Southeast Journal of Educational Administration

Volume 16
Issue 1
Spring 2016

A referred journal sponsored and published by the
Southern Regional Council on Educational Administration

Southeast Journal of Educational Administration

Volume 16, Issue 1
Spring 2016
ISSN 2689-307X

Angela Farmer, Editor
Mississippi State University

Contents

Female Principals Leading in 21st Century Urban Schools: Instructional Leadership, Supervision and Evaluation in the Era of Accountability

Lavetta Henderson
Florida A&M University
Sheila D. Moore
Florida A&M University

Abstract

This study examined the leadership practices of female principals in a 21st century urban school district. The purpose of this study was to describe the present status of instructional leadership, teacher supervision and evaluation in the era of No Child Left Behind (NCLB) as experienced by ten female school principals in a Florida school district. The study compared the findings from the literature in the areas of instructional leadership, supervision and evaluation. The literature cited focused on the ways that female principals enacted the role of an instructional leader when supervising and evaluating teachers. The study took the form of a case study in order to provide a detailed description of the leadership practices of female principles in a single school district in the state of Florida. Interview questions were constructed based on the research question. Each interview was transcribed and content analysis was employed to identify commonalities in the data. Common themes were identified for the research question based on the responses of the principals. The study revealed profound consistency between the information cited in the literature and the information reported by the ten principals in the areas of instructional leadership, supervision and evaluation. The study also revealed the potential conflicts between the beliefs of the principals and the NCLB legislation and the effects of NCLB on the practices of the principals.

Keywords: female leadership, instructional leadership, supervision and evaluation

Southeast Journal of Educational Administration, Vol. 16, No. 1—Spring 2016
ISSN 2689-307X © 2016 Southern Regional Council on Educational Administration

Female Principals Leading in 21st Century Urban Schools

No Child Left Behind (NCLB) altered the landscape of public education and transformed the practice of school principals. Principals are a key element in school improvement efforts. The emphasis on accountability brought on by the No Child Left Behind legislation, required principals to work in prevention reality. They worked to prevent their schools from being placed on a warning list or school improvement plan for failing to meet adequate yearly progress (AYP) that NCLB mandated. There was an era when principals had time to experiment with innovative programs; however, now principals were bound by the considerations of scientifically based research when determining what programs to implement. Principals once could look at the individual strengths and weaknesses of students to gauge their progress in the school's curricula, now they must analyze student progress based on the results of standardized test scores and their contribution to meeting AYP.

The principal's work with teachers was also impacted by NCLB since schools were evaluated and possibly disciplined based on student achievement. Kaplan and Owings (2004) observed that improving teacher effectiveness was the center of educational reform. Increasingly, research confirmed that teacher and teaching quality are the most powerful predictors of student success. Teacher effectiveness was cited as one of the most decisive factors in student achievement, (Kennedy, Peters, & Thomas, 2012; RAND Corporation, 2012; Sanders, 2004; Wenglinsky, 2002). After nearly 30 years of revamping increased graduation requirements, curriculum standards, and high-stakes testing, stakeholders concluded that unless changes occurred inside the classroom with improved teaching and learning, educators cannot prepare all students for proficiency in advanced education and work. In short, principals enabled higher student achievement by assuring better teaching (Wenglinsky, 2002). Due to the increasing demands on the principal's time for both instructional and managerial activities, as well as the ever-present threat of NCLB sanctions, one questions if the gender of the principal is a determining factor in his or her school's success.

This study was based on the assumption that a principal's gender may influence his or her practices as an instructional leader. The literature cited focused on the ways that female principals enact the role of instructional leaders when supervising and evaluating teachers. How the gender of the principal influenced instructional leadership, supervision and evaluation was a key study since teacher quality was a critical factor in the accountability movement.

Characteristics of Female Principals

Researchers depicted the lives of female principals (Stanley, 2002; Growe & Montgomery, 2001; Shakeshaft, 1989) as being child and achievement centered, more focused on teaching and learning, motivated by building and maintaining relationships, and more visible in schools.

In her extensive study of female administrators, Shakeshaft 1989 found that "women possess characteristics that are conducive to good schooling" (p. 200). She found that female principals focus on instructional and educational issues. In addition, female principals stressed achievement within a supportive environment. Other indicators of female principals were that

they stressed cooperation, facilitated vision into action, and monitored and intervened more than men. Shakeshaft and Nowell (2000) suggested that female principals evaluated student progress more frequently than men, managed more orderly schools, and encouraged participation in decision-making.

As a result of her research, Shakeshaft described schools headed by a female as child centered, small, nonhierarchical, and marked by shared decision making. She also concluded that the style of female principals was motivated by a focus on building community, establishing relationships, and improving teaching and learning. Shakeshaft added that female principals spend more time interacting with others in order to improve their schools. In describing the female principal's work day, she noted, "Women spend more time with people, communicated more, care more about individual differences, are concerned with teachers and marginal students, and motivated more" (Shakeshaft, 1989 p. 197). Shakeshaft concluded that female principals viewed their role as master teacher or educational leader.

Conceptual Underpinnings

Instructional Leadership

Smith and Andrews (2008) identified four roles for an instructional leader: resource provider, instructional resource, communicator, and visible presence. In the role of instructional resource, "the instructional leader supervises the staff, using strategies that focus on the improvement of instruction" (p. 14). They added that when a principal displays strong instructional leadership, "Teacher evaluation is characterized by frequent classroom visitation, clear evaluation criteria, and feedback" (p. 8). In addition, Blasé and Blasé (2007) concluded that positive instructional leadership improved teacher performance and student learning.

Although there are many roles involved in instructional leadership, the primary focus of a principal was to improve teaching and learning through completing formal and informal observations and discussing teaching and learning with teachers. These are two important activities in instructional leadership that affected teaching and learning. By being visible in classrooms via informal and formal observations, the instructional leader can improve teaching and learning through supervision. First, the instructional leader can monitor the curriculum. Wiles and Bondi (2000) contended that "the primary purpose of instructional leadership is to improve classroom teachers' link between the planned curriculum and the learning experienced by the student" (p. 234). In order to achieve this purpose, the principal must have extensive knowledge of learning and teaching theory to recognize what was seen and not seen in classrooms. The knowledge principals hold must include understanding of planning elements, learning activities, and evaluation (Wiles & Bondi, 2000).

By being visible in classrooms, principals used their knowledge of learning and teaching theory to improve instruction. This required the principal to stay up to date with the latest research on teaching and learning. Visibility in classrooms allowed principals to motivate teachers, monitor instruction, be accessible, provide support, and keep informed (Blasé & Blasé, 2007). Therefore, being visible in classrooms—an essential component of instructional leadership—enabled principals to help improve teaching and learning. By discussing teaching

and learning with teachers through informal conversations and post-observation conferences, principals used their role as supervisor to improve teaching and learning. Mainly, principals shared their knowledge of teaching and learning theory. Since "most teachers expand their teaching range only with carefully designed support and assistance," the conference was a pivotal element in improving student achievement (Blasé & Blasé, 2007, p. 19). Next, supervisors accomplished many tasks through conferencing. The conference allowed the principal to give feedback, model good instruction, use inquiry, make suggestions, and solicit advice and opinions about instruction (Blasé & Blasé, 2007; Springer, 2006).

Florida Instructional Evaluation System

The supervision and evaluation of teachers was a key task for principals which became increasingly important under the mandates of NCLB. States were more involved in the evaluation process. The Florida Department of Education (FLDOE) reviewed and approved each school district's instructional personnel evaluation systems and any substantial revisions subsequent to initial approval. In addition, the Department assisted districts in monitoring the fidelity of implementation of each district's evaluation system for compliance with the law. The FLDOE approved the following Teacher Evaluation Instructional Practice Models: State Model based on the research and meta-analyses of Robert Marzano, and the Danielson Model based on the research of Charlotte Danielson (http://www.fldoe.org/profdev/adpes.asp).

Florida Statute 1012.34 (2011) mandated that evaluation systems support effective instruction and student learning growth, and that performance evaluation results must be used when developing district and school level improvement plans. This rule also dictated that evaluation performance levels must differentiate among four levels:

1. Highly Effective;
2. Effective;
3. Needs improvement or, for instructional personnel in the first 3 years of employment who need improvement, Developing; and
4. Unsatisfactory.

According to the rule, performance evaluations were to be based upon sound educational principles and contemporary research in effective educational practices in three major areas: Performance of Students, Instructional Practice, and Professional and Job Responsibilities. In measuring performance of students, Florida utilized a value-added model that measured the impact of a teacher on student learning, by accounting for other factors that may impact the learning process. The value-added model considered a student's prior performance, current performance, and predicted performance. An advantage of value-added models is that they leveled the playing field by accounting for differences in the proficiency and characteristics of students assigned to teachers. The value-added model also recognized that there was an independent factor, the school component, that impacted student learning which is taken into account when "leveling the playing field" (American Institutes for Research, 2011). The Florida Value-Added Model is one part of a multi-faceted teacher evaluation system in the state.

Limitations

Only elementary and middle school female African American principals in a local school district in the state of Florida were included in this study, and results may not be generalizable to other school districts in the state of Florida, other states, and nations. However, the findings may have relevance to researchers in these settings as they examine leadership practices in schools and foster student performance.

Research Question

The following research question guided this study: How do female principals construct their roles as instructional leaders, supervisors and evaluators of teachers in the era of No Child Left Behind?

Methodology

Research Design

Since the few studies that exist on the supervision and evaluation of teachers conducted by female principals are dated, qualitative research was appropriate. Qualitative research was also appropriate for this study because of its intent to describe the meaning female principals ascribed to their experiences as instructional leaders and in the supervision/evaluation process. This study took the form of a case study. According to Bogdan and Biklen (1998), "A case study is a detailed examination of one setting or a single subject ..." (p. 54). Yin (2003), noted that the case study was the best form to use when seeking the answers to "how" and "why" questions. The case study was most appropriate when the researcher seeks to provide a detailed description of a single subject using a variety of sources, such as interviews, observations, and documents (Yin, 2003; Bogdan & Biklen, 1998; Merriam, 1998; McMillan & Schumacher, 1997). It brought new discoveries, broadened experiences, and confirmed previous knowledge (Merriam, 1998). In education, the case study allowed specific problems to be articulated and defined (Merriam, 1998).

Interviewing is an appropriate technique when behavior and feelings cannot be observed directly, when past events are being studied, and "when conducting case studies of a few selected individuals" (Merriam, 1998, p. 72). Yin (2003) adds, "Interviews are an essential source of case study evidence because most case studies are about human affairs" (p. 92). An initial interview occurred with each principal to collect baseline data; a second interview occurred after the researcher reviewed the documents and transcripts; final interviews seeking clarification occurred during the writing of the case study.

Interview questions were constructed to answer the research question. Interviews in qualitative studies were usually more open ended and less structured and flowed like conversations (Yin, 2003; Merriam, 1998). "The largest part of the interview is guided by a list of questions or issues to be explored, and neither the exact wording nor the order of the questions is determined ahead of time," advised Merriam (1998, p. 74). Two interview protocols were created: initial questions derived from the literature for female principals, and follow-up

questions derived from the interview transcripts of the principals. Questions were modified when necessary to become clearer to the audience. Additional probing questions were devised to gain deeper responses. Interviews were conducted, taped, and transcribed with the participants. The transcripts became a data source for future analysis (Merriam, 1998). Each transcript was sent to the appropriate participant to ensure accuracy. One participant made subtle changes in wording to her transcript.

Data Analysis

Each interview, along with related documents from the site, was analyzed after it was conducted: "The right way to analyze data in a qualitative study is to do it simultaneously with data collection" (Merriam, 1998, p. 162). This process enabled the researcher to pose clarifying questions when needed and helped to develop common and uncommon patterns in the data.

After each interview was transcribed, the data were examined for common patterns and irregularities and coded based on those patterns since "typically, qualitative research findings are in the forms of themes, categories, typologies, and concepts" (Merriam, 1998, pp. 7-8). Content analysis was employed to review the data and categorize it according to the research questions. Content analysis enabled large amounts of data to be reduced into smaller chunks to create meaning (Weber, 1990). Merriam (1998) noted, "Devising categories is largely an intuitive process, but it is also systematic and informed by the study's purpose, the investigator's orientation and knowledge, and the meanings made explicit by the participants themselves (p. 179).

The data were initially coded as they related to specific research questions, however new categories emerged from the data (Merriam, 1998; McMillan & Schumacher, 1997). Initial categories were determined by the researcher after a review of the transcripts. Microsoft Word software was also used to find keywords in the transcripts based on the research questions. The search command was also used to determine the frequency of words in order to develop categories.

Context

One school system was selected to participate in the case study. The criteria for selection of female principals were that the district had at least three African American female principals and these principals had at least three years of administrative experience. Over 20,000 students attended the elementary, middle school, and high schools located in this Florida school district. The district served a diverse population of students from various social, economic, and cultural, backgrounds. The district described itself as offering a rigorous educational program to help each child maximize his or her potential to achieve success. The district has won many awards and was characterized by high achievement. The district's dedication to high achievement was directed by its strategic plan, which addressed goal areas in student achievement, curriculum and instruction, technology, global competence, collaborative leadership, and Common Core Standards. These areas have influenced the district's focus on instruction and student achievement, dedication to various forms of professional development, and its development of a teacher supervision and evaluation model.

Ten female principals participated in the study. The principals reported a range of 3–37 years of principal experience and were former teachers (see Table 1).

Table 1
2012–2013 Principal Profile as Reported by Principals

Participant	Years of Experience as Principal	Years of Experience as Principal at School	Years of Teaching Experience	Highest Degree	Ethnicity
Principal A	22	17	10	M.Ed.	African-American
Principal B	3	3	15	M.Ed.	African-American
Principal C	30	28	10	M.Ed.	African-American
Principal D	37	3	10	Ph.D.	African-American
Principal E	10	10	10	M.Ed.	African-American
Principal F	17	7	15	M.Ed.	African-American
Principal G	3	3	19	Ed.D.	African-American
Principal H	12	12	8	M.Ed.	African-American
Principal I	6	3	11	M.Ed.	African-American
Principal J	8	4	14	M.Ed.	African-American

In addition, all of the principals participating in this study reported that they served in schools that were high poverty and high minority (see Table 2). Also, the principals participating in this study were African American females (see Table 2).

