup>c</sup>*Georgia Center for Assessment, University of Georgia, Athens, GA, USA*

The article examines the tensions one superintendent in the USA experienced as he evaluated principals in a high-stakes environment that had undergone numerous transformations at the central office. Using qualitative methods, primarily, shadowing techniques, observations and debriefing, the following tensions emerged and were examined in light of the work of the superintendent evaluating principal performance: (1) discrepancies between principal performance when compared to performance data, (2) length of time in the principalship compared to results, (3) finding the right balance between student achievement data and other indicators of principal performance, (4) what types of achievement data are important and when these data are made available, (5) credence paid to complaints about structural changes implemented by the principal, (6) balancing the principal self-evaluation rating scores with the final evaluator scores and (7) accounting for personal factors such as relationship to principals and knowledge about principal capabilities. Each of these tensions contributes to the difficulty a superintendent may feel when conducting the principal evaluation process.

**Keywords:** leadership; school improvement; educational change; principal evaluation

### Introduction

There is increasing concern about the quality of educational leadership as schools strive to meet the challenges of educating students in high-accountability environments (Darling-Hammond 2007; Mendels 2012; Suraya and Yunus 2012; Wallace Foundation 2013a; Zepeda, Bengtson, and Parylo 2012). This concern has been challenged by research results illustrating school leadership has substantial impact on student achievement (Honig et al. 2010; MacNeil, Prater, and Busch 2009; Leithwood et al. 2004; Louis et al. 2010; Robinson, Lloyd, and Rowe 2008). Busher and Harris (1999) underscore that ‘research from diverse countries and school contexts’ maintains ‘one of the fundamental tenets of school effectiveness and school improvement ... concerns the powerful impact of leadership’ (306). At the vortex of this work are superintendents and principals whose attention must continue to focus on increasing opportunities for children to experience measurable gains in their achievement.

Superintendents who understand the core technology of schools – learning – are able to focus their energy to provide coherence in leading teachers, leaders, students and parents to approach the work of schooling differently (Jackson 2000). The work

---

\*Corresponding author. Email: [szepeda@uga.edu](mailto:szepeda@uga.edu)

associated with leadership for school improvement is messy, recursive and fraught with tensions from internal and external constituents. One such tension associated with accountability is the evaluation of the principal and the interplay of data and other factors superintendents must examine throughout the process of assessing the impact of principals related to school improvement and student achievement.

Convincing evidence indicates principals hold responsibility for the success or failure of their schools and in supporting student achievement in both direct and indirect ways (Hallinger 2003; Hallinger and Heck 1998; Heck and Hallinger 2009; Leithwood and Jantzi 2000, 2005; Riley and MacBeath 2003). But their work is increasing in complexity, as have accountability expectations for student achievement and school improvement (Harris 2002; Harris et al. 2002; Lortie 2009).

Principals do not, however, need to stand alone in their efforts to increase student achievement (Honig 2013; Louis et al. 2010; Zepeda 2013). There is strong evidence that in school systems in which superintendents embrace core values about teaching and learning and provide the supports necessary to reach school and system-wide improvement targets, student achievement can increase (Honig 2013; Honig et al. 2010; Louis et al. 2010; Spanneut and Ford 2008). Superintendents and central office leaders play prominent roles orchestrating the conditions to promote transformational leadership practices across systems (Cordeiro and Cunningham 2013; Honig 2013). In a meta-analysis involving 2817 districts and the achievement scores of 3.4 million students, Waters and Marzano (2006) reported significant correlation between district leadership and student achievement.

One of the many leadership responsibilities of the superintendent is to evaluate principals who lead school-improvement efforts that support student achievement (Honig 2013; Normore 2005, 2010). The Wallace Foundation (2013b) offered specific 'key actions' to, 'develop fair, reliable performance evaluations to help principals improve their work and [to] hold them accountable for their students' progress' (17). Regardless of how fair or reliable, there are tensions ranging from the instruments, processes and procedures, to the types of data, the metrics used and the approaches of the superintendent who must enact principal evaluation.

There are numerous tensions inherent with creating more rigorous approaches to the evaluation of principals in light of heightened accountability predicated on school improvement and student achievement in schools where principals lead (Normore 2004, 2005, 2010). One tension is the controversy about leader evaluation in the current political climate in the USA (Conklin-Spillane 2012). Another tension is an overall lack of empirical research about principal supervision and evaluation conducted by superintendents given myriad variables associated with the context, climate and culture of the school systems in which superintendents lead system-wide change and transformation (Mattingly 2003; Normore 2005).

One article (Normore 2005) examined personnel evaluation in light of school-improvement efforts and an empirical study examined the congruence between principal evaluation and performance standards (Catano and Stronge 2006). Much has occurred during this nine-year period; yet, there is a dearth of research. The lack of empirical research could be due to the fact that the shifts in accountability have not allowed the time necessary to study the connections between the work of the principal and the relationship to school improvement coupled to the work of the superintendent related to the evaluation of principals (Clifford and Ross 2011). The paucity of research is further exacerbated with another tension predicated by local, state and federal accountability

systems that are often at odds with what occurs at the local level, in which the superintendent is faced with addressing competing interests that are important but may not necessarily relate to the direction and speed of movement in which a system must move to address school improvement focusing the work of the principal on student achievement.

With the heightened awareness created by federal and state mandates, superintendents must be instructional leaders consistently moving the system to increase student achievement, close learning gaps and holding those who oversee the instructional programme accountable for site and system-level improvements (Wallace Foundation 2013b). The passage of the 1965 Elementary and Secondary Education Act (ESEA) was one of the levers that opened the door for federal and state intervention in schools in the USA (Gottfried et al. 2011). The ESEA led to the No Child Left Behind Act of 2001 (NCLB 2002) as a means of ‘holding states, school districts and schools more accountable for improving the academic performance of each student regardless of economic status, race, ethnicity, proficiency in English or disability’ (NSBA 2013, para. 1).

In 2009, the Race to the Top (RTTT) programme was initiated, as part of the American Recovery and Reinvestment Act (ARRA), funded by the US Department of Education. In the first round, ARRA provided \$4.35 billion for competitive grants to states to encourage education innovation and reform centring on ‘great teachers and leaders’. RTTT gave higher scores to states that linked teacher and principal evaluations to student performance (Clifford and Ross 2011; Lohman 2010).

### **Purpose of the study**

The purpose of this intrinsic case study was to examine in a more granular way the tensions one superintendent faced as he assessed principals based on an evaluation system that was tied to the results of school-improvement targets at the site levels. One major question sparked us in examining these tensions: How does the superintendent reconcile vast amounts of data about student achievement, school-improvement progress, and then tie this information to the overall performance of the principal? This is a question worthy of examination given the fast-pitched manner in which superintendents and principals must consistently push forward to increase student achievement; therefore, this study is timely and might be able to serve as the basis for further research and to add meaning about the tensions that superintendents could possibly experience while conducting principal evaluations.

### **Literature review**

Honig (2013) underscored the complexities of school system leadership. Superintendents feel the sense of urgency to act, to empower and to provide the necessary supports for building-level principals (i.e., autonomy) in ways that lead to transformation of beliefs and practices as ‘federal and state policies place challenging demands on US school district central office(s)’ (1). One challenge is principal evaluation, a responsibility that falls to the superintendent or a designee from the central office (e.g., associate superintendent). The research about the role and work of the superintendent related to the evaluation of principals continues to lag in comparison to what we know about the supervision and evaluation of teachers (Ponticell and Zepeda 2004; Zepeda 2006); however, the evaluation of principals by superintendents warrants systematic study

commensurate with the complexities that school systems face as they are held responsible for increased levels of accountability at the local, state and federal levels (Clifford and Ross 2011; Normore 2005; Parylo, Zepeda, and Bengtson 2012).

The heightened importance of school leadership has steadily, yet slowly, expanded to closely examine the relationship of the superintendent and central office personnel to student achievement (Gurr, Drysdale, and Mulford 2006; Honig et al. 2010; Louis et al. 2010; Wallace Foundation 2013b). Waters and Marzano (2006) pointed directly to the leadership of the superintendent to permeate the system, leading not only to school improvement by increasing student achievement but also for the superintendent to lead district-wide reforms in which the entire central office personnel push forward, providing the necessary guidance to principals as they lead in their schools (Honig 2013; Honig et al. 2010; Wallace Foundation 2013b).

Numerous researchers report that superintendents who set high standards for academic achievement help facilitate student success (Honig et al. 2010; Leithwood et al. 2004; Petersen 2002; Peterson, Murphy, and Hallinger 1987; Wallace Foundation 2013b). In like fashion, 'whether by design or chance, superintendents communicate their beliefs about what is important educationally and the roles they expect their principals to fulfill', according to Spanneut and Ford (2008, 28).

In a meta-analysis to examine the influence that superintendents had on student achievement and to identify the characteristics of effective superintendents, Waters and Marzano (2006) reported correlations ranging between  $r = .24$  and  $r = .33$  between two variables – district leadership and student achievement across six leadership responsibilities, including (1) goal-setting process, (2) non-negotiable goals for achievement and instruction, (3) board alignment with and support of district goals, (4) monitoring goals for achievement and instruction, (5) resources to support the goals for achievement and instruction and (6) defined autonomy. Arguably significant, these skills need to permeate across the system with a focus on students and the leadership needed to ensure precision on what keeps a system moving towards school improvement and the conditions to realise student achievement. The tensions here are not only what to include in a principal's evaluation but also more importantly, how does a superintendent evaluate a principal based on metrics that align with the key leadership responsibilities expected of the superintendent to move the system towards improvement?

