

A Tripartite Framework for Leadership Evaluation

Broad scale organizational transformation demands that leaders nurture individual agency and build collective capacity to support fundamental change. Key in this regard is the leader's ability to inspire teachers, parents, school community leaders and students around a common vision of reform. -Anthony Bryk et al., 2010

University of California, Berkeley
Graduate School of Education
Leadership Connection

Lynda Tredway
Daphannie Stephens
UC Berkeley, Leadership Connection

Carol Hedgspeth
Cynthia Jimes
Rudy Rubio
ISKME

The Leadership Connection in the Graduate School of Education at the University of California, Berkeley supports leadership development in urban schools and districts through its multiple programs: the Principal Leadership Institute (PLI) preparation program, the Leadership Support Program (LSP) induction program, and customized professional development and coaching services to districts.

Rebecca Cheung is the Coordinator of the PLI and LSP. W. Norton Grubb is Faculty Coordinator of the PLI and LSP. The Graduate School of Education and the University Office of the President support PLI and LSP.

The development of the Leadership Connection Rubric and this working paper was supported by a grant from the Stuart Foundation to the Leadership Connection at the Graduate School of Education, University of California at Berkeley. The Griot Project, a three-year inquiry for experienced urban school leaders, informed the rubric development, the working paper, and the processes for professional development of urban leaders described in this paper.

The working paper was developed in conjunction with the Leadership Connection Leadership Rubric.

Suggested citation: Tredway, L., Stephens, D., Hedgspeth, C., Jimes, C., & Rubio, R. (2012). *A tripartite framework for leadership evaluation*. Berkeley, CA: Graduate School of Education.

Lynda Tredway was the founding Coordinator of the Principal Leadership Institute and the Leadership Support Program. She retired from UC Berkeley in 2012 and is supporting a project in Oakland USD to use the processes in this paper for leadership evaluation and facilitating workshops for school leaders on race and equity.

Daphannie Stephens is the Coordinator of Public Programs for the Leadership Connection. She was the Project Coordinator for the Stuart Foundation grant.

The research and evaluation work supported under the grant was conducted by the Institute for the Study of Knowledge Management in Education (ISKME). Cynthia Jimes, director of research at ISKME, served as the PI on the research and evaluation project for Leadership Connection's Stuart Foundation grant. Carol Hedgspeth, senior research associate at ISKME, and Rudy Rubio, research assistant at ISKME, co-managed the research and evaluation project. Jimes, Hedgspeth and Rubio were all co-authors on this paper.

Project Director: Daphannie Stephens

Editing and Design: Leadership Connection and IKSME

Copyright 2012 © Leadership Connection, Graduate School of Education, University of California, Berkeley.

Permission to reproduce with the copyright notice is hereby granted.

Table of Contents

Executive Summary	4
Premises and Recommendations	5
I. INTRODUCTION: LEADERSHIP EVALUATION OVERVIEW	6
Leadership Evaluation: What to Do & What Not to Do	7
Toward Multiple Metrics and Robust Evidence	9
Leadership Evaluation History	10
Federal and State Policies Require Effective Leadership.....	11
Leadership Evaluation Approaches.....	12
II. THE TRIPARTITE LEADERSHIP FRAMEWORK	14
The 360° Feedback Tool.....	14
The Leadership Accountability Report Card.....	16
Evidence from Rubric and Cycle of Inquiry	19
The Leadership Connection Rubric: Elements of Leadership Effectiveness	21
Evidence from Leadership Practice	23
Evidence from an Asset-based Cycle of Inquiry	24
III. IMPLICATIONS FOR THEORY AND PRACTICE.....	28
APPENDIX A: Supporting Tables – Instrument and Rubric Review	29
APPENDIX B: Rubric Development – Design Principals and Lessons Learned	33
APPENDIX C: Leadership Connection Rubric - Short Form.....	37
REFERENCES.....	40
ACKNOWLEDGEMENTS	45

Executive Summary

The Tripartite Framework for Leadership Evaluation provides a comprehensive examination of the leadership evaluation landscape and makes key recommendations about how the field of leadership evaluation should proceed. The chief concern addressed by this working paper is the use of student outcome data as a measurement of leadership effectiveness. Since leaders do not have direct influence over student achievement and mediate instructional influence with students through teachers, we find that the use of student test scores as a measurement of leadership effectiveness is neither fair nor useful. The collective actions of teachers and the leader lead to improved student achievement, but the leader should concentrate his or her work on teacher working conditions and teacher motivation as a part of that collective responsibility (Leithwood, 2012).

A second concern in our work with urban leaders is the absence or surface treatment of race and equity in nearly all evaluation instruments or processes. We believe that to countermand the historical predictability of achievement outcomes for our most vulnerable students, attention should be directed specifically to the roles that urban school leaders play in fostering consistent attention to the structural issues of race and poverty. We contend that often the attention to race and equity do not get sufficient credit in leadership evaluation for the ways that school leaders build trust and keep equity concerns at the forefront of their work.

Finally, we call for an overhaul of the conventional cycle of inquiry, which is based largely on needs analysis and leader deficits, and incomplete use of evidence to support recurring short cycles within the larger yearly cycle of inquiry.

In this working paper, we recommend a tripartite framework for leadership evaluation that includes:

- A valid and reliable 360° perceptual survey instrument that depends on feedback from multiple respondents who know the principal's work firsthand and report their perceptions objectively.
- A leadership accountability report card (LARC) that identifies quantitative metrics as leading indicators for improved student outcomes over which the principal has substantial influence or control. These include such metrics as attendance, teacher assignments, and discipline data.
- Evidence-based practice using a rubric that undergoes calibration and agreement among its users on the levels of quality used to determine the rubric rating and a revised cycle of inquiry that fosters asset observations and short-term outcomes and the use of evidence to inform next steps.

In sum, we present key findings from our analysis, research, and practice work with coordinating recommendations. The metrics we recommend in the leadership accountability report card require validation, and the evidence-based and asset-based cycle of inquiry requires further study. Thus, our intent is not to suggest we have a complete answer to leadership evaluation. Rather, we intend to spark a conversation about leadership evaluation that will enrich the field and provide fair and multiple metrics for effectively assessing leaders – and supporting them to improve. Based on reading of the research about leadership and analysis of current methods of assessment, we contend that basing leadership on the metrics we suggest will be more useful in

understanding exactly what it takes for leadership practice to impact teachers and, in turn, student achievement.

Premises and Recommendations

Premise: No one leadership evaluation tool or metric is sufficient in determining leadership effectiveness.

Recommendation: Use a combination of research-based tools and processes that support a tripartite framework for effective leadership evaluation: a 360° feedback survey, a leadership accountability report card, and evidence-based use of a rubric and a “revised” cycle of inquiry.

Premise: Leadership evaluation that uses student outcomes as a proxy for leadership outcomes obscures the ability to see which leadership actions lead to increased supports for achievement. Leaders do not have direct influence over student achievement and impact it through teachers.

Recommendation: Use multiple metrics for leadership evaluation that rely on quantitative measures over which the leader has direct influence.

Premise: Leadership evaluation tools that rely on perceptual responses need to be reliable and valid so that they can mitigate the research-based concerns with perceptual evaluation tools.

Recommendation: Use a reliable and valid tool as one of the components of effective leadership evaluation, but combine it with other metrics to ensure fairness.

Premise: Multiple research-based metrics provide quantitative data that include metrics over which the leader has direct influence and can be used to assess leader effectiveness.

Recommendation: Put research efforts behind developing a set of quantitative tools that can serve as a Leadership Accountability Report Card in order to, in combination with 360° feedback and evidence-based process for leader goal-setting, provide a full portrait of leadership effectiveness.

Premise: Evidence-based practice is a vital component of effective leadership evaluation.

Recommendation: Use the asset model of inquiry by planning strategic modeling and training for coordinators, leaders, and facilitators.

I. INTRODUCTION: LEADERSHIP EVALUATION OVERVIEW

Effective school leaders are the connective tissue in school reform, and substantial consensus among researchers verifies the importance of school leadership in influencing teacher practice to improve student outcomes (Branch, Hanushek & Rivkin, 2012; Bryk et al., 2010; Darling-Hammond et al., 2007; Knapp et al., 2003; Knapp et al., 2010; Louis et al., 2010; Leithwood & Louis, 2012; Leithwood et al., 2004; Leithwood, 2010; Wallace Foundation, 2012; Waters, Marzano & McNulty, 2003). However, debate persists regarding the key factors that should be measured to gauge and report on effectiveness. Many state evaluation systems, prompted by accountability under NCLB and Race to the Top requirements, have chosen student test scores as a part of the formula for evaluating leaders, but a growing body of research demonstrates that the assessment of leadership effectiveness should concentrate on factors over which the leader has more direct control – like attendance (Bryk et al., 2010) or teacher working conditions (Leithwood, 2012).

The Leadership Connection at UC Berkeley’s Graduate School of Education – through its preparation, induction, and professional programs for school leaders – seeks to inform the school leadership evaluation discussion. Our work has been specifically focused on understanding the work of urban school leaders who are serving our most vulnerable student communities. As a result of preparing leaders in the Principal Leadership Institute (PLI), providing induction for leaders in the Leadership Support Program (LSP), and coaching and providing professional development to leaders in multiple urban school districts, we have developed a response to what constitutes effective leadership evaluation. That response is informed by practice and is evidence-based.

This working paper provides an introduction to the current landscape of leadership evaluation and an argument for a tripartite framework for evaluation that includes the combined use of:

- A validated and reliable **survey instrument** (often termed 360° feedback) that gauges principal effectiveness based on responses from multiple constituents.
- A **leadership accountability report card** that includes quantitative, research-based metrics that serve as prerequisite indicators for improving student achievement (e.g., attendance, teacher retention, discipline referrals);
- **Evidence** of leadership improvement on specific outcomes that result from targeted use of a **leadership rubric** and a **‘revised’ cycle of inquiry process**.

By reviewing the strengths and challenges of current evaluation practices, we offer a theoretical frame and methodology for consideration. Addressed to researchers, policymakers, school leaders, leadership preparation programs, and districts and states engaged in making decisions about the cause-effect relationship between leadership and school improvement, the paper offers guidance for effective leadership evaluation in the preparation, induction, ongoing support, coaching and supervision of school leaders.

Section I, the introduction, discusses the complexities of leadership evaluation and provides the argument for a different approach to leadership evaluation. In analyzing the historical and current leadership evaluation landscape, we advocate in *Section II* for a 360° feedback tool, and a set of quantitative metrics that comprise what we term as a leadership accountability report card (LARC). The discussion of these two components of the tripartite framework for leadership

evaluation are followed by a presentation of the third component, the use of a rubric and a revised cycle of inquiry, which provide qualitative evidence for leadership effectiveness.

In *Section III*, we conclude by discussing the implications of the framework for the theory and practice of leadership evaluation. Leadership evaluation is indeed a complex and perplexing undertaking that has a straightforward question driving its conceptions: How should school leaders be evaluated as they seek to improve teacher practice and student learning outcomes?

Leadership Evaluation: What to Do & What Not to Do

The current panorama of leadership evaluation recommendations communicates conflicting understandings of what matters. Public perception supports principal evaluation determined by using the school's student achievement scores. In this section we propose what to do -- broaden the metrics used for leadership evaluation -- and what not to do -- use student test scores as a proxy for leadership effectiveness. A recent study by the Regional Educational Laboratory (REL) identified student performance data as a key criterion for principal evaluation even though principals exercise only indirect effect on student outcomes, operating largely through teachers (White et al., 2012). The study found that 43 percent of California school districts used principal evaluation to make decisions about removal, and 35 percent of districts used evaluation results for retention decisions (White et al., 2012). However, since most principals (83 percent) received the highest ratings on the performance scales, the purported use of student performance data for removal or retention of principals applies to only a few individuals (White et al., 2012).

However, the REL report raises a central dilemma in principal evaluation: *Is the evaluation primarily for development and support, or is it primarily disciplinary and attached to job security?* The distinction highlights a need for an evaluation approach that is both supportive and developmental with an expectation that the principal can have a significant impact on specific measurable outcomes that in turn produce student achievement outcomes. Thus, we support the use of student achievement as one measure for school effectiveness when there is collective responsibility of all school constituents for outcomes, but we do not support the use of this metric alone for measuring leadership effectiveness.

The tripartite evaluation framework presumes that the constituents in a school have adequate information to respond to surveys (360° feedback instruments) as one component of comprehensive leadership evaluation. In contrast to the use of student achievement data as a metric for leadership evaluation, our framework identifies potential quantitative metrics that are identified in the research as leading indicators or prerequisites of improving student outcomes (Bryk et al., 2010; Louis et al., 2010; Leithwood, 2012; Grubb, 2009; Branch et al., 2012). Finally, we propose evidence -- supported by a rubric, a professional development process and a revised cycle of inquiry -- be used in conjunction with the survey and a leadership accountability report card to support a multiple metric approach to assessing the school leader's effectiveness.

