Research Report

Teacher Leadership: An Assessment Framework for an Emerging Area of Professional Practice

Teresa Jackson

Jeremy Burrus

Katherine Bassett

Richard D. Roberts

December 2010

ETS RR-10-27

Listening. Learning. Leading.®

Teacher Leadership:

An Assessment Framework for an Emerging Area of Professional Practice

Teresa Jackson, Jeremy Burrus, Katherine Bassett, and Richard D. Roberts

Center for New Constructs, ETS, Princeton, New Jersey

December 2010

As part of its nonprofit mission, ETS conducts and disseminates the results of research to advance quality and equity in education and assessment for the benefit of ETS's constituents and the field.

To obtain a PDF or a print copy of a report, please visit:

http://www.ets.org/research/contact.html

Technical Review Editor: John Sabatini

Technical Reviewers: Caroline Wylie and Richard Tannenbaum

Copyright © 2010 by Educational Testing Service. All rights reserved.

ETS, the ETS logo, and LISTENING. LEARNING. LEADING. are registered trademarks of Educational Testing Service (ETS).

Abstract

This paper examines various definitions and frameworks that have been used to justify the emergence of a new category for the teacher professional: teacher leader. The emergence of this new professional category may lead to greater retention levels, and improved knowledge management and transfer within the teaching profession. Various key dimensions of this profession are examined, allowing us to highlight some key personal skills that would appear requisite for the teacher leader. An agenda for developing and validating assessments of teacher leadership is then proposed. It is argued that these assessments have the potential of legitimizing research within this field, as well as providing the opportunity to better understand what it takes to become a successful professional in this new domain of teaching practice.

Key words: teacher leader, teacher leadership assessment, teacher personal skills, teacher mentoring, teacher personality, teaching as a profession

Acknowledgments

This paper was supported, in part, by both the Research and Development and Higher Education and School Assessment Divisions of ETS. We would like to thank the following individuals for supporting this research in a number of ways, from funding, to project management, to preparation of the manuscript: Anthony Betancourt, Pamela Coleman, Sally Gillespie, Patrick Kyllonen, Ida Lawrence, Mary Lucas, Walter McDonald, Jennifer Minsky, Steven Schreiner, Linda Tyler, Gregory Vafis, Cathy Wendler, and Susan Woodward. We are also grateful to John Sabatini, Rick Tannenbaum, and Caroline Wylie for a number of helpful comments and suggestions on an earlier version of this report.

Table of Contents

Background 1

Defining Teacher Leadership and its Components 5

Key Roles of Teacher Leaders: Coaching and Learning Team Facilitators 6

What are the Personal Skills of Teacher Leaders? 7

Work ethic. 8

Teamwork. 8

Leadership. 8

Openness. 8

Vision. 8

Positive affect. 8

Risk taking. 9

Teaching related skills. 9

Cultural and Structural Conditions that Influence Teacher Leadership 10

A Review of the Empirical Research on Teacher Leadership 12

An Assessment Framework for Teacher Leadership 12

Traditional Assessment Methods 13

Self-assessment. 13

Other ratings. 15

Situational Judgment Tests (SJT) 15

Authentic Assessment 16

Setting standards.. 16

Creating tasks.. 16

Defining criteria. 18

Creating rubrics 18

Knowledge Assessment 18

Validation of Teacher Leadership Assessments 19

Scoring. 21

Generalization 22

Extrapolation. 23

Implication 25

Validation Examples 25
Authentic Assessment Validation Example: Video Demonstration of Mentoring Skills 26
Traditional Assessment Validation Example: Self-report of Work Ethic 27
Other Validation Considerations 28
Conclusion 29
References 30
Appendix A 36
Appendix B 40

List of Tables

Table 1. Definitions of teacher leadership 2

Table 2. The potential benefits of teacher leadership (York-Barr & Duke, 2004) 3

Table 3. Personal skills associated with teacher leadership..9

Table 4. Cultural and structural conditions that influence teacher leadership (after Institute for 5 - Educational Leadership, 2008) 11