Table 2
2012–2013 School Demographics as Reported by Principals

Principal	School Population	% Free/Reduced Lunch	Florida School Grade	Type of School	Student % Ethnicity
Principal A	469	83%	A	Elementary K-6	87% African American 9% Hispanic
Principal B	568	95%	A	Middle 6-8	93% African American 4% Caucasian
Principal C	284	69%	A	Elementary K-6	82% African American 13% Caucasian
Principal D	527	77%	B	Elementary K-5	99% African American
Principal E	398	82%	B	Elementary K-5	99% African American
Principal F	347	92%	A	Elementary K-6	86% African American 4% Hispanic 4% Asian 6% Caucasian
Principal G	346	58%	B	Elementary K-6	77% African American 21% Caucasian
Principal H	291	52%	A	Elementary K-6	92% African American 7% Caucasian
Principal I	247	59%	A	Elementary K-6	99% African American
Principal J	295	58%	B	Middle 7-8	99% African American

Results/Discussion/Conclusions

The purpose of this study was to describe the present status of instructional leadership, teacher supervision and evaluation in the era accountability as experienced by ten female school principals in a Florida school district compared to the findings from the literature in the areas of instructional leadership, supervision and evaluation. The study took the form of a case study because the researchers sought to provide a detailed description of a single subject using a variety of resources (Yin, 2003; Bogdan & Biklen, 1998; Merriam, 1998; McMillan & Schumacher, 1997). This particular case took place in a single setting, a school district in Florida. The researchers connected the practices experienced by the principals involved in the study with discoveries from the related literature to demonstrate similarities and differences between current practice and the practices as described in the literature. Common themes were identified for the research question based on the responses of the principals.

The research question asked: How do female principals construct their roles as instructional leaders, supervisors and evaluators of teachers in the era of No Child Left Behind? The district's supervision model required the principals to frequently observe in classrooms and interact with teachers to improve teaching and learning. Visibility in classrooms allowed supervisors to motivate teachers, monitor instruction, be accessible and provide support, and keep informed (Blasé & Blasé, 2007). The four domains of teaching/evaluation were explored, allowing teachers a fair and comprehensive evaluation. They observed the teachers not only in classroom settings, but in their interactions with parents and in professional development and committee work settings. They also supervised teachers by reviewing lesson plans and providing feedback to the teachers on these documents. Smith and Andrews (2008) explained that when a principal displays strong instructional leadership, "Teacher evaluation is characterized by frequent classroom visitation, clear evaluation criteria, and feedback" (p. 8). Principal sample responses follow:

Principal Three: I'm in and out [of classrooms] frequently in order to see what's going on. I may only stay ten minutes. I focus on a different area during different parts, and I try to align it with their goals.

Principal One: Most of my time is with the teachers, talking to them about issues when I'm walking around. If I feed and water the teachers, then they'll feed and water the kids.

Principal Seven: With this model, I am in and out of the teachers' rooms. We rarely have a set time for me to come and observe.

The following themes emerged: classroom expectations, defining teacher quality, collaborative supervision, providing feedback, promoting teacher growth, and visibility.

Classroom Expectations

The principals based what they expected to see teachers and students doing in classrooms on a student-centered constructivist learning environment and approach to curriculum. Although each principal was interviewed separately, their responses were remarkably similar when describing what they expect to see in classrooms. The common themes included interaction (between teacher and student, between student and student, and among the class), movement, use of a variety of resources, integration of subjects, focused inquiry, engagement, and emphasis on the whole child (social, academic, and emotional aspects). The principals examined the learning environment to see that classrooms were bright and inviting, with standards and student work posted on the walls. By emphasizing the whole child, the principals believe that a teacher's instructional delivery and the classroom learning environment should allow students to be happy to be in the class and to see learning as fun.

In addition to identifying what they expected to see in classrooms, the principals detailed what they hoped not to see. They did not want to see a lot of teacher talk or a majority of the class devoted to direct instruction. Instead, they preferred the teacher to act as facilitator and the children to view each other as experts. They wanted the students to know the goals of the lesson, why they were learning what they are learning, and how they would be assessed. Principal sample responses follow.

Principal Six: The classroom should not look the same at the beginning of the lesson as it does at the end of the lesson.

Principal Two: There needs to be a lot of checks for understanding with the kids during instruction, so that the teachers can get a real feel for where the children are individually and as a whole group.

Principal Ten: I expect to be able to go up to students and ask them what they're doing and have them tell me about the assignment, why they're doing it, and how they know if it's good work. I expect them to have criteria that they're using to judge their own work and take responsibility for it.

Principal One: The teacher should say 'This is the expectation as to why we are doing this. This is how it's connected to the standards.' Those are the kinds of things I'm looking for.

Defining Teacher Quality

The district's model for teacher supervision and evaluation, along with the principals' descriptions of what they expected to see in classrooms, were consistent with the research on what comprises teacher quality. Researchers identified the teacher's knowledge of content and pedagogy, the teacher's skills and classroom practices in delivering the curriculum, and the teacher's relationships with students and other members of the school community when defining teacher quality (Wenglinsky, 2002).

Collaborative Supervision

The district's model and the principals' resulting approach were consistent with the positive research associated with a collaborative approach to supervision. Collaborative supervision enabled teachers to evaluate their own instructional effectiveness (McBride & Skau, 1995) and to solve instructional problems (Nolan & Francisa, 2008; Beck, 2004; Zepeda, 2003). Collaboration placed responsibility for teacher learning and growth on the teacher and the principal equally (Grimmett, Rostad, & Ford, 2002). They believed that a teacher's practice can only improve when teachers play a central role in their evaluation process. Drago-Severson (2004) noted, "The central goal of reflective practice is improving one's teaching ... Creating a context wherein teachers are encouraged to engage in reflection promotes (and models) risk taking" (p. 105). Just as they hoped teachers would encourage their students to take risks and try new approaches to learning in their classrooms, the principals modeled this process in their interactions with teachers. Principal Two explains the process:

> [The goal] could be something that they do already and expand or learn more about. It's supposed to help them in their role as teacher. I usually let them pick an area of interest, and then I'll encourage them if I see something that they're doing that's great. I'll support that, and I'll want them to share that.

Whether the teacher was marginal or satisfactory, these principals believed that the teacher/principal relationship was critical to the teacher's improvement. They remembered their own classroom experiences as teachers, the good and the bad, and realize that teachers have these same good and bad days. Focusing on the positive allowed these principals to maintain relationships with teachers so that teaching and learning was constantly improving in their buildings. Shautz (2005) affirmed this view that the principals have of their teachers: "Female

principals generally viewed the teachers with whom they worked as being professional, dedicated individuals. Female principals placed a great deal of trust in teachers" (p. 212). Principal sample responses follow.

Principal Five: Because I bring fifteen years of classroom experience with me, I understand that sometimes your very best lesson is when no one is there. I do give them the opportunity to invite me in for something special, because sometimes I think teachers plan and work really hard and think, 'Oh I'd love for somebody to see this.' I give them that opportunity.

Principal One: A lot of my belief about supervision of teachers has more to do with the kind of relationship that you build with them than it has to do with the actual supervision.

Principal Nine: I think this is maybe a gender related way of feeling. You can say ten good things. You make one comment that's negative and that's what people focus on and feel badly about. So I don't want to do that.

Providing Feedback

The principals' focus on the positive in teacher performance characterized their verbal and written feedback to teachers. This feedback occurred not only after a formal observation but through informal interactions and conversations. The principals maintained visibility and made themselves available to teachers as part of their supervisory practices. This approach was consistent with what Nolan and Francis (2008) described as ideal: "Supervisors must see themselves not as critics of teaching performance, but rather as collaborators with teachers in attempting to understand the problems, issues, and dilemmas that are inherent in the process of learning and teaching" (p. 58).

Promoting Teacher Growth

The principals in this study reported the strategies they used to promote teacher growth as providing professional development, modeling teaching behaviors that they hoped to see in classrooms and working with the grade level teaching teams. The principals indicated that teacher learning and continued growth were the keys to student learning and growth. The principals' view of their role in promoting teacher growth was consistent with a study conducted by Drago-Severson (2004) in which she concluded: principals have a key role in supporting teacher learning and a responsibility to develop a clear vision of how school contexts can better support this learning; leadership supportive of teacher development made schools better places of learning for children; and schools needed to be places where the adults as well as the children are growing. Each principal devoted a substantial amount of time to informal and formal methods of professional development. Principal Two reports,

> This district that I'm in enables us to provide a lot of professional development activities in the district as well as in the building. As a result, there's this environment, a culture of learning, which is real important to have a safe, nurturing community for everyone.

Each principal had at least two professional development meetings a month in place of the weekly staff meeting, preferring to handle announcements via email. Literacy, math, and technology coaches were also available to help teachers connect their lessons to standards. Other

methods of professional development that the principals described were showing engagement in data analysis, videos of model instruction, encouraging teachers to observe each other, sharing what they see in classrooms, and distributing and discussing research. Blasé and Blasé (2007) noted that effective principals facilitated professional development, shared professional readings, and discussed teaching with teachers. The principals mentioned building wide, formal in-service experiences that were designed to promote teacher growth.

Visibility

All ten principals discussed the importance of visibility as part of their supervisory style. The principals had an open door, open calendar policy for teachers and were frequently out in their schools to promote accessibility. Principal Three noted that she kept her calendar available to her teachers so that they can easily schedule appointments. The principals indicated that their visibility allowed their teachers to view them as ready and available to help them solve problems, address instructional concerns, or just listen. Blasé and Blasé (2007) explained that being available and providing an open, friendly, supportive environment were keys to teacher empowerment. Principal One reported that this informal learning, where individual concerns were addressed, were critical to a teacher's development and indicated:

> The real learning and the real growth comes out of inspiring people to want to be better.... The real learning comes when they own it, they choose it, they know what it is that they want to learn, and I help facilitate their movement along that learning continuum.

The principals in this study reported that when not in formally scheduled meetings, they were roaming the halls in order to have opportunities to speak to teachers and students, observing informally in classrooms, and connecting personally with the teachers and students, practices consistent with those identified by Drago-Severson (2004). They indicated that their accessibility allowed their staff to view them as supports in the teaching and learning process, particularly if suggestions about performance needed to be made in the future.

Summary and Connections to Related Literature

Based on the data reviewed for this research question, the presentation was organized into the following categories through data analysis, as delineated by the interview questions posed to the principals: purposes of supervision, the school district's model of supervision and evaluation, classroom expectations, defining teacher quality, collaborative supervision, providing feedback, promoting teacher growth, and visibility. The principals in the study reported that they supervised teachers in a variety of ways including utilizing the district's evaluation structure to promote teaching and learning through the implementation of a learning teams, devoting staff meetings to professional development, sharing resources and their own educational expertise, visiting and observing in classrooms, being visible in the school, and meeting with teachers about educational issues. These methods were consistent with those articulated by Zepeda and Ponticell (1998), and Beerens (2010). The principals reported a belief that the supervisors of teachers should examine teachers through a variety of lenses, not just the lens of a single classroom observation. This belief was similar to the views of differentiated supervision espoused by Danielson and McGreal (2000) and Glickman (2002).

Conclusion

Working in a high performing school district created unique advantages and challenges to meeting the mandates in this era of accountability. The principals explained that the pressure to constantly improve achievement was ever present in their daily lives as administrators. This pressure was complicated by the high levels of achievement currently present in their schools. The principals reported that it was harder to attain continuous improvement and to motivate their staff to attain higher levels of achievement when their FCAT scores were well above state averages. The principals indicated that as their schools moved closer to one hundred percent proficiency as mandated in this era of accountability, it became harder and harder to achieve adequate yearly progress. The parental pressure that comes with working in a high performing district compounded the problem. Parents in this district, while supportive, expected their schools to be the best. Since their schools were high poverty, high minority, and high achieving, they had to continually find new ways to motivate their staff and convince them that they cannot just work to maintain the status quo.

As the principals continued to work with their staff to improve student achievement, each of them employed a particular approach to motivate their staff. One principal relied on constant visibility in hallways and classrooms to ensure that she was available for students and teachers as needed to discuss teaching and learning. She viewed these informal learning opportunities as ways to improve achievement. Another principal constantly showed caring with teachers and students so that they would do their best work. The sixth principal employed the committee structures in place in her building to build community and a shared vision of student success in her school. While all three of the principals utilized each of these approaches, the emphasis of each of them appeared to be related to their particular school context. Because she was new to her school, visibility helped one of the principals to bond with her staff. Since her school was perceived by some members of the community to be weaker than the other elementary schools, one of the principals used her caring to make the teachers and students feel special. As a result of her extended experience with her staff, another of the principals was able to rely on the committee structures to further improve her school.

Fortunately for this district, they had resources available to assist them in reaching the one hundred percent proficiency mandate. This school district closely emulated best practices as described in the research. Policymakers who are interested in ensuring that all schools are high performing would be wise to study these practices and include them in any revisions to NCLB or subsequent legislation. For instance, the district utilized models, such as differentiated supervision, which is associated with high achievement. Literacy, math, and technology coaches were employed to assist teachers in their classrooms. Professional development was a priority in this district so that the teachers can continually refine their skills. There was an abundance of supplies and resources for use in classrooms. The teachers in each school served as important resources as they often took on leadership roles; these roles assisted the principals in completing the many tasks they must perform. Finally, there was a great deal of parental involvement and support in the district. These advantages could help the principals as they strive to meet proficiency mandates.

Recommendations for Further Research

This study only focused on the perspectives of ten African American female elementary and middle school principals and their experiences in the areas of instructional leadership, supervision and evaluation in comparison to the literature and reality in the era of NCLB, a number of suggestions were made for further research. First, the study could be repeated with female middle school or high school principals and teachers. Since middle schools and high schools are often more content driven, than child centered, this perspective may yield different findings. The study could be repeated with male elementary principals. Since the interview questions were drawn from literature related to female principals, a study of how male elementary principals coped with the mandates of NCLB may also yield different results. A comparison study between male and female elementary principals and their methods of addressing NCLB would prove insightful. This study indicated if there were still differences and what kinds of differences in their approach to meeting the mandates of NCLB. A final area of study would be to repeat this study with female elementary principals and teachers in a school that is on a warning or school improvement list. Such a study would reveal the perspective of teachers and principals who feel a greater sense of urgency in meeting the requirements in this era of accountability.