The call for the trio of evidence sources is buttressed by findings from recent reports that offer guidance about developing leadership evaluation processes (Clifford et al., 2012a; Clifford et al., 2012b; Leon et al., 2011; Davis et al., 2011). One report acknowledges that "the impact perspective (measuring student achievement) is appealing because of its apparent simplicity," yet cautions that "sophisticated statistical models are necessary" to accurately associate principal leadership practices with student achievement (Clifford et al., 2012a, p. 6). This caution is echoed by the National Research Council on the use of value-added modeling for teacher effectiveness, which reports that value added models are complex statistical models that attempt to "isolate

teacher, or program effects, at least two years of students' test scores are taken into account, sometimes along with other student and school-level variables, such as poverty, family background, or quality of school leadership" (Braun et al., 2010, p. 4). However, the recommendation from the NRC indicates that models that do not "control" for these factors cannot be used as a means to determine progress unless modeling can incorporate some of these limitations. They argue that value added models used for making high stakes decisions like teacher effectiveness must be held to higher standards of reliability and validity" and could only be used in combination with other indicators.

Therefore, basing principal evaluation on student outcomes (a yearly measure, one step removed from teacher and teacher variables) is even more questionable—both in common sense and statistical terms, despite the Texas study by Branch, Hanushek and Rivkin (2012) that employs a value-added approach linking leadership to achievement. A second report underscores how principals mediate effects on student learning through these variables: quality of teachers (controlled by hiring and facilitating the departure or development of less competent teachers), teacher distribution (determined by teacher assignment and scheduling), and teacher practice (influenced through clarified expectations, supervision, observation and evaluation) (Clifford et al., 2012b). A third set of reports from WestEd, based on a literature review of 68 scholarly and professional articles, developed three key questions and thirteen key features to guide leadership evaluation decisions and processes (Leon et al., 2011; Davis et al., 2011; Sanders, Kearney & Vick, 2011):

1. Why Evaluate Principals? (Purpose, Mission Alignment, Professional Growth, and District Policy Considerations)
2. What Should Be Evaluated? (Expectations, Research and Standards, Principal Participation in Choices)
3. How Should Principals Be Evaluated? (Principal and Supervisor Participation, Procedures, Evaluator Training and System Review (Leon et al., 2011, p. 1).

In addition to these efforts, several instruments propose emphasizing research-based variables for leadership performance, while others call for input from multiple constituents on a wider range of variables (Halverson & Kelley, 2011; Porter et al., 2008; NLNS, 2009; NBPTS, 2010; Bottoms & Schmidt-Davis, 2010; Marzano, Waters & McNulty, 2006). The multiple approaches to leadership evaluation highlight a truism: the supervision and evaluation process for school leaders, like the work of leadership itself, is a complex process and responsibility.

In addition, the school leadership research literature deepens our understanding of what matters for leadership evaluation – particularly in curricular and instructional guidance, organizational management, and accountability (Wallace, 2012; Louis et al., 2010; Bryk et al., 2010; Leithwood, 2012; Leithwood et al., 2004; Knapp et al., 2010). Yet, while these variables are important, they do not fully paint an accurate picture of what matters if a school leader is to achieve durable outcomes for students – particularly in the urban contexts in which our work has focused. First of all, the role of relationships as a vital resource and metric of effective leadership and component of school reform is often shortchanged even though there is substantial evidence to indicate its foundational and critical importance (Grubb, 2009; Bryk et al., 2010).

And more importantly, the central role that equity and advocacy play in leaders' ability to address achievement challenges in schools with entrenched patterns of historically underperforming groups is typically missing. Both in research and in most leadership rubrics and 360° survey instruments we have investigated, equity is either overlooked or addressed in general terms that are inconsistent with our understanding of its crucial role in urban school

leadership. It is not enough to assess, as the VAL-ED instrument does (which will be discussed in depth later), how well a principal “advocates for a culture that respects the diversity of all students” (2008, p. 12). The leader needs to be assessed on the ability to have courageous conversations precisely about the ways in which specific school practices in disciplinary practices, for example, are not only inequitable but constitute racial profiling of students of color (Singleton & Linton, 2006; Ferguson, 2000). Any leadership evaluation framework, therefore, should include specific attention to the foundational elements of leadership practice -- the so-called soft skills (presence, attitude, identity, relationships, and resilience) and the profound ways that equity and advocacy play out in the lives of urban school principals.

Granted, the cause-effect relationship between these leadership dimensions is more complex to measure¹. However, relationships and trust (Kruse et al., 1994; Bryk et al., 2010) as well as cross-cultural understanding and focused attention on closing opportunity gaps have long been cited as important contributors to success in schools recognized for increased achievement (Darling-Hammond, 2010; Boykin & Noguera, 2011). Therefore, the absence or lack of depth of these fundamental elements in much of the leadership research, reports or rubrics impedes school leadership evaluation that will impact what matters most – increased achievement for all students. If we are to achieve our goal of positioning school leaders to greatly affect and advance the achievement of our most vulnerable students, we must institute evaluation systems and tools that incorporate a broader understanding of the foundational elements of practice, multiple metrics and a robust collection of leadership evidence.

Toward Multiple Metrics and Robust Evidence

Thus, as this working paper proposes, specific changes are needed in the current evaluation framework. To increase leader morale, retention, and capacity, emphasis must be placed on valuing development over discipline in leadership evaluation. Methods must privilege the complexity of the work, and strive against the temptation to be satisfied with naming single and efficient metrics as a substitute for effective evaluation. It is also crucial that we understand leadership areas of direct influence and pay attention to the leadership chain of effect (Figure 1) in relation to school improvement outcomes.

The leadership chain of effect refers to the causal relationship between leadership and student outcomes. The outcomes are always mediated through teachers and often through assistant principals, instructional coaches, and/or department chairs. While the principal is a critical organizational actor, the result of authentic distributed leadership means that he or she may be spending the bulk of his or her time on creating, with other administrators and teacher-leaders, the conditions for school reform and not necessarily spending as much direct time on instructional observation, which is often the purview of other support personnel. This is particularly true of schools that are in the “low-performing” or “program improvement” status in which they have additional resources to support coaching. Thus the leader has no direct influence on student outcomes.

This last point cannot be over-emphasized. As a field we need to be more disciplined about separating leadership actions and effectiveness from school level effectiveness and student

¹ In particular, see the discussion of the difficulty of drawing a causal relationship between leadership dimensions and student outcomes in a study of the Pittsburgh effort to link principal support and evaluation to incentives by Hamilton et al. (2012)

outcomes. This distinction should yield specific metrics for leader effectiveness as he or she guides and supports the improvement process.

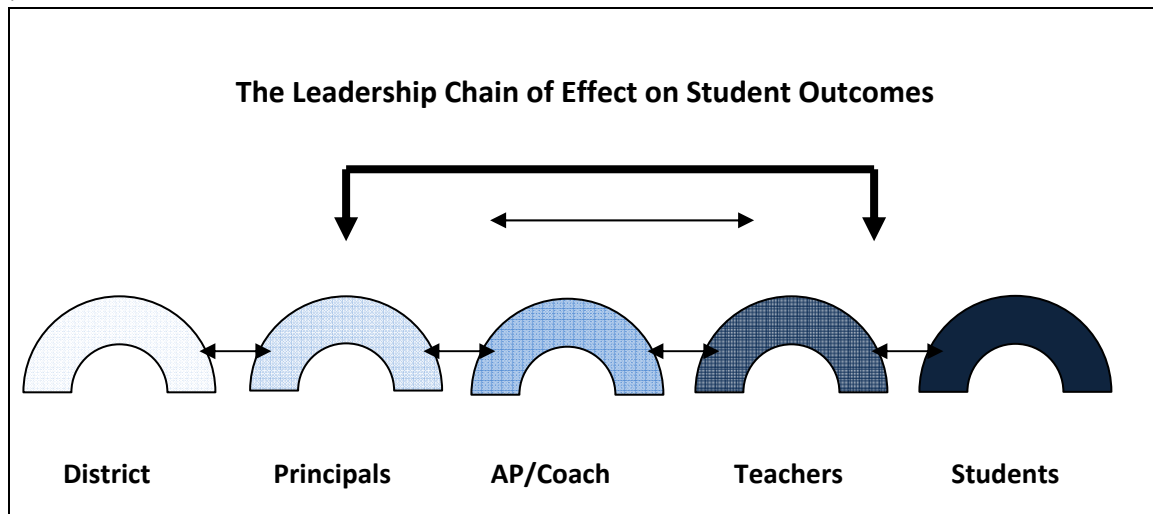


Figure 1. The Leadership Chain of Effect

Our track record in the field of leadership preparation, induction, coaching novice school leaders, and ongoing work with school and district leaders led us to foster the tripartite framework. The benefits of the new framework respond to a key criterion about “what to do” in the reports referenced above: fairness. Without multiple metrics, the evaluation of leaders based on student test scores as the major or key criterion is simply unfair. Traditional methods of evaluation – what not to do – are often used for the purpose of discipline rather than development, hierarchical as opposed to collaborative, one-dimensional and stagnant rather than dynamic. The traditional method is out of touch with the new educational environment, virtually ignores the depth of equitable practices a leader needs to support genuine “turnaround”, and is not supportive of the leaders we need for urban and rural schools where our most vulnerable children are not successful. The new methodology we propose – what to do – is inclusive, responsive to content, collaborative, objective, multi-dimensional, and evidence-based.

Leadership Evaluation History

The current state of leadership evaluation results from a history of “piling on” increasing responsibilities to the role of principal-teacher as schools and districts became more hierarchical. The “modern” principal’s duties are described in a 1941 book, *Duties of School Principals* (Jacobsen et al., 1941). On the coattails of the thirty-year call for scientific management in school organizations (Taylor, 1911), their evidence indicates that secondary principals in the 1930s spent their time on 59 different functions. However, the highest frequency responses indicate these personal qualities as critical: “ability to get along with people”; personality; and a general category termed “leadership”. These criteria related to presence and professionalism populate the remaining qualities: tact and diplomacy; good judgment and common sense; professional attitude; character; appearance; poise and emotional stability; and health, energy and vigor. Only three of the criteria indicate attention to teaching and learning, one related to management, and one related to interest in community affairs. The Horng, Klasik and Loeb (2009) study of the ways

principals spend their time concurs and indicates that principals perform more than 40 different tasks in each day.

This recitation from the 1941 book may indicate that “the more things change, the more they stay the same”. However, the point in presenting these results is to reiterate how the principalship in 21st century America with additional responsibilities for accountability, technology, partnerships, site-based budgeting, and political and legal knowledge for policy implementation requires a more effective method for assessing how the leader can attend to multiple duties and be accountable for school outcomes, given the current policy context.

Federal and State Policies Require Effective Leadership

Federal and state policy has prioritized assessing school leader quality to address No Child Left Behind (NCLB) requirements. As one result, federal and state education agencies instituted policies that prioritize assessment of leadership *effectiveness*, marking a shift from measuring educational inputs – for example degrees and other indicators of leader qualifications – to outputs including student performance. Because current policies emphasize the use of “rigorous measures” for leadership evaluation (see [Indiana’s legislative guidance doc](#) as an example of this language) states and districts have chosen a variety of leadership effectiveness evaluation approaches. However, what is measured and how it is measured vary greatly, leaving gaps in our knowledge about what principal leadership effectiveness includes, how it relates to leadership qualifications, how it can be improved, and how ultimately it relates to student, school, and district success.

In efforts to systematize what practitioners and others understand as constituting principal leadership effectiveness, more than 40 states adopted the Interstate School Leaders Licensure Consortium (ISLLC) standards, or some iteration of them, as a uniform foundation for principal assessment. Many adopted the revised 2008 ISLLC standards (based on 1995 original standards), which, many argue, are grounded in behaviors linked to improving student achievement and are intended to bolster the capacity for principals to facilitate necessary changes in their schools (CCCSO, 1996, 2008; Porter et al., 2008). However, standards typically use normative and broad language that is not easily translated into observable and documentable performance outcomes or evidence.

In fact, according to the Wallace Foundation’s 2009 report about new directions in principal assessment, few processes effectively utilize the application of new standards. Some states – notably Kentucky, Iowa and Delaware – use the ISLLC standards as a basis for assessing leaders, improving their performance, and redesigning programs that prepare them for their jobs (Wallace Foundation, 2009). A review of existing assessment instruments in 44 districts and states for which there is sufficient documentation revealed that nearly half fail to provide leaders with clear feedback on how to improve teaching and learning (Goldring et al., 2008). Moreover, efforts in evaluation are inconsistent with the content and frequency of professional development that principals receive to help them improve once areas of improvement are identified (Wallace Foundation, 2009).

The next section discusses the evaluation approaches. This discussion explores the complexities in the broader principal assessment terrain. While emphasizing that there is no single “right” process or tool, we assess the strengths and concerns about each approach and suggest that a composite of evaluation tools offer a more effective direction in principal assessment.

Leadership Evaluation Approaches

According to Sanders, Kearney, and Vince (2012), recent developments in the measurement of principal knowledge and performance emphasize the need to collect and use multiple forms of evidence to capture the scope and complexity of new expectations. Each variation in form, collection method or measurement type adds a nuance to data analysis. The examination of the current panorama of leadership evaluation instruments and processes (Clifford et al., 2012a; Davis, et al., 2011; Porter, et al., 2008) reveals three general approaches: perceptual, value-added, and evidence-based.

Perceptual data approaches take two forms. One is the use of a reliable and validated survey based on constituent (teacher, supervisor, parent, student) responses about leader's capacity. This approach satisfies the criteria of input from a range of persons who work with the principal, as well as validity and reliability, (Davis et al. 2011), and is best represented by two tools: Vanderbilt Assessment of Leadership in Education/VAL-ED tool with substantial validity testing and reliability (Porter et al., 2008), and the Comprehensive Assessment of Leadership for Learning/CALL (Halverson & Kelley, 2011²). A second perceptual approach, more common in our experience in districts and of more concern, has not been well-researched or documented but we feel that it is prevalent and want to name it as a key focus of research. This perceptual evidence is concerned with unsystematic observations that a supervisor or others may conduct in the course of visiting or observing in a school. This practice typically results in perceptions that have little or no evidentiary basis.