Table 5. Interpretive argument for teacher leadership (adapted from Kane, 2006)......................21

Table 6. Validity evidence for teacher leadership example assessments.....................................26

Background

It is a widely agreed-upon and highly intuitive proposition that students learn best when they have high quality teachers (Institute for Educational Leadership, 2001). Although few would disagree with this principle in the abstract, the process of filling a school with high quality teachers in order to promote student learning is more difficult in practice. First, there must be preparation programs—traditional and alternate (e.g., Teach for America) routes—that prepare teachers with the skills they need to succeed in a variety of school settings with diverse, 21st century learners. Then, schools must locate and hire new teachers with the potential to become high quality teachers. Next, they must provide professional development opportunities to assist these new teachers in increasing their pedagogical knowledge and honing their teaching skills. Finally, once teachers have been hired and trained, they must be retained. However, teacher retention is proving to be a significant problem. In the United States, 40 to 50 percent of teachers leave within their first five years of teaching (e.g., Grissmer & Kirby, 1997; Ingersoll & Smith, 2003).

The promotion of *teacher leadership* (see varying definitions in Table 1) as one way to improve teacher quality and retention is an idea that has gathered recent momentum across the United States. Teacher leadership has become central in the discussion of possible ways to further professionalize teaching and reform schools (Billingsley, 2007; Little, 2003; Smylie, 1995). For example, in the 1990s, both the National Commission on Teaching and America's Future (1996) and the Council for State School Officers (Interstate School Leaders Licensure Consortium, 1996) suggested that more leadership opportunities be provided for teachers, with the intention also to provide greater incentives to the professional.

There is increasing interest in emphasizing roles for teacher leaders within schools (e.g., Danielson, 2006; Killion & Harrison, 2006). In fact, the Kansas State Department of Education has recently developed standards for creating a licensure endorsement for teacher leaders within schools (see Appendix A). This opportunity provides an alternative for teachers from the track that might otherwise be associated with school building or district leadership roles (i.e., being a principal or superintendent). In addition, a consortium of organizations, institutions of higher education, state agencies, teachers, principals, and superintendents has convened to study the concept of teacher leadership and to draft model standards for teacher leadership. Finally in recent months, a number of states, including Arkansas, California, Connecticut, Georgia,

Kansas, Illinois, and Louisiana have either begun or completed legislating teacher leader positions or established criteria for endorsing, certifying, or credentialing teacher leaders.

Table 1

Definitions of teacher leadership

Definition	**Source**
Teachers become leaders when they function efficiently in professional learning communities to impact student learning, contribute to school improvement, inspire excellence in practice, and empower stakeholders to participate in educational improvement.	Childs-Bowen et al. (2000)
Teacher leadership facilitates principled action to foster whole-school success. Teacher leaders transform teaching and learning and tie the school and community together, and advance the community's social mission and quality of life.	Crowther et al. (2002)
Teacher leadership is a set of skills demonstrated by teachers who are able to influence students outside of the classroom and beyond.	Danielson (2006)
Teacher leadership encompasses "inter-related domains of commitment and knowledge." Thus, there is a commitment to moral purpose, continuous learning, knowledge of learning processes, as well as an understanding of the educational context and change processes.	Fullan (1994)
Teacher leadership is a mobilization of the available attributes of teachers to strengthen student performance at the ground level. Teacher leaders work toward collaboration and shared leadership in the daily activities in the school.	Institute for Educational Leadership (2001)
Teacher leaders lead within and outside of the classroom. A teacher leader is a member and contributor to a community of teacher learners. They are influential in the continued improvement of educational practice.	Katzenmeyer & Moller (2001)
Teacher leaders hold influence outside of the classroom and are autonomous in their own work. Still, they do not engage in managerial or supervisory tasks.	Murphy (2005)
Teacher leaders have the ability to encourage other teachers and colleagues to change and begin to think about taking part in things they ordinarily would not consider.	Wasley (1991)
Teacher leadership is an idea that emphasizes that teachers hold an important and central position within the schools.	York-Barr & Duke (2004)

Although little empirical research documenting the benefits of teacher leadership exists, the overall consensus is that teacher leadership has many potential benefits. York-Barr and Duke (2004) summarized these benefits into four main categories. Each of these benefits (and two additional benefits), along with a brief comment, is provided in Table 2.