References

American Institutes for Research (2011). *Florida's Value-Added Model*. Retrieved from http://www.fldoe.org/committees/pdf/air-intro-valueaddedmodel.pdf.

Beck, L. G. (2004). *Reclaiming educational administration as a caring profession*. New York, NY: Teachers College Press.

Beerens, D. R. (2010). *Evaluating teachers for professional growth: Creating a culture of motivation and learning*. Thousand Oaks, CA: Corwin Press.

Blasé, J., & Blasé, J. (2007). *Empowering teachers: What successful principals do*. Thousand Oaks, CA: Corwin Press.

Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative research in education: An introduction to theory and methods* (3rd ed.). Boston, MA: Allyn and Bacon.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.

Drago-Severson, E. (2004). *Helping teachers learn: Principal leadership for adult growth and development*. Thousand Oaks, CA: Corwin Press.

Florida Department of Education. (2013). *2012–13 Florida teacher evaluation instructional practice models*. Retrieved from http://www.fldoe.org/profdev/adpes.asp

Florida Department of Education. (n.d.). *Approved district performance evaluation systems*. Retrieved from http://www.fldoe.org/profdev/adpes.asp

Florida Statutes, 1012.34 (2011). *Personnel evaluation procedures and criteria*. Retrieved from http://www.flsenage.gov/laws/statutes/2011/1012.34

Glickman, C. D. (2002). *Leadership for learning: How to help teachers succeed*. Alexandria, VA: ASCD.

Grimmett, P. P., Rostad, O. P., & Ford, B. (2002). The transformation of supervision. In C. Glickman (Ed.), *Supervision in transition: 1992 yearbook of the Association for Supervision and Curriculum Development* (pp. 185-202). Alexandria, VA: ASCD.

Growe, R., & Montgomery, P. (2001). Women and the leadership paradigm: Bridging the gender gap. *National Forum*. Retrieved http://www.nationalforum.com/12growe.htm

Kaplan, L. S., & Owings, W. A. (2004). Introduction to special issue: Teacher effectiveness. *NASSP Bulletin*, *88*(638), 1-4.

Kennedy, K., Peters, M., & Thomas, M. (2012). *How to use value-added analysis to improve student learning: A field guide for school and district leaders*. Thousand Oaks: CA: Corwin.

McBride, M., & Skau, K. G. (1995). Trust, empowerment, and reflection: Essentials of supervision. *Journal of Curriculum and Supervision*, *10*(3), 262-277. Retrieved from http://www.ascd.org/otb/cs/95spring/mcbride.htm

McMillan, J. H., & Schumachu, S. S. (1997). *Research in education: A conceptual* i*ntroduction.* New York, NY: Longman.

Merriam, S. B. (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Jossey-Bass.

Nolan, J., & Francis, P. (2008). Changing perspectives in curriculum and instruction. In C. Glickman (Ed.), *Supervision in transition: 2008 yearbook of the Association for Supervision and Curriculum Development* (pp. 44-60). Alexandria, VA: ASCD.

RAND Corporation. (2012). *Teachers matter: Understanding teachers' impact on student achievement*. Santa Monica, CA: Author. Retrieved from http://www.rand.org/pubs/corporate_pubs/CP693z1-2012-09

Sanders, W. L. (2004, June). *A summary of conclusions drawn from longitudinal analysis of student achievement data over the past 22 years*. Paper presented to Governors Symposium, Asheville, NC.

Shakeshaft, C. (1989). *Women in educational administration* (2nd ed.). Newbury Park, CA: Corwin Press.

Shakeshaft, C., & Nowell, I. (2000). Gender and supervision. In *The Jossey-Bass reader on educational leadership* (pp. 257-266). San Francisco, CA: Jossey-Bass.

Shautz, D. (2005). Women supervisors have a greater understanding of what takes place in classrooms than men. *Education*, *116*(2), 210-214.

Smith, W. F., & Andrews, R. L. (2008). *Instructional leadership: How principals make a difference.* Alexandria, VA: Association for Supervision and Curriculum Development.

Springer, S. J. (2006). *Teacher perceptions of principal behavior related to student achievement* (Doctoral dissertation). Available from ProQuest Dissertation and Theses Global database. (UMI No. 9700419)

Stanley, R. B. (2002). *Relationships between gender and teachers' perceptions of principal effectiveness in Georgia schools* (doctoral dissertation). Available from ProQuest Dissertation and Theses Global database. (UMI No.3053925)

Weber, R. P. (1990). *Basic content analysis.* Newbury Park, CA: Sage Publications.

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performances. *Education Policy Analysis*, *10*(12).

Wiles, J., & Bondi, J. (2000). *Supervision: A guide to practice* (5th ed.). Upper Saddle River, NJ: Merrill.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Zepeda, S. J. (2003). *Instructional supervision: Applying tools and concepts*. Larchmont, NY: Eye on Education.

Zepeda, S. J., & Ponticell, J. A. (1998). At cross-purposes: What do teachers need, want, and get from supervision? *Journal of Curriculum and Supervision*, *14*(1), 68-87. Retrieved from http://www.ascd.org/otb/jcs/98fall/zepeda.html

Follow the Leaders: A Team Approach to MBTI

Angela Farmer
Mississippi State University

Abstract

"Nature or nurture"," instinct or learned behavior"," lead or follow", these are some of the debates regarding the nature of behavior. In the children's game of Follow the Leader, participants have to do what the leader says or does until such time as a new leader arises. Then, the new leader takes the helm. Similar innate behaviors are witnessed among flocks of birds, schools of fish, and troops of soldiers. In collegiate athletics, with groups of students these behaviors are also evident. The most successful teams appear to operate as one unit. This study examines how a leader, (in this case a collegiate coach's Myers Briggs Type Indicator [MBTI]) impacts variability in the selection of staff, coaches and players on the team in a Southeastern Conference university from that of a normal sample population.

Southeast Journal of Educational Administration, Vol. 16, No. 1—Spring 2016
ISSN 2689-307X © 2016 Southern Regional Council on Educational Administration

Follow the Leaders

Myers-Briggs Type Indicator (MBTI), is a personality inventory which adopts the theory of types first described by Dr. Carl Jung, and applies it to people's lives. The tool seeks to allow one to better understand how seemingly random variability in one's behavior is instead rather predictable and orderly. The behaviors are primarily due to innate differences in the ways in which individuals use their perception and judgment (Briggs Myers, McCaulley, Quenk, & Hammer, 2009). It was the author's intent that this analysis of type dynamics of a unique subgroup of athletes and coaches in a defined setting would lend clarity to the concept of team dynamics as it relates to type and the influence that an effective leader has upon a group's effective functionality.

According to Hunter (2006), assessment of psychological type allows "attention to student characteristics, needs, behaviors, and experiences." For student athletes to begin and maintain a successful career as both a student and an athlete, understanding themselves and others around them more fully is viewed as a positive action toward enhancing individual and team dynamics. As Coach Mike Krzyzewski, Coach of Duke University Men's Basketball once stated, "Talent is important. But the single most important ingredient after you get the talent is the internal leadership." Perhaps a link may be discovered that helps unveil why teams chosen by certain MBTI personalities may be heavily influenced by those individual types in order to be productive and effective dynamic presences on and off the field. Fortunately, sports participation allows for a plethora of "psychological, social, and development opportunities like learning to perform under pressure, dealing with adversity, developing self-confidence and decision-making strategies, and learning communication skills" (Williams & Krane, 2013).

Clearly, these multifaceted opportunities are companioned by a myriad of complex interactions within the essence of the team dynamic. Therefore, the MBTI insight may provide assistance in translating and streamlining actionable behaviors and pre-actionable behaviors to both avoid conflict and enhance program efficiency translating to an enhanced program effectiveness.

The underlying premise of Myers-Briggs Type Inventory is to recognize that while individuals have unique processes for consistent daily life, they typically follow predictable "polar configurations" (Sanborn, 2013). In referencing type theory, credits are to Carl Jung's theory, interpreted by Isabel Myers and Katharine Briggs as the MBTI personality inventory (Briggs Myers et al., 2009) whereby personality types are divested based on functional pairs.

These include one's preference for Introversion or Extraversion (I vs. E). The four basic mental functions detailed by MBTI scholars are Sensing (S) and its opposite, Intuition (N); and Thinking (T) and its opposite preference of Feeling (F). The final inventory is to select between Perception (P) and Judgment (J).

Individuals are assessed via a personality tool called, Form M (Briggs & Briggs Myers, 1998), administered by a certified MBTI trainer under controlled conditions, allowing the participants to select from a series of questions traits or characteristics which best describe their

personality. At the end of the survey, the participants quantitatively assess their strengths, arriving at a four lettered Type, aligned with one of the 16 MBTI Personality Types.

Individuals who prefer Introversion draw energy from the environment and internalize the experience, allowing them to focus on the internal state and think about things before discussing them. Conversely, individuals who prefer Extraversion are compelled to objects and individuals in the environment and prefer to "talk things out" (Briggs Myers et al., 2009).

The perception types, Sensing and Intuition, are differentiated by Jung's work as follows: Sensing preferences prefer to focus on the immediate experiences available to their five senses while Intuitive types prefer to perceive what is beyond immediately perceptible to the senses and include possible future opportunities. The judgment paradigm focuses on whether individuals prefer linking ideas together via logical connections, Thinking, or arriving at decisions based on values and merits of the decision, Feeling (Briggs Myers et al., 2009).

The final dichotomy reflects how participants prefer to orient themselves to the Outer World, rather as Judging or Perceiving. However, its analysis and evolution is the work of Katharine Briggs not Carl Jung. Its incorporation into Type Theory classification is essential to fully appreciate one's orientation toward the Outer World. Judgment types are seen as individuals who prefer seeing closure, planning operations or organizing activities. Perceiving types, conversely, are acclimated to incoming information upon which they may modify or change their opinion or resulting action (Briggs Myers et al., 2009).

To fully explain type dynamics for each of the 16 MBTI Personality Types is a work into itself; however, a general insight into the types is provided on Table 1.

Table 1
Contributions by Type (Briggs Myers et al., 2009, p. 38)

		Sensing with		Intuitive with	
		THINKING	FEELING	FEELING	THINKING
Introverts	Judging Types	ISTJ	ISFJ	INFJ	INTJ
		Concentration	Concentration	Concentration	Concentration
		Fact Focused	Fact Focused	Possibilities	Possibilities
		Logic	Warmth	Warmth	Logic
		Organization	Organization	Organization	Organization
	Perceiving Types	ISTP	ISFP	INFP	INTP
		Concentration	Concentration	Concentration	Concentration
		Fact Focused	Fact Focused	Possibilities	Possibilities
		Logic	Warmth	Warmth	Logic
		Adaptability	Adaptability	Adaptability	Adaptability
Extraverts	Perceiving Types	ESTP	ESFP	ENFP	ENTP
		>Interests	>Interests	>Interests	>Interests
		Fact Focused	Fact Focused	Possibilities	Possibilities
		Logic	Warmth	Warmth	Logic
		Adaptability	Adaptability	Adaptability	Adaptability
	Judging Types	ESTJ	ESFJ	ENFJ	ENTJ
		>Interests	>Interests	>Interests	>Interests
		Fact Focused	Fact Focused	Possibilities	Possibilities
		Logic	Warmth	Warmth	Logic
		Organization	Organization	Organization	Organization

A quantitative study resulted from the series of trainings conducted with the group and facets of the group over a period of four months. The Myers-Briggs instrument (Form M) was used to collect data from the participants. The participants included the following: the athletic team (comprised of 119 student athletes & coaches) and 1 head coach was administered the MBTI.

In June 2015, the MBTI (Form M) was given independently to each half of the team. Due to the size of the teams that the researcher needed to split the delivery of the instrument and the presentation regarding the use and limitations of the tool. All student athletes and coaching staff participated, including the head coach.

The resulting types were grouped, sub-grouped and analyzed based upon their particular team functions. Further dialog and additional information was shared with each facet to ensure that an enhanced appreciation for MBTI was developed as well as tools were understood for expansion of the participants range of skills to know themselves and one another better to enhance team synergy and effectiveness. For the research interest, the types were grouped and compared to national norms to evaluate whether this unique group, directly selected by a given MBTI type, would present as significantly variant from a normally selected population.

Observed types were compared as relevant percentages of the entire sample population as well as to the normal population percentiles. Differences were detailed as well. Finally, SRTT analyses were calculated whereby the observed population's relative percentages were divided by the normal or expected population's percentages. Detailed analyses are provided.

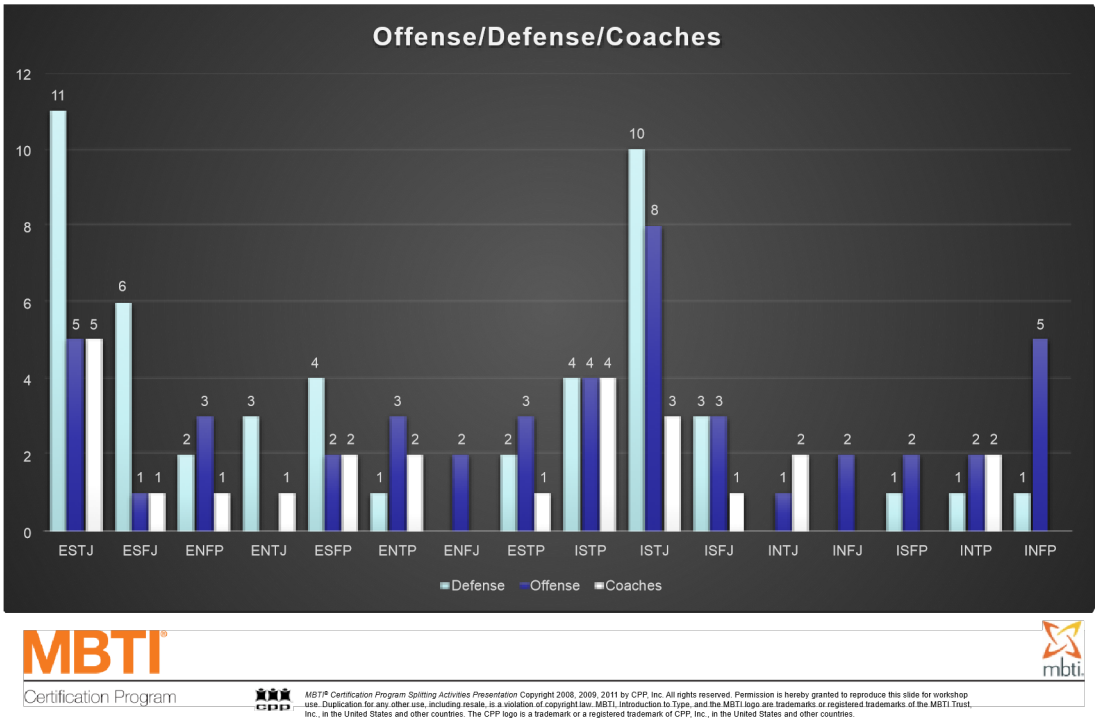
The researcher suspected that the leader's profile would influence the composition of the team and staff such that the types indicated would not be reflective of the typical percentages in the United States.