Adding harm to the absence of systematic evidence, random conversations about principals and schools that occur at the district office often lead to inaccurate perceptions of the individual or the organization based on hearsay. It is the case that a few unsubstantiated "stories" about a principal or school often influence perceptions and then decisions about the leader's or school's effectiveness. We contend that the second type of perceptual data can be organized to be more systematic and evidence-based, and that is the objective of the professional development process attached to the rubric and the revised cycle of inquiry we discuss in the next section.

While perceptual data often serve as crucial input about a principal's leadership role and effectiveness, several studies indicate challenges in conducting quality evaluations using even reliable and valid perceptual data tools (see Brandt et al., 2007; Danielson & McGreal, 2000; Little et al., 2009). These studies suggest that lack of evaluator training can be a threat to the reliability of the evaluation and objectivity of the results. An untrained observer may introduce bias into observations; the observer's expectations may influence the observation to a greater degree than the observed principal behaviors (Mujis, 2006). Thus, any use of perceptual data should be accompanied by training for the respondents.

Value-added approaches for leadership evaluation are commonly based on student test scores and teacher evaluations and use psychometric formulas that attempt to control for factors such as student demographics. Our reading of the research indicates that this approach does not meet the fairness standard. As previously mentioned, such data are typically used as an inaccurate proxy for leadership effectiveness.

² The VAL-ED (Vanderbilt Assessment of Leadership in Education) instrument is now available from <http://valed.discoveryeducation.com/>. The CALL (Comprehensive Assessment of Leadership for Learning formative evaluation instrument is used for secondary principals. The CALL websites are <http://www.call.wceruw.org/index.html> and <http://www.call.wceruw.org/resources.html>

One example highlights the issue. Memphis City School District uses a value-added approach in evaluating principal effectiveness, required by the state and its Tennessee Instructional Leadership Standards (TILS) Appraisal Instrument. Regional superintendents evaluate school administrators once per year, and measure effectiveness in three components: data on student growth from the Tennessee value-added assessment system (TVAAS), student achievement indicators, and observations of practice using the TILS rubric. In addition, Memphis City uses the VAL-ED survey to measure principals' effectiveness. This multi-instrument value-added approach that includes pertinent data from the principal, peers, the supervisor, and other school staff has strengths: the use of qualitative (rubric-guided observations) and quantitative (survey effectiveness scores and sub scores) data and the use of formative and summative methods (survey results and yearly evaluations by superintendent).

The How of Assessment 360° Feedback

360° feedback has been used in business and industry as a self-development tool for a number of years, and is designed to gather and process multi-rater assessments on an individual's performance, and then provide feedback on the results (Bookman, 1999). The process includes tuning into the observations and perceptions of those who work most closely and are potentially most affected by the individual, and who are in a position to observe behavior and skills. The critical characteristic of 360° feedback is to identify gaps between perception and desired performance (Wilson, 1997). The process typically includes a self-assessment and peer appraisals.

However, reliance on evaluation data or yearly student test scores (50% of the total evaluation in Tennessee) as a measure of leadership effectiveness produces an invalid measure. The evidence supporting the validity and reliability of value-added modeling is sufficiently weak, and such results are not yet appropriate to use as the primary metrics for leader effectiveness (Baker et al., 2010). Similarly, the standardized exams used in most value-added assessment systems are not "instructionally sensitive" (Stumbo & McWalters, 2011). Popham (2007) defines "instructional sensitivity" as "the degree to which students' performances on a test accurately reflect the quality of instruction specifically provided to promote students' mastery of what is being assessed" (p. 146).

The example is not to disparage the processes on which Memphis and countless other districts rely heavily to make decisions. Rather, by highlighting a particular example, we call attention to the complications in the current evaluation landscape. The example draws attention to the importance of considering multiple evaluation approaches and underscores the difficulty in creating an evaluation framework that is equitable and rigorous for all leaders. Originating in the medical field, the *evidence-based practice approach* is an emerging area of research in health, education and psychology and refers to how practitioners enact what they know, what they do, and what they believe when interacting with constituents (Metz et al., 2007).

The evidence-based approach to leadership evaluation requires more study, but its format, which we have field-tested, relies on a rubric that is used in conjunction with a refined cycle of inquiry (COI) that includes qualitative analysis of leadership observations. This process will be more fully discussed in our leadership rubric, the asset-based cycle of inquiry and the evidence from specific targets in a cycle of inquiry in Section II. We briefly propose another form of the evidence-based approach that shows promise and needs more study: the portfolio. Its use is premised on leadership as dynamic rather than static, on evidence as targeted instead of broad, and on opportunities for sustained self-reflection.

The Effective Practice Incentive Community (EPIC) Professional Learning model of evaluation from New Leaders for New Schools (NLNS) supports the use of a portfolio of evidence to demonstrate effective practice in a targeted area. Yet, since the EPIC process relies on an initial choice linked solely to test score improvement, this particular tool raises the question of how the focus of the portfolio is chosen – again, relying on school improvement as measured through test scores rather. However, the use of a portfolio requires further attention as a possible process in a multi-dimensional evaluation framework. Each of these key tools will be introduced and discussed in turn in Section II: the survey using a 360° feedback instrument, the leadership accountability report card, and evidence garnered from use of a rubric and asset-based cycle of inquiry.

II. THE TRIPARTITE LEADERSHIP FRAMEWORK

The 360° Feedback Tool

In order to provide a critical view of the leadership evaluation landscape, we reviewed a representative group of twelve assessment tools, rubrics, and manuals from states, organizations and universities, as well as accompanying documentation about several evaluation approaches currently in use. Our goal was to determine how these approaches differed in terms of the following: respondents; categories of assessment (i.e., the types of indicators of leadership being evaluated); application of assessment feedback (developmental and/or corrective) and delivery (formative and/or summative); evidence of measurement rigor (e.g., validity and reliability of the tools); alignment with industry standards within the assessment tool; and requirements for evidence use. In addition, there are approaches in some tools that could be used for future validation and consideration. However, our analysis concludes that two 360° feedback tools are reliable and valid for use: Vanderbilt Assessment of Leadership in Education/VAL-ED tool with substantial validity testing and reliability (Porter et al., 2008), and the Comprehensive Assessment of Leadership for Learning/CALL (Halverson & Kelley, 2011). Tables summarizing the analysis of the rubrics and instruments are in Appendix A.

In addition, our aim was to understand how (or if) common evaluation tools are used explicitly to assess principal leadership in high need or urban contexts. By comparing the rubrics and instruments side by side, the organizational values and the underlying theory regarding the leader's roles within the school and expectations around the leader's impact and effectiveness are apparent. The conclusions of this analysis include the following:

- Participation expectations vary. Typically, respondents for the instruments in particular were to include a combination of the leaders themselves (self-analysis), students, teachers, other school staff, parents, supervisor and/or parents.
- Both formative and summative assessments lead to principal accountability, particularly when using a comprehensive or 360° feedback tool. Additionally, use of formative and summative assessments can provide data that link to strong instructional leadership practices, and that data can help target professional development (Portin et al., 2006).
- The assessment categories included personal dimensions of a leader, (resilience, honesty, ethics); the leader's ability to create and sustain a school vision; and the leader's capacity to serve as an instructional leader. The weight of each of these categories, however, varies by instrument, although each plays a role in determining the "quality" of the school leader.

- Nearly all of the rubrics – including CALL, New Leaders’ Principal Leadership Actions Rubric, and KIPP’s Leadership Proficiency Roadmap – and all of the instruments assess categories that refer to the leader’s ability to manage operational duties.
- Few instruments require respondents to provide any tangible artifacts or other evidence of the principal’s leadership practice or behaviors.
- Two instruments of those we analyzed -- the KIPP Leadership Proficiency Roadmap and NYC’s Leadership Performance Planning tool – show promise for indicating the degree to which school leaders navigate the special complexities of urban school contexts.
- None of the evaluation approaches analyzed use portfolios as an observable, qualitative data collection tool.
- Some tools support a developmental approach to leadership evaluation that considers the leader’s level of experience.

Finally, in one important area, equity and advocacy, the tools and rubrics are by and large silent or insufficient. The references to equity (or diversity) and/or advocacy (e.g., in the VAL-ED and the TILS instruments) are inconsistent with our understanding of the ways urban leaders need to understand and communicate how race and equity as structural dimensions of their work influence achievement outcomes and teacher working conditions (Theoharis, 2009; Leadership Learning Community, 2010). There are two main differences between the evaluation tools in our examination in respect to diversity and equity: 1) the weight given to equity or diversity in the leader’s assessment, and 2) the depth to which equity and respect is defined for the leader. Tennessee’s Instructional Leadership Standards, for example, include diversity and equity as a broad category of assessment. The Massachusetts Educator Evaluation Regulations rubric, on the other hand, embeds diversity and equity within their standards, implying that diversity and equity leadership serve as an interim benchmark for an overall leadership effectiveness metric.

The extent to which diversity and equity is defined within the evaluation tools also varies. While Tennessee’s rubric defines the terms through the use of several standards and competences, other rubrics allow for a wide range of interpretation in determining how increased diversity and equitable outcomes are accomplished within the school. Given the variation among schools and the importance of paying close attention to how equity issues are broached, this critical component of a comprehensive evaluation needs further attention and more specific guidance to raters. For example, how does the leader use disaggregated student discipline referral data to facilitate school faculty discussions, which almost always indicates serious racial and gender disparities?

Overall, the analysis of the tools points to two considerations. One, the measures of principal effectiveness seminal to urban, high-needs schools would represent a key addition to the leadership evaluation landscape. Measuring indicators of how a school leader enacts his or her job attitudinally and professionally, how s/he advocates for equitable access for all students, and how s/he maintains the trust of school constituents are essential to understanding key strengths and places for growth an urban school leader. These factors reflect real-world, on-the-ground experience of leaders, and speak to the contextual richness of urban schools, communities and districts typically overlooked in the leadership literature and in evaluation practices. Two, acknowledging the level of the leadership experience (total years of experience and/or experience in a particular school) is vital to understanding where the school leader is now and to identifying the professional development steps to take foster growth.

Building on the analysis, we argue that a valid and reliable 360° feedback tool is essential to the leadership evaluation process. Despite the issues with perception, two perceptual instruments merit wider use. As previously identified, these are: the Vanderbilt Assessment of Leadership in

Education/VAL-ED (Porter et al., 2008) and the Comprehensive Assessment of Leadership for Learning/CALL (Halverson & Kelley, 2011), a formative assessment for secondary school principals. While neither requires that the respondent provide actual evidence to support the assessment of the leader's summative (VAL-ED) or formative (CALL) evaluation, there is sufficient attention to the multiple factors that constitute effective leadership, and the validity studies for each instrument indicate their use. The use is limited to a perceptual picture of leadership effectiveness that relies on constituent feedback (Portin et al., 2008).

To its credit, VAL-ED asks respondents to cite evidence they use to make perceptual judgments, and the survey results may be useful in reflecting on practice and deciding on goals for improvement. However, one of those evidence sources is "reports from others", which reinforces the "hearsay" issue we previously described. And, without analysis of specific qualitative or quantitative evidence that supports the perceptions, the leader is often in a quandary about precisely what to do or how to accomplish change.

In addition, while VAL-ED has a specific category for advocacy as indicated earlier in this paper, and that includes that the principal advocate for a "culture of learning that respects diversity" (p. 12). However, that is simply insufficient as a direction to the respondent and is inconsistent with our experience that an urban school leader needs to address deep structural issues of race and poverty. Because every tool or instrument has a specific purpose and cannot represent the multi-faceted ways we need to view and assess effective leadership, we recommend that the results from VAL-ED or CALL be combined with the other metrics we recommend in this paper. That will result in a more robust leadership evaluation framework.

The Leadership Accountability Report Card

As a result of our investigation and analysis of rubrics, tools and manuals, and in addition to the 360° survey, we propose a second component of the tripartite leadership evaluation framework: the Leadership Accountability Report Card (LARC). The LARC is meant to assist leaders and their supervisors in better assessing leadership performance through quantifiable metrics that comprise indicators of effective leadership and prerequisites for student achievement improvement.

The development of the LARC metrics emerged from a review of the literature and from the observations and evidence that leaders provided during a three-year professional development project with experienced PLI graduates (described in Section II on evidence and asset-based cycle of inquiry). We used this project as a workbench for developing the leadership rubric and professional development process detailed in the next section. As a result, we identified a set of research-based quantitative metrics over which the leader has substantial control and responsibility, such as school climate, school structure, and teacher satisfaction.

One example highlights how we approached the work from both research and practice. Bryk, et al. (2010) documented the importance of attendance as a necessary pre-condition for improved school achievement, and a principal we observed was analyzing data from attendance charts he had used at a staff meeting. The analysis indicated issues primarily in kindergarten and 5th grade. The principal identified several actions within his direct responsibility that would support the school in meeting its attendance goals, including regularly analyzing data at staff meetings and engaging teachers in talking to families. In contrast to student achievement, the principal has direct influence over working with parents and teachers to set up attendance and tardiness

expectations, set in motion systems for recognizing improvement, and keep this critical precondition for student outcomes in the forefront of his expectations for the school community.