The systematic analysis of the work of the superintendent in principal evaluation is elusive; however, principal evaluation 'is an important goal for leaders in almost all districts and schools aiming to improve student learning' (Leithwood et al. 2004, 12). In a research synthesis, the Wallace Foundation (2009) reported the shortcomings of leader evaluation for principals. Namely the evaluations lacked emphasis and in most systems, 'they were not based on standards that reflected current understandings of the work of the principal as a leader of learning, first and foremost' (2). Explicitly, Clifford and Ross (2011) reported that the push for stringent principal evaluation systems was, 'prompted by RTTT ... redesigning principal evaluation systems as one means of providing support for principals' work. Ideally, principal evaluation systems reflect new goals and priorities for principals and support improved school leadership' (2).

At the federal level, there is increased scrutiny and demand for more highly qualified teachers and principals with the enactment of NCLB. The focus on teacher and leader evaluation is explicit in the RTTT programme in the USA (Clifford and Ross 2011; Wallace Foundation 2009). 'Improving teacher and principal effectiveness based on performance' clearly outlines that states who receive RTTT funding must, according to

the Race to the Top Executive Summary (2009), support leader and teacher evaluation, in part, to use data to:

- (1) (a) judge effectiveness using multiple rating categories with data on student growth being a significant factor and (b) are designed and developed with involvement by teachers and principals;
- (2) provide teachers and principals with data on student growth for their students, classes, and schools. (US Department of Education 2009, 9)

In a study about district-wide leadership, Leithwood et al. (2004) pointed to the necessity to evaluate school personnel. The key attributes reported by the Wallace Foundation (2013b) as well as the Joint Committee on Standards for Educational Evaluation as reported by Yarbrough et al. (2011) and the research of Clifford and Ross (2011), outline the overall aspects needed to develop a principal evaluation system. The call is for principal evaluation systems to focus ‘primarily on the instructional aspects of a leader’s performance’, a primary driver of school improvement (Wallace Foundation 2009, 3).

A common thread across studies, translations of research and the popular literature is to elevate the necessity of district-wide leadership to transform the ways in which they support leaders through specialised professional learning opportunities (Honig 2013; Honig et al. 2010; Wallace Foundation 2013b), to emphasise the evaluation processes and products used to assess the principal (Amsterdam et al. 2003; Clifford and Ross 2011; Leithwood et al. 2004; Lohman 2010; Wallace Foundation 2013b) and to examine the relationship between school improvement and leader evaluation (Mattingly 2003; Normore 2005, 2010; Parylo et al. 2012).

## **Methodology**

Although the principal researcher spent almost three years in the field site, we focus only on an eight-month period of time in the US school system, in the Clarke County School District (CCSD), in Athens, Georgia. Following a qualitative design and a constructivist theoretical perspective (Yin 2003), this intrinsic case study was conceptualised to help us develop an understanding of the tensions experienced solely by the superintendent as he enacted principal evaluation within a context of system-wide transformations that began five years prior to this study. It is important to know that the momentum from this transformation was the impetus for the development of the leader evaluation system in the CCSD; hence, the principal evaluation system did not serve as the impetus to change behaviours and beliefs. The CCSD superintendent was resolute that the development and codification of a school and system-wide school-improvement process, as well as the implementation of a teacher evaluation system, made the development of a leader evaluation the next logical step in aligning and steering the work of the system.

## **Research design and research methods**

To capture the real-life context of a superintendent evaluating principals as a way to study the tensions that arose during this process, a single and intrinsic case study method was chosen (Yin 2003). We chose an intrinsic case study approach because we wanted to understand more fully the tensions a superintendent experiences as he conducts principal evaluation in a high-stakes learning environment where the principal was expected to

have a positive influence on student achievement. We did not set out to build theory; we wanted to better understand the case – the superintendent as the unit of analysis and the tensions he experienced conducting principal evaluations. We were interested in studying the evaluation process and this superintendent because of the transformations the system had undertaken in a five-year period, the focus on accountability and the superintendent’s beliefs that evaluation was a growth-oriented process. Metaphorically, we believed that the context provided a blank canvas to examine the data to see if any discernible patterns surfaced to help us further our thinking about the superintendent’s work evaluating principals, school improvement, accountability and the tensions associated with ‘reckoning data, how data are reported, when data are reported, and using this data’ as a baseline to evaluate principals, according to the superintendent.

This case study was bounded (Corbin and Strauss 2008; Merriam 1998; Yin 2003) by the processes and procedures of the principal evaluation system in place in the CCSD as depicted in Table 1 that illustrates the time of the year in which these processes occurred across eight months and the locations where the processes and procedures unfolded.

For example, in August, the principals engaged in a process called a pre-conference in which they met with the superintendent to develop goals based on the overall needs of the school as found in the gaps identified in the school improvement plan, the perceived leadership needed to support closing the identified gaps, and the types of artefacts and evidence that would show narrowing the gaps. On average, principals identified two to three goals to focus attention for the year. In contrast, during mid-year reviews, the superintendent met with the principals at their school site after an impact check, which in brief is a meeting where school improvement progress is examined through myriad sources including student achievement data on standardised tests, pre- and post-test results and other site measures. At the end of the year, the principal and the superintendent met at the central office to review progress towards identified goals and their achievement and overall movement in meeting school-improvement targets.

This case was also bound by the context of the CCSD that is amplified later in the description of the research setting. The superintendent visited each school site (21 schools) on a regular basis, frequently interacting with principals, and he met monthly co-directing the principal professional learning communities (PLCs) sessions (full day of professional

Table 1. Time-frame of leader evaluation system – processes, procedures and locations.

Time-frame	Processes and procedures	Locations
Fall (August/September)	Pre-conference – setting goals for the school related to the principal and the school-improvement goals	Superintendent’s office
February–March two-part process	Part 1: mid-year review based on the impact check of the progress of the school in meeting its targeted school-improvement goals Part 2: Self-assessment towards meeting goals set in the fall by the principal against the school’s performance in key areas targeted in the school-improvement plan and the monitoring and redirection performed by the principal	Individual school Individual school – immediately following the impact check
End of March	Summative evaluation meeting	Superintendent’s office



learning based on the needs of the school system), and in all likelihood, these interactions influenced the superintendent, perhaps even added to the tensions of evaluating principals against the backdrop of accountability. The primary researcher observed the superintendent co-direct monthly PLC meetings and shadowed the superintendent while he made approximately 20 school visitations. These data were important for understanding the context of the system related to accountability, but they are not included in the data-set of the present study.

### ***Research methods***

As we framed the methods of this study, there were certain ethical considerations that had to be reconciled – the nature of the content of principal evaluation and that legally matters of personnel are confidential. Audio-recordings could not be used given the ethical considerations involved in matters of personnel (e.g., confidentiality), and moreover, we believed that audio-recording these meetings would have stifled, even promoted a chill effect between, the principals and the superintendent.

Although the researcher had entrée into the CCSD, each principal was asked if he or she had an objection to the presence of a third person in the room. No principal had objections to the researcher's presence. There appeared to be no instances where the presence of the observer disrupted the normal flow (Lofland and Lofland 1999) of the superintendent evaluating the principals. Finally, no data related to the content of any principal's conversations or artefacts shared were included in the findings as the study centred solely on the superintendent and the tensions he experienced.

The research methods included shadowing (Gilliat-Ray 2011; McDonald 2005), observing the superintendent conducting principal evaluations (Merriam 1998), debriefing conversations (Onwuegbuzie, Leech, and Collins 2008) and analysing artefacts (Corbin and Strauss 2008). The primary researcher was able to gain entrée to the site because of a two-year association with the system, serving as a professor-in-residence through the University of Georgia, College of Education and its partnership programme with the CCSD.

### ***Shadowing***

For eight months, we shadowed the superintendent as he was the only person in the system who conducted each phase of the evaluation system with 20 of the 21 principals in the CCSD. Shadowing is 'a research technique which involves a researcher closely following a member of an organization over an extended period of time' (McDonald 2005, 456). We shadowed the superintendent as he conducted the various processes and procedures associated with the principal evaluation outlined in Table 1. The shadowing method situated the primary researcher in the same room and space of the superintendent as he conducted principal evaluations, yielding first-hand, detailed data rather than the superintendent's account of what happened while he was conducting the various processes or following the procedures of evaluating principals. Table 2 highlights how the 180 hours of time was spent in the field shadowing, observing and then engaging in conversations while debriefing with the superintendent about each phase of the principal evaluation process portrayed in Table 1.



Table 2. Procedures and time spent in the evaluation of principals.

Data collection process	Hours with each individual principal	Hours invested to complete process
Pre-conference	1	20
Pre-conference – debriefing	1	20
Mid-year review – impact check	2	40
Mid-year review – debriefing	1	20
Self-assessment	1	20
Self-assessment – debriefing	1	20
Summative evaluation	1	20
Summative evaluation – debriefing	1	20
Total hours invested to complete evaluation of principals		180

Note: Number of principals = 20.

### *Observations*

Shadowing included observations. Observation of the superintendent as he interacted with principals during their evaluation was a ‘particularly helpful strategy for understanding ill-defined phenomena’ (Merriam 1998, 96). The observations occurred as the superintendent interacted with principals while the evaluation process was being enacted with each principal including discussing goals each principal set, reviewing data showing attainment of goals, or discussing data indicating gaps in student performance.