The results showed a variety of interesting outcomes, many of which were both remarkable and unanticipated. The MBTI trainer further anticipated that, given the athletic prowess required for participation in such a challenging sport, similar types would be found among like positional assignments. For statistical references, the *Introduction to Type and Leadership* by Sharon Lebovitz Richmond, (2008) was used exclusively. It is a valuable tool for MBTI certified practitioners and part of the CPP's exclusive tool kit.

The team dynamics listed below show the types preferred by the participants as a whole. It is paramount to note that the head coach reflected a MBTI of ISTJ or Introverted, Sensing, Thinking and Judging. This type is present in the normal population at a percentage of 11.6% and in an executive leadership position 15.2%.

Table 2
Team Presentation

Offense/Defense/Coaches



Differences in Myers-Briggs preferences for the group versus the total population are detailed. Comparisons are made based on data from the team, total population data, the difference between the two as well as SRTT analysis whereby the percentage of the sample population is divided by the percentage of the normal population sample which shows the over or under representation of a research sample compared to a national base type preference sample.

Interestingly, the group dynamics (team plus coaches) presented a very skewed data set when compared to the normal population. For example, the whole represented 120 individuals. Of that population, the head coach identified as an Introverted, Sensing, Thinking and Judging type or ISTJ. His particular type was represented by the whole with 21 members of the whole. This sample, however, reflected a frequency of 17.5%, 5.9% greater than a normal population would reflect of 11.6%. The SRTT index or ration at 1.5 showed ISTJ to be dramatically more prevalent in this group than the normal population.

Furthermore, the second most frequent type identified was Extraverted, Sensing, Thinking and Judging or ESTJ. It was represented by 21 individuals or 17.5% of the sample population but appears in only 8.7% of the normal population. The SRTT ratio reflected 2.0 or double the expected quantity from a normal population.

Collectively, ISTJ and ESTJ represented 42 of the 120 individuals from the entire group. Furthermore, this represents 35% of the total with just 2 of the 16 types possible to identify. Clearly, there was a dramatic overlap to have types whose variance only transgressed the Extraverted versus Introverted type to have identified with such a large population of the whole.

Tracking from left to right, Extraverted, Sensing, Feeling, and Judging or ESFJ was nearly non-existent among the defensive players and the coaches. For clarity, special teams were compiled with the defensive group to prevent identification of individuals due to the special teams group having less than 10 members. Among offensive players, it presented in only 4 out of 120 players or 3.3%. Within a normal population, however, one would have expected to see 12.3%. The SRTT ratio shows 0.27, a dramatically lower sample from the group than would be expected from the normal population showing less than 30% of what would be expected.

Extraverted, Intuition, Feeling, and Perceiving or ENFP presented minimal representation as well with only 6 of the total of 120 participants or 5%. This was significantly below what would have been represented within a normal population where 12.3% would be anticipated to be ENFP. The SRTT ratio reflected a value of 0.41, significantly below what one would have expected from the normal population.

Extraverted, Intuition, Thinking, and Judging or ENTJ type representation was entirely absent from representation among the defensive group and only reflected among the offensive and coaching groups with one individual per group. This ENTJ type in a normal population would reflect 1.8% while among the entire team and coaching group it only represented 0.83%. The SRTT ratio identified a 0.46 value, less than half the personality type concentration of ENTJ that would be expected in the normal population.

Extraverted, Sensing, Feeling, and Perceiving or ENTJ type representation was entirely absent from representation among the defensive group and only reflected among the offensive and coaching groups with one individual per group. This ENTJ type in a normal population would reflect 1.8% while among the entire team and coaching group it only represented 0.83%. The SRTT ratio identified a 0.46 value, less than half the personality type concentration of ENTJ that would be expected in the normal population.

Conversely, Extraverted, Intuition, Thinking, and Perceiving (ENTP) was indicated in 6 of the total 120 population or 5% while normal population statistics would have indicated 3.2%. When analyzed by SRTT, the ratio of 1.6 was generated, indicating that ENTP was significantly overrepresented in this population with over one and a half times as many individuals presenting as this type as would be expected in the normal population.

Extraverted, Intuition, Feeling, and Judging (ENFJ) was another rare type reflected. With only 2 individuals from defense reflecting ENFJ and no members of either the offensive or coaching group identifying with this type. While a normal type distribution would have expected 2.5% or 3 individuals versus the 2 or 1.6% which actually materialized, this value is not significantly out of the range of what would be considered normal. The SRTT ratio revealed a value of 1.16, indicating that ENFJ is overrepresented in this population when compared to a normal population.

Extraverted, Sensing, Thinking, and Perceiving (ESTP) represented 6 individuals or 5% of the total group. Within that limited population, offense demonstrated 2 players to identify with ESTP while 3 defensive players and only one 1 coaching member identified with it. This type was reflected close to a standard population since a normal population one would have expected 4.3% to versus the 5% found with this group as a whole. SRTT data revealed a ratio of 1.16, indicating a substantially overrepresented type when compared to what would have been expected in the normal population.

Interestingly, Introverted, Sensing, Thinking, and Perceiving (ISTP), was consistently represented across all 3 subgroups with 4 individuals from offense, defense, and coaching presenting with ISTP. This represents 9.9% of the entire group compared with 3.3%, which would be expected in a normal population. SRTT revealed a ratio of 3.0, tripling individuals within the group identified with ISTP as compared with the normal population.

Introverted, Sensing, Feeling, and Judging or (ISFJ) was represented by 7 of the total group population or 5.8% versus 13.8% for a normal population. Only one of the coaching groups presented with this type. SRTT data revealed a ratio of .42, indicating a substantially underrepresented type when compared to what would have been expected in the normal population.

Introverted, Intuition, Thinking, and Judging (INTJ) was represented by only 3 of the total group or 2.5% of the group population. Furthermore, no offensive players presented with this type followed by only one defensive player and no coaching staff. A normal population would have predicted 2.1% compared to the 2.5% discovered. The SRTT ratio produced a value of 1.19, indicating an overrepresented INTJ in the sample population.

Introverted, Intuition, Feeling, and Judging or INFJ identified with only 2 of the total population. No offensive players or coaches represented this type, and only 2 of the subgroup defensive players identified with INFJ. In a normal population 1.5% would have been expected; similarly, this group reflected only 1.6%. The SRTT value generated for INFJ of 1.07 showed the type to be closely aligned with the normal population.

Introverted, Sensing, Feeling, and Perceiving (ISFP) identified 3 of the total population with only one of the subgroup offense and 2 individuals from each of the defense and coaching subgroups identifying with ISFP. This represents 2.5% of the population compared to 8.8% which would have been expected in a normal sample. This represents a 6.3% variance from the normal population. SRTT analysis produced a ratio of 0.28, which revealed a significantly underrepresented group type for ISFP.

Introverted, Intuition, Thinking, and Perceiving (INTP) identified 5 of the 120 participants or 4.1%. This compared to the 3.3% one would expect for a normal population. The SRTT ratio produced for INTP was 1.24, indicating that INTP types are overrepresented in the sample group from what one would expect from the normal population.

Introverted, Intuition, Feeling, and Perceiving or INFP represented a total of 6 individuals of the collegiate group. One offensive player, 5 defensive players and no coaches identified with this group. In a normal population one would have expected 4.4% versus the 5% reflected in this study, a variation of only 0.6%. INFP's were slightly overrepresented in the group, revealing a SRTT ratio of 1.14.

Supplemental data utilized in observing MTBI preferences for subgroups:
The following graphic depicts the expected values, the reported values, the variance noted between the sixteen MBTI types within the normal population compared to the sampled population as well as the SRTT ratio. A graphic representation is provided to show the individual variances between the group and normal United States population data for each of the 16 MBTI preferences.

Table 3
MBTI Outcomes

MBTI Type	Observed (Sample)Population Percentage	Expected (Normal) Population Percentage	Percentage Difference: O-E	SRTT: O/E
ISTJ	17.5	11.6	+5.9	1.50
ISFJ	05.8	13.8	-8.0	0.42
INFJ	01.6	01.5	+0.1	1.07
INTJ	02.5	02.1	+0.4	1.19
ISTP	09.9	03.3	+6.6	3.00
ISFP	02.5	08.8	-6.3	0.28
INFP	05.0	04.4	+0.6	1.14
INTP	04.1	03.3	+0.8	1.24
ESTP	05.0	04.3	+0.7	1.16
ESFP	06.6	08.5	-1.9	0.78
ENFP	05.0	12.3	-7.3	0.41
ENTP	05.0	03.2	+1.8	1.60
ESTJ	17.5	08.7	+8.8	2.00
ESFJ	03.3	12.3	-9.0	0.27
ENFJ	01.6	02.5	-0.9	0.64
ENTJ	00.8	01.8	-1.0	0.46

The most notable variances between the sample population and the normal population are in the following areas where the sample group was significantly greater than one would expect of a normal population:

ESTJ reflects an 8.8% higher prevalence in the sample population.
ISTP reflects a 6.6% higher prevalence in the sample population.
ISTJ reflects a 5.9% higher prevalence between the sample population.

The converse situation was also noticed where the sample population was markedly below what one would expect of a normal population.

ESFJ reflects a 9% lower presence in the sample population.
ISFJ reflects an 8.0% lower presence in the sample population.
ENFP reflects a 7.3% lower presence in the sample population.
ISFP reflects a 6.3% lower presence in the sample population.

The three highest variances for the out of range sample population occur along the ST dichotomy where all three ESTJ, ISTP, and ISTJ are present in significantly higher quantities

than one would expect in a normal population, 17.5, 9.9, and 17.5, respectively. In simplest terms, nearly 45% of the entire team and coaching set best identify with ISTJ, ISTP, or ESTJ personality types.

SRTT analysis revealed that ISTJ, ENTP, ISTP, and ESTJ were significantly above (1.5 times or greater) than what would be expected in the normal population where ESFJ and ISFP. Conversely, SRTT ratios revealed that ESFJ, ISFP, ISFJ, ENFP, and ENTJ's each represented less than half the numbers that would be expected in a normal population.

If one were to focus exclusively on a remarkably present dichotomy, the ST or Sensing and Thinking set, shows variations in excess of normal presentations for every type containing an ST pair with 3 of the 4 overrepresented SRTT ratios to reflect types with ST pairs.

Conversely, the most notable types with significantly lower presentations, ESFJ, ISFJ, ENFP, and ISFP all have consistency with the presence of the F or feeling type indicator. These types reflect 9%, 8%, 7.3%, and 6.3% less than would be expected in a normal population, respectively. Further, they represented less than half the representation one would have expected for a normal population according to the SRTT ratio.

Therefore, the data indicates that the prevalence of the ST (sensing and thinking) dichotomy, especially when a subset of either ESTJ, ISTP, or ISTJ personas appears to be a dominant strength of the group. The fourth dominant subgroup, ENTP, while absent the sensing type, consistently presented with the thinking type, T.

Conversely, the presence of the F (feeling) trait appeared to be significantly absent in the group dynamics when compared to the normal population. It was recognized in the 5 personality types which were least reflected in the sample population, reflecting SRTT ratios below 0.5.

Given that the head coach identifies most closely with the ISTJ personality profile, while not conclusive evidence, it lends emphasis to the thought that his selection of both coaches as well as players is influenced by his ST dichotomy. Further, and perhaps just as significant, is the notable absence of feeling types, F, throughout the group, especially with regard to the coaching subgroup.

It is reasonable to conclude, therefore, that one's identified personality profile through MBTI, highly influences his perception of both ideal coaches and players which he selects/recruits for his team or group. Given that these subgroups selected were heavily skewed to toward the ST dichotomy, one would reasonably infer that these personality preferences are considered key to acclimating to the team's dynamic.

Further research into additional team preferences with different coaches and in various sports has the potential to be most validating in further exploring the likelihood that a coach's Myers Briggs Personality Inventory influences who he views as the best fit for his game and best suited to following his lead.

Much gratitude is due to the coaches and athletes for their patience and cooperation in this multi-phase analysis of group dynamics, which began as a curiosity and evolved into an observational pattern of behavior, unanticipated in the researcher's original plan. As with many a great investigation, the reward was not in the inception or the presentation but rather in the exploration of type with a remarkable audience.

References

Briggs Myers, I., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (2009). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator instrument* (3rd ed.). United States of America: CPP, Inc.

Briggs, K. C., & Briggs Myers, I. (1998). Form M. *MBTI Self-Scorable Form M*. CPP, Inc.

Hunter, M. S. (2006). Lessons learned: Achieving institutional change in support of students in transition. *New Directions for Student Services*, *114*, 7-15.

Richmond, S. L. (2008). *Introduction to type and leadership.* United States: CPP, Inc.

Sanborn, D. K. (2013). *Myers-Briggs preferences and academic success in the first college semester*. Iowa State University.

Williams, J. M., & Krane, V. (2013). *Applied sport psychology* (Vol. 7). New York, NY: McGraw-Hill.

The View of Principals on Standards-Based Evaluation Systems and Their Role as Instructional Leaders

James Allen Zoll
University of North Georgia

Abstract

Using a modified Delphi method, this paper examined how the implementation of Georgia's Teacher Keys Effectiveness System (TKES), challenges the principal's role of instructional leadership. Principals expressed concern about the time required for evaluation. However, they perceive they are better instructional leaders and they are in classrooms more often. There was great concern with calibration and the TKES platform itself. A large majority stated that the number of observations should be dependent on the teacher. Ideas for further research including conducting a similar study three years from now when there is more student achievement data available.