In another example, a novice high school principal, in analyzing attendance data, observed a relationship between demographic composition of attendance/tardiness and achievement. He gained consensus to allocate resources to hire an Attendance Dean to improve attendance. As a result of these leadership actions, the school garnered significant resources and addressed issues of equity in the attendance of African American students.³ From these examples, it is clear that assessing improvement in attendance (and similar measures) offer quantitative indicators of potential change in school achievement success and constitute metrics over which the leader has substantial influence.

As such, we focused our search on research-based metrics over which the leader can have primary and direct impact rather than outcomes like test scores over which s/he has only indirect influence. In addition, what the leader does to affect improvements on those metrics is the subject of the third component of the tripartite approach when the leader sets goals and analyzes iterative evidence to determine which leadership actions lead to improvements.

Table 1, on the following page, lists a variety of quantifiable metrics that could be used or developed to assess leadership effectiveness. In deciding how to organize the metrics, we made decisions based on research and observations. The metrics are by no means exhaustive, and we recommend that the leadership field cull through observable and quantifiable metrics to glean a set of protocols, surveys, and rating scales that could offer reasonable quantitative metrics for assessing leadership effectiveness. For example, it would be important to have specific metrics that can isolate how collaborative leadership affects teacher working conditions and motivation (Leithwood, 2012) and how the major factors of instructional leadership – like the ratio of insider/outsider professional development facilitation and teacher retention -- have a cascading effect on student outcomes (Robinson et al., 2008).

The metrics are in two forms: data that exist within most school and are collected quantitatively, but need to be disaggregated for analysis (i.e., attendance or discipline data), and data that will typically need to be collected and converted to a metric (i.e., relational trust, teacher capacity). Some metrics apply to more than one area and are co-listed in the table. Some elements of leadership effectiveness that we have identified in the next section, including attributes like presence and attitude, do not necessarily lend themselves to quantitative metrics and have not been included in this list.

Thus, by combining existing data with additional survey data or data collected from other assessment tools, a leader better understands his/her work in a particular area. In addition to attendance data, for example, a leader would need to have a deep understanding of how the school's curriculum connects to both state and national standards in order to more fully comprehend the strength of his or her organizational leadership. For example, after examining the percentage breakdown of diversity in classes within the school, a leader would also need to know what concerted efforts are being made to *alter* diversity within the school so that class rosters – particularly for gatekeeper classes like Algebra or AP or elementary gifted and talented pullout classes -- more closely resemble the student body racial/ethnic makeup. Thus, pre and post data or longitudinal data would need to be collected and analyzed to make assertions about

³ In California, state funding is tied to Average Daily Attendance (ADA) calculations; the increased revenue in the year that the principal instituted an Attendance Dean position was \$255,000. The overall increase for students was 1.6% with an increase of 9300 total days; the increase for African-American students was 2.3%.

leader improvement. In addition, we are aware that the data must be interpreted in light of the school's demographics and may be then compared to similar schools. However, the individual leader can use the information from these metrics to address key areas that contribute to student achievement.

Clearly, we believe that further development of a comprehensive evaluation framework is necessary. In the interim we are pushing away from quantitative measurement of student achievement as an indicator of effective leadership performance and moving toward qualitative and quantitative metrics that better capture and measure effective leadership. We believe that the important long-term quantitative directives put upon principals and school leaders (e.g., long-term goals to decrease suspensions by 25 percent or increase student attendance by 10 percent) may be met best when we are better able to measure leader effectiveness using the qualitative, formative and summative interim data throughout the year. To that end, we offer an approach to principal leadership evaluation that embraces the quantitative methods while emphasizing the equal importance of the qualitative metrics described in the next section.

Table 1: Metrics for Leadership Accountability Report Card

Existing Data. <i>These data are typically collected on a regular basis in aggregate form in most schools and are or can be disaggregated and analyzed.</i>		Data to Collect and Analyze*. <i>These data are typically available, but not regularly collected and analyzed; they could be collected through surveys or other instruments and converted into metrics.</i>	
Identity and Relationships			
Note: Schools do not routinely collect or analyze relational trust data.		<input type="checkbox"/>	Relational trust data
		<input type="checkbox"/>	Climate and culture data
Equity and Advocacy			
<input type="checkbox"/> Student placement as evidenced by class schedules or student groupings		<input type="checkbox"/>	Levels of cultural awareness
<input type="checkbox"/> GATE/IB/ AP class rosters		<input type="checkbox"/>	Teacher attitudes toward students and families
<input type="checkbox"/> Discipline data regarding percentage of subgroup referrals		<input type="checkbox"/>	Student perceptions of equity
<input type="checkbox"/> Retention in grade level with regard to subgroup ratios		<input type="checkbox"/>	Teaching assignments, including the distribution of experienced teachers who demonstrate results
Curriculum and Instruction			
<input type="checkbox"/> Retention in grade level		<input type="checkbox"/>	Survey responses of teachers to leadership coaching and support
<input type="checkbox"/> High quality teacher levels		<input type="checkbox"/>	Amount of time per week spent in formal or informal observations and conversations with teachers
<input type="checkbox"/> Teacher retention rates		<input type="checkbox"/>	Ratio of internal (teacher-facilitated) PD to external (outsider-facilitated) PD
Organization and Systems			
<input type="checkbox"/> Student attendance and tardiness data		<input type="checkbox"/>	Time-coding and analysis of use of time per week
<input type="checkbox"/> Student truancy data		<input type="checkbox"/>	Class schedules
<input type="checkbox"/> Referrals and suspension rates		<input type="checkbox"/>	Teacher schedules
<input type="checkbox"/> Teacher retention rates		<input type="checkbox"/>	Teaching assignments

Change and Coherence	
Note: These data are not typically collected, because data on change and coherence typically involve the teaching staff and principals, with some input from other staff.	<input type="checkbox"/> Indicators of trust and support from principal among teachers <input type="checkbox"/> Principal knowledge of crosswalk between national standards and school curriculum <input type="checkbox"/> Levels of cohesion in school <input type="checkbox"/> Shared or collective leadership among teachers, staff <input type="checkbox"/> Occurrence and frequency of teacher learning communities

* Note: The metrics on the VAL-ED survey include many of these items.

Evidence from Rubric and Cycle of Inquiry

One key component of an effective evaluation framework relies on the use of evidence-based practice that includes use of a leadership rubric, which can be coupled with a cycle of inquiry or used by itself as a “stand-alone” metric. In the latter, the rubric raters need to agree on the levels of quality required for each rubric level, as is used in the Tennessee state evaluation system discussed earlier. Raters in that system have to agree what constitutes the level of performance of the principal in one of these four levels: aspiring, beginning, professional or exemplary. For each of the Tennessee standards, there are specific indicators for each of the respondents (self, director/supervisor, and peer/coach) to rate. To agree on measurement for each leadership standard in a rubric, the developers and users of the rubric engage in calibration of what they consider the level of effectiveness. We support this level of calibration, developed first as a response to writing rubrics in the 1970s and used widely throughout education since that time to determine ways of assessing performance and giving guidance through the standards about how to improve.

Griots

Griots are praise singers who have a repository of knowledge and the ability to extemporize on current events, weaving the past into the present – much like experienced urban leaders. They hold a repertoire of ways to relate to people and support the life of the community. In the African tradition, the praise singer provides support and a careful eye for the good of the community – in other words, a critical friend to the community. We believe that many of these qualities exemplify what the successful urban leader must embody in his/her practice.

However, rubric use can be broadened to guide the analysis of qualitative evidence that is embedded in a cycle of inquiry and a professional development process. We propose this process in Section II, in addition to the 360° feedback survey and the leadership accountability report card, as the third part of the tripartite leadership evaluation framework. The tripartite framework emphasizes a developmental, rather than a disciplinary conception of evaluation. We briefly introduce the Leadership Connection Rubric. The rubric elements respond to our analysis of what key dimensions or standards deserve more emphasis in evaluation, and those elements inform the qualitative analysis. The revised asset-based cycle of inquiry (COI) and the analysis required in engaging in the COI yield formative and summative evidence from addressing specific outcomes (See Figure 2 on p.23).

In developing the Leadership Connection Rubric, we engaged in field-testing the rubric by observing and analyzing the leadership work of 25 experienced PLI graduates. Concurrently, some of our graduates were engaged in a professional development project. The aim of the Griot

Project (See box, right) was to develop a supportive process that was rooted in the reality of leaders' everyday work. It included use of the rubric and developed over three years to include observations that resulted in evidence. In turn that evidence informed each leader's focus in the cycle of inquiry. During the last year of the project, we transferred the process to a group of experienced leaders in a partner school district as we were interested in testing out whether this process could be used as one part of evaluation. At this writing, that process in the district will continue with the district customizing it as a part of leadership evaluation.

The Leadership Connection Rubric: Elements of Leadership Effectiveness

Understanding the organization of the Leadership Connection Rubric is fundamental to the third component of the leadership evaluation framework and how we can gain substantial qualitative evidence from its use. The seven research-based rubric elements of the Leadership Connection Rubric identify the large conceptual components of the leadership chemistry and, in the right combination, produce a personal “formula” for effective leadership. We use the term elements as a normative standard. Each of the seven elements of the Leadership Connection Rubric includes three descriptors of practice (DoP). The DoPs further explicate the essential components of that element or standard, but the descriptors are too broad and normative to constitute an outcome that can be measured. Therefore, each DoP includes 2-5 indicators of practice (IoPs) in which detail or outcome is sufficient for the leader or the rubric rater to determine the level of effectiveness and determine observable outcomes that yield specific evidence.

Table 2 identifies two examples from the rubric to illustrate how the descriptors and indicators of practice delineate each element. The details of the indicators of practice (IoP) serve as outcome targets. Note that the developing level pertains to novice leaders and the practicing level pertains to experienced leaders. The presumption is that all practicing level leaders will incorporate the developing level outcomes. The full description of the rubric development process and the lessons learned from that process is in Appendix B.

Elements One, Two and Three of the Leadership Connection Rubric focus on essential personal and professional perspectives and attributes. However, we highlight some considerations that are critical to our rubric content and to the work of urban school leaders. The first three elements are vital prerequisites for leading schools and recognize that out-of-school factors—including an increasing wealth gap and a lack of basic services in under-resourced communities—have a bearing on student and school performance (Rothstein, 2004; Berliner, 2009). The school leader’s ability to incorporate specific and routine actions in these areas was a striking finding of field-testing the rubric, as these foundational elements were often more frequently represented in the qualitative data than other elements. The foundational three elements are: Element One—Presence and Attitude; Element Two -- Identity and Relationships; and Element Three—Equity and Advocacy.

We again call attention to the importance of equity as a key component of any rubric. In order to successfully interrupt the historical predictability of the effects of race and class on achievement outcomes, it is not enough to be a “no excuses” principal or repeat the “close the achievement gap” message. We believe that: “Gaps don’t just happen; there is a generative element inside them, a welling motion as when cold waters shoulder up through warm oceans” (Ryan, 2000, p. 24), and that the structural conditions of race and poverty and our inability as a society to address these issues sufficiently contribute mightily to the achievement gap. We support activist leaders who commit themselves to the principles of social justice and equity in their daily work and are aware of the opportunity gaps rooted in the large-scale and historical inequities and the absence of advocacy that created the conditions of unacceptable achievement outcomes for our most vulnerable students (Darling-Hammond, 2010; Boykin & Noguera, 2011).

We are fully aware of the ways that schools can seek to interrupt these historical patterns by unwittingly lowering standards and expectations of students, and basing their instructional choices on the familiar pedagogy of poverty. We are also aware of schools that respond with curricular and pedagogical choices that fully support students academic, civic and social-emotional outcomes, by holding consistently high expectations for students and emphasizing cognitive complexity in instruction. Those schools support opportunity to learn standards that operate side-by-side with rigorous core curricular standards. Equity and high expectations can and should operate hand in hand. Leadership evaluation systems that comprehensively embrace equity acknowledge opportunity gaps and provide guidance to leaders about how to discern the appropriate way to address them. Leaders and those who support their development

need a clear understanding about how increased attention to opportunity gaps at the school site can lead to increased achievement for all students.