Table 2

The potential benefits of teacher leadership (York-Barr & Duke, 2004)

Benefit	Comment
Engagement and commitment	Teacher leaders inform management and engage in decision making. Participating at this level can increase the teacher's ownership and commitment to the profession and school goals.
Knowledge, skills, and learning transfer and growth	Teacher leaders advance teaching and learning by modeling effective practice to other teachers and sharing their knowledge and skills with others in the field. Teacher leaders themselves continue to learn and grow as they lead and work with others.
Recognition and rewards lead to retention and advancement	Additionally, the recognition, rewards, and opportunities that accompany the title teacher leader are thought to retain, motivate, and recruit teachers
Student benefits	Teacher leaders provide students with a positive example of leadership by modeling democratic leadership and collective responsibility in a community environment
Recruitment of Generation Y teachers	Generation Y employees value careers that have opportunity for advancement.
Retention of Generation Y teachers	Inadequate opportunities for advancement and feelings of being unprepared contribute to new teacher attrition

The first benefit is that teacher leadership encourages teachers to begin to actively engage in, contribute to, take responsibility for, and become accountable for what is happening in their schools. For example, when teachers are able to participate in the decision-making process, they are likely to become committed to these decisions and work diligently towards implementation, thus fostering teacher empowerment (e.g., Barth, 2001). Further, the conversations among teachers that such leadership encourages will tend to unite teachers and

create a professional work environment that is based on informed decision-making (York-Barr & Duke, 2004).

Second, teacher leaders serve as models and mentors to other teachers, and in the process continue to learn themselves (Barth, 2001; Darling-Hammond, 1988; Ovando, 1996; Porter, 1986; Ryan, 1999). To date, perhaps the strongest effects of teacher leadership are on the teachers themselves. The impact is often exhibited by increases in the teachers' leadership and organizational skills (e.g., Ryan, 1999). Growth has also been reported in additional areas, such as instruction, due to exposure to new information and opportunities (Ovando, 1996; Porter, 1986; Smylie, 1994).

Third, teacher leadership acknowledges and rewards teacher's expertise and dedication, increasing teacher retention and opportunities for advancement (Hart, 1995). The important contributions that teachers make are more prominent and easily recognized by administrators and their peers when they take on teacher leadership roles. For example, in a survey of 76 successful principals (defined as such because their schools had made *adequate yearly progress*), these principals stated that their schools would be less successful if they did not have teacher leaders (Institute for Educational Leadership, 2008). Such recognition can lead to career advancement or a sense of renewal (Steffy, Wolfe, Pasch, & Enz, 2000). Ultimately, the benefit accrues to students, not only through improved instruction, but also by providing students with positive examples of leadership and models of collective responsibility in a community environment.

The above benefits may be especially important in the recruitment and retention of so-called *Generation Y* teachers, or teachers born after about the year 1977. For example, research has indicated that factors such as "inadequate opportunities for advancement" and "feelings of being unprepared," conditions directly addressed by teacher leadership, contribute to new teacher attrition (Behrstock & Clifford, 2009, p. 6). Behrstock and Clifford (2009) note that many private sector strategies that have been effective in the recruitment of Generation Y employees could translate to the education sector. These include merit-based compensation, job rotation into other departments, opportunities for career advancement, as well as early career mentoring.

Despite the momentum and potential benefits, there are a number of issues and obstacles that have not been addressed nor resolved adequately concerning teacher leaders. Specifically,

there is little consensus on the definition of teacher leadership, and a consequent lack of clarity about roles for teacher leaders. In addition, the cultural and structural conditions that would support the fostering of teacher leadership as a formal (or indeed, informal) career track within schools have only been discussed, not empirically evaluated. Many of the sources cited in this paper only represent theoretical discussions of teacher leadership frameworks and descriptive studies of teacher leaders and are not empirical in nature. Therefore, more work is needed to organize them into a framework that could serve as a foundation for training opportunities. Finally, if teacher leadership is to be fostered, and eventually become a formal career track with training, licensing and/or evaluation components, then valid assessments need to be conceptualized and developed. This report is organized to address each of these points and therefore serve as a foundation for moving the concept of teacher leadership forward.