Southeast Journal of Educational Administration, Vol. 16, No. 1—Spring 2016
ISSN 2689-307X

The View of Principals on Standards Based Evaluation Systems

The responsibilities of the principal continue to change from management tasks to duties more aligned with improving instruction. The literature strongly asserts that the most important role of the principal is that of instructional leader. However, Stronge (1988, 2005) argues that while the principal participates in many activities during the day, only 11 % of them relate to instructional leadership. May, Goldring, and Huff (2012) raise that percentage to nineteen. The implementation of No Child Left Behind (NCLB, 2001) legislated more accountability and testing. More recently, Race to the Top (RT3, 2009), calls for stronger evaluation methods for teachers in order to improve school achievement. Race to the Top calls for rigorous evaluation procedures that influence professional development, compensation, retention, and tenure.

Strong school leadership is imperative for school success. Leadership has significant effects on student achievement, second only to classroom instruction (Leithwood, 2008; Marzano, Waters, & McNulty, 2005). A review of the literature on school leadership (Leithwood, Day, Sammons, Harris, & Hopkins, 2008) concludes that there is no evidence of a school that has successfully improved student achievement without strong leadership. The influence of the principal is indirect in that the actions of the principal affect the actions of the teacher and that affects student achievement (DuFour & Marzano, 2011; Leithwood et al., 2008; Marzano et al., 2005). These indirect actions fall under the umbrella of the instructional leadership role of the principal.

Though the use of the term instructional leader is widespread, there are many definitions. In his synthesis of meta-analyses of student achievement, Hattie (2009) says the term refers to principals who focus on a climate free of distractions, high expectations for students and teachers, and clear objectives for teaching. According to Smith and Andrews (1989), the principal is a resource provider, communicator, instructional resource, and "visible presence" (p. 7).

Hallinger and Heck (1996) conclude in their empirical review of research that the principal influences internal processes indirectly by: creating policies and norms, setting high academic expectations, developing mission and vision, protecting learning time, and organizing instruction. "For the purposes of this article, we can do more than note that this function of the principal—sustaining a schoolwide purpose focusing on student learning—does receive empirical support" (p. 38). Looking at teachers' perceptions of instructional leadership, Blase and Blase (1999) say instructional leadership has two major themes. They state these as "talking with teachers to promote reflection and promoting professional growth" (p. 367). This evolved into their RG model with the R standing for Reflection and the G for Growth.

One of the key responsibilities of the instructional leader involves the evaluation of teachers in order to assure students have the best teachers in the classroom. Traditional evaluation methods face criticism for the time they require and the evidence that most teachers receive satisfactory rankings. A 2007 survey of over 1000 teachers found little support for the formal evaluation process (Duffett, Farkas, Rotherman, & Silva, 2008). Only 28% felt the tool helped them to be a better teacher

Several high-profile reports offer criticism of teacher evaluation. The Measures of Effective Teaching project (MET) funded by the Bill and Melinda Gates Foundation issued an initial report (2010) with an introductory statement that summarizes the concerns surrounding teacher evaluation: "In too many schools, principals go through the motions of visiting classrooms, checklist in hand. In the end, virtually all teachers receive the same 'satisfactory' rating" (p. 3). The report recommends teacher evaluation to include student achievement gains, specific feedback on benchmarks and feedback from students. Similarly, the New Teacher Project report, The Widget Effect (Weisberg, Sexton, Mulhern, & Keeling, 2009) found that 99% of teachers get satisfactory ratings in evaluation programs that have two rating choices. The number drops to 94%, if there are three choices. The concern is that if such a large percentage of our teachers are satisfactory, why is student achievement lagging in so many schools. Because of these concerns, new standards-based evaluation methods incorporate multiple measures including classroom visitations, student surveys and value-added growth on student achievement measures.

The solution to these checklist type or drive-by evaluations (Toch & Rothman, 2008) is the use of comprehensive evaluation systems that measure instruction with the goal of improvement. Many of these tools evolved from the work of Charlotte Danielson (1996, 2007). She divides teaching into four major categories and 21 themes. The categories are planning and preparation, classroom environment, instruction, and professional responsibilities. From this, she devised scoring rubrics to give administrators details on how to rate teachers as unsatisfactory, basic, proficient and distinguished.

The State of Georgia, like many other states, utilizes a multiple measure evaluation process that looks at the many roles of a teacher and ranks their performance on a four-level rubric based on the work of Danielson (2007). Teachers receive rankings of unsatisfactory, needs development, proficient and exemplary in the new Teacher Keys Evaluation System (TKES). Under this system, principals conduct two formal observations and four informal observations during the school year. Formal observations must be at least thirty minutes long and informal observations must be at least ten minutes long. Pre and post conversations also take place for the planned formal observations. There are five domains on the instrument: Planning, Instructional Delivery, Assessment Strategies, Learning Environment, and Professionalism and Communication. Under each domain, there are two standards. Other measures that go into the evaluation include student surveys and student achievement growth. Student achievement as measured by the growth of a student from one year to the next based on standardized testing is known as a value-added measurement (VAM).

There is support and caution in the literature regarding the effectiveness of these evaluation tools. Darling-Hammond, Amrein-Beardsley, Haertel, and Rothstein (2012) support the use of standards-based evaluations but caution the use of value-added models due to their inconsistency. They point out that value added performance is often a condition of the students assigned to the teacher and that a teacher who is effective one year may be unsatisfactory the next using this approach.

Noting that these types of evaluations have only been in place for 15 years and that there are not many studies on their effect, Murphy, Hallinger, and Heck (2013) call for caution in

determining that these evaluations enable schools to improve. Administrators are "poorly positioned to make teacher evaluation work well" (p. 351) as they lack the credibility in content and instructional strategies that teachers value. They also note the concerns of inter-rater reliability, context of the classroom and the use of value-added measures with test scores as deterrents to teacher evaluation being a tool to improve schools. Papay (2012), states that inter-reliability may be hard to achieve since administrators have different standards.

Jacob and Lefgren (2005) found that principals are effective in identifying the top and bottom ten percent of teachers but not as effective with the remaining 80%. Their study took place in a mid-sized school district in the western United States. Elementary principals were asked to rank teachers who teach core subject matter on several categories on a scale from one to ten. The purpose of the study was to look at the use of subjective principal assessment and value-added measures as tools for compensating teachers. Jacob and Lefgren's (2005) study indicated that principals are effective in rating excellent and poor teachers using subjective evaluative tools, but not as effective with the many teachers who fall in the middle of these two categories.

Taylor and Tyler (2011, 2012) find "suggestive evidence" (2012, p. 82) that teachers do improve during the school year in which they are evaluated and that these effects linger past the year of evaluation. They conducted their research using the Cincinnati Public Schools Teacher Evaluation System (TES). This is one of the earlier implementations of an evaluation system based on Danielson's framework. Although the Georgia system (TKES) relies most heavily on administrator observations, the TES involves highly trained peer evaluators for three of the four observations, with the fourth conducted by the principal. Another difference is that Cincinnati's teachers are evaluated every few years depending on their tenure status and years of service.

Claiming theirs is the first empirical study of the effects of standards-based evaluation systems; Taylor and Tyler (2011, 2012) used a quasi-experimental analysis comparing the student achievement data of teachers before the year of evaluation, during the year, and after the year of evaluation. They focused on midcareer teachers who were in the system before TES began.

Student achievement data came from math test scores in grades 4–8. The comparison of achievement comes from the students the teacher teaches during the year of evaluation compared to the students she taught the year before and the year after. The authors argue this helps to control for teacher change over time and that they control for student characteristics such as previous achievement, classification and racial/ethnic background. They found the average teacher score 0.05 standard deviation higher on summative math assessments during the year of evaluation and that this rises to 0.11 in the year following the evaluation, leading to the conclusion that the instrument does add positive effect on teacher quality.

Though noting some of these positive findings, Hill and Grossman (2013) caution against the use of a one size fits all instrument. They note that reform efforts tend to fail when they attempt to change the work of teaching, and add new layers onto existing layers. They say these new evaluation systems do exactly that and are destined to fail without changes. Their concerns include the generic nature of the instruments, the lack of content specificity, and the lack of quality feedback.

Teachers of all subjects are evaluated using the same tool. “The current systems ask us to believe that teaching kindergarten requires the same set of practices and knowledge needed to teach high school algebra” (Hill & Grossman, 2013, p. 6). On top of this, many principals are generalists and do not know all the content they are observing. Finally, feedback is often given too late or not at all, and is often not content related. The author suggests that more time be spent in evaluation of those teachers who need the most improvement and that rich and specific feedback is given to those teachers so that they can improve. They also suggest the use of content experts in the evaluation process such as department chairs or instructional coaches.

Hallinger et al. (2014) ask if the time and resources poured into teacher evaluation lead to a “robust pathway for school improvement” (p. 2). They first separate the definition of teacher evaluation from instructional supervision noting that evaluation has the ultimate goal of judging the teachers’ capability while instructional supervision has the goal of coaching the teacher towards better performance. They note that the new evaluation processes based on Danielson’s (2006) framework, are intended to show a causal relationship between teacher evaluation and student growth, but often leave out other factors most notably the students family background which the Coleman report (1966) sited as the number one impact on student achievement.

The authors are critical of reports such as the Measures of Effective Teaching Project (Gates Foundation, 2011) because they leave out family and school variables calling the results overstated and the evidence weak and inconsistent to conclude that these evaluation methods increase teacher effectiveness. Finally, they scoured the literature on topics such as school improvement, instructional leadership, data-based decision making and school reform with the notion that if teacher evaluation is a significant way to improve instruction, it should appear in the literature as a solution. They found the existence to be “conspicuously absent” (p. 14).

While the effectiveness of these comprehensive evaluation systems is mixed, the time on the part of the principal to conduct the evaluations is demanding (Kimball & Milanowski, 2009; Papay, 2012). Halverson, Kelley, and Kimball (2004) indicate that a standard based evaluation system required up to 25% of the principal’s time. Toch and Rothman (2008) suggest that these evaluation systems cause principals to treat evaluation seriously but agree that they are time consuming. Marshall (2003) says the tools cause an “impossible workload” (p. 50) for principals and questions the validity of the instruments.

Research Problem

Principals are called to be instructional leaders (Hattie 2009; Hallinger & Heck, 1996, McEwan, 2003). The implementation of standards-based evaluation systems requires much of a principal’s time to be diverted to teacher evaluation, which is only one of the many components of instructional leadership (DuFour & Mattos 2013; Kimball & Milanowski, 2009; Papay, 2012). Principals must use the same instrument for every teacher and have no choice in the duration and amount of observations conducted. The one-size-fits-all process will affect other responsibilities of the principal and is contrary to the loosely coupled nature of a school (Elmore, 2000; Sergiovanni, 2004). Moreover, there is little to be found in the literature about how principals perceive this new system as affecting their job performance.

Research Questions

How is the implementation of the TKES system changing the role of principal?
How do principals perceive the effectiveness of TKES as it relates to instructional leadership?
How has the implementation of TKES affected the other responsibilities of a principal as outlined in LKES?

Procedures

This will be a qualitative study using the Delphi Method. This method was first used in 1948 to forecast advancements in technology and was the catalyst for formation of the RAND corporation (Bourgeios, Pugmire, Stevenson, Swanson, & Swanson 2006; Skulmoski, Hartman, & Krahn, 2007). This method has four components. First, participants are anonymous. Though they may see the responses of others, they do not know who is responding. The second key feature is iteration meaning that participants get a chance to refine their views during the process. Controlled feedback is the third component, which allows participants to clarify and change their views in light of the other participant's responses. Finally, statistical aggregation of responses can allow for a quantitative analysis (Skulmoski, Hartman, & Krahn, 2007).

Principals with at least three years of experience will be invited to participate in this study. The three-year limit will ensure that the participants have experience with TKES and the former evaluation system. These principals will serve at elementary, middle and secondary schools in a school system containing both urban and suburban schools. This will allow for gathering feedback from principals in a variety of school settings. Sixty principals will receive invitations with the goal of having at least 25 participants. Round one questioning will be broad in order to gather feedback about their perceptions about the Teacher Keys program. Subsequent rounds will become more specific based on the feedback of the principals.

The researcher used purposeful sampling (Creswell, 2007) to select principals who have at least three years of experience in the school system or in a system implementing TKES for at least one year. This is designed so participants have knowledge of both the current evaluation tool and the previous tool. In asking for all principals within the system with three years of experience, the researcher hoped to gain information from principals working in a wide diversity of schools in order to enrich the conversation and to refrain from producing "a cozy group of like-thinking individuals which excludes mavericks and becomes a vehicle for inbreeding" (Turoff & Linstone, 2002, p. 568).

The Delphi method is a group facilitation technique that allows of participants to anonymously respond first to open ended questions and then have a chance to clarify or add to their responses in multiple stages. While the traditional Delphi has four or more stages, two or three rounds are preferred especially with the constraints of time and survey fatigue on the part of the participants (Hasson, et.al. 2000; Sumsion, 1998).

The researcher followed the steps of the process beginning with the selection of experts in the field (Gordon, 1994). Experts met the following four criteria: knowledge and expertise,

willingness and ability to participate, time to participate and the ability to communicate effectively (Skulmonski et al., 2007). Principals were selected for this study because they are experts in the field and work in it on a daily basis (Gordon, 1994). Once the experts were selected, the researcher contacted them individually by email with a full description of the paper including the objectives, anticipated time commitment, promise of anonymity, and a way they can confirm their acceptance to participate.

The researcher then developed the first round of questions. These questions defined the problem and asked participants to list concerns, ideas and solutions surrounding the problem. Careful attention went into developing these questions so there were no ambiguities and to assure they are "sharp and answerable" (Gordon, 1994). Questions derived from the literature review and the problem statement. Results from round one were compiled, consolidated, and used to form the round two questionnaires.

Round two included the consolidated responses of all participants from round one and gave participants a chance to respond. These questions allowed the participant to rank their level of agreement to statements on five-point scale (strongly agree, agree, undecided, disagree, strongly disagree). Again, the information was compiled and, a third round took place to get closer to consensus. While there is no clear definition on what consensus is, a 51% to 80% is a suggested range (Hasson et al. 2000). Nevertheless, the study should end when consensus is obtained or "the law of diminishing results sets in" (Hasson et al., 2000, p. 1010). Sumsion (1998) suggests a response rate of anywhere from 50% to 70%. Seventy percent will be used with this study.