### *Conversations and debriefings*

Onwuegbuzie et al. (2008) described five different types of debriefing. The debriefing strategy used in the present study occurred between the researcher and the superintendent ‘on a regular basis ... in a ... private location’ so we could identify and examine the tensions that the superintendent experienced as he conducted the principal evaluations (5).

After each observation, the researcher and the superintendent debriefed about the experience. There was no script for the debriefing or pre-determined questions developed prior to any of the processes being observed as the researcher shadowed the superintendent. At times, the primary researcher would ask questions about some aspect that came up in the discussion between the superintendent and the principal, or conversely, the superintendent would ask for a perspective about something that unfolded between him and a principal. The questions were spontaneous and occurred in the moment.

Open-ended field notes from shadowing the superintendent, observing the superintendent and the conversations during the debriefing sessions were the primary sources of data. To counter a primary shortcoming of data overload during shadowing and observations, the primary researcher ordered field notes (Lofland and Lofland 1999) to bring clarity to data analysis that was simultaneously occurring during data collection.

### *Document review*

Records, documents, artefacts and archival information constitute a particularly rich source of information about settings (Marshall and Rossman 2006; Patton 2002).

Table 3. Performance area 1: leading change for school improvement (LCSI). The leader consistently implements an organisational vision and school improvement planning process that motivates others to take action to improve processes, staff performance and student learning. These processes situate the leader to do the following.

Elements	Exemplary	Proficient	Emerging	Unsatisfactory/ not evident
LCSI 1: Drive and sustain change in a collegial environment				
LCSI 2: Build buy-in from faculty and staff				
LCSI 3: Establish a school-improvement leadership team that regularly monitors and adjusts the school improvement plan				
LCSI 4: Use multiple data sources in developing, monitoring and adjusting the school improvement plan				
LCSI 5: Regularly communicate progress in meeting the goals of the school improvement plan				
LCSI 6: Ensure that the school meets and/or shows progress towards meeting district and state student performance targets				

Numerous documents from within the system were examined including memos about principal evaluation, the Leader Evaluation System including the forms and the policies associated with principal evaluation. The system-wide *School Improvement Handbook* and school-improvement plans were also examined. Of particular interest was the leader evaluation instrument that contained eight performance leadership areas with associated elements. [Table 3](#) illustrates the performance area, school improvement and its corresponding elements.

### ***Data analysis***

After observing the superintendent during each phase of the principal evaluation process, the primary researcher debriefed with the superintendent for approximately one hour about the experience. After each debriefing session, the primary researcher began organising and ordering notes (McDonald 2005). After ordering field notes, analysis started by reviewing the field-notes, notating through memoing techniques the ‘ahaaa’ insights, and then developing preliminary codes for recurring items that emerged as the data collection continued throughout the eight-month period in the field. This process assisted the researcher to frame any ideas to ponder about at a later date during debriefing meetings with the superintendent and then, with other members of the research team.

By the middle of the annual goal-setting conferences, there were some tensions emerging. For example, numerous principals wanted to focus on ‘their leadership development’ associated with targets outside the realm of where their school needed to improve. For Dr Lanoue, the superintendent of the CCSD, ‘Leadership is School Improvement’, not ‘an item to be added onto a particular initiative – all aspects of school improvement based on a school’s gaps need a focus from the principal’s leadership to the

Table 4. Analysing raw data to get to categorisation of tensions.

Code	Source of tension	Factors associated with the tensions	Overall tension
DPP/data	Is the individual or the school being assessed during the principal evaluation?	(1) Performance data (2) Ability to articulate and enact change (3) Difficulty focusing the lens on specific causes by looking at data alone	Discrepancies between principal performance data when compared to student performance data

student's desk'. Another emerging tension included the realisation by Lanoue that he had to 'rely on data and which direction the needle moved on student achievement' as a large basis for rating a principal – although the principal 'might grow tremendously' and there are 'other important improvements occurring in the school or extenuating circumstances'.

As the year progressed, the ideas associated with tensions were catalogued for ongoing and later analysis of the field notes that numbered in length 157 typed, single-spaced pages. Through iterative discussions, we came to agreement to render a common understanding about the codes found in the data, moving us into developing categories to cull further responses, phrases, words and metaphors associated with the categories as illustrated in Table 3. Through this process of examining data, we were able to isolate the overall tensions, attach names to describe them and find, in many instances, the inter-related nature of factors that ran across more than one tension (see Table 4).

### Context and research setting

The context of the research site is paramount to understanding this intrinsic case study examining principal evaluation in light of the numerous factors, most notably the transformations that had occurred as a result of the superintendent reorganising the central office to support teaching and learning and in the same vein, establishing clear expectations for principals, teachers and others involved with the instructional programme. First, we offer an overview of the context of the CCSD; second, a snapshot of the upwards slopes in student achievement that have resulted from the efforts of the superintendent, central office personnel and principals; and third, a timeline of events and activities leading up to the principal evaluation system.

### Overview of the context of the CCSD

Located in Athens, Georgia, the CCSD serves 12,750 students. Ethnically, 51% of the students are African-American, 23% are Hispanic, 20% are White and 2% are Asian. Nearly 12% of the students have English as their second language and 11% are special needs students. There are 2341 employees – 58% White, 38% African-American and 3% Hispanic. Students benefit from the expertise of nearly 1200 teachers – with over

60% having advanced degrees, 250 certified in gifted education, 16 National Board Certified teachers and 7 Georgia Master Teachers.

As a community, Athens-Clarke County, the seat of the CCSD, has the third highest poverty rate among US counties with populations between 65,000 and 249,000. Over 30% of children in Clarke County live in poverty – around 82% of students are eligible for the federal meal programme, 49% live in single-parent homes and 19% of adults (>25 years) did not complete high school – all significant risk factors that can keep students from graduating from high school. Despite these factors, today, the district's current graduation rate is 71.6, which is above the state average.<sup>1</sup> The CCSD was named the 2010 Title I Distinguished School District (Large) for closing the achievement gap between economically disadvantaged and non-disadvantaged students. All of these efforts are meaningless without results showing improvement in student academic achievement.

### *Upwards slopes in student achievement*

Table 5 shows the average percentage of students that met state standards in reading/ELA over the three-year period before and after the current superintendent assumed the role for all students within the district and across various subgroup populations. The policies and practices implemented by district and school leaders have resulted in dramatic increases in student achievement for all students combined and for each subgroup population with the exception of Asian students. The percentage point increase of all students meeting standards in reading/ELA was 8.2 percentage points. For subgroup populations, the largest percentage point increase was 15.2 percentage points for students with disabilities (SWD), a 14.1 percentage point increase for English language learners (ELL) and a 10.7 percentage point increase for Hispanic students, respectively.

Table 6 represents the average percentage of students that met state standards in mathematics over the three-year period before and after the current superintendent assumed the role for all students within the district and across various subgroup populations. The percentage point increase for all students meeting standards in mathematics was 13.7. For subgroup populations, the largest percentage point increase was seen with SWD at 22.0, followed by a 16.9 percentage point increase for

Table 5. Three-year average of the percentage of students meeting standards in reading/ELA for AYP (adequate yearly progress) purposes across subgroups pre- and post-superintendency.

	2007–2009	2010–2012
All students	71.1	88.7
Black	75.3	84.7
White	93.4	96.5
Hispanic	79.5	90.2
Asian	98.2	95.4
SWD	53.2	68.4
ELL	67.9	82.1
EDS	76.1	86.4

Note: Superintendent assumed position in July 2009.

Table 6. Three-year average of the percentage of students meeting standards in mathematics for AYP purposes across subgroups pre- and post-superintendency.

	2007–2009	2010–2012
All students	66.1	79.8
Black	56.7	73.1
White	87.3	93
Hispanic	67	82.9
Asian	97.3	95.9
SWD	37.8	59.7
ELL	56.4	73.3
EDS	59.3	76.2

Note: Superintendent assumed position in July 2009.

economically disadvantaged students (EDS), a 16.8 percentage point increase for ELL and a 16.3 percentage point increase for Black students, respectively.

#### *Timeline of events and activities leading up to the principal evaluation system*

Prior to the superintendent's arrival, a new associate superintendent for the Division of Instructional Services was hired (2007–2008) and a system-wide walk-through protocol was developed to align instructional practice with district objectives, to provide data to schools regarding the level of implementation of standards-based instructional practices and to ensure implementation of state curriculum requirements. The implementation of the system-wide walk-through practices was a *door opener* to the focus on instruction for the newly hired superintendent (summer, 2009).

Underlying his reforms was a concerted effort for the system to support teachers and leaders in focusing their efforts on high-yield instructional strategies, the transparent use of data across the system to gauge success and the expectations of principal performance and appraisal to be more aligned with the work that was expected of teachers. The reforms were aimed to establish a culture of mutual accountability which, in the words of the superintendent, occurs 'one conversation at a time'. The data in Tables 5 and 6 and the graduation rates illustrate the urgency in moving towards more accountability and embracing a principal evaluation system that would 'squarely put principals at the centre of school improvement' in their buildings, according to the superintendent.