Table 2: Leadership Connection Rubric: Examples of Descriptors and Indicators of Practice

Element 3: <u>Equity and Advocacy</u>: Advocate for equitable academic, civic and social-emotional outcomes for students who have been historically underserved by schools and society		
Descriptors of Practice (DoP) <i>3.1 Personal Commitment</i>	Indicators of Practice (IoP) <i>3.1.1 Equity Framework</i>	
<ul style="list-style-type: none"> • Enunciate a deep understanding of equity and social justice in the school and community (historical and present) in its many forms (race, class, gender, sexual orientation, cultural difference, ableness) • Map the school and community assets to surface strengths and frame possibilities • Assess the school situation and the readiness of constituents to engage in conversations 	<i>Developing level</i>	<i>Practicing level</i>
	<ul style="list-style-type: none"> • Develop systems for addressing staff practices and beliefs that are consonant with a socially just organization • Engage in conversations with staff based on an appraisal of the school's current assets • Support the implementation of changing practices in alignment with changing beliefs, adapting to needs of a particular school 	<ul style="list-style-type: none"> • Analyze and provide feedback on alignment of beliefs and practices • Assess readiness of staff to promote equity-driven outcomes • Routinely engage the constituents in fostering practices and beliefs that are consonant with a socially just organization • Frame all conversations and actions based on recurring appraisal of the assets and areas of needed change
Element 4: <u>Curriculum and Instruction</u>: Cultivate high expectations and ensure durable academic, civic and social-emotional learning outcomes for students and adults		
Descriptors of Practice (DoP) <i>4.3 Professional Growth</i>	Indicators of Practice (IoP) <i>4.3.1 Adult Learning</i>	
<ul style="list-style-type: none"> • Consistently demonstrate the value of informal and formal professional development for improving teacher practice and support staff outcomes • Model and facilitate exemplary professional development practices 	<i>Developing level</i>	<i>Practicing level</i>
	<ul style="list-style-type: none"> • Facilitate and enhance organizational structure and small and large group meetings that support formal and informal professional development • Co-develop systems to allocate time and resources to address professional development needs • Provide opportunities for formal learning for individual, grade/department/team levels, parents, staff and whole school development and informal structures for learning from one another • Model best practices for learning in adult meetings and professional development 	<ul style="list-style-type: none"> • Fully implement learning community structures that are premised on distributed leadership and shared facilitation • Collaboratively design and implement short and long-terms plans based on identified needs • Share responsibility for facilitation, leadership, engagement and implementation • Support teacher-leaders to use best practices for adult learning in meetings and professional development • Monitor structures to assess effectiveness of implementation

The foundational elements represented in Elements One-Three are the backbone of Elements Four-Seven, which are the common standards found in most leadership rubrics that have received more attention in the research (Louis et al., 2010; Wallace, 2012; Leithwood, 2010). Element Four encompasses Curriculum and Instruction; Element Five includes Organization and Systems; Element Six is Change and Coherence and Element Seven presents an equity focus to Assessment and Accountability. Our conception of Curriculum and Instruction, however, makes a key distinction that is counter to the current forcefulness with which we concentrate almost solely on the economic goals of schooling (Grubb, 2004). This element attends to the wider goals of schooling as preparing students for work, citizenship and family life by indicating that we need durable academic, civic and social-emotional learning outcomes for students and adults. In Organization and Systems, we are particularly interested in the leader's ability to focus on equitable student learning by aligning systems and operations that foster a culturally responsive environment. Under the banner of this element, the leader ensures processes for internal self-analysis and exhibits the ability to monitor and improve inequitable systems. The final element in this second group, Assessment and Accountability, involves a persistent focus on teacher and student learning outcomes by implementing an equity-driven assessment system. In responding to multiple accountability policies and mandates, the leader holds fast to using multiple measures to assess academic, civic and social-emotional outcomes. To reiterate, a rubric can be used as a stand-alone instrument when the users of the rubric engage in calibration activities to determine common ratings. However, we recommend that the rubric be extended for use in professional development, and the evidence from that use provides a key component of a comprehensive and fair evaluation framework.

Evidence from Leadership Practice

The professional development process that results in qualitative evidence is related to specific goals that are based on observations of a leader through videotaping and analyzing his or her practice. Because a "principal's job is complex and multifaceted, and the effectiveness of principals depends □ on their level of experience, their sense of efficacy on **particular kinds of tasks**, (emphasis added) and their allocation of time across daily responsibilities" (Rice, 2010, p. 1), the evaluation or performance assessment needs to focus on a specific task that builds efficacy. And particularly since the leader is charged with being an instructional leader, but actually spends only 13% of his or her day on instructional leadership (Hornig et al., 2009), we address the efficacy consideration by an alternate approach. The resulting process offers specificity and focus that can be one part of the comprehensive leadership evaluation framework.

We termed this process *in situ* professional development, as its focus is the leaders' daily work as the context and text for the PD content.⁴ In this section, we highlight how the revised cycle of inquiry results in formative and summative evidence related to goal setting and what we term temporal targets – monthly interim outcomes that are in service of a yearly or long-term goal.

In all the districts in which our graduates serve, principals are engaged in goal setting with a district supervisor. They typically use a cycle of inquiry and often use SMART goals as prescribed by Schmoker (1999; 2011); however these goals – even the short term "wins" of Schmoker were largely related to school level improvement goals. Leaders told us repeatedly that the conventional cycle of inquiry – focused on deficits or needs analysis – was not helpful to

⁴ In 2011-12, a second group of leaders from a school district group provided an additional group of 20 leaders engaged in asset-based observations and cycle of inquiry professional development. The evidence and stories provided in this section are from both projects.

either achieving observable results or their sense of efficacy. The goals were too broad and visionary and related primarily to the school plans, not the specific leadership actions that might contribute to those goals. An example is: *improve literacy scores on CST (California state tests) by engaging in professional development on the literacy curriculum*. The goal for the leader in this example is directly and incorrectly connected to student achievement and conflates leadership outcomes with student outcomes. A better approach is to focus on specific leadership actions related evidence from observations instead of broad outcomes related to full school improvement. Plus, the action statement is too broad. Leaders reported that they completed the goal statements, put them away until spring when they met with their supervisors again, and gathered just-in-time evidence to show how they had met their goals. This hastily gathered evidence was meant to “prove” that the leader had progressed on the goal instead of using evidence regularly to reflect on or inform decisions about practice.

As one example, a PLI graduate in the induction program (Leadership Support Program) expressed the need for more tangible goals that provided regular and concrete evidence on which he could reflect and make decisions. That leader suggested a process for time-coding his daily work as a mechanism to examine his practice, similar to the methodology used for assessing principal use of time in the 2009 study (Hornig et al., 2009). We relied on this representative story as an impetus to respond to many leaders’ frustrated attempts to reach the goals they set at the start of each school year. Through additional observations and conversations with leaders, we developed a professional development process based on collecting and analyzing meaningful evidence. The revised professional development process, rooted in cooperative inquiry practices (Cooperrider & Whitney, 1995) and activist research methodology (Hale, 2002), is described below.

Evidence from an Asset-based Cycle of Inquiry

A second issue was the deficit model of the conventional cycle, based on needs analysis or areas of improvement. Learning theory tells us that building on the current knowledge or skill and pushing the learning within the current zone of proximal development provides the most successful transfer (Vygotsky, 1978; Bransford et al., 2000). Thus, we developed an assets-based cycle of inquiry.

This revised cycle of inquiry relies on an initial observation of the leader’s assets in his or her leadership practice. An asset is the leader’s current level of practice on any given leadership element or dimension. An asset may be a strength that can be parlayed to other areas of leadership development, or it may be the current level of performance on a leadership dimension. The asset observation provides the evidence for decisions about a long-term goal but more importantly strategically designed temporal targets or short-term actions. Figure 2 conveys the asset-based cycle of inquiry. The starting point for this cycle is an asset observation—

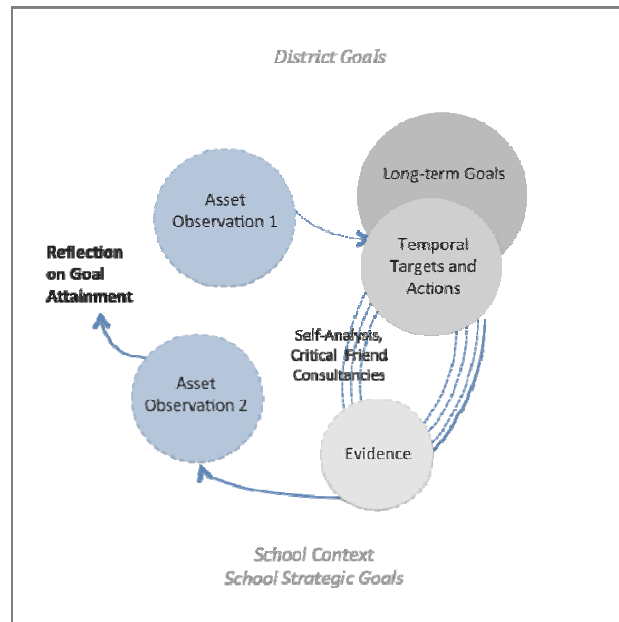


Figure 2. Asset-based Cycle of Inquiry

or observation of leadership practices resulting in low inference evidence that names (codes) the leader's practices. Qualitative evidence is collected in support of these initial questions:

- What assets do I have as a leader?
- What am I doing with my time?
- How can I apply those assets to setting a long-term goal with clear **temporal targets** and focus my **actions** on reaching those targets?
- How can I use the rubric elements, descriptors of practice (DoP), indicators of practice (IoP) and analysis of leader practices using names/codes of practice to drive my goals and targets?
- What **evidence** can I use from observations and other sources to ascertain progress?

To analyze the initial observation evidence, the leader has a conversation with the professional development facilitator, supervisor, and/or a critical friend (peer). As a result, the leader sets goals and temporal targets and reviews additional evidence in regular conversations. These routine opportunities for reflections inform iterative development and revision of temporal targets and actions. A second observation specifically related to the goal (typically conducted in the spring of each school year) provides outcome evidence. In our PD model, the leaders meet each month, bring evidence to review, reflect on the outcomes they have set, and revise as needed through conversations with their critical friend peer and facilitators.

The asset-based cycle of inquiry is situated in the “surround sound” of multiple contexts: the school, the district strategic plan and the school strategic plan, as well as the leader's own self-analysis and experiences. Of course, the leader has an eye toward long-term goals, but this PD process interrupts the old paradigm of long-term goal setting that has frustrated leaders with short-term temporal targets. The monthly targets keep the goal in mind, but support the leader to act and to collect evidence of actions that lead to outcomes. The conventional goals, even the “short win” goals of Schmoker (1999), are elusive; in contrast, these temporal targets – immediate and actionable -- are situated in the regular work of the leader, bolstered by support from a critical friend and supervisor or facilitator. By linking goals to assets and by concentrating on temporal targets that can be addressed immediately with concrete actions and produce evidence that can be used for reflection and revision, our experience suggests leaders are more efficacious and present in focusing on an important outcome and collecting and analyzing useful evidence.

Two examples elucidate the process and results. A third is represented in the story in a side box (p. 26). A high school principal, who had been an assistant principal (AP), was in his second year as principal. He supervised and had regular check-ins with four assistant principals who in turn supervised specific academies in the school. The initial asset observation included a conversation with one of the APs. As a result of the evidence analyzed from the observation, he recognized his assets: a clear equity focus on teaching practices related to standardizing writing assessments that would better support success for underachieving students; strong interpersonal relationships and trust with the AP; and humor. So how could he use those assets to address a focus area? In the video and analysis, he talked too much (approximately a 4:1 ratio of P:AP). The principal related to the AP how he might address the AP's issue at hand instead of asking questions to support the AP's problem-solving and leadership style. Thus, the long-term goal was to use his relational trust, focus on equity, and humor to: *Discern and use the appropriate stance for coaching APs in their work – that is, move from direct informational/instructional to facilitative to transformative coaching), including asking more questions instead of telling.* After a month when he had accomplished the initial targets so he could distinguish the stances and questions, the next target was to list typical formats or questions that he could use as a cheat-sheet in a meeting with the AP and try them out. One example was: *This is what the large picture or outcome is. What is the process? What supports do you need from me?*

There were multiple steps between the goal-setting and second observation, including revised targets and conversations with his critical friend and facilitator. In the final video, which provided qualitative evidence, there were observable shifts in practice. When he asked questions, he observed that the AP had ideas and took more responsibility for choosing actions. When he reverted to his former style of telling, instead of asking, the AP also reverted. This informed his future actions, as he wanted to be true to his vision of a collaborative administrative team structure and his core beliefs about distributed leadership. The equity focus was apparent in the final conversation; he had the core direction, but he could see from concrete qualitative evidence that his goal was only partially met and that he still needed to shift his practice if shared decision-making and collaboration was to move from espoused to enacted theories of practice (Simons, 1999).

A second story involves a middle school principal – ten years into his tenure and still confronting the specter of how he was or was not an effective “instructional leader”. When he reviewed the analysis of the evidence from the first observation of his leadership practice that included names or codes of practice for each of his leadership actions in 1.5 hours, he was astounded by the number of codes related to his considerable assets as a leader (especially Element Two which included visibility and interpersonal communication) and was intrigued by the question: How can I use my leadership assets to address instructional leadership? His long-term goal was: *Develop capacity for instructional support to teachers that fits my leadership style.* What emerged as an initial temporal target was using informal observation time differently: being more intentional about instructional conversations as he moved through the building having informal conversations. He began by time-coding his conversation content each 15 minutes so that he would have baseline evidence of what he actually did and said.

The second temporal target was to use his assets in visibility (an indicator of practice in Element Two of the rubric) to intentionally cross-pollinate instructional practices as he checked in with teachers. This changed his perception of what constituted instructional leadership from more structured observations conducted in a specific way to understanding these brief conversations as intentional instructional leadership using what he did best – cross-pollinate ideas among teachers and use his visibility in the halls at passing to talk to teachers in focused ways. In addition, he recognized that if he conducted brief observations, checked in with several students, and concentrated on several key practices in each classroom, he had a stronger hand on the pulse of instruction across the school. He began focusing on the use of academic vocabulary across all classrooms as well as ratio of teacher to student talk. His final leadership video included a post-conference with a first year teacher that was less formal than he had conducted before, but was based on multiple sets of evidence from the teacher’s classroom.