The first round was analyzed using content analysis, which is an appropriate technique when analyzing a large quantity of text and "content that is difficult to see or document with causal observation" (Neuman, 2011, p. 363). Similar items were grouped together noting the importance that wording of the participants was not changed and no items were added (Hasson et.al., 2000). Qualitative software, NVivo 10 was utilized to place the participant comments into nodes or themes. The software was also used for word frequency queries.

Second round information is of a quantitative nature in that it will show where consensus is occurring and allow participants to clarify or change their opinion. Since a five-point scale was utilized, measures of central tendency (mean, median and mode) were analyzed along with the determination of the standard deviation. The analysis of data was conducted using SPSS. The responses were coded as follows: Strongly Agree = 1; Agree =2; Undecided = 3; Disagree = 4; Strongly Disagree = 5. Some questions were coded as: Already Use = 1; Will Use = 2; Unsure = 3; Will not use = 4. According to Gordon (1994), the most important role of the Delphi is to be able to synthesize the judgment of a group rather than consensus.

The third round questionnaire was sent on June 4, 2014, in order to determine if consensus, set at 70%, could be found on twenty-three items. A four-point scale was used on this round, removing the choice of undecided causing a forced choice on the part of the participant. Data was analyzed using SPSS. This research took place between April and June of 2014.

In looking at the schools that responded to the survey, the largest number, eighteen, were from the elementary school principals. Nine middle school principals responded and only four high school principals completed the survey. Due to retirements, promotions and moves, many of the high school principals have less than the required three years of experience. Of the thirty-two schools, fourteen receive Title 1 funding, meaning they receive federal funds based on the percentage of students in poverty that attend their schools. As a comparison, more than 50% of the schools in the study cite receive this funding.

Twenty of the participants are female and eleven are male. Twenty-four have worked at the site for more than five years as teachers or assistant principals. One is retiring at the end of the school year of this study, two were promoted to other positions, and one left the system. Two were moved to other schools within the same system. Seventeen have children still living at home and several have grown children. Most of the respondents are involved in the system on many levels, serving on committees, superintendent's advisory, and various other positions. Several are close to being able to retire while many are very early in their career. Ages of the participants range from the low thirties to the early sixties.

First Round Results

The first round of the Delphi study contained open-ended questions. The first two questions dealt with the time spent by principals on walk-through and formative observations including the actual visit, write-up and conferencing. For walk-throughs, the times listed range from 15 minutes to 60 minutes. The mean time was 36 minutes and the mode was 30 minutes. Principals must conduct four walk-throughs on every teacher each year. If every principal uses the average amount of time, that would equal to 2.4 hours per teacher.

Formative observations require a minimum of 30 minutes in the classroom on the part of the observer. There are two formative observations each year. The time range for the time in the classroom, writing the instrument, and holding a conference ranged from 40 minutes to a high of two and a half hours. The mean time was 94 minutes and the mode was 2 hours.

Changing Roles

The next question asked how the implementation of the Teacher Keys Evaluation System (TKES) has changed their role as principal. Using a word frequency query in NVivo, time was cited in the first round survey 88 times and similar words such as hours, 29 times and minutes 39 times. Comments such as "there is not enough time in the day" were frequent and questions about how it changed their role fell into several categories. First, and most prominent, was the impact on work and home balance most expressing they are spending late hours at work and spending weekends to fulfill their responsibilities. "The amount of time we spend writing up and assessing teachers on 10 different areas makes it impossible to have any work/life balance."

Several statements indicated that principals felt they had less time to meet with important groups such as the leadership team, PTA, and community groups. A few also mentioned that they are delegating non-instructional roles to others. One expressed the concern by stating:

> We often talk about how much time we spend assessing students that there is not time to teach, the same is true for TKES. We have to put our heads down and make sure we are getting to all of these evaluations that we can't afford our teachers the gift of having time to really improve on areas for growth from the last observation before we are there again for the next one.

Respondents were asked to give examples of how TKES has made them more or less effective as an instructional leader. Most of the respondents were positive about the impact on their role as an instructional leader. They feel they are better at monitoring instruction and conferencing with teachers and providing specific feedback based on the standards. Some expressed they have a better view of what is happening in their classrooms and two stated they felt closer to their staff. The data that TKES provides was seen as positive and was a useful tool in planning professional development. The more negative comments went back to the issue of time and the process itself with one saying they were spending more time managing the process and keeping their head above water. Another applauds the increase of instructional discussions but laments it is at the expense of student issues and community engagement.

Effects on Other Principal Roles

In responding how TKES has affected their other principal roles, almost every respondent said they spend more time in classrooms because of the TKES process. The word feedback appears 20 times in the survey as does the word conversations. Principals perceive that their conversations with teachers are more focused on the standards and that they are able to give specific feedback. It has also forced them to have those difficult conversations where teachers are not meeting the standards. One principal states it "has affected my roles/responsibilities by: making me more conscious of the time I spend in classrooms-changing the conversations I have with my administrative team regarding teacher performance-allowing my teachers to see my knowledge as a former teacher."

Training and Calibration

The word adequate appears most often when describing the training in the next question. Calibration seems to be a concern with the notion that results on TKES might look differently from school to school which was reinforced by one respondent who noted that participants in training were not consistent in their ratings. Secondary principals express concern about the ability to observe the content in all classes and there was some concern with rating teachers in non-core disciplines.

Process Changes

Fourteen comments about changes to the process revolved around the number of observations required especially when it came to the number of required observations especially for those teachers who are ranked as proficient or exemplary. One suggested that this be based on the years of experience but most suggested reduction on the effectiveness of the teacher. One suggested that this would allow more time to work with struggling teachers.

Tools for Success

Principals were asked how they find successful ways to work with the process. Responses included the need for planning ahead including developing year-long calendars with frequent check points to make sure timelines are being met. Staying in the room for the minimum amount of time required and actually entering information into the portal during the observation were other suggestions.

The final question asked if there were other items they would like to share about the process. Again, there seems to positive support for the instrument and concern about the time. Several expressed concern about the possibility of tying teacher compensation to the TKES process as it would "create a tension that will be very hard to overcome." The actual portal, the names of the levels, and the need for calibration were also mentioned.

Second Round Results

In round two of the process, comments from the open-ended question were sent out in a Likert-type rating scale. Responses of Strongly Agree and Agree were combined as well as those of Strongly Disagree and Disagree. 70 % or higher response rate to the combined responses was used to determine consensus. Forty-three of 66 statements reached consensus.

Several statements in the survey regarded the time involved in using the process. Item four had a consensus of 88% agreement (standard deviation = .98) that the TKES process causes principals to spend more time at work after hours. The response was similar with having to spend more time on weekends, with 88.4% agreement (standard deviation = 1.0). Increased stress as indicated in item 11 received 92.3% agreement (standard deviation = .75). Principals feel they fall behind on other tasks 76.9% agreement (standard deviation = 1.07) and struggle to keep up with emails and other district responsibilities with 96.1 % agreement (standard deviation = .58).

Almost 77% (standard deviation = 1.13) of respondents believed they could not attend as many school functions during the day and that their time to visit classrooms in a non-evaluative was limited 84.0% (standard deviation = 1.01). Principals are delegating non-instructional roles with 84.6 % agreement (standard deviation = .98). 73.1% (standard deviation = 1.13) agree with the statement that it takes too much time to conduct six observations and required conferences. Items 6, 12, and 14 did not receive consensus. These dealt with being less visible to stakeholders 65.4% agreement (standard deviation = 1.13) and having less time to spend with teachers and students with 56% agreement (standard deviation = 1.32). Item fourteen asks whether the time is worth it when teachers use the feedback for growth. This received agreement of 61.6% (standard deviation = .96).

Only the third statement received consensus in round three with 90% (standard deviation = .61) agreeing the time is worth it when principals see teachers implementing feedback for further growth. 60 % (standard deviation = .72) agree they are not as visible to stakeholders while 50% (standard deviation = .85) say they have less time to spend with teachers and students.

Instructional Leadership

Several questions revolved around the instructional leadership role and how TKES affects the role of principal. Exactly, 80.7% (standard deviation = .86) agree they spend more time in the classroom and 76.9 % (standard deviation = .93) feel they are more in tune with what is happening in their classrooms. Almost all, 96.2% (standard deviation = .53) know the strengths and weaknesses of the staff and 76.9% (standard deviation = 1.09) that the process has changed the way they observe teaching and learning. On the question on whether TKES has forced the principal and staff to reflect more on the effectiveness of instructional strategies, 84% agree (standard deviation = .79).

TKES forces them to be more of an instructional coach according to 73.1% (standard deviation = .90). TKES has increased coaching and conversations about teaching and learning according to 88.4% of respondents (standard deviation = .76) and 100% (standard deviation = .40) agree that conferences are anchored in honest conversation on standards. Bringing conversations back to the standards was important to 88.4% (standard deviation = .67). A large percentage, 88.5% (standard deviation = .80) state they are able to provide frequent in-depth feedback and 80.7% (standard deviation = .86) are more effective in giving feedback that changes practice. They also agree that TKES has identified pervasive needs that require specific professional development 84.6% agree (standard deviation = .82). 76.9% (standard deviation = .75) say that the process has made them more interactive with data collection and analysis.

At the 70% cut off, several items went to the third round. Three items had high percentages of disagree. 44% (standard deviation = 1.3) disagree with the statement that TKES has made them closer to their staff, but 36% agree. That changed little in round three, with 55% disagreeing with the statement (standard deviation = .51). On whether, they have a better understand of student learning, 42% agree (standard deviation = 1.1), but 38% disagree. This changed dramatically in round three with 75% agreement (standard deviation = .52). 50% disagree (standard deviation = 1.1) that they are more connected to current trends in teaching and learning with 38.4% agreeing. The only change in round three was that 50% now disagree and 50% agree (standard deviation = .60). While 44% (standard deviation = .76) agree that TKES has improved student achievement, 44% were undecided, possibly because the process is new and correlation to test scores in Georgia has not taken place. This reached consensus in round three with 75% agreement (standard deviation = .45), with no one selecting strongly agree and one comment that there was no evidence yet to support this idea.

The remaining items had more than 60% agreement and went to the third round. 61% (standard deviation = 1.2) felt the process placed unrealistic demands on the principal to provide instructional leadership and 65.3% (standard deviation = 1.2) feel they are more effective instructional leaders because of TKES. Exactly, 65.4 % (standard deviation = 1.2) have become more intentional about the vision and use of professional learning communities in their school and 69.2% (standard deviation = .91) feel they spend more time on monitoring curriculum and instruction.

All three items reached consensus in round three. In regards to placing unrealistic expectations to provide instructional leadership, 75% agreed (standard deviation = .69). Once the

undecided choice was removed, 85% (standard deviation = .51) state they are more effective as instructional leaders. They also reached consensus on the idea that they spend more time monitoring curriculum and instruction with 80% (standard deviation = .65) in agreement. With 70% (standard deviation = .62) agreement, principals reach consensus on being more intentional about the vision.

TKES makes data related to instructional practices accessible to the principal according to 65.3% (standard deviation = .90) and 61.5% (standard deviation = 1.0) say it has increased their professional knowledge about instruction. In round three, 75% (standard deviation = .44) agreed that the accessibility of data has improved. Eighty percent (standard deviation = .41) state it has increased their professional knowledge.

Time appears as a factor again with 64% (standard deviation = 1.2) saying they have less time to work with school teams and 65.4% (standard deviation = 1.2) having less time to lead staff development. 61.6% (standard deviation = 1.3) have less time to work with teachers who need the most assistance. Both items went to round three and neither reached consensus level. In terms of having less time to work with school teams, 65% (standard deviation = .64) agree with the statement and with having less time to lead staff development, 60% (standard deviation = .66) agree. 65 % (standard deviation = .41) have less time to spend on teachers who need the most support.

The TKES Process

When responding about the instrument itself, 88.4% (standard deviation = .87) agree that the process is much better than the previous instrument. However, 65.4% (standard deviation = .1.1) perceive the standards are concise and easy to communicate, 61.5% (standard deviation = .1.0) believe the evaluation system paint a holistic picture of a teacher's performance, but 76.9% (standard deviation = 1.0) feel teachers should see growth in their performance over the course of a year.

Principals feel the number of required observations is overwhelming (73.1%, standard deviation = 1.1) and 84% (standard deviation = 1.1) suggest principals should have the ability to differentiate the number of observations based on the needs of their staff. This is similar to the statement that the system should require fewer observations for teachers based on the needs of their staff with 76.9% in agreement (standard deviation = 1.1).

Training and calibration were addressed with 79.2% (standard deviation = 1.1) in disagreement with the statement that the training was adequate and 77% (standard deviation = 1.0) disagreeing that calibration was affectively covered in the training, 80.8% (standard deviation = .79) state that TKES results look different from school to school with 76.9 % (standard deviation = .80) feeling concern that some schools give the rating of exemplary too hastily. Agreement was not reached with the comment that the toughest part of the process is the difference between exemplary and proficient with 61.5% agreement (standard deviation = 1.3). That did not change in round three with 63.2% (standard deviation = .60) agreement. 69.2% (standard deviation = 1.0) felt calibration was consistent in their school. This dropped in round three to 60% (standard deviation = .60). The district should conduct a calibration exercise to

assure consistency according to 92.3% (standard deviation = .75) of respondents and 92.2% (standard deviation = .72) request more support on how to objectively score the standards.

While 73% (standard deviation = 1.2) state it is difficult to see the standards at work in non-core subjects, 69.2% (standard deviation = .89) are confident with use of TKES with all teachers. This reached consensus of 85% (standard deviation = .45) in round three. 53.8% (standard deviation = .85) were undecided about the concern among secondary principals lacking content knowledge to effectively rate all teachers. In the third round, 40% (standard deviation = .50) agree with 60% disagreeing. The potential of tying TKES results to teacher compensation was a worry among 80.8% (standard deviation = 1.1) of the respondents.