Defining Teacher Leadership and its Components

Over the years, there have been many conceptions about how to define teacher leadership (Katzenmeyer & Moller, 2001; Killion & Harrison, 2006). Because it is considered by many to be an umbrella term encompassing a variety of teacher roles, little consensus exists on how to define teacher leadership (York-Barr & Duke, 2004). Several examples of definitions of teacher leadership are listed in Table 1.

For example, teacher leadership has been defined as "[T]he process by which teachers, individually or collectively, influence their colleagues, principals, and other members of school communities to improve teaching and learning practices with the aim of increased student learning and achievement" (York-Barr & Duke, 2004, p. 287–288). Danielson (2006), who espouses an informal teacher leader role, has defined teacher leadership as "a set of skills demonstrated by teachers who continue to teach students but also have an influence that extends beyond their own classrooms to others within their own school and elsewhere. It entails mobilizing and energizing others with the goal of improving the school's performance of its critical responsibilities related to teaching and learning" (p. 12). Danielson (2006) also notes that a teacher leader's first responsibility is to his/her students. Katzenmeyer and Moller (2001) describe teacher leaders as "teachers who are leaders within and beyond the classroom, identify with and contribute to a community of teacher learners and leaders, and influence others toward improved educational practice" (p.9). In essence, teacher leadership is an emphasis on teachers

holding an important and central leadership position within the dynamic organizational system that is the school.

Teachers often emerge as teacher leaders after they achieve success in the classroom through their experiences and expertise (York-Barr & Duke, 2004). Teacher leadership occurs in all aspects of school life—it informs school-wide policies and programs, teaching and learning processes, as well as engagement and community relations. It is exhibited in classrooms, by departments and teams, and at the school, district, state, and national levels. Teacher leaders often coordinate and manage events and meetings, complete school/district curriculum work, participate in the professional development of others, lead collaborations with parents and the community, make professional contributions to the field, and/or assist in the preparation of pre-service teachers or mentoring of novice teachers.

Teacher leaders can be formally acknowledged, or can emerge spontaneously from exercising leadership when a need, possibility, or opportunity arises. Formal leadership roles (e.g., coordinators, parent/community involvement, preservice teacher partnerships, board members, specialists) tend to dominate the field. Nonetheless, the role, title, and perception of being a teacher leader can be earned through informal work and interactions that occur among teachers, their students, and their colleagues, slowly transforming into a formal role (Danielson, 2006).

Key Roles of Teacher Leaders: Coaching and Learning Team Facilitators

Two key roles of teacher leaders that have emerged are *coaching* and *learning team* facilitator. Coaching has been found to enhance other forms of teacher professional development. For example, Joyce and Showers (1995) found that demonstrations, presentations, and practice had the greatest impact on teacher professional development (i.e., increased teaching knowledge and skill) when they were combined with coaching and other support. Coaches provide support and training to teachers in and out of the classroom; aid colleagues in expanding their skills, knowledge, and processes; and support and encourage colleagues to reflect, plan, and adapt their practices when necessary (Killion & Harrison, 2006; Neufeld & Roper, 2003).

Although many roles that teacher leaders pursue emerge informally and organically, coaches are usually formally selected into the role. The success of coaching primarily depends on having the right coach working in the right school. Coaches are selected based on factors

such as the desired outcome of coaching, the work the coach will do, and the attributes necessary to effectively coach (Killion & Harrison, 2006).