These stories represent the use of specific evidence that leads to goal attainment and tangible changes in leadership practice. But how can concentrating on a narrow focus be a part of a framework from leadership evaluation? A principal in our partner district is clear about how: *What you actually need in this work is to be self-reflective and be able to identify your own strengths and areas of improvement. This process coaches people to be able to do that. That is my argument about how you could evaluate a person holistically on their whole practice by taking a slice of the work. What you are evaluating is their ability to be self-reflective on one part of their work so that can transfer to other parts of their work.”*

These stories, on some level, are not astounding school leader tales. However, they reveal that for the leaders involved, regular success in targeted areas affected other parts of their work. The importance lies in the full process and the sense of efficacy that results – how important it becomes to ground one’s goals in assets, routinely use specific evidence to document and reflect

on practice, and then institute systems that are visible to other constituents as critical improvements leadership practice. The evidence from the leaders' stories informed the leadership rubric architecture and content and provided clear evidence of targeted outcomes that constituted observable improvement.

Building Collective Responsibility: The Story of a Middle School Leader

For a middle school leader in a large urban district, the **first asset observation** that was conducted produced a set of results related to organizational and management assets. The leader was aware from conversations with teachers and the instructional coach that she was deemed well organized, but that she missed the relational glue necessary for a sense of mutual belonging and collective responsibility for outcomes. The leader talked about how to use her assets to address relationships, and she knew that she had to plan for those interactions and for including more community-building time on staff agendas.

The leader's **long-term goal** was to: *Be intentional about micro-political conversations with teachers to validate them and build common understanding of the vision.* Her overall actions and evidence included: Using feedback forms at each staff meeting regarding her accessibility, videotaping subsequent meetings, sending memos or thank you notes to teachers and keeping copies to analyze frequency and results, and examining agendas, notes and feedback forms for changes in practice.

The initial **temporal targets** were scheduling and preparing one agenda for a staff meeting. She proceeded to set monthly targets in reflection with her critical friend, a final set of feedback from teachers, and informal conversations with selected teachers and the instructional coach to assess her progress. These tangible "**evidence moments**" throughout the year marked forward progress. She shared this goal with her district supervisor and sought input from her as well. By the conclusion of the year, this second year principal had used her assets to "schedule" building relationships, and she as well as those she checked in with reported that she had made important strides in communicating a common vision and building teacher-administrator relationships. She reported that she had a stronger sense of efficacy and that teachers initiated more informal conversations with her.

The **second asset observation** indicated that she had met the goal for the year. Her reflection indicated that setting a clear and observable target with the regular meetings to reset the temporal targets had contributed to her ability to move forward. Her regular meetings with a critical friend was a key factor in keeping faith with herself to address her role as instructional leader.

III. IMPLICATIONS FOR THEORY AND PRACTICE

In the working paper, we have argued for a tripartite leadership evaluation framework that includes a trio of instruments and processes: a reliable and validated 360° feedback tool, an analysis of multiple quantitative metrics, and the use of a rubric in combination with an evidence-based professional development process. We believe it is not useful to debate whether instructional leadership, transformational leadership or managerial leadership is critical. As Cuban (1987) indicated, leadership is all of the above and more. Leaders must be inspirational and collaborative as well as technical data managers. They must, above all, stay true to the original role of school leadership: principal-teacher. As tempting as it is to summarize leadership effectiveness through student outcomes only, that obscures the ability to see which leadership actions lead to increased supports for achievement.

An important epiphany from our rubric work is that the ability to engage in collective leadership -- what Spillane, Diamond and Halverson (2002) term distributed leadership -- means that the principal as single person does not need to have the highest level of performance on every rubric component. In fact, she or he can build on his or her assets and make certain that the complex functions of school improvement, including leadership in various aspects that may not be the strengths of that particular leader, can be shared. What we came to understand is that a leader who had a composite of assets across the elements (or standards) did not need the highest level of performance in every descriptor or indicator of practice to be "successful" if s/he had the ability to engage in collective leadership and draw on the cognitively distributed function of leadership among all the adults in the school community. While a leader should continue to engage in improvement, using his or her assets, on the research-based areas that will have the strongest effect on student outcomes, s/he does not have to superwoman or superman to be effective.

The evaluation framework was designed to manifest a particular programmatic theory of action and mission. We are well aware that what gets counted matters. If we do not expand and deepen what matters in our understanding of evaluation, we may dangerously neglect to develop in school leaders the most significant skills and capacities they need to make a difference in academic achievement for all students. Our findings are not only rooted in research; they are tried and tested through "on the ground" engagement, feedback from the beneficiaries of the tools and examination against results from leader goals. What we are proposing has already made a difference to multiple leaders. Our framework has become an integral component of our pipeline of support for leaders. We offer this paper in the hopes that our work may inform other preparation programs, district offices and researchers. But most of all, we hope that the process informs school principals who are on the frontlines of school reform. They need to be validated for the work they do well, understand better how to address what they need to improve in their practice, and have more supportive evaluation tools to assess how they parlay their assets to rigorous outcomes that support teachers and improve student outcomes.

APPENDIX A: Supporting Tables – Instrument and Rubric Review

The rubrics and instruments included in the review were identified through an examination of leadership evaluation literature and through conversations with experts working in the leadership development space. In total, six rubrics and six instruments were reviewed. The focus of the review was to assess the rubrics and instruments on the following parameters:

- Categories of assessment – What elements of leadership practice does the rubric or instrument assess?
- Indicators of rigor – Is there evidence that the instrument or rubric was tested for validity and reliability, was field-tested, etc.?
- Specific to high needs – Was the rubric or instrument designed for high need school settings?
- Levels of leadership – Are various levels of leadership addressed – ranging from, e.g., novice to experienced?
- Evidence collected – For evaluation instruments only, what types of evidence or data, if any, are collected to support the evaluation?
- Respondent – For instruments only, who are the intended respondents? Teachers, principals, supervisors, other stakeholders?

Table 1, below, details the results of the review of the rubrics. The rubrics reviewed include (title, author, year):

1. Comprehensive Assessment of Leadership for Learning (CALL), Rich Halverson and Univ. of Pittsburg Institute for Learning (2005)
2. Connecticut School Leadership Standards, Connecticut Department of Education (2012)
3. Leadership Performance Planning, New York City Leadership Academy, Inc. (2010)
4. Leadership Proficiency Roadmap, KIPP (2011)
5. Principal Leadership Actions Rubric, New Leaders (2009)
6. Reeve's Leadership Performance Matrix, Douglas Reeves (2011)

Table 2 details the results of the review of the instruments. The instruments reviewed include (title, author, year):

1. Comprehensive Assessment of Leadership for Learning (CALL), Wisconsin Center for Education Research, Univ. of Wisconsin (2009)
2. Education Leadership Self Inventory, Connecticut State Board of Education (2001)
3. Learning Centered Leadership Policy, Tennessee State Board of Education (2011)
4. Massachusetts Educator Evaluation Regulations, Massachusetts Department of Elementary and Secondary Education (2011)
5. Tennessee Instructional Leadership Standards Appraisal Instrument, Tennessee State Board of Education (2011)
6. Vanderbilt Assessment of Leadership in Education (Val-Ed), Andrew Porter, Joseph Murphy, Ellen Goldring, Stephen N. Elliott (2005)

1. Rubrics

Category of Analysis		CALL	CT School Leadership Standards	Leadership Performance Planning	Leadership Proficiency Roadmap	Principal Leadership Actions Rubric	Reeve's Leadership Performance Matrix
Assessment Categories	Personal Dimensions		③	③	③	③	③
	Ethics		③				③
	Communication			③			③
	Vision and Mission		③		③	③	
	Relationships, Partnerships						
	Equity and Diversity						
	Curriculum and Instruction	③					③
	Operations Management	③	③	③	③	③	
	Culture and Climate	③	③		③	③	
	Community Outreach		③				
	Teaching and Learning	③	③	③	③	③	③
	Professional Development					③	③
Type of Feedback	Developmental				③	③	
	Corrective	③	③	③			③
	Summative						
	Formative				③	③	③
Indicators of Rigor*	Tested for Validity	③					
	Tested for Reliability	③					
	Aligned to ISCC	③	③				
	Field Tested						
	Research Based	③	③				③
Specific to High Needs?	Yes			③	③		
	No	③	③			③	③
Levels of Leadership	Yes				③	③	
	No	③	③	③			③

*For those rubrics left blank for indicators of rigor, there was not sufficient information to assess them on each of the indicators.

2. Instruments

Category of Analysis		CALL	Education Leadership Self Inventory	Learning Centered Leadership Policy	MA Educator Evaluation Regulations	TN Instructional Leadership Instrument	VAL-ED
Assessment Categories	Personal Dimensions			③			
	Ethics			③	③	③	
	Communication		③		③		③
	Vision and Mission		③		③		
	Relationships, Partnerships		③		③		
	Equity and Diversity		③	③		③	
	Curriculum and Instruction		③		③	③	③
	Operations Management	③	③	③	③	③	
	Culture and Climate	③	③		③	③	③
	Community Outreach		③		③		③
	Teaching and Learning	③	③	③		③	③
	Professional Development	③	③	③		③	③
Respondent	Leader	③	③	③	③	③	③
	Supervisor	③		③			③
	Teacher				③		③
	Other Administrator					③	
Type of Feedback	Developmental	③	③	③		③	③
	Corrective	③			③		③
	Summative		③	③	③	③	③
	Formative	③		③	③	③	③
Supporting Evidence Collected	Self Assessment Data		③	③	③	③	
	Observation Data	③		③	③	③	③
	Student Achievement Data				③	③	
	Other**						③
Indicators of Rigor*	Tested for Validity	③					③
	Tested for Reliability	③					③
	Aligned to ISCC	③					③
	Field Tested						③

	Research Based	③					③
Specific to High Needs	Yes						
	No	③	③	③	③	③	③
Levels of Leadership	Yes	③		③		③	
	No		③		③		③

- For those instruments left blank for indicators of rigor, there was not sufficient information to assess them on each of the indicators.

** Other includes state approved survey of schools' climate, school data analysis, and student and staff surveys

APPENDIX B: Rubric Development – Design Principals and Lessons Learned

The context of the rubric development was a set of university-based preparation, induction, coaching, and professional development programs for urban leaders. In the preparation and induction programs, which provide credentials, candidates are assessed but not evaluated as schools and districts do. Therefore, we chose a developmental approach and identified three levels of leadership development: emerging (leaders in preparation, which is typically 1-2 years); developing (leaders in an induction or novice phase, which is typically 1-3 years after taking a leadership position, although it requires a longer process from novice to expert than three years); and practicing (leaders in years 3 and above as they move to veteran phase of their work).

The rubric development process was rooted in participatory action research, which requires that relevant stakeholders are involved in the research process, through examination of their own activities to inform possible collective action on problems (Freire, 1970). Specifically, we drew upon Hale's (2002) activist research frame and the following propositions: (1) deeper empirical knowledge from “on the ground” sources leads to better outcomes and more informed research and (2) research outcomes, as well as action outcomes, are better improved when tensions are surfaced and used to inform subsequent actions – in this case subsequent rubric development.

The rubric development process drew on relevant literature and systematic observations of over 400 Bay Area school leaders' experiences, including 40 leaders who offered direct input. This process was iterative and collaborative, as new research and reports were continuously examined and fed into the rubric design. Leaders provided feedback during early design phases and rubric field-testing.

Over the four-year rubric development process we tested assumptions about the key elements of effectiveness in the reality of daily leadership actions and results. Specific points of inquiry included:

- Reviewing relevant and recent literature on leadership practice and criteria for leadership evaluation
- Developing and revising the seven elements, the descriptors of practice (DoPs) and the indicators of practice (IoPs)
- Field-testing the rubric components by observing and coding observations of urban school leaders;
- Engaging in conversations and feedback from the practicing leaders and district level partners;
- Implementing a concurrent professional development for veteran leaders that included a qualitative analysis of leadership effectiveness; and
- Aligning our preparation and induction programs with the rubric.

The architecture of rubric construction is complex and requires reflection and multiple iterations. What became apparent in analyzing other rubrics was the need to emphasize the “on the ground” evidence of leadership actions and results that was missing from other tools. By developing processes that operated in “real time” with leaders, we hoped to strengthen the research-theory-practice connection and inform the rubric development through implementing and testing an analytical process that could be replicated by others.

Design Principles

Building a thoughtful rubric requires iterative attention to several construction principles: (1) the elements (or standards) that are chosen should be based on the organizational values about what constitutes strong and effective leadership; (2) explicit and observable language should be used to communicate expectations of performance; (3) decisions about assigning relative value to observed leadership actions requires field-testing; and (4) levels of quality need to be determined. Each of these is discussed in turn, below.

Elements that communicate organizational values. By choosing key standards, or elements, the organizational entity (state, district, principal preparation program, or non-profit) communicates a philosophical stance and expresses the core values of the organization. As previously discussed, the first three elements were the foundation and the additional four should be viewed through an equity lens. With substantial input and multiple drafts, the process promotes organizational coherence, which is one of the goals of the rubric development: to communicate an organizational purpose and direction, via rubric content.

Explicit and observable language. – Some rubrics suffer from descriptive language that is vague and subject to multiple interpretations. While that is not entirely avoidable in documents that forward a set of norms for leadership behaviors, attention to observable actions and specific actions is key. Using Element 7 Assessment and Accountability in the Leadership Connection Rubric as an example, at the level of element, the language is largely normative, broad and visionary: Exhibit a persistent focus on teacher and student learning outcomes by developing, aligning, and monitoring an equity-driven assessment system. The actual leader actions/behaviors used for assessment are at the Indicator of Practice (IoP) level and are written in behavioral language that communicates observable and documentable actions. Therefore, the specific rubric language guides what to do and at the same time, an observer can assess the effectiveness of the leader conducting observations of actions.