#### *The context for evaluation*

The CCSD developed its own teacher and principal evaluation system because it could not wait for the state of Georgia to render a valid teacher and principal evaluation system in response to the RTTT (2009) policies in which Georgia became one of the many RTTT states that legislated a statewide teacher and leader evaluation system. Moreover, the major shifts in instruction, the development and implementation of a common instructional framework across the system, the adoption of non-negotiable practices and the professional development provided to principals to be instructional leaders able to support teachers propelled the movement to develop the CCSD Principal Evaluation

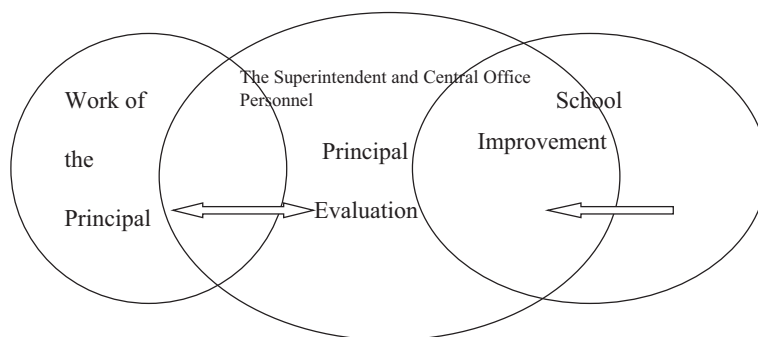


Figure 1. Relationship between the work of the principal, school improvement, leader evaluation, and the superintendent and central office personnel.

System because the work expected to bring about school improvement was not present in the statewide principal evaluation system.

There was a sense of urgency for this system to continue its focus on accountability given the upwards slopes in student achievement as illustrated in [Tables 5](#) and [6](#) and the marked increases in graduation rates. According to the superintendent, ‘the system could not go backwards – we had to keep moving all systems to align with the transformations’. The superintendent believed it was an ‘obligation’ to align the work required of principals to support teachers, to steer the direction of school improvement to ‘each child’s desk’ and to be able to monitor and change direction mid-course if results were not moving the ‘achievement needle’ in the right direction. The existing state evaluation system for principals did not take into account school improvement and lacked coherence between what the CCSD was asking teachers and principals to do and what the system felt was necessary to assess principal performance. Moreover, there were several stalls in the development and roll out of the ‘new’ state evaluation system with no clear indication of when the final leader evaluation system would be finalised. The CCSD had ‘to keep moving forward’ with the momentum of the changes, successes and lessons learned from the transformations that had been implemented.

[Figure 1](#) frames the relationship between the work of the principal, school improvement, leader evaluation and the overarching support from and monitoring of the superintendent and key central office leaders as evidenced in the CCSD Leader Evaluation System.

### ***Participant profile***

The single participant in this study was the superintendent of the CCSD, Dr Philip Lanoue. The year in which we conducted this study, Lanoue had been the superintendent for three years, and he was, by board of education policy, directly responsible for evaluating the principals in the system. Prior to assuming the superintendency of the CCSD, he was Area Assistant Superintendent in another district and his duties included supervising and evaluating principals at the elementary, middle and high school levels. Prior to entering administrative positions at the district level, Lanoue served as a high school principal for 18 years in 4 high schools in Massachusetts and Vermont, and he taught biology at the high school level.

Lanoue's prior experiences as an evaluator have undoubtedly shaped his beliefs and actions related to accountability, the need for systems that allow leaders 'to grow, develop, and experience success'; however, he was acutely aware that principal evaluation was part of the accountability equation that included 'a preponderance of data along with artifacts and evidence about the school in which the principal leads'. Lanoue hires principals who are or can quickly become instructional leaders capable of leading school improvement. [Figure 1](#) shows explicitly the nexus between school improvement and principal evaluation.

The metaphors that Lanoue uses reveal a great deal about his beliefs about accountability, his beliefs about supporting people as they develop as professionals, and his beliefs that 'every action of the principal must be razor-focused' granularly to 'each child's desk'. Through his candid discussions emerged tensions illustrating how Lanoue's actions and responsibilities can, at times, conflict with what he might know about a particular situation surrounding a principal or how an issue in a community, or a change or shift in personnel can 'stretch you tight at opposing sides of thinking about a principal's overall rating in a particular area'. Other tensions such as the uses of data as a marker of principal performance or what he calls the 'elephant in the room', the time in office, and other variables often created an interplay in which one tension could possibly be interrelated with or bound within another tension.

Although Lanoue reported he spends approximately three hours analysing, reflecting, and then writing the final evaluation report for each principal, there were times he pondered whether or not he spent 'enough time' examining artefacts and evidence about a particular principal's performance or whether he was consistent across the board, or whether his particular mood affected the way in which he finalised a principal's evaluation. Lanoue was confident, although the principal evaluation system had a formative (growth) and a summative (final) component, that the system itself, 'promoted growth, that the processes ensured that principals could seek assistance as the year progressed, and that the central office was poised to assist them' in their work as school leaders.

Lanoue's conversations surrounding principal evaluation had one very prominent thread – he cared about education because, 'everything I do, including evaluating the performance of principals relates to the quality of education that our students get'. Lanoue was resolute that there are, 'no excuses in Clarke County, and we expect that all of our students can learn, will be able to apply what they have learned – through a variety of assessments'. However, Lanoue is realistic, knowing there are many variables associated with the children from 'our neighborhoods'. He further elaborated, 'We will never let socio-economic status [SES] influence expectations. The students who come to school are the students we educate, and it is our responsibility to narrow the achievement gap'. The reality, however, is that students may be three or four years behind grade level, but Lanoue is emphatic that 'we must embrace this tension to do what is absolutely best for kids'. Lanoue examines data – several points – including, for example, student characteristics, the demographics that encompass the neighbourhood school, SES and trend data – as he is evaluating a principal because, 'all of these variables mean something about the leadership of the principal, especially monitoring the instructional program on a daily basis', especially as these relate to school improvement – 'and this is what the principal must own'.

The primary observer shadowed Lanoue conducting 20 cycles of principal evaluation (see [Table 2](#)) with as many principals. The principals are 45% Caucasian and 55%



African-American. Of the 20 principals, 7 (35%) are male and 13 (65%) are female. The years of experience as principal at their current schools range from 1 to 12 years, and the mean number of years is 4.85.

## **Results**

### ***Superintendent tensions in the principal evaluation process***

Data analysis revealed seven tensions that were, at times, interrelated in accurately rating principals by the superintendent. These tensions are (1) discrepancies between principal performance when compared to performance data, (2) length of time in the principalship compared to results, (3) finding the right balance between student achievement data and other indicators of principal performance, (4) what types of achievement data are important and when these data are made available, (5) credence paid to complaints about structural changes implemented by the principal, (6) balancing the principal self-evaluation rating scores with the final evaluator scores and (7) accounting for personal factors such as relationship to principals and knowledge about principal capabilities. Each of these tensions contributes to the difficulty a superintendent may feel when conducting the principal evaluation process.

There are also some other ideas to ponder when reading our findings. First, a majority of these tensions can be traced to discrepancies (e.g., in performance judged against data). Second, many tensions are inter-connected to factors found across the tensions; hence, the tensions are multi-faceted (e.g., length of time in the position and data). Third, the nature of school improvement, accountability and data as a foundational aspect of principal evaluation does not paint a complete picture of a principal's performance. The following tensions are examined and where appropriate the interconnected relationship across tensions is examined as found in the data from this study.

#### *Tension 1: discrepancies between principal performance data when compared to performance data*

While principals may have in place school processes, protocols and systems that fully support the work required in the leadership performance areas, data often indicate that while progress may be made in performance areas, schools often are far from being fully operational. Rating principals on the *results* of the implementation of new protocols or systems often does not align to the evaluation of the *implementation* of the new protocol or system. This tension is often seen when an experienced principal moves to a school to create change and the systems are developed, but the 'results are not there at that time'. Lanoue elaborated, 'The leader could be masterful at turning around a school', but in most situations, 'the results will not be seen in the first two or three years'. Lanoue shared, 'although it is tough for me, patience, has become part of my own skill set related to checking for results'.

This first tension refers to discrepancies between actual principal performance and performance data for the school as a whole. The discrepancies between principal performance and the school performance data can be evident in either direction, meaning a principal could be very highly rated and receive a lower school rating or a lower-rated principal could be at a school which receives higher scores for school performance.

Notes illustrated that this tension resides in the discrepancy of two elements that reflect: (1) the ability of leaders to design and to articulate a change process and (2) the data do not show immediate results. For example, classroom walk-through data are strong regarding the pervasive implementation of an instructional practice in the system-wide non-negotiable practices, for instruction; however, benchmark scores and other standardised scores remain ‘essentially static’. From this discrepancy comes the tension in which two fundamental questions emerge. The first question is, ‘Is this discrepancy a function of the change process being too superficial and not monitored beyond walk-through data?’ The second question is, ‘Are the data on the use of research-based instructional practices incomplete (or accurate from the observer who is sitting in the classroom)?’ These tensions unfold to a bigger one – are the walk-through data stable as these data are provided by various members of the walk-through team?

Other issues cloud this discrepancy as well, including student or family tragedy, teacher absenteeism, or when key teachers have been out for an extended time, ironically participating in professional learning opportunities. Lanoue explained that ‘these types of situations impact small neighborhood schools’. Lanoue elaborated, ‘For me, the real tension is, was the principal responsible for leading or not leading through with the most effective leadership strategies in relationship to impacting results?’