Coaching can take many forms. Killion and Harrison (2006) identified ten specific roles that coaches may have within the school. They note that coaches can be resource providers, data coaches, instructional specialists, classroom supporters, learning facilitators, mentors, liaisons to school leaders, catalysts for change, and most importantly, *lead learners*. Lead learners are professionals who focus on learning in three areas – themselves, their work, and the field of education (Killion & Harrison, 2006). They identify what they want to learn, take charge of their own learning, put their new knowledge into practice in their classroom, and spend time reflecting on their experiences (Mizell, 2004).

Another way teacher leaders can help to create a culture of collaboration is through the facilitation of *learning teams*. As the term implies, a learning team is a group of teachers working together to focus on important issues within the school, such as addressing the meeting of student needs (e.g., Richardson, 2009). Typically, a principal creates learning teams when a need is identified, for example, when test scores are lower than expected. The principal then provides the learning team with information about learning and collaborating, and guides teacher research on alternative teaching strategies they can use in the pursuit of improved student learning. The role of convening and facilitating learning teams, however, need not fall solely on the principal; a teacher leader can also convene or facilitate learning teams. This can provide crucial support for principals and facilitate buy-in from teachers who may be more likely to support initiatives that are teacher-led, rather than principal-led.

What are the Personal Skills of Teacher Leaders?

Several authors have suggested the requisite skills that might be necessary to fill the role of a teacher leader. We do not discuss pedagogical skills and knowledge here because, to most who discuss teacher leadership, it is assumed that teacher leaders already possess these particular skills. In Table 3, we organize the proposed skills by construct and identify the researcher that posited the necessity of the skill. Although we organize skills by construct, many of the skills identified can be placed into several categories. However, it is our belief that organizing this information into a very general framework might help to inform the assessment of such skills. These skills are briefly described below:

Work ethic. Teacher leaders have been described as perseverant, resourceful, action-oriented committed, and passionate (York-Barr & Duke, 2004). The passion they have for their mission allows teacher leaders to find the courage to persist in the face of adversity and obstacles (Danielson, 2006). Of note, this skill, with its close links to Conscientiousness, has been demonstrated in meta-analytic studies to be one of the best predictors of workplace performance across a range of occupations (see Kyllonen, Lipnevich, Burrus, & Roberts, 2010).

Teamwork. Because teacher leaders must work with many constituencies (to improve teaching practice and promote positive change within the larger learning community) they should have good teamwork skills. They must possess the ability to build solid relationships with colleagues, parents, students, administration, and the community (York-Barr & Duke, 2004). To build such relationships, they must be able to engender trust, work well with colleagues, communicate effectively, and resolve conflict (Danielson, 2006; Killion & Harrison, 2006).

Leadership. Closely related to teamwork skills are leadership skills, as, often, leadership is necessary to promote good teamwork. Teacher leaders can lead by engaging, inspiring, and motivating others to improve and become better through their actions (Bascia, 1996). They are able to lead by effectively communicating with colleagues and informing them of their goals in ways that garner support for their vision for the school (Danielson, 2006).

Openness. Teacher leaders are adaptable, open-minded, and creative. They are open to exploring options to gather the necessary resources to improve the state of education (York-Barr & Duke, 2004). Furthermore, teacher leaders are able to adapt and adjust to situations through their creativity and flexibility (Danielson, 2006; Killion & Harrison, 2006).

Vision. Related to the concept of openness to new ideas, a good teacher leader has enough vision to be able to identify opportunities for improvement or to fix problems within the school. They actively seek out such opportunities rather than simply waiting for them to appear (Danielson, 2006). They also have the ability to see the big picture and how what they are doing fits into the larger goal of student learning (York-Barr & Duke, 2004).

Positive affect. Positive emotions have been demonstrated to have a number of beneficial outcomes. For example, the presence of positive emotions can help people to think more creatively, deal with stressful situations, be more engaged in activities, and build social relationships (e.g., Fredrickson, 1998). As such, it is not surprising that successful teacher

leaders often succeed with the help of positive affect. Their tendency to display optimism, enthusiasm, and confidence leads others to think creatively and work together to solve problems (Danielson, 2006).

Risk taking. Teacher leaders are willing to take risks in order to achieve their goals. They are willing to take the chance to attempt new and innovative initiatives despite the fact that what they are doing may end up failing and expose them to external criticisms (Danielson, 2006).