Relative Value Much like grammar subscribes to the principle of parallel construction, the relative weight of the contents requires discussion and calibration of the rubric (element, DoP, IoP). We used the visual term “bucket” to represent the elements in our rubric. The term “grain size” (see sidebar) helped us understand if the categories of leadership work had the same relative value or importance. Thus, the questions were: What are the large buckets of performance? What are the subcategories of those buckets? Is the grain size appropriate for similar leadership descriptors of practice or indicators of practice? The rubric construction process assumes the reader or user recognizes the order and breakdown as subsets of the larger standard or bucket of the leadership performance. However, the relative weight or grain size of these sub-categories must maintain a consistent strength or value. The names of these sub-categories are critical, as the sub-category typically communicates the details and visible actions of the standard.

Grain Size

Imported as a term of art from agronomy and metallurgy, grain size in the rubric development process was used to understand the relative weight of an element, descriptor of practice and the subset of the descriptor, the indicator of practice.

After drafting the rubric elements and descriptors and indicators of practice, we used a qualitative analysis process for field-testing the rubric. By scripting and coding 25 one-two hour observations, we were able to determine the degree to which the draft rubric was represented in

leadership practice. For example, the practice of “overlapping” was not identified as an indicator of practice in the original rubric; however, after the analysis of multiple observations, it was clear that the leader’s day involved facilitating meetings, observations, conversations and general movement in the building. The leaders were, therefore, exhibiting a degree of multi-tasking that we termed “overlapping” because the multiple actions were occurring simultaneously. Although research informed initial rubric development, the on-the-ground documentation and analysis of leadership actions was absolutely vital in the co-construction and co-ownership of the rubric content.

Levels of quality. Rubrics indicate levels or degrees of quality. The quality in leadership rubrics is typically expressed as developmental or standards-based. A development assessment responds to this question: Where is the leader on the novice to expert continuum? For example, the Tennessee Instructional Leadership Standards Appraisal Instrument uses the categories of Aspiring, Beginning, Professional and Exemplary as four levels of development of leadership assessment. A standards-based assessment rates the level of performance or quality, regardless of experience, responding to this question: What is the level of performance according to the standard, as specified on the rubric? In contrast, the Indiana Department of Education Principal Effectiveness Rubric uses four ratings: Highly Effective (4), Effective (3), Improvement Necessary (2), or Ineffective (1).

For the Leadership Connection Rubric, we chose a developmental model, with three levels: Emerging (in preparation), developing (novice leaders from 1-3 years) and practicing (beyond 3 years). Within these developmental levels, we did not use a rating of quality, but there could be levels of performance attached to career experience levels of leaders. For example, within the novice or developing level, the leader could be rated as approaching, meets standard or exceeds standard.

In summary, the discussion of rubrics and their construction establishes what values are held by the state, district or organization. The Leadership Connection Rubric represents our vision of leadership development with an equity focus as the core of the leadership mission and organizes the entire rubric on that principle. The language at the level of indicator of practice establishes leadership actions in observable verbs. The rubric is by design developmental, rather than evaluative, and the relative value assigned to each level of element, descriptor of practice, and indicator of practice has been calibrated among the designers and field-tested using a professional development process that we will discuss.

Lessons Learned

Multiple methodological and practical questions arose during the design process, which led to important lessons informing our work. These include:

Collaboration is Necessary. As previously indicated, we used a field-testing process that included observing 25 leaders and naming (coding) the leadership evidence. The naming was intended to develop a common language for engaging in conversations about leadership evidence. Reciprocity between the rubric developers and the school leaders was key, and required careful attention to terminology that would validate the observable leadership actions, in common language, helping to diminish the outsider-insider dichotomy of many efforts in school improvement (Honig & Hatch, 2004). An example of a leadership practice that was collaboratively coded related to the ways in which leaders change their tone/language, throughout the day, as they talked – face-to-face, by phone, or via email – to parents, students, teachers, staff or the community. This shift in tone and word choice is often undertaken without conscious thought. The standard reply when this practice was observed and noted to leaders

was: “I guess I do that” or “Really? I never thought about that before; I just do it.” We named (coded) this practice “register” and began to understand that giving voice to these daily authentic actions meant that leaders could become more intentional about how and when to employ “register” in their work.

Responsible implementation of terminology.

However, because rubric designers were engaged in crafting the rubric as well as analyzing the research and the aggregate observations, they often had a more nuanced and layered understanding of the appropriate name for a leadership action. If asked about the usefulness of a name (code) too early in the process, the principals might not decide it was useful. For example, the name “buffering” (see sidebar) was not familiar to most of the leaders with whom we worked. Intuitively, as leaders discussed, used and defined the term “buffering,” analyzing examples from observations, they participated in sense-making about its meaning. Only after its consistent use in the professional development process, however, did leaders start to use it as a way to describe their work, finding a useful way to describe their “middle manager” role. Thus, responsible implementation included decisions about when and how to use the process of naming as a professional development opportunity, when to rely on the greater expertise of those more deeply engaged in the development of the rubric and the research literature, and when to rely on feedback that the name was not conducive to the leaders’ experiences (Dewey, 1938).

Buffering

The leader acts as a middle manager who a) makes decisions about what information to share and what information to keep confidential for the strength of the community vision and actions; b) articulates a system for filtering information, including evaluating whether it is relevant to an equitable vision; c) consistently employs optimizing discourse that fosters a “both-and” approach; and d) urges constituents to focus on long term and productive solutions and maintain an appropriate level of transparency with one another.

Research-based. A third lesson learned involved the question, “How should research literature inform the rubric?” The literature served as an ongoing source of direction for rubric development. However, much of the literature omitted foundational aspects of leadership observed in practice. The literature provided only the most generic standards related to the civic and social-emotional goals of schooling, and virtually ignored a focus on equity. Ignoring race, culture, language, ethnicity and poverty sidesteps underlying structural issues and is inconsistent with the “on the ground” reality of urban and rural principals, whose schools serve a high numbers of students of color, many of them poor and/or non-native English speakers. Thus, we made an intentional decision to use a broader literature base for rubric construction, particularly for enunciating the equity considerations as an organizing principle for the entire rubric (Theoharis, 2009; Leadership Learning Community, 2009; Sklra et al., 2004).

In sum, the particular questions and points of inquiry mentioned above forced contemplation of how to best incorporate terminology, how to assess research criterion and how research literature could inform the rubric. While these various dilemmas brought challenges to the process of rubric development, they ultimately sharpened thinking on the part of developers and leaders and honed the rubric. The rubric development, relying on an activist research methodology and the field-testing process, began to inform how to use the rubric for professional development and formative assessment of leadership effectiveness.

APPENDIX C: Leadership Connection Rubric - Short Form

The Leadership Connection Rubric assesses the effectiveness of school leaders at different levels/stages of their career (from preparation to novice leader to expert). While the LCJE Rubric is primarily focused on the individual leader and his or her elements of leadership, the LARC is designed to place the leader within the context of the school. Since school leadership occurs both in the individual as he or she relates to the context, it is influenced by the context in which the leader operates. Thus the LARC addresses the contextual variable. A leader's action space varies depending on the school context.

The rubric is premised on seven key elements, highlighted in boxes. Elements 1-3 are professional and personal domains that affect the ability of the leader to focus on Elements 4-7. Element 4 is the cornerstone of the work: the leader as the curricular and instructional guide. Elements 5, 6, and 7 are fundamental to achieving durable student and adult outcomes, the focus of Element 4.

Each element is designed with three **descriptors of practice** and two to five **indicators of practice**. All elements, descriptors of practice and indicators are research-based. On the full rubric the research is referenced. In addition, each of these is aligned to leadership growth and development beginning with leaders in preparation (emerging) and moving to induction years (developing) and finally to the fully practicing administrator, who continues to develop levels of expertise. The organization of the elements, descriptors of practice and indicators are diagrammed on the following pages and represent a "short form" of the full rubric.

Element 1: PRESENCE AND ATTITUDE Communicate a compelling presence and a steadfast belief in the power of the possible.		
PROFESSIONAL IMPRINT Core Values Vision Resiliency	FLEXIBILITY Reframing Buffering Brokering	DEMEANOR Emotional Acuity Cultural Consonance Stance and Tone
Element 2: IDENTITY AND RELATIONSHIPS Demonstrate personal and professional self-awareness and nourish trusting relationships in a culturally and racially diverse learning organization.		
SELF-ASSESSMENT Identity Power and Privilege Initiative	INTEGRITY Discernment and Action Decision-making Reflection	INTERDEPENDENCE Modeling Optimizing Rituals and Celebrations Re-engagement
Element 3: EQUITY AND ADVOCACY Advocate for equitable academic, civic and social-emotional outcomes for students who have been historically underserved by schools and society.		
PERSONAL COMMITMENT Equity Framework Equity Knowledge and Action Equity Data	DEMOCRATIC LEADERSHIP Civic Goal Participatory Structures: Adults Participatory Structures: Students	ADVOCACY Individual Advocate Role Community Mapping Collective Advocacy
Element 4: CURRICULUM AND INSTRUCTION Cultivate high expectations and ensure durable academic, civic and social-emotional learning outcomes for students and adults.		
LEARNING ENVIRONMENT Expectations Courageous Conversations Knowledge Base	STANDARDS PEDAGOGY Content Standards Opportunity to Learn Standards Equitable Content of Curriculum Evidence-based Practice	PROFESSIONAL DEVELOPMENT Adult Learning Stance Formative Feedback Teacher Outcomes

Element 5: **ORGANIZATION AND SYSTEMS**

Align systems, structures, and resources that sustain
a culturally consonant environment in the service of student learning.

PERSPECTIVE

Ecological Mapping
Organizational Readiness
Systems Leadership

STRUCTURES

Operations
Governance and Legal Structures
Resource Management
Discipline Systems
Family and Community Engagement

MONITORING

Organizational Identity
Information Flow
Organizational Balance

Element 6: **CHANGE AND COHERENCE**

Engage all adults in change efforts that respond collectively and coherently
to the assets and challenges in schools and communities.

SCHOOL READINESS

Preparation for Change
Approach to Learning
Building Capacity

COLLABORATIVE GOALS & ACTIONS

Problem-posing Stance
Maintaining Focus

REFRAMING PERSPECTIVE

Optimizing Potential
Situational Awareness
Creativity, Flexibility, and Resilience

Element 7: **ASSESSMENT AND ACCOUNTABILITY**

Exhibit a persistent focus on teacher and student learning outcomes
by developing, aligning, and monitoring an equity-driven assessment system.

**PERSPECTIVE AND
KNOWLEDGE BASE**

Standard-Aligned and Equity-Based
Assessment Literacy
Integration

**MANAGEMENT AND
ACCOUNTABILITY**

Organizing Systems
Technology Literacy
Monitoring Implementation
Communicating Outcomes

**EVALUATION OF
EFFECTIVENESS**

Impact on Learning
Impact on Instruction
Supervision and Evaluation

REFERENCES

- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R.J., & Shepard, L.A. (2010). "Problems with the use of student test scores to evaluate teachers." (Briefing Paper No. 278). Washington, DC: Economic Policy Institute.
- Berliner, D. (2009). "Poverty and Potential: Out-of-School Factors and School Success." Tempe, AZ: Education and the Public Interest Center & Educational Policy Research Unit.
- Bottoms, G. & Schmidt-Davis, J. (2010). *The three essentials: Improving schools requires district vision, district and state support*. New York: Wallace Foundation.
- Boykin, A.W., & Noguera, P. (2011). *Creating the Opportunity to Learn: Moving from Research to Practice to Close the Achievement Gap*. Princeton, NJ: Association for Supervision and Curriculum Development.
- Branch, G. F., Hanushek, E. A., & Rivkin, S. G. (January 2012). "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals." (NBER Working Paper No. 17803.) Washington, D.C.: National Bureau of Economic Research.
- Brandt, C., Mather, C., Oliva, M., Brown-Sims, M., & Hess, H. (2007). "Examining district guidance to schools on teacher evaluations policies in the Midwest Regions." (Issues & Answers Report No. 30). Washington, D.C.: U.S. Department of Education.
- Bransford, J., Brown, A., & Cocking, R. (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, D.C.: National Academy Press.
- Braun, H., Chudowsky, N., & Koenig, J. (2010). *Getting Value Out of Value-Added: Report of a Workshop*. Washington, D.C.: Committee on Value-Added Methodology for Instructional Improvement, Program Evaluation, and Educational Accountability.
- Bryk, A., Sebring, P., Allensworth, E., Luppescu, S., & Easton, J. (2010). *Organizing Schools for Improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.
- Council of Chief State School Officers. (2008). "Educational Leadership Policy Standards: ISLLC 2008." Washington, D.C.: CCSSO.
- Clifford, M., Menon, R., Gangi, T., Condon, C., & Hornung, K. (April 2012). "Measuring School Climate for Gauging Principal Performance: A Review of the Validity and Reliability of Publicly Accessible Measures." Washington, D.C.: American Institutes of Research.
- Clifford, M., Sherratt, E. B., & Feters, J. (May 2012). "The Ripple Effect: A Synthesis of Research on Principal Influence to Inform Performance Evaluation Design." Washington, D.C.: American Institutes for Research.
- Cooperrider, D. L., & Whitney, D. (2005). *Appreciative inquiry: A positive revolution in change*. San Francisco, CA: Berrett-Koehler Publishers, Inc.
- Cuban, L. (1987). *The managerial imperative and the practice of leadership in schools*. Albany, NY: SUNY Press.