However, other tensions can cloud issues of principal performance including, for example, the length of time in the position. Although it is not common in the CCSD for principals to ‘move around from one building to the next’, there are instances where a principal will move from one school to another or a ‘newly hired’ principal takes the helm of a school that has been ‘established by the past principal’. The length of time in the principalship, results, evaluation and the tensions that these factors present are examined next.

### *Tension 2: length of time in the principalship compared to results*

This tension raises three questions to ponder. The first question is ‘How long does it take for a principal to see results?’ In the leadership evaluation process, the second question emerges: ‘How much leeway is given for the school to turn around and for systems and practices to result in change?’ Conversely, the third question is, ‘How do you rate a principal entering a school that is fully operational when the previous leader’s systems were likely the reason for success?’ These tensions create complexity, but these complexities are based on the reality of the situations in which superintendents must evaluate the performance of the ‘principal sitting in the office – at the present time’, according to Lanoue.

The tension is really about the length of time in the principalship and school performance results. Primarily, this refers to the subjective feeling of a superintendent that is manifested in this question: ‘How long is long enough to see results?’ The possible outcomes of this tension could be both positive and negative in terms of the principal evaluation score. One possible result of this tension could be that a superintendent may artificially inflate the evaluation of a newly hired principal by not holding him/her accountable for poor school performance because ‘their plans have not had time to reap results’. Likewise, a superintendent may artificially deflate the evaluation of a newly hired principal at a school which receives high performance scores because the credit for these results is attributed to the initiatives and the culture which preceded the current principal at the school. According to Lanoue, ‘I am confident when I hire a principal,

I hold confidence in the work we are doing, and the professional development we provide', but he offers that he 'walks the tight-rope' with 'his new principals', holding the 'safety-net' all the while holding 'accountability and school improvement at the centre of what great leaders do to improve student achievement'.

While the preceding outlines the possible results for newly hired principals, those with more experience might also receive 'lower ratings' if the school performance results are lacking. This is likely due to the expectation that as the number of years in the principalship increases, so too should the performance of the school that could be linked to the initiative of the principal and his/her efforts with implementation. According to Lanoue, 'we expect our principals to become more skilful as time in the principalship increases'. Lanoue shared this tension occurs 'at any time you put a new principal in place'.

The years of experience of principals at their current schools range from 1 to 12 years. According to Lanoue, because of this range, system-wide efforts are predicated 'on strong professional learning, and that is why we have a monthly, Principal Professional Learning Community'. Lanoue states, 'we do not talk about discipline; we put our minds around improving instructional practices, studying the variability in our student populations, and understanding how our leadership has to ratchet up'. The tension, according to Lanoue, is 'How does a superintendent look at the developmental differences across the principals and factor that into the rating?' He also struggled with a couple of interrelated issues such as, 'Do I subconsciously let the length of time in the principal's chair in one building influence my rating of the principal?' and 'Is there anything wrong about this?' Lanoue reflected that regardless of years in the position, the targets change and 'so too do the expectations for increased principal performance'.

The data presented in [Tables 5](#) and [6](#) certainly illustrate an upwards spiral related to student achievement, but 'yeomen work has been done, most likely in assessing the overall culture and climate in the school, establishing routines, and getting faculty on the same page – instruction matters'. In a 'principal turn-around' situation, there could also be other tensions such as 'teachers leaving, and in droves; we can track that through human resource data'. The tension is that, if some teachers leave, 'it may be a good thing, but some may perceive it as a hostile environment'. How much time in as a 'turn-round principal' is needed to see results? 'My guess is two to three years', shared Lanoue. With this time-frame in mind then, Lanoue questioned, 'What do I do to factor that variable in a principal evaluation in the first year, then the second year and then after that?'

Lanoue's tension about the relationship to school improvement and the work of the principal related to time and progress was presented with this metaphor: 'sleep, creep, and leap'. Lanoue believes school-improvement strategies and processes very much like the performance of the principal during school improvement, are manifested in this metaphor. Lanoue is cognisant of the varying levels of skills his principals have. He consistently spoke of the CCSD Principal Evaluation being a developmental model with one intent – 'help principals grow so they can, in turn, grow in their professional roles as instructional leaders in the buildings'. Lanoue shared that he will walk through schools at any given time of the year and randomly ask different teachers, staff members, and even students a pivotal question: 'Who are the instructional leaders in your school?' He chuckled that if he does not hear 'principal as one of the top answers', he knows some more work might be needed to help that principal gain, 'not only the public perception but more importantly, the confidence and skill set necessary to be the instructional leader in the school'.

Lanoue shared a converse tension to the ‘turn-around principal’ situation that he has had to face in past principal evaluations. The tension emerges when, ‘a new principal who goes into an existing school that is high performing, and will not likely change results, but it may not mean they are or are not leading’. More questions that illustrate the complexity of this tension included, ‘What if the staff is at a natural plateau? What if their students are exceeding in just about every aspect?’ Lanoue reflected, ‘Again it goes back to the tension of having the right practices in place and whether these practices are pervasive and consistent’. He continued, ‘Likely, we will see the most accurate change in principals who have gone into a school that has mixed practices and through short-term interventions and where practice becomes more consistent, student performance increases’. Lanoue made it clear that in any situation, he does not believe he has ever been ‘left “best-guessing” based on data from primary sources’, and the data helps to assuage, to a degree, this tension. Lanoue was very cautious when he spoke about the tensions associated with measuring principal effectiveness. A principal’s effectiveness ‘can and should be measured by more than mere measures of student outcomes – a principal’s influence extends [beyond] test scores’, explained Lanoue. However, data are ‘the coin of the realm’ and with what a school and ultimately what a principal’s leadership will be measured by external constituents. ‘It’s a balancing act’ for me, Lanoue repeatedly came back to, throughout the duration of the study.

*Tension 3: finding the right balance between student achievement data and other indicators of principal performance*

According to Lanoue, ‘the large “visible elephant in the room” is test scores or the quantitative metrics used for assessing school and leader performance’. He continued, ‘the bottom line in the new world of accountability is that a school leader should be evaluated on the performance of the school’. Lanoue’s tension rests in a question he raised: ‘What metrics are used to determine school performance?’ This is an important question that often defies a clear-cut answer. Lanoue elaborated the source of this tension summarising, ‘the measures used for NCLB and now the new College and Career Ready Performance Index [CCRPI]<sup>2</sup> are clearly geared to test scores, but the new accountability system has more indicators’. Balancing growth in performance scores or even declines in test scores as part of the evaluation process is ‘delicate at the least’, and must be ‘examined in light of so many other factors’.

Lanoue opined more about this tension, ‘we know that other indicators contribute to improved performance, but we do not have an accurate timeline’. In the state of Georgia, the Georgia House Bill 251<sup>3</sup> provides for parents and guardians to have public school choices within local schools. The tension here is a political one as well in that ‘HB 251 can impact school performance though this is not included in the evaluation – nor should it’, according to Lanoue who also shared, ‘I have seen, first-hand where a situation caused by an external force can be overcome’. Lanoue shares where ‘a new principal was brought in who was able to rally the support of the community to rebuild, and the school changed’. The tension is ‘Would this have occurred without this community buy-in?’ and, ‘How do I quantify this outcome on the leadership of the principal? The community? Or the teachers?’ Lanoue pressed the question, ‘Would this principal have been able to pull off the change if he had no prior experience as a principal?’ Lanoue explains he embraces this tension because ‘something is working and that principal is keeping things moving upward every year’.

Other tensions abound related to test scores and the evaluation of principals. Lanoue questioned, ‘If a leader experiences increases in test scores in four areas but a substantive decline in two other areas, it becomes difficult to pinpoint the issues through the process outlined in leaders’ performance areas’. Lanoue explained that this scenario becomes even more of a tension in high-transience schools where school leaders ‘may have up to 50% of their students change schools throughout the year’. Although the new CCRPI system has changed the criteria for counting full-time equivalent students, this is still a dilemma, adding to the tensions of evaluating a principal who leads in a building with a highly transient population. Or, even in a building where there is high teacher or leader turnover?

Furthering the tensions about the relationship of test scores on principal evaluation, Lanoue adds, ‘In addition, the performance metrics for a school may not be aligned to ones outlined for the state especially in special needs or non-traditional programmes’. The CCSD has in some schools ‘big pockets of special needs and non-traditional programmes to serve the needs of students’.

This tension refers to finding the right balance between student achievement data and other indicators of principal performance because in the world of accountability, much of the focus has been on standardised test scores. Prioritisation and balance are one way in which Lanoue brings into focus the tensions by examining all performance areas, knowing that principals will make strides in time over all areas. However, when a leader cannot do this, then the focus has to be on the larger picture of the school, and perhaps, even ‘different types of conversations’ need to occur. Lanoue emphasised that for schools that are considered as ‘turn-around’ with a turn-around leader, ‘time is important and the circumstances warrant urgency’.

*Tension 4: what types of achievement data are important and when these data are made available*

Timing is everything. In many states, standardised test score results are not available until after the leader or teacher evaluation cycle. According to Lanoue, ‘this lag in reporting does not leave many options if this metric was to be included in the principal evaluation process’. Lanoue elaborated more on the struggle he has when the only data available are, ‘Benchmark scores available throughout the year, but we have touted them as a formative measure – not a summative measure’. Lanoue explained, ‘I can’t change the rules in the middle of the game and moreover, the use of formative assessments as summative ones would wreak havoc in our system’. Additionally, Lanoue pointed out that ‘CRCT data<sup>4</sup> are only available for the previous year’. Lanoue shared, it would be a disaster having two annual evaluation scores – ‘one before the results of standardised testing are in and then another score one month later when scores are available’. The tension goes beyond the timing issue and what data to use.