Teaching related skills. Teacher leaders are also good teachers. They display outstanding teaching skills, hold a comprehensive philosophy of education, and are relatively altruistic (Katzenmeyer & Moller, 2001). Furthermore, they have a deep understanding of various theories and practices of teaching (Killion & Harrison, 2006). Teacher leaders are willing to expand their role as teacher to develop their career, and have the time, energy, and patience required to assume such a role (York-Barr & Duke, 2004).

Table 3

Personal skills associated with teacher leadership

Work Ethic	**Teamwork**
Achievement-oriented	*Agreeableness*
Achievement-Oriented (a, f)	Interested (f)
Commitment	Intuitive and sensitive to the needs of others (c, e)
Committed (a)	*Collaboration*
Decisive (a)	Able to build trust (f)
Dedicated (a)	Collaborative (a, d)
Persistence	Develops strong relationships with colleagues (d)
Perseverant (a)	Patient (f)
Persistent (a)	Supports colleagues (f)
Passionate (a)	Works well with others (c)
Resourcefulness	*Communication*
Resourceful (a)	Effective communicator (a, e, f)
Works through processes (f)	*Conflict Resolution*
Responsible	Able to resolve conflict and mediate (e, f)
Responsible (e, f)	
Leadership	**Openness**
Influence	*Adaptability*
Engaging (a)	Adapts and adjust to situations (c)
Influential (d)	Expands role to enhance career/skills (f)

Leadership (cont.)	Openness (cont.)
Motivating (a)	Flexible (a)
Leadership Preparedness	*Creativity*
Has a readiness to assume leadership roles (b)	Creative (a, e)
Positive Affect	*Open-minded*
Confident (a)	A learner (e)
Enthusiastic (a, e)	Open-minded (a)
Optimistic (a)	Reflective (e)
Risk Taking	**Vision**
Courageous (a)	Able to identify and assess student, teacher, and school needs (d, f)
Risk-Taker (a)	Sees the big picture (f)
Teaching Related	
Attitude	
Has a teacher-first attitude (a)	
Pedagogical Knowledge	
Able to articulate a thorough philosophy of education (b)	
Holds a deep understanding of theory and practice (c)	
Knowledgeable about profession and content (d, e)	
Teaching Skills	
Excellent teaching skills (b, f)	
Experienced (e)	
Successful in the classroom (f)	
Viewed as a superior educator (d)	

Note. Personal skills mentioned in the following sources: a = Danielson (2006); b = Katzenmeyer & Moller (2001); c = Killion & Harrison (2006); d = Sherrill (1999); e = Yarger & Lee (1994); f = York-Barr & Duke (2004).

Cultural and Structural Conditions that Influence Teacher Leadership

In order for teacher leadership to be successfully implemented, it is crucial to have certain cultural and structural conditions intact. The skills and dispositions mentioned above are best maximized when administrators actively support the contributions of teacher leaders, honor their contributions, and promote the development of these leaders. Schools should be organized in such a way that teachers are made aware of opportunities available and given the chance to engage in these leadership activities (Danielson, 2006). In Table 4, we list some of the conditions identified by researchers.

Table 4

Cultural and structural conditions that influence teacher leadership (after Institute for Educational Leadership, 2008)

Condition	Examples	Cited From
Social Culture & Context	• School establishes standards and expectations for leadership • Teachers made aware of leadership activities • Teachers encouraged to take initiative, practice teamwork, and share responsibility • School-wide focus on learning and reflection • Teachers are valued and respected as examples and models for other teachers in the profession • Teachers are valued and respected as examples and models for other teachers in the profession • Principals who foster, support, and model teacher leadership.	• Danielson (2006), Katzenmeyer & Moller (2001), Little (1988), Ovando (1996), York-Barr & Duke (2004)
Interpersonal	• Positive relationships between teacher leaders, colleagues, principals, and administrative staff • Effective teamwork between colleagues	• York-Barr & Duke (2004)
Structures	• Access to resources, available time and space	• LeBlanc & Shelton (1997)

In addition, it is imperative that the school culture itself is such that the role of the teacher leader is embraced. This will mean several things in terms of change:

1. Training of principals to understand and facilitate the role of teacher leaders;
2. Training of teacher leaders in working with adult learners and in other constructs of leadership not typically taught in teacher preparation programs; and
3. Training of teachers to understand the role of teacher leaders.