Suggested Improvements

Principals had many suggestions for how to improve the process. In terms of who should evaluate, 84.6% (standard deviation = .85) should have more staff allowed to do observations, and 81% (standard deviation = 1.1) felt other support staff should be available to conduct observations such as administrative assistants.

Process questions included combining the assessment uses and strategies standards with 80.8% (standard deviation = 1.1) agreement. Many, 84.6% (standard deviation = .94) felt that the summative evaluation should self-populate from the previous instruments. Agreement was not found with 42.3% (standard deviation = 1.5) agreeing that the names of the levels (exemplary, proficient, needs development, ineffective) be changed. Finally, there were several concerns with the TKES portal itself with 92.3% (standard deviation = .74) saying it is too cumbersome and requires too many steps and 100% (standard deviation = .44) saying the portal often runs slow or shuts down.

Research question one asks how the Teacher Keys Evaluation System (TKES) is changing the role of the principal. Time was the word that appeared the most in a word frequency test using qualitative software and principals' responses confirm the time-consuming (DuFour & Mattos, 2013; Kimball & Milanowski, 2009; Marshall, 2012; Papay, 2012; Rothman, 2008;) nature of these instruments.

Under the current process, evaluators must conduct four walk-throughs a year for each teacher. When including the time it takes to observe, conference, and write results, the time spent per walk-through ranged from 15 minutes to 60 minutes, with a mean time of 36 minutes and a mode of 30 minutes. As an example, a principal has 950 students and completes TKES evaluations on thirty teachers. If this principal uses the average time, 72 hours is spent on conducting walk-through observations.

Two formative assessments take place per year on every teacher and must be a minimum of thirty minutes. The range of time including observation, conferencing and writing the results ranged from 40 minutes to 2.5 hours, with a mean time of 94 minutes and a mode of 2 hours. Going back to the principal with 30 teachers to observe, using the average time, the principal would spend 47 hours on each round of formative observations. Using the average time, the

principal spends about 3 hours and 15 minutes per teacher to conduct evaluations. In all, 176 hours, or 23 days of the principals time is spent on teacher evaluation.

Maintaining high visibility is a component of instructional leadership (Hallinger & Murphy, 1985; McEwan, 2003; Smith & Andrews, 1989) and is a component of standard eight of LKES (2014) communication and community relations. Reaching consensus, 77% of respondents indicate the implementation of TKES has decreased the time they are visible at school functions, and over half stated they have less time to spend time with staff, students, and to a greater degree, stakeholders. While, principals perceive they are in classrooms more, large percentages state they have less time to visit the classroom for non-evaluative reasons. McEwan (2009) says one role of the instructional leader is to maintain positive relations with staff, students and community. When asked if TKES has made them closer to their staff, 55% disagreed with the statement. It appears the implementation of TKES is challenging many principals in this area.

Beyond the time concerns, there was positive consensus about the change in their role. They spend more time in the classroom and the process for how they observe teaching and learning has changed. There is more of a focus on the effectiveness of instructional strategies and principals believe they are more of an instructional coach. Principals are more confident with conferencing and giving specific feedback based on the standards. Principals agree that when they see teachers using that feedback for their own growth, the time concerns are worth the effort, though this did not reach consensus until round three.

Sample comments from the round one questionnaire supporting the question include the following:

> "I like being in the classroom but the paperwork and conferences are far too consuming"
> "I am in classrooms more and have more direct discussions with teachers about matters pertaining to teaching and learning"
> "My feedback to teachers improved, and my time in the classroom greatly increased"
> "I am in the classroom more often. I cannot attend as many school functions during the day. I am not as available to meet with parents as timely as in the past."

Research Question 2

Question two asks principals if they perceive TKES has made them more effective, less effective or had no effect on their role of instructional leader. The most basic question of whether they are more effective instructional leaders because of TKES did not receive consensus in round two. The Hallinger and Murphy (1985) model defines instructional leadership as defining the school mission; managing the instructional program, of which teacher evaluation is a component; and developing the school learning climate program.

Items regarding time not reaching consensus in round two indicate principals perceive they have less time to spend with teachers and students and to attend daily events. Having less time to visit classrooms for non-evaluative purposes is also a concern. These concerns potentially affect the school climate component of instructional leadership.

Reaching consensus in round three was the idea that principals are more effective as instructional leaders. There were also other positive comments surrounding instructional leadership. This was especially evident with the conferencing and feedback portion of the survey. Leithwood, et al. (2004, 2010) cite providing instructional guidance as part of instructional leadership. TKES has made principals more aware of the strengths and weaknesses of their staff. They are more effective in giving specific feedback based on the standards. This statement seems to contradict participants' agreement that the process places unrealistic demands on administrators to provide instructional leadership.

Identifying professional development needs is another component of instructional leadership (Blase & Blase, 1999; Eberts & Stone, 1998; Hallinger & Murphy, 1985). Principals find the TKES process enables them to identify pervasive needs in their staff. They believe the instrument forces staff to reflect on the effectiveness of instructional strategies. Reaching consensus in round three, 80% stated the process has increased their own professional development. Though professional development is cited by participants as an area for improvement, 60% of principals believe they do not have the time to lead staff development.

The LKES (2014) process asks that principals monitor the academic achievement of students, also a component of Hallinger and Murphy's (1985) model of instructional leadership. Principals agree TKES has increased the amount of time they spend monitoring curriculum and instruction. Respondents were also in agreement that they are more interactive with data collection and analysis.

Principals could not reach consensus on whether TKES was improving student achievement, in the second round, but did reach consensus in the third round. This is most likely due to the fact that the process has not been in place long enough to correlate it with student achievement. This could also be problematic in that the state of Georgia is implementing a new testing system in the 2014-2015 school year, making longitudinal comparisons virtually impossible. While there are documented cases where researchers have deemed that an evaluation system has improved achievement with the teachers in their study, (Milanowski et. al, 2004; Taylor & Tyler, 2011, 2012), it may be years before the state can determine if TKES has improved student achievement.

Only one question surfaced in round two dealing with the key component of instructional leadership, that of setting a vision and mission (Hattie, 2009; Hallinger & Heck, 1996; Halinger & Murphy, 1985). The statement asked if they were more intentional about the vision and use of professional learning communities in their school and did not reach consensus in the second round. However, 70% agreed with the statement after round three. Providing strong and collaborative environments where teachers have a voice in decision-making is important for a positive school climate (Blase & Blase, 1999; DuFour, DuFour & Eaker, 2011; Greenlee & Brown, 2009; Seashore et al., 2010), making it curious that this was only mentioned once as a component of instructional leadership. This might be because the survey specifically focused on the teacher evaluation component of instructional leadership.

Sample comments from the round one questionnaire supporting the second research question include the following:

> "I have to develop teacher leaders more effectively as much of my time is spent in classrooms and completing write-ups and conferencing with teachers."
> "TKES has made me more effective as an instructional leader primarily through its specific standards."
> "TKES has provided a system where everyone evaluating should have the same view of instructional expectations."
> "This process requires a one size fits all, which makes me less effective when it comes to having the time to spend on individuals who could really benefit from more support to actually improve rather than focusing just on evaluating." "I do not have any evidence to suggest that LKES has made me either more or less effective as an instructional leader."

Research Question 3

Question three asked what other roles as included in the Leader Keys Evaluation Instrument (LKES) were impacted by the implementation of TKES. Teacher evaluation is one component of LKES and is a separate standard from instructional leadership. The eight standards are Instructional Leadership; School Climate, Planning and Assessment; Organizational Management; Human Resources Management; Teacher/Staff Evaluation; and Communication and Community Relations (2014).

In round two, respondents agreed that TKES causes them to distribute noninstructional tasks to other people. They express that they fall behind on other tasks and that keeping up with emails and other district responsibilities is more challenging. In the open-ended questionnaire, one principal indicated that he had less time to spend on community outreach and that they were less visible to stakeholders. Those comments did not reach consensus in the second level or the third. Another principal states that they have less time for informal conversations with staff and that they have seen "breakdowns in other areas due to the lack of capacity to complete the requirements in TKES and in all other areas with great integrity and fidelity."

Results indicate that the time involved conducting these evaluations is taking away from other responsibilities aligned with the Leader Keys Evaluation System (LKES). Principals need to spend time on the organizational management of the school (Horng & Loeb, 2010; Leithwood et al., 2010) including the hiring of effective teachers and supporting those that struggle and removing those that are ineffective. This is also covered by the organizational management and human resources component of LKES. Respondents in the survey indicate concerns in this area, especially with the statement by 65% in round three that they have less time to focus on those teachers who need the most help.

Sample comments from the round one questionnaire supporting the question include the following:

> "Many of the tasks associated with running a large school cannot be accomplished during the school day"
> "I have to delegate more tasks and I find it harder to keep up with daily tasks." "There is just no time to do it all, and do it well. I do my best to schedule all meetings before or after school, knowing that any situations that take place during the school day will take me away from observations."

"There are certainly times when I fall behind on other tasks because of the time devoted to TKES, but I believe it is a correct order of priority."

Concerns of Calibration and Content Knowledge

Another concern and a threat to the reliability of the TKES instrument is that of calibration. Research indicates that inter-rater reliability is a concern (Meyer & Rowan, 1987; Murphy, 2013; Papay, 2012.) Jacob and Lefgren (2005) state that principals are successful in identifying the top and bottom performing teachers but are inconsistent with those in the middle. Respondents to this survey express great concern about calibration between schools. What is exemplary at one school might not be so at other schools. Principals felt somewhat comfortable with the calibration within their school, but not across the district. In round three, 60% agreed that calibration was consistent in their school.

Training for the utilization of the TKES process was inadequate according to the respondents. Again, the largest concern was the issue of calibration. However, respondents also state it can be difficult to make the distinction between proficient and exemplary. They agree that they need more support on how to objectively score the standards.

Knowledge of the content and instructional strategies was a concern in the literature (Grubb & Flessa, 2006; Hill & Grossman, 2013; Horng & Loeb, 2010; Murphy, 2013; Spillane et al., 2007). Principals had varying views on this question. When asked if they felt comfortable using the tool for all teachers there was strong agreement. However, there was some agreement, especially among secondary principals that they may not have the content knowledge to support the process with all teachers. This statement did not reach consensus in round three possibly because more of the respondents were from the elementary level.

Concerns about the Process

Though not specific to the research questions, respondents expressed many concerns about the TKES process beyond the amount of time required. In round three, 80% state that TKES paints a holistic picture of a teacher's performance. They could not reach consensus, however, on whether the standards are concise and easy to communicate. While they overwhelmingly agree this is a better instrument then the previous instrument, there was great concern about the platform itself, how it requires many steps, is slow, and often shuts down. Combining certain standards such as two that deal with assessment and having a summative evaluation that self-populates from the previous six observations would be helpful and save time. Revisions to the procedural components as suggested by the principals would reduce the time involved in conducting the evaluations.

Conclusions

This study analyzed how the use of standards-based evaluation systems, in this case the Teacher Keys Evaluation System used by the state of Georgia, impact the principal in their role of instructional leader and their other roles as principal. It allows for the voice of the principal to be heard which is conspicuously missing from the literature. The study was viewed from the lens

of the theory of loose coupling, a theory that allows for an institution to be both rational and indeterminate. Standards-based tools are what Elmore (2000) would call the solution to the problem of loose coupling. These evaluations are another solution to finding best practices that provide specific prescriptions for how to improve a school. In this case, mainly through the work of Danielson (1996), specific standards and indicators for quality teaching are prescribed.

There is no doubt that principals' see this as a more effective tool for evaluation than the previous model. As instructional leaders, they are in the classroom more often and they are far more comfortable giving standards-based feedback to the teachers. They are better at monitoring instruction and they know the strengths and weaknesses of their teachers better than before. These are all strong indicators that in some ways the tool allows them to be a stronger instructional leader.

At the same time, these tools require an inordinate amount of time on the part of the evaluator. Principals state they are spending far more time after school and on weekends to fulfill their responsibilities and that the additional stress is affecting their home and work life. Working beyond an eight-hour day is standard for principals and this new process adds more to the plate of already busy leaders.

Loose coupling suggests that when a system is placed into the school organization, adaptations are made to the process to make it fit into the organization rather than having the organization change to match the process. These results bear that out in several ways. For example, while all agree they are time consuming, when asked how long a principal spends on the entire formative observation process, the answers went form 45 minutes to 2.5 hours. Principals look to their environment and their own style to decide the amount of time they will spend. For example, instead of holding conferences, they may send emails for feedback. Some do not leave the room until they have filled out the form while others wait until they get back to the office. Many hold themselves to the minimum time required in the classroom.

There are definite concerns that other key responsibilities are affected by the time-consuming nature of these instruments. Many principals, using the idea of distributed leadership, delegate some of these responsibilities to others. In many ways, this is a good idea and is a way to develop teacher leaders. However, it is important that principals are visible to stakeholders and are able to attend crucial meetings. It is important that they meet with teachers in non-evaluative formats and attend planning meetings. While others can deliver professional learning, principals should have time to carry on this important task, as they are the instructional leaders of their schools.

State and district leaders should listen to the voice of principal as they attempt to make sense of an instrument that in many ways they support. Requiring a teacher who consistently receives scores of proficient and exemplary to have the same number of observations as a struggling teacher is an obvious change that can be implemented. Principals should be able to differentiate the number of observations teachers receive based on the teachers' competence. Of course, to make this work, there would probably have to be a standard formula.

Simple changes to the process would reduce the time required. The platform itself is slow and complicated. Some of the standards could be combined and some only need to be ranked on an annual basis or by exception, most notably professionalism and communication. Currently, the summative evaluation process requires the principal to complete another form that summarizes the year before having the final conference. Having the system populate that form would save much time and frustration. As this process is involving, many of these issues, especially the platform issues, should improve.

There continue to be inherent concerns with the TKES process especially in the areas of calibration and content knowledge. It is important that more work be done to improve the inter-rater reliability of the evaluator. This is more critical as the system and state move towards using the TKES data as part of a compensation model, on which principals express much concern. It might be interesting to have evaluative teams visit other schools and compare results to those rankings of the school evaluators.

Although respondents in this survey could not agree there was concern at the secondary level about the content knowledge of the evaluator, most felt comfortable using this tool with all teachers. Might that change if elementary evaluators observe a pre-calculus class when their instructional background is social studies? Is it reasonable to expect a high school principal to have all the content knowledge necessary to evaluate all the teachers in the school? Perhaps involving content experts in the evaluation process can alleviate this concern.