According to Lanoue, the standardised data tension ‘may be the most difficult and most problematic for school leaders’. He elaborated, ‘While other assessments can be used such as benchmark or pre–post-test assessments, which are formative in nature, raises questions about their use’. Looking at this tension, Lanoue begs the question, ‘Do you want to make formative measures more high stakes?’

*Tension 5: credence paid to complaints about structural changes implemented by the principal*

What is the impact of outside noise on principal evaluation? It is common knowledge that ‘principals that engage in hard level change’ often make difficult decisions that ‘shake up the organisation’. Lanoue reconciles this tension in this way, ‘If one gets lots of complaints or the same complaint from several individuals who are very vocal, then you need to look at other data which are typically survey data’. However, Lanoue backs down on his belief that he can easily reconcile this tension with an attitude of ‘ignore the vocal ones’ because he does not know if the, ‘survey data (climate survey) can give the correct information to analyse the situation which may shed light on the “real” situation’. Lanoue shared, ‘even the timing of the noise is a factor – if not current, then did the school leader resolve the matter? This is a critical question!’

Lanoue spent a great deal of time detailing ‘outside noise’ that can come from inside the schools, and that this noise follows several predictable patterns. He gave numerous examples of noise, their sources and how the media has intensified the messages within the outside noise. Noise often comes from parents because, ‘many have a favourite teacher they really like, but that teacher is rarely engaged in a standards-based teaching or does not like to collaborate, but the teacher made the parents feel good’. Daily, Lanoue is filtering noise that ranges from the, ‘correspondence I get from parents, citizens from within the community, letters written to the board and shared with me, and through social media (e.g., Facebook, Twitter)’. He also said, ‘And of course, the bloggers who can hide but say some awful things about our leaders, teachers and me, too’. This tension refers to the superintendent being able to wade through the complaints when a principal implements new programmes or organisational changes to determine if, and with how much credence, these should factor into the principal’s evaluation.

*Tension 6: balancing the principal self-evaluation rating scores with the final evaluator scores*

Principals engage in a self-assessment before they have their summative evaluation with the superintendent. The superintendent has the principals use the annual evaluation form to conduct their self-assessment, and they sit and discuss the self-assessment based on the school improvement plan and the data that show gains, gaps, or no movement towards student achievement. In this tension, the overriding question is, ‘Is perception reality?’ Data from myriad sources are presented during the impact check – formative assessments, walk-through data, pre- and post-test results, benchmark data and the continual actions of the school, the school improvement leadership team and the results of vertical and horizontal efforts of the school are examined. The impact check gives the superintendent the opportunity to look beyond the data by examining the strategies and approaches that the school has undertaken to narrow gaps, and for Lanoue, the opportunity to see the ‘artefacts and evidence’ including the leadership of the principal – ‘is the principal monitoring progress and supporting necessary changes that align with what the story the data tells’.

Lanoue explained, ‘balancing self-evaluation data from the school leader and the superintendent can often present conflicting ratings’. Lanoue explained that when a conflicting assessment rating occurs, ‘this situation raises questions about the evaluator and school leaders’ definition or perception of the work – and more so, if you do not use

clear rubrics'. Balancing the principal self-evaluation and the final evaluator ratings can present challenges. Lanoue shared that he often struggles with the connection between leadership, implementation of sound strategies to support school improvement, and the results of these efforts, balanced against gains in student achievement. When Lanoue finds himself struggling with this tension, he shared that he reflects deeply, asking, 'Do I let the principal know the work is good, but maybe it's the wrong work given the needs of the school and its students?' Lanoue further shared that he must examine data and then continually reflect on the 'totality of the situation – how many years has the principal been in the building? What kind of internal and external noise is being made, etc.' These questions reflect tensions already examined but illustrate that factors within one tension overlap with factors associated with other tensions.

Closely associated with this tension is how far off are the principal's perceptions of effectiveness. Lanoue shared about human nature, 'often people are far more critical about their own performance than are others'. It is always a tension when there is not agreement, and this tension is often difficult to reconcile, especially since those data provide evidence. In the instances where there are gaps between the perceptions of the principal and Lanoue, he makes it a practice, to 'examine other evidence, artifacts, and data collected throughout the year'. He looks at the preponderance of available evidence and data and then engages in framing more overarching questions to talk through the differences.

The mid-year conference following the impact check is a reality check – for both the superintendent and the principal. This extra step, the mid-year self-assessment and then the month before the summative evaluation conference gives the superintendent and the principal the 'time and space' to re-review data, and think through other variables related to the school, and the principal's leadership. Lanoue believes time is critical as it allows him to 'examine, re-examine, and reflect' on data and evidence so he can 'hold confidence in his own assessment and understanding' of the principal's self-assessment before the final summative meeting with the principal.

*Tension 7: accounting for personal factors such as relationship to principals and knowledge about principal capabilities*

The final tension is simply the human side of the process. Knowing your principals and their capabilities can cloud decisions, especially if the evaluator sees a drastic change. This tension is one that exists in any rating situation where the evaluator personally knows those he/she is charged with evaluating. Personal relationships, knowledge about capabilities of the principals and other factors such as years until retirement or awards given to principals all could impact the rating provided by an evaluator. Knowing how to navigate these issues in a professional way to not show favouritism is the key to reconciling this tension. As a solution, Lanoue tries to 'muffle the noise' by focusing more on concrete data.

Lanoue also shared that a, 'superintendent [who] knows principals' strengths or admires their work in a specific area may interpret results differently'. Further elaborating on this tension and the possibility of bias, Lanoue shared that it could be that the, 'principal is an effective educator and willing to compromise or help when asked; therefore, they are easier to score higher than a principal who was difficult and defiant foisting oppositional strategies to subvert system progress'. Lanoue acknowledged that he 'had to be on guard' about this tension, recalling a time he was evaluating a leader who



‘really was not as proficient in many areas’; I ‘had some really tough conversations with this leader’, leading to a ‘temporary uncomfortable working relationship’.

## **Discussion**

In addition to identifying the tensions, we begin to see that it is critical to embrace the tensions surrounding the evaluation of principals because ‘somewhere in these tensions’ are the ‘opportunities’ for leaders to develop more fully in their work and skills as leaders. What permeate the overall evaluation process of principals are tensions between the evaluation systems (i.e., mandated accountability) and understanding that the school system comprises neighbourhood schools in which each school building and its leadership serves diverse communities that hold onto beliefs about ‘their home school’. Lanoue ‘strongly’ believed the success of principals reflects the success of schools and to the totality of the success of the district in which he holds ‘ultimate responsibility’. Given the present press for principal evaluation, we could not find empirical studies that identified and described the tensions a superintendent experiences while evaluating principals – even though it is crucial to understanding the increased and pervasive focus on accountability and its relationship to school improvement, and student achievement (e.g., Riley and MacBeath 2003), nor could we find studies in which researchers shadowed a superintendent conducting principal evaluations.

The tensions emerge from this superintendent’s absolute belief that it was his ‘duty’ to ‘first build capacity’ in each principal by ‘tapping into their individual skills and talents’ and then ‘aligning the collective work of all principals in meeting the goals of the district’. Lanoue defined his role as superintendent as ‘to first ensure that the work of leaders is clearly understood’, ‘to provide the needed supports and guidance necessary for leaders to grow’, and ‘to engage in conversations that are framed around the great problems of practice’ that principals encounter as they enact ‘change and transformations’ in their own buildings that simultaneously ‘align with the overall direction, scope and depth of those that are priority targets for the system’. This is and continues to be an annual challenge of school district leadership.

Although Lanoue spends a great deal of time communicating with various internal and external constituents, the real communication for him occurs when he is framing ideas, asking probing questions and getting ‘principals to think and reflect on their own skill sets’ and more importantly, to broach unabashedly the big question, ‘Does your skill set serve the best interest of the kids in your school?’ These types of conversations occur frequently, but most intensively, during the four times a year that the superintendent meets one-on-one with the principal to engage in the evaluation process (see [Table 1](#)). These ‘conversations’ focus the superintendent’s balanced attention on school improvement, the impact of the principal and the data and evidence to support the principal’s overall assessment.

An overarching observation about the tensions that this superintendent experienced and how he embraced them centres on the uses of data. Lanoue offered a metaphor to help the principals understand that in school improvement where change and transformations need to occur there is a lag time in which change first sleeps, then creeps and then leaps; however, this metaphor also applied across the tensions. Lanoue referenced the use of data to be able to assess both the ‘perception’ and the ‘reality’ of principal evaluations related to their overall performance. The effective and accurate evaluation of principal performance is dependent on many variables. The importance and contribution of each variable to a

principal's overall rating is usually based on the professional judgement of the superintendent charged with conducting the evaluation. The discussion of the results attempted to shed light into the overall tensions this superintendent experienced throughout the process of evaluating principals. The superintendent held the belief that 'tensions are to be embraced as they shed light' and in this context, on the complexities involved in supervising and evaluating principals related to a host of variables 'that must be contextualised' for each principal and his/her school site.