- Danielson, C. & McGreal, T. (2000). *Teacher Evaluation to Enhance Professional Practice*. Princeton, NJ: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. (2010). *The Flat World and Education: How American's Commitment to Equity will Determine our Future*. New York, NY: Teacher's College Press.
- Davis, S., Kearney, K., Sanders, N., Thomas, C., & Leon, R. (2011). "The Policies and Practices of Principal Evaluation: A review of the literature." San Francisco, CA: WestEd.
- Dewey, J. (1938) *Experience and Education*. New York, NY: Collier Books.
- Ferguson, A. A. (2000). *Bad boys: Public schools in the making of black masculinity*. Ann Arbor, MI: University of Michigan Press.
- Freire, P. (1970). *Pedagogy of the Oppressed*. New York, NY: The Seabury Press.
- Goldring, E., Cravens, X. C., Murphy, J., Elliott, S. N., Carson, B., & Porter, A. C. (March 2008). "The Evaluation of Principals: What and How do States and Districts Assess Leadership?". Paper presented at the American Educational Research Association.
- Grubb, N. & Lazerson, M. (2004). *The Education Gospel: The Economic Power of Schooling*. Cambridge, MA: Harvard University Press.
- Grubb, W.N. (2009). *The Money Myth: School Resources, Outcomes, and Equity*. New York, NY: The Russell Sage Foundation.
- Hale, C. (2002). "Does Multiculturalism Menace? Governance, Cultural Rights and the Politics of Identity in Guatemala." *Journal of Latin American Studies*, 34(03), 485-524.
- Halverson, R. & Kelley, C. (2011). *Comprehensive assessment of leadership for learning (CALL)*. Retrieved from <http://www.call.wceruw.org/resources.html>.
- Hamilton, L.S., Engberg, J., Steiner, E.D., Nelson, C.A., & Yuan, K. *Improving school leadership through support, evaluation, and incentives*. Pittsburgh, PA: RAND Corporation.
- Hill, T. (2011). "Every Closed Eye Ain't Sleep: African American Perspectives on the Achievement Gap." Lanham, MD: Rowman & Littlefield Education.
- Honig, M.I., & Hatch, T.C. (2004). "Crafting coherence: How schools strategically manage multiple, external demands." *Educational Researcher*, 33(8), 16-30.
- Horng, E., Klasik, D., & Loeb, S. (2009). "Principal Time-Use and School Effectiveness." (CALDER Working Paper 34). Washington, DC: The Urban Institute.
- Jacobson, P., Reavis, W., & Logsdon, J. (1941). *Duties of School Principals*. New York, NY: Prentice-Hall.
- Knapp, M., Copland, M., Ford, B., Markhold, A., McLaughlin, M., Milliken, M., & Talbert, J. (2003). "Leading for Learning Sourcebook: Concepts and Examples." Seattle, WA: Center for the Study of Teaching and Policy – University of Washington.

- Knapp, M., Copland, M., Honig, M., Plecki, M., & Portin, B. (August 2010). "Learning-focused Leadership and Leadership Support: Meaning and Practice in Urban Systems." Seattle, WA: Center for the Study of Teaching and Policy – University of Washington.
- Kruse, S., Louis, K. S., & Bryk, A. (1994). "Building professional community in schools." Madison, WI: Center on Organization and Restructuring of Schools – University of Wisconsin.
- Leadership Learning Community. (2010, July). *Leadership & race: How to develop and support leadership that contributes to racial justice*. Oakland, CA: LLC.
- Leithwood, K., Louis, K.S., Anderson, S., & Wahlstrom, K. (2004). "How leadership influences student learning." St. Paul, MN: Center for Applied Research and Educational Improvement – University of Minnesota.
- Leithwood, K., Harris, A., & Strauss, T. (2010). *Leading School Turnaround: How Successful Leaders Transform Low-Performing Schools*. San Francisco, CA: Jon Wiley & Sons.
- Leithwood, K. & Sun, J. (2012). "The Nature and Effects of Transformational School Leadership: A Meta-Analytic Review of Unpublished Research." *Educational Administration Quarterly*, 48, 387-423.
- Leithwood, K. & Louis, K.S. (2012). *Linking Leadership to Student Progress*. San Francisco, CA: Jossey-Bass.
- Leon, R., Davis, S., Kearney, K., Sanders, N., & Thomas, C. (2011). "Key features of a comprehensive principal evaluation system." San Francisco, CA: California Comprehensive Assistance Center – WestEd.
- Little, O., Goe, L., & Bell, C. (2009). "A Practical Guide to Evaluating Teacher Effectiveness." Washington, DC: National Comprehensive Center for Teacher Quality.
- Louis, K. S., Leithwood, K., Wahlstrom, K. L., & Anderson, S. E. (2010). "Investigating the Links to Improved Student Learning: Final Report of Research Findings." St. Paul, MN: Center for Applied Research and Educational Improvement – University of Minnesota.
- Metz, A. J. R., Bowie, L., & Bandy, T. (2007). "The Role of Frontline Staff in the Implementation of Evidence-Based Programs: An Exploratory Study." Washington, DC: Child Trends.
- Mujis, D. (2006). "Measuring teacher effectiveness." *Educational Research and Evaluation*, 12(1), 53–74.
- New Leaders for New Schools. (2009). "Principal Effectiveness: New Principalship to Drive Student Achievement, Teacher Effectiveness, and School Turnarounds with Key Insights from the UEF." New York, NY: NLNS.
- National Board of Professional Teaching Standards. (2010). "National Board Certification for Principals: A New Way to Develop, Recognize & Retain Top School Leaders." Arlington, VA: NBPTS.
- Popham, J. W. (2007). "Instructional insensitivity of tests: Accountability's dire drawback." *Phi Delta Kappan*, 89(2), 146–155.
- Porter, A. C., Murphy, J., Goldring, E., Elliott, S. N., Polikoff, M. S., & May, H. (2008). *Vanderbilt*

- Assessment of Leadership in Education: Technical manual*. New York: Wallace Foundation.
- Portin, B. S., Feldman, S., & Knapp, M. S. (October 2006). *Purposes, Uses, and Practices of Leadership Assessment in Education*. Seattle, WA: Center for the Study of Teaching and Policy – University of Washington.
- Rice, J. (2010). "Principal Effectiveness and Leadership in an Era of Accountability: What Research Says." (CALDER Brief 8). Washington, D.C.: Urban Institute.
- Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (September 2008). "The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects on Leadership Types." *Educational Administration Quarterly*, 44(5), 635-674.
- Rothstein, R. (2004). *Class and schools: Using social, economic, and educational reform to close the black-white achievement gap*. Washington, D.C.: Economic Policy Institute.
- Ryan, K. (2000). *Say uncle: Poems*. New York: Grove Press.
- Sanders, N. M., Kearney, K., & Vince, S. (2012). *Using Multiple Forms of Data in Principal Evaluations: An Overview with Examples*. San Francisco, CA: Integrated Leadership Development Initiative – WestEd.
- Schmoker, M. (1999). "Realizing the Promise of Standards-Based Education." *Educational Leadership*, 56(6), 17-21.
- Schmoker, M. (2011). *Focus: Elevating the Essentials to Radically Improve Student Learning*. Princeton, NJ: Association for Supervision & Curriculum Development.
- Simons, T. (1999). "Behavioral integrity as a critical ingredient for transformational leadership." *Journal of Organizational Change Management*, 12, 89-104.
- Singleton, G.E., & Linton, C. (2006). *Courageous Conversations about Race: A Field Guide for Achieving Equity in Schools*. Thousand Oaks, CA: Corwin Press, Inc.
- Spillane, J. P., Halverson, R., & Diamond, J. B. (2001). "Investigating school leadership practice: A distributed perspective." *Educational Researcher*, 30(3), 23-28.
- Sklra, L., Scheurich, J. J., Garcia, J., & Nolly, G. (2004). "Equity audits: A practical leadership tool for developing equitable and excellent schools." *Educational Administration Quarterly*, 40 (1), 133-161.
- Stumbo, C., & McWalters, P. (2011). "Measuring Effectiveness: What Will It Take?" *Educational Leadership*, 68(4), 10-15.
- Taylor, F. (1911). *The Principles of Scientific Management*. New York, NY: Harper and Brothers Publishers.
- Theoharis, G. (2009). *The School Leaders our Children Deserve: Seven keys to equity, social justice and school reform*. New York: Teachers College Press.
- The Wallace Foundation (2009). *Assessing the Effectiveness of School Leaders: New Directions and New Processes*. New York, NY: The Wallace Foundation.
- The Wallace Foundation. (2012). *School principal as leader: Guiding schools to better teaching and*

learning. New York, NY: Wallace Foundation.

Waters, T., Marzano, R., & McNulty, B. (2003). *Balanced Leadership; What 30 Years of Research Tells us About the Effect of Leadership on Student Achievement*. Aurora, CO: Mid-continent Research for Education and Learning.

White, M. E., Makkonen, R., Vince, S., & Bailey, J. (March 2012). "How California's local education agencies evaluate teachers and principals." (REL 2012 – No. 23). San Francisco, CA: REL West.

Vygotsky, L. (1978). "Interaction between learning and development." In M. Gauvain and M. Cole (Eds.), *Mind and Society*. New York: W.H. Freeman and Company.

ACKNOWLEDGEMENTS

The Leadership Connection at U.C. Berkeley acknowledges the multiple contributions of its staff and our partners and colleagues. The Leadership Connection Rubric referenced in this working paper is the work of Lynda Tredway, Daphannie Stephens, Linda Leader-Picone and Janette Hernandez with substantial input from the current coordinator for the Principal Leadership Institute (PLI) and the Leadership Support Program (LSP), Rebecca Cheung.

Daphannie Stephens, Coordinator of Public Programs at the Leadership Connection, and Norton Grubb, PLI and LSP Faculty Coordinator provided feedback on multiple working paper drafts. Our partners at the Institute for the Study of Knowledge Management in Education (ISKME), including Cynthia Jimes, Rudy Rubio, and Carol Hedgspeth, provided analysis and input on substantial portions of the working paper and are co-authors. Special thanks to Ben Grandy, the Program Assistant for PLI, and the Leadership Connection for technical support for videotaping and formatting documents.

The Griot Project for experienced leaders, which provided the context for developing many of the concepts in this paper, was supported by funding from the Stuart Foundation. We thank the Stuart Foundation Board and the Advisory Board for their funding and in particular the program officers who supported the project from the inception: Susan Little and Greg John.

The Griot Council participants are the backbone of the project. Their consistent and important work in schools is the foundation of the rubric. The participants in Years I, II, and/or III of Griot and district administrators, principals and assistant principals from eleven schools districts and include : Matin Abdel Qawi, Karling Aguilera-Fort, Theresa Armada, George Bullis, Katy Cobb Van Husen, Mary Lou Cranna, Brent Daniels, Glenn Dennis, Victor Diaz, Gini Dold, Mary Dybdahl, Leo Fuchs, Jerome Gouridine, Jeremy Hilinski, Matt Huxley, Greg Ko, Cheryl Lana, Moraima Machado, Lucius McKelvy, Irma Muñoz, Robert Patrick, Greg Santiago, Pasquale Scuderi, Ron Smith, Angela Stevens-Stevenson, and Pam Vandekamp.

The Griot facilitators contributed a willingness to experiment and provided important feedback: Jonetta Leek, Susan Audap, Ann Rarden, Jane Stern, Margot Fontes, and Carole Robie.

As well, our partners in Oakland Unified School District, including Anne Prozan, Kyla Johnson-Trammell and Maria Santos, and 20 leaders in OUSD provided substantial support on the design of the professional development and the use of the process for evaluation of leaders in the district. We acknowledge Gia Truong's foresight in designing the task force work in Oakland. We particularly thank our PLI graduates who are Oakland leaders: Sondra Aguilera, Michael Hatcher, Jerome Gouridine, Edith Jordan, Monica Moreno Bowie, Eyana Spencer and their OUSD colleagues: Sheilagh Andujar, Zarina Ahmad, Fulton Brinkley, Katherine Carter, Moyra Contreras, Leroy Gaines, Carol Johnson, John Melvin, Alicia Romero, Mel Stenger, Brandy Stewart, Kevin Taylor, Natalie Tovani-Walchuk, Asali Waters.

We acknowledge the support of the U.C. Berkeley and Chancellor Birgeneau, the Graduate School of Education at U.C. Berkeley, and former Deans Eugene Garcia and P. David Pearson and current Dean Judith Warren Little, and the University Office of the President and Nina Moore and Russ Rumberger. Kenneth E. Behring supported the scholarships for thirteen cohorts (2000-2013) of the Principal Leadership Institute. The Development Office of the GSE, including former director Janine Sheldon, current director Teresa McGuire, and Claire Raymond has consistently supported the work of Leadership Connection.

The guidance from Karen Kearney from West-Ed and Christopher Thomas from University of San Francisco as well as Jessica Rigby, who as a doctoral candidate served as the Graduate Student Researcher for the project, was invaluable.

Finally, the rubric and the working paper would not have been possible without a commitment to the collective leadership for which we advocate in the working paper. Leaders are organizational actors who work in concert with others to achieve durable and significant outcomes for students, and we at the Leadership Connection have a commitment to "walk our talk", remain hopeful in the face of substantial

obstacles for urban leaders, and align our work with the daily practice of over 400 graduates of the Principal Leadership Institute, without whom this work would not be possible.

Leadership Connection
Graduate School of Education
University of California, Berkeley

© 2012 Leadership Connection,
Graduate School of Education,
University of California, Berkeley.

August 2012