On the Leader Keys Evaluation Instrument, Instructional Leadership is a separate standard than that of Teacher Evaluation, where in most models, teacher evaluation is a component of instructional leadership. This study asked principal specific questions about how TKES affects their role as instructional leader. Understandably, most of the responses surrounded the teacher evaluation component of instructional leadership. The majority of responders say they are much better instructional coaches. While this is important, it should not be confused with being a better instructional leader. Principals must also have time to communicate goals, set and implement a vision, and build a strong learning climate.

References

Anderson, V. A. (2011, June). *Defining the EdD and PhD in education: A Delphi study* (Doctoral dissertation). Retrieved from http://gradworks.umi.com /34/56/3456042.html

Bill and Melinda Gates Foundation. (2010). *Learning about teaching: Initial findings from the Measures of Effective Teaching Project*. Seattle, WA: Author.

Blase, J., & Blase, J. (1999). Principals' instructional leadership and teacher development: Teachers' perspectives. *Educational Administration Quarterly*, *35*(3), 349-378. doi:10.1177/0013161X99353003

Bourgeios, J., Pugmire, L., Stevenson, K., Swanson, N., & Swanson, B. (2006). *The Delphi Method: A qualitative means to a better future.*

Brewer, D. J. (1993). Principals and student outcomes: Evidence from U.S. high schools. *Economics of Education Review*, *12*(4), 281-292. doi:10.1016/0272-7757(93)90062-L

Bridges, E. M. (1982). Research on the school administrator: The state of the art. *Educational Administration Quarterly*, *18*(3), 12-33. doi:10.1177/0013161X82018003003

Bridges, E. M. (1967). Instructional leadership: A concept re-examined. *Journal of Educational Administration*, *5*(2), 136-147. doi:10.1108/eb009614

Bush, T., Bell, L., & Middlewood, D. (2010). *The principles of educational leadership & management*. SAGE.

Camburn, E. M., Spillane, J. P., & Sebastian, J. (2010). Assessing the utility of a daily log for measuring principal leadership practice. *Educational Administration Quarterly*, *46*(5), 707-737. doi:10.1177/0013161X10377345

Cresswell, J. (2007). *Qualitative Inquiry and Research: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage

Danielson, C. (2007). *Enhancing professional practice a framework for teaching*. Alexandria, Va.: Association for Supervision and Curriculum Development.

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, *93*(6), 8-15

Donaldson, M. (2009). *So long, Lake Wobegon? Using teacher evaluation to raise teacher quality* (pp. 1–25). Washington, DC: Center for American Progress.

Donaldson, M. L., & Donaldson, G. A., Jr. (2012). Strengthening teacher evaluation: What district leaders can do. *Educational Leadership*, *69*(8), 78-82.

Duffett, A., Farkas, S., Rotherham, A., & Silva, E. (2008). *Waiting to be won over: Teachers speak on the profession, unions, and reform* (p. 25). Washington, DC: Education Sector. Retrieved from http://www.educationsector.org/publications /waiting-be-won-over

DuFour, R., & Marzano, R. J. (2011). *Leaders of learning: How district, school, and classroom leaders improve student achievement*. Bloomington, IN: Solution Tree Press.

DuFour, R., & Mattos, M. (2013). How do principals really improve schools? *Educational Leadership*, *70*(7), 34-40.

Eberts, R. W., & Stone, J. A. (1988). Student achievement in public schools: Do principals make a difference? *Economics of Education Review*, *7*(3), 291-299. doi:10.1016/02727757(88)90002-7

Eggers, R. M., & Jones, C. M. (1998). Practical considerations for conducting Delphi studies: The oracle enters a new age (Undetermined). *Educational Research Quarterly*, *21*(3), 53.

Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, *17*(1), 101-128. doi:10.1023/A:1025083214622

Elmore, R. (2000). *Building a new structure for school leadership* (p. 40). Washington, DC: The Albert Shanker Institute.

Firestone, W. A. (1984). *The study of loose coupling: Problems, progress, and prospects*. Retrieved ERIC database. (ED247264)

Fusarelli, L. D. (2002). Tightly coupled policy in loosely coupled systems: Institutional capacity and organizational change. *Journal of Educational Administration*, *40*(6), 561.

Georgia Department of Education. (2013). *Teacher key effectiveness* system. Retrieved from https://www.gadoe.org/School-Improvement/Teacher-and-LeaderEffectiveness/Pages/Teacher-Keys-Effectiveness-System.aspx

Goldring, E., Huff, J., May, H., & Camburn, E. (2008). School context and individual characteristics: What influences principal practice? *Journal of Educational Administration*, *46*(3), 332-352.

Gordon, T. J. (1994). The Delphi method. *Futures research methodology*, 1-33.

Greenlee, B., & Brown, J. J., Jr. (2009). Retaining teachers in challenging schools. *Education*, *130*(1), 96-109.

Grissom, J. A., Loeb, S., & Master, B. (2012). *Effective instructional time use for school leaders: Longitudinal evidence from observations of principals* (p. 47). Presented at the Association for Public Policy Analysis and Management, Baltimore, MD.

Grubb, W. N., & Flessa, J. J. (2006). "A job too big for one": Multiple principals and other nontraditional approaches to school leadership. *Educational Administration Quarterly*, *42*(4), 518-550. doi:10.1177/0013161X06290641

Hallinger, P. (1992). The evolving role of American principals: From managerial to instructional to transformational leaders. *Journal of Educational Administration*, *30*(3), 8.

Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools*, *4*(3), 221-239.

Hallinger, P. (2011). Leadership for learning: lessons from 40 years of empirical research. *Journal of Educational Administration*, *49*(2), 125-142. doi:10.1108/09578231111116699

Hallinger, P., & Heck, R. H. (1996). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly*, *32*(1), 5-44. doi:10.1177/0013161X96032001002

Hallinger, P., & Murphy, J. F. (2013). Running on empty? Finding the time and capacity to lead learning. *NASSP Bulletin*, *97*(1), 5-21.

Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability*, 1-24. doi:10.1007/s11092-013-9179-5

Hallinger, P., Wang, W.-C., & Chen, C.-W. (2013). Assessing the measurement properties of the Principal Instructional Management Rating Scale: A meta-analysis of reliability studies. *Educational Administration Quarterly*, *49*(2), 272-309. doi:10.1177/0013161X12468149

Halverson, R., Kelley, C., & Kimball, S. (2004). Implementing teacher evaluation systems: How principals make sense of complex artifacts to shape local instructional practice. In W. Hoy & C. Miskel (Eds.), *Research and theory in educational administration* (Vol. 3). Greenwich, CT: Information Age Publishing.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008-1015. doi:10.1046/j.13652648.2000.t01-1-01567.x

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

Hill, H., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review*, *83*(2), 371-384.

Horng, E. L., Klasik, D., & Loeb, S. (2010). Principal's time use and school effectiveness. *American Journal of Education*, *116*(4).

Horng, E., & Loeb, S. (2010). New thinking about instructional leadership. *Phi Delta Kappan*, *92*(3), 66-69.

Jacob, B. A., & Lefgren, L. (2005). *Principals as agents: Subjective performance measurement in education* (Working Paper No. 11463). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w11463

Kimball, S. M., & Milanowski, A. (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly*, *45*(1), 34-70. doi:10.1177/0013161X08327549

Leithwood, K., Anderson, S., Mascall, B., & Strauss, T. (2010). School leaders' influences on student learning: The four paths. In T. Bush, L. Bell, & D. Middlewood (Eds.), *The Principles of Educational Leadership & Management* (pp. 13-30). SAGE.

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership & Management*, *28*(1), 27-42. doi:10.1080/13632430701800060

Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning. Review of Research*. New York, NY: The Wallace Foundation. Retrieved from ERIC database (ED485932)

Lincoln, Y. S., & Guba, E. (1985). *Naturalistic inquiry*. Newbury Park, CA: SAGE.

Louis, K. S., Leithwood, K., Anderson, S., Mascall, B., & Wahlstrom, K. (2010) *Learning from leadership: Investigating the links to improved student learning*. Center for Applied Research and Educational Improvement, University of Minnesota.

Marshall, K. (2003). Recovering from HSPS (Hyperactive Superficial Principal Syndrome): A progress report. *Phi Delta Kappan*, *84*(9), 701.

Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, *86*(10), 727-735.

Marshall, K. (2013). *Rethinking teacher supervision and evaluation: How to work smart, build collaboration, and close the achievement gap*. San Francisco, CA: Josey-Bass.

Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, *70*(3), 14-19.

Marzano, R. J., Waters, T., & McNulty, B. A. (2005). *School leadership that works: From research to results*. Alexandria: VA: Association for Supervision and Curriculum Development. Retrieved from ERIC database (ED509055)

May, H., Huff, J., & Goldring, E. (2012). A longitudinal study of principals' activities and student performance. *School Effectiveness and School Improvement*, *23*(4).

May, H., & Supovitz, J. A. (2011). The scope of principal efforts to improve instruction. *Educational Administration Quarterly*, *47*(2), 332-352. doi:10.1177/0013161X10383411

McEwan, E. (2009). *10 traits of highly effective schools*. Thousand Oaks, CA: Corwin Press.

McKenna, H. (1994). The Delphi technique: a worthwhile research approach for nursing? *Journal of Advanced Nursing*, *19*(6), 1221-1225. doi:10.1111/j.13652648.1994.tb01207.x

Medley, D., & Coker, H. (1987). The accuracy of principals' judgments of teacher performance. *Journal of Educational Research*, *80*(4).

Meyer, H.-D. (2002). From "loose coupling" to "tight management"? Making sense of the changing landscape in management and organization theory. *Journal of Educational Administration*, *40*(6), 515.

Meyer, J., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, *83*(2), 340-363.

Milanowski, A., Kimball, S. M., & White, B. (2004). *The relationship between standards-based teacher evaluation scores and student achievement: Replication and extensions at three sites* (pp. 1-32). Madison, WI: Consortium for Policy Research in Education.

Murphy, J. F., Hallinger, P., & Heck, R. H. (2013). Leading via teacher evaluation: The case of the missing clothes? *Educational Researcher*, *42*(6), 349-354.

Neuman, W. L. (2011). Nonreactive research and secondary analysis. In *Social research methods: Qualitative and* Quantitative Approaches (7th ed.) (pp. 358-380). Boston, MA: Pearson.

Neumerski, C. M. (2013). Rethinking instructional leadership, a review. What do we know about principal, teacher, and coach instructional leadership, and where should we go from here? *Educational Administration Quarterly*, *49*(2), 310-347. doi:10.1177/0013161X12456700

O'Donnell, R. J., & White, G. P. (2005). Within the accountability era: Principals' instructional leadership behaviors and student achievement. *NASSP Bulletin*, *89*(645),56-71. doi:10.1177/019263650508964505

Odden, A. (2004). Lessons learned about standards-based teacher evaluation systems. *Peabody Journal of Education*, *79*(4), 126-137. doi:10.1207/s15327930pje7904_7

Orton, J. D., & Weick, K. E. (1990). Loosely coupled systems: A reconceptualization. *Academy of Management Review*, *15*(2), 203-223. doi:10.5465/AMR.1990.4308154

Papay, J. P. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, *82*(1), 123-141,167.

Peterson, K. D. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices*. Corwin Press

Roberts Evaluation. (n.d.). *Delphi technique*. Retrieved from http://www.robertsevaluation.com .au/index.php?option=com_content&task=view&id =49

Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (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

Seashore, L, K., Dretzke, B., & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement*, *21*(3), 315-336. doi:10.1080/09243453.2010.486586

Seifert, E. H., & Vornberg, J. A. (2002). *The new school leader for the 21st century: The principal*. Rowman & Littlefield.

Sergiovanni, T. J. (1984). Leadership and excellence in schooling. *Educational Leadership*, *41*, 4-13.

Sergiovanni, T. J. (2006). Toward a new theory. In *The principalship: A reflective practice perspective* (5th ed.). Boston, MA: Pearson.

Skulmoski, G., Hartman, F., & Krahn, J. (2007). The Delphi Method for graduate research. *Journal of Information Technology Education*, *6*, 1-21.

Smith, W. F., & Andrews, R. L. (1989). *Instructional Leadership: How principals make a difference.* Alexandria, VA: ASCD.

Spillane, J. P., Camburn, E. M., & Pareja, A. S. (2007). Taking a distributed perspective to the school principal's workday. *Leadership and Policy in Schools*, *6*(1), 103-125.

Stronge, J. H. (2005). Teacher evaluation and school improvement: Improving the educational landscape. In *Evaluating Teaching: A guide to current thinking and best practice* (2nd ed.). Corwin Press.

Supovitz, J. A. (2000). Manage less lead more. *Principal Leadership: Middle Level Edition*, *1*(3), 14-19.

Sumsion, T. (1998). The Delphi technique: An adaptive research tool. *British Journal of Occupational Therapy*, *61*(4), 153-156.

Taylor, E. S., & Tyler, J. H. (2011). *The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-career teachers* (Working Paper No. 16877). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w16877

Taylor, E. S., & Tyler, J. H. (2012). Can teacher evaluation improve teaching? *Education Next*, *12*(4).

Taylor, K. C. (2007). *A study of principals' perceptions regarding time management* (Unpublished doctoral Dissertation). Kansas State University, Kansas.

Toch, T. (2008). Fixing teacher evaluation. *Educational Leadership*, *66*(2).

Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education*. Retrieved from http://www.edexcellence.net/commentary/educationgadfly-weekly/2008/february-7/rush-to-judgment-teacher-evaluation-in-publiceducation.html

Turoff, M., & Linstone, H. (Eds.), (2002). *The Delphi method: Technique and applications*. Boston, MA: Addison-Wesley.

Waters, T., Marzano, R. J., McNulty, B., & Regional, E. L. M.-C. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement.* A Working Paper.

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, *21*(1), 1-19.

Weick, K. E. (1982). Administering education in loosely coupled schools. *Phi Delta Kappan*, *27*(2), 673-676.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect*. Brooklyn, NY: The New Teacher Project. Retrieved from http://gcpstv.org/gcpsmainweb01.nsf/092DF14366B4598F8525788C0067CF48/$file/TNTPTheWidgetEffect.pdf