The tensions experienced by this superintendent were in many ways interrelated as discussed in the findings. The interrelated nature of these tensions brings forward the idea that like school improvement, the evaluation of principals is a messy, recursive process in which there are no absolutes. Within these complexities reside the tensions. There was uniformity in the implementation of the processes and procedures found in the district policies about principal evaluation. However, each evaluation was individualised, tailored by such factors as longevity in the present principalship, overall years in the principalship, and the professional learning offered by the central office during the monthly principal learning community in which principals participated in a full day of activities focused on leadership development. By embracing this emphasis on principal growth and development, the tensions associated with myriad uses of data gave way to a concerted effort to increase the skills the principals needed to monitor the overall instructional programme.

As one would expect, the superintendent experienced, as do most evaluators, the tug-of-war between the formative and the summative aspects of principal evaluation to 'promote growth and development', to focus on 'what's most important – kids learning' and to 'see the relationship between what the principal does at the site and how these [principal] behaviors either enhance or impede school improvement'.

Given the proliferation of data available on any given school related to student achievement, there are numerous other sources of data that must also be examined. In this school system, data points ranged from observation and walk-through data designed to measure the quality of the teachers in a principal's school, to student achievement data and data illustrating the ability of the principal to meet the professional development needs of his/her teachers, and the list could continue to include climate surveys, pre- and post-test results and the overall school-improvement plan which would include all of these data points and many more. The superintendent had an expectation that leaders would be in a position to drill deeper finding root causes of problems, examine alternative actions/strategies, implement and assess, repeating a cycle of drilling deeper into data and their meaning related to the leadership strategies the principals used to monitor the school-improvement plans.

This superintendent was able to embrace tensions, grappling with questions that often illustrated the issues and their underlying associations surrounding the evaluation of principals. The recurring theme was the tensions associated with principal evaluation centred on how much data to factor into the evaluation and overwhelmingly, that a superintendent must be both knowledgeable about each school and even more so about the strengths and weaknesses of each principal.

This superintendent openly embraced tensions during his conversations with principals, and then more reflectively during the debriefing sessions, and he was all about the conversations and how these conversations strengthened his relationships with his principals. At the heart of these conversations about school and principal performance were the translations of data applied to the leadership practices associated with results as

a mechanism to help principals make sense of the changes needed in the practices that were impeding students and the school from reaching targeted improvement. However, these conversations allowed the superintendent to affirm leadership behaviours that could be linked to student achievement.

This superintendent was able to individualise the evaluation process with each principal through one seemingly simple but complex strategy – having the right conversations tailored to the context of the school and the system and the overall ability to identify areas to focus more attention. The superintendent framed data points around the performance areas and elements found within the elements of the Clarke County School District Leader Evaluation System. The right conversations included a series of questions that helped the principal to frame his or her thinking through the discourse with the superintendent. He was able to frame discussions with the big picture in mind, and by doing so, he was able to engage in larger discussions about the totality of needs in the CCSD and the leadership skills needed for the principal to be a part of the processes of school improvement at the site and across the system.

### Concluding thoughts

In anything worthwhile, tension can be debilitating or empowering. By embracing the tensions in the evaluation of the principal, the tensions become a means to an end and system-wide trend data indicated student achievement. In this context, the tensions served to boost performance with marked increases in student achievement. The tensions, when not embraced, can ironically become the focus and in doing so, the task of evaluating the principal can get lost in the work not necessarily related to building school, system and leader capacity. This would then lead to the principal evaluation being clouded to the point of losing its effectiveness as a tool to support leaders in ‘getting better at what they do’.

Keeping in this spirit, superintendents need to carefully craft questions to ensure principals understand the work and have a level of ownership in how this work should be accomplished – and each principal’s *quality* of work. Superintendents need to balance this approach with wise uses of data. The biggest lesson we learned through examining these tensions was the high priority that this superintendent placed on principal evaluation emanated ‘one conversation at a time’ to developing capacity for the system through honest conversations with a focus on the present but an eye moving to the future.

### Notes

1. Following federal regulations, the State of Georgia changed the method of calculating graduation rates beginning with the 2012 school year from the ‘Leaver Method’ (those who leave) to what is now called the Cohort Method (those who stay to graduate in four years with their beginning cohort). In 2009, the state graduation ‘leaver rate’ was 78.8% and CCSD’s graduation ‘leaver rate’ was 63.3%, a 15.5% difference. The district focus was to ensure that every student would graduate in four years or for some students five years. Both calculations are used in the new Cohort Method. In 2013, the state cohort graduation rate for five-year students was 71.6% and the cohort graduation rate for five-year students in CCSD was 71.4%, representing less than 1% difference from the state average. The CCSD closed the gap from the state average by over 14%. In 2012, CCSD exceeded the state graduation rate for four-year students for the first time since the implementation of the NCLB.
2. The College and Career Ready Performance Index (CCRPI) is a recently instituted school improvement and accountability index designed to assess college and career readiness for

Georgia public school students. The CCRPI is a score primarily composed of student achievement (70%), including standardised tests performance, and also other factors such as student growth (15%) and the school's ability to close the achievement gap (15%). Under NCLB, minimum achievement targets, which increase each academic year, are set as benchmark values to determine if schools have successfully made adequate yearly progress (AYP). While examinations used to determine AYP vary by state, this determination is based on state-standardised examinations.

3. Under HB 251, students eligible for transfer under the unsafe school choice option and students in needs improvement schools that must offer public school choice under NCLB must get first priority for available seats at those schools in the district that are not in needs improvement (2009, Georgia Education News, Georgia HB 251, para 5).
4. The Criterion-Referenced Competency Tests are standardised examinations geared at determining the level of mastery of students of the state curriculum for which they were designed. These tests are in reading, English/language, arts, mathematics, science and social studies and are used to assess the quality of education within the state.

### Notes on contributors

Sally J. Zepeda is a Professor at the University of Georgia in the Department of Lifelong Education, Administration, and Policy, and she is a Fellow in the Institute for Interdisciplinary Research in Education and Human Development. Dr. Zepeda's research spans teacher and leader evaluation, instructional supervision, and professional development as supports to organisational and leadership development. She was the inaugural Master Professor for the University Council for Educational Administration.

Philip D. Lanoue is the Superintendent in the Clarke County School District and was recently named a finalist for Georgia Superintendent of the Year by the Georgia School Superintendents Association. Dr. Lanoue was named as one of the Top 50 Technological Innovators in Education by the Center for Digital Education and Converge.

Noris F. Price was the Deputy Superintendent in the Clarke County School District and now is the Superintendent of Baldwin County Schools. Dr. Price serves on the advisory board for the Principal's Center at Georgia State University and is a co-principal investigator for a Race to the Top grant.

Albert M. Jimenez is a Research Professional with the Georgia Center for Assessment at the University of Georgia. Dr. Jimenez's research interests include teacher and leader evaluation, test and observation instrument construction, and item response theory.

### References

- Amsterdam, C. E., R. L. Johnson, D. M. Monrad, and S. L. Tonnsen. 2003. "Collaborative Approach to the Development and Validation of a Principal Evaluation System: A Case Study." *Journal of Personnel Evaluation in Education* 17 (3): 221–242. doi:10.1007/s11092-005-2981-y.
- Busher, H., and A. Harris. 1999. "Leadership of School Subject Areas: Tensions and Dimensions of Managing in the Middle." *School Leadership & Management* 19 (3): 305–317. doi:10.1080/13632439969069.
- Catano, N., and J. H. Stronge. 2006. "What Are Principals Expected to Do? Congruence between Principal Evaluation and Performance Standards." *National Association of Secondary School Principals Bulletin* 90 (3): 221–237. doi:10.1177/0192636506292211.
- Clifford, M., and S. Ross. 2011. *Designing Principal Evaluation Systems: Research to Guide Decision-making – Executive Summary of Current Research Prepared for the National Association of Elementary School Principals*. Baltimore, MD: Johns Hopkins University.
- Conklin-Spillane, C. 2012. "Potential Impact of Controversial New Teacher & Principal Evaluation System." *River Journal*. <http://www.riverjournalonline.com/schools/2271-potential-impact-of-controversial-new-teacher-a-principal-evaluation-system.html>
- Corbin, J., and A. Strauss. 2008. *Basics of Qualitative Research*. 3rd ed. Thousand Oaks, CA: Sage.

- Cordeiro, P. A., and W. G. Cunningham. 2013. *Educational Leadership: A Bridge to Improved Practice*. 5th ed. Boston, MA: Allyn and Bacon.
- Darling-Hammond, L. 2007. "Standards, Accountability, and School Reform." In *Facing Accountability in Education: Democracy and Equity at Risk*, edited by C. Sleeter, 78–111. New York: Teachers College Press.
- Gilliat-Ray, S. 2011. "‘Being There’: The Experience of Shadowing a British Muslim Hospital Chaplain." *Qualitative Research* 11 (5): 469–486. doi:10.1177/1468794111413223.
- Gottfried, M. A., B. M. Stecher, M. Hoover, and A. Brown Cross. 2011. *Federal and State Roles and Capacity for Improving Schools*. Santa Monica, CA: The RAND Corporation. [http://www.rand.org/pubs/technical\\_reports/TR989.html](http://www.rand.org/pubs/technical_reports/TR989.html)
- Gurr, D., L. Drysdale, and B. Mulford. 2006. "Models of Successful Principal Leadership." *School Leadership & Management* 26 (4): 371–395. doi:10.1080/13632430600886921.
- Hallinger, P. 2003. "Leading Educational Change: Reflections on the Practice of Instructional and Transformational Leadership." *Cambridge Journal of Education* 33 (3): 329–351. doi:10.1080/0305764032000122005.
- Hallinger, P., and R. Heck. 1998. "Exploring the Principal’s Contribution to School Effectiveness: 1980–1995." *School Effectiveness and School Improvement* 9 (2): 157–191. doi:10.1080/0924345980090203.
- Harris, A. 2002. *School Improvement: What’s in It for Schools?* New York: Routledge Falmer.
- Harris, A., C. Day, D. Hopkins, M. Hadfield, A. Hargreaves, and C. Chapman. 2002. *Effective Leadership for School Improvement*. London: Routledge.
- Heck, R., and P. Hallinger. 2009. "Assessing the Contribution of Distributed Leadership to School Improvement and Growth in Math Achievement." *American Educational Research Journal* 46 (3): 659–689. doi:10.3102/0002831209340042.
- Honig, M. I. 2013. *From Tinkering to Transformation: Strengthening School District Performance. Education Outlook, 4*. Washington, DC: American Enterprise Institute. <http://education.washington.edu/areas/edlps/profiles/faculty/honig/2013%20HONIG%20AEI%20Outlook%20From%20Tinkering%20to%20Transformation.pdf>
- Honig, M. I., M. A. Copland, L. Rainey, J. A. Lorton, and M. Newton. 2010. *Central Office Transformation for District-wide Teaching and Learning Improvement*. Washington, DC: Center for the Study of Teaching and Policy, University of Washington.
- Jackson, D. S. 2000. "The School Improvement Journey: Perspectives on Leadership." *School Leadership & Management* 20 (1): 61–78. doi:10.1080/13632430068888.
- Leithwood, K., and D. Jantzi. 2005. "A Review of Transformational School Leadership Research: 1996–2005." *Leadership and Policy in Schools* 4 (3): 177–199. doi:10.1080/15700760500244769.
- Leithwood, K., and D. Jantzi. 2000. "The Effects of Transformational Leadership on Organizational Conditions and Student Engagement with School." *Journal of Educational Administration* 38 (2): 112–129. doi:10.1108/09578230010320064.
- Leithwood, K., K. S. Louis, S. Anderson, and K. Wahlstrom. 2004. *How Leadership Influences Student Learning*. New York, NY: Report to the Wallace Foundation. <http://www.wallacefoundation.org/knowledge-center/school-leadership/key-research/Pages/How-Leadership-Influences-Student-Learning.aspx>
- Lofland, J., and L. Lofland. 1999. "Data Logging in Observation: Fieldnotes." In *Qualitative Research*, edited by A. Bryman and R. Burgess, 3–12. London: Sage.
- Lohman, J. 2010. "Comparing No Child Left Behind and Race to the Top." OLR Research Report R-0235. <http://www.cga.ct.gov/2010/rpt/2010-R-0235.htm>
- Lortie, D. C. 2009. *School Principal: Managing in Public*. Chicago, IL: University of Chicago Press.
- Louis, K. S., K. Leithwood, K. Wahlstrom, and S. Anderson. 2010. *Investigating the Links to Improved Student Learning: Final Report of Research Findings*. New York, NY: Wallace Foundation.
- MacNeil, A. J., D. L. Prater, and S. Busch. 2009. "The Effects of School Culture and Climate on Student Achievement." *International Journal of Leadership in Education: Theory and Practice* 12 (1): 73–84. doi:10.1080/13603120701576241.
- Marshall, C., and G. B. Rossman. 2006. *Designing Qualitative Research*. Thousand Oaks, CA: Sage.

- Mattingly, M. 2003. *A Study of Superintendents' Practices of Principal Supervision and Evaluation: A Contrast of Low Performing and Performing Schools*. Athens, GA: University of Georgia.
- McDonald, S. 2005. "Studying Actions in Context: A Qualitative Shadowing Method for Organizational Research." *Qualitative Research* 5 (4): 455–473. doi:10.1177/1468794105056923.
- Mendels, P. 2012. "The Effective Principal." *The Journal of Staff Development* 33 (1): 48–52.
- Merriam, S. B. 1998. *Qualitative Research and Case Study Applications in Education*. San Francisco, CA: Jossey-Bass.
- No Child Left Behind Act of 2001 (NCLB), Pub. L. No. 107-110, § 115, Stat. 1425. 2002.
- Normore, A. H. 2004. "Performance Appraisals of School Administrators in One Canadian School District: A Contemporary Model." *Alberta Journal of Educational Research* 50 (3): 283–298.
- Normore, A. H. 2005. "Integrating Personnel Evaluation in the Planning and Evaluation of School Improvement Initiatives." *American Journal of Evaluation* 26 (3): 348–351. doi:10.1177/1098214005278757.
- Normore, A. H. 2010. *The Development, Preparation, and Socialization of Leaders of Learning and Learners of Leadership: A Global Perspective*. Bingley, UK: Emerald Group.
- NSBA. 2013. "Reauthorization of the Elementary and Secondary Education Act (Issue Brief)." <http://www.nsba.org/Advocacy/Key-Issues/NCLB>
- Onwuegbuzie, A. J., N. L. Leech, and A. M. T. Collins. 2008. "Interviewing the Interpretive Researcher: A Method for Addressing the Crises of Representation, Legitimization, and Praxis." *International Journal of Qualitative Methods* 7 (4): 2–17.
- Parylo, O., S. J. Zepeda, and E. Bengtson. 2012. "Principals' Experiences of Being Evaluated: A Phenomenological Study." *Educational Assessment, Evaluation, and Accountability* 24 (3): 216–238. doi:10.1007/s11092-012-9150-x.
- Patton, M. Q. 2002. *Qualitative Research & Evaluation Methods*. 3rd ed. Thousand Oaks, CA: Sage.
- Petersen, G. J. 2002. "Singing the Same Tune: Principals' and School Board Members' Perceptions of the Superintendent's Role as Instructional Leader." *Journal of Educational Administration* 40 (2): 158–171. doi:10.1108/09578230210421114.
- Peterson, K. D., J. Murphy, and P. Hallinger. 1987. "Superintendents' Perceptions of the Control and Coordination of the Technical Core in Effective School Districts." *Educational Administration Quarterly* 23 (1): 79–95. doi:10.1177/0013161X87023001006.
- Ponticell, J. A., and S. J. Zepeda. 2004. "Confronting Well-learned Lessons in Supervision and Evaluation." *The NASSP Bulletin* 88 (639): 43–59. doi:10.1177/019263650408863905.
- Riley, K., and J. MacBeath. 2003. "Effective Leaders and Effective Schools." In *Leading and Managing for Effective Education: Effective Educational Leadership*, edited by N. Bennett, M. Crawford, and M. Cartwright, 173–185. London, UK: Paul Chapman.
- Robinson, V. M. J., C. A. Lloyd, and K. J. Rowe. 2008. "The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects of Leadership Types." *Educational Administration Quarterly* 44 (5): 635–674. doi:10.1177/0013161X08321509.
- Spanneut, G., and M. Ford. 2008. "Guiding Hand of the Superintendent Helps Principals Flourish." *Journal of Staff Development* 29 (2): 28–33.
- Suraya, W. H., and J. N. Yunus. 2012. "Principal Leadership Styles in High-academic Performance of Selected Secondary Schools in Kelantan Darulnaim." *Journal of Independent Research and Studies* 1 (2): 57–67.
- United States Department of Education. 2009. "Race to the Top Program Executive Summary." <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>
- Wallace Foundation. 2009. *Assessing the Effectiveness of School Leaders: New Directions and New Processes*. A Wallace Foundation Perspective, 7–8. New York, NY: Wallace Foundation. <http://www.wallacefoundation.org/knowledge-center/school-leadership/principal-evaluation/Documents/Assessing-the-Effectiveness-of-School-Leaders.pdf>
- Wallace Foundation. 2013a. *The School Principal as Leader: Guiding Schools to Better Teaching and Learning*. Expanded ed. New York, NY: Wallace Foundation. <http://www.wallacefoundation.org/knowledge-center/school-leadership/effective-principal-leadership/Pages/The-School-Principal-as-Leader-Guiding-Schools-to-Better-Teaching-and-Learning.aspx>

- Wallace Foundation. 2013b. *District Matter: Cultivating the Principals – Urban Schools Needed*. New York, NY: Wallace Foundation. <http://www.wallacefoundation.org/knowledge-center/school-leadership/district-policy-and-practice/Pages/Districts-Matter-Cultivating-the-Principals-Urban-Schools-Need.aspx>
- Waters, J. T., and R. J. Marzano. 2006. *School District Leadership that Works: The Effect of Superintendent Leadership on Student Achievement*. Denver, CO: Mid-Continent Research for Education and Learning.
- Yarbrough, D. B., L. M. Shulha, R. K. Hopson, and F. A. Caruthers. 2011. *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users*. 3rd ed. Thousand Oaks, CA: Sage.
- Yin, R. K. 2003. *Case Study Research: Design and Methods*. 3rd ed. Newbury Park, CA: Sage.
- Zepeda, S. J. 2006. "High Stakes Supervision: We Must Do More." *The International Journal of Leadership in Education* 9 (1): 61–73. doi:10.1080/13603120500448154.
- Zepeda, S. J. 2013. *The Principal as Instructional Leader: A Practical Handbook*. 3rd ed. New York, NY: Routledge.
- Zepeda, S. J., E. Bengtson, and O. Parylo. 2012. "Examining the Planning and Management of Principal Succession." *Journal of Educational Administration* 50 (2): 136–158. doi:10.1108/09578231211210512.