Of these, the training of principals, and the building into principal preparation programs of knowledge of working with teacher leaders is particularly critical. If the principal does not embrace this role, it is highly unlikely that teachers will do so.

A Review of the Empirical Research on Teacher Leadership

The existing literature is limited by small, homogenous samples and poor study designs that make it difficult to draw conclusions (York-Barr & Duke, 2004). To date, much of the research conducted in this area is descriptive and policy driven, with a focus on program descriptions, roles, and implementation. Most studies do not derive their research questions and study design from formal theories, and change has tended to be measured in indirect ways, leaving the reader to determine whether or not the implementation of teacher leadership programs was effective or ineffective. In many cases, the view of teacher leadership presented is limited by narrow perspectives, specifically attending to one type of leadership and focusing only on the teacher as a leader isolated from their other roles within the school (Smylie, 1995).

In sum, available research tends to be segmented, making it difficult to study the effectiveness of teacher leadership empirically. More research is required to address even basic research questions (York-Barr & Duke, 2004). However, to conduct any empirical study, there is a need for measures of teacher leadership skills and knowledge. Below, we describe an assessment framework for teacher leadership that could be used to develop such measures.

An Assessment Framework for Teacher Leadership

How might researchers seek to understand what personal skills, knowledge, and attributes contribute to teacher leadership? One approach is represented in the work of the Kansas Department of Education, who are aiming to develop the nation's first assessment of teacher leadership. As can be seen in Appendix A, they have developed a preliminary set of eight standards of teacher leadership. Interestingly, for teacher leaders to meet these standards, they should possess most, if not all, of the skills listed in Table 3. For example, for a teacher to be able to meet Standard 1, "The teacher leader is able to apply strategies of adult learning across teacher leadership activities," he or she should have the pedagogical knowledge, collaboration, communication, and conflict resolution skills listed in Table 3, among others.

Below we describe and evaluate three possible methods for assessing teacher leadership such as described in the Kansas standards. The first method will be familiar to those versed in

psychological assessments of knowledge, skills, and attitudes. To distinguish it from the second method discussed, this method is labeled *traditional assessment methods*, although many of the types of assessments within this method may not be considered traditional (e.g., Gulikers, Bastiaens, & Kirschner, 2004). The second method, which has gained recent popularity in educational research, is referred to as *authentic assessment* (e.g., Darling-Hammond & Snyder, 2000; Gulikers et al., 2004; Mueller, 2009). The third method we will discuss briefly is *knowledge assessment*. Below we describe each method, give advantages and disadvantages that are associated with each technique, and conclude with our own recommendations for the use of each methodology in the assessment of teacher leadership.

Traditional Assessment Methods

We use *traditional assessment methods* as a term to refer to assessment methods that dominate the social sciences, such as survey items, cognitive tests, and other reports (e.g., Gulikers et al., 2004; Kyllonen et al., 2009; Mueller, 2009). Typically, respondents are provided with statements or questions and are asked to indicate their response to each by choosing one or more options provided to them either in a paper-or-pencil or computerized format. There are several ways one can assess the skills, attitudes, and knowledge required of teacher leadership with traditional methods. Three forms of assessment are described below, along with some potential problems with the use of each one, and some potential solutions to these problems.

Self-assessment. One of the most simple, easy, and cost-effective ways to assess the skills, attitudes, and knowledge of potential teacher leaders is to have them report on their own characteristics. This can be done in several ways, but is most commonly assessed with Likert-type scale items. In an example of a Likert-type item, a teacher leader would respond to a statement concerning their skills, attitudes, or knowledge (e.g., "I have a sound work ethic") on a scale ranging from one to five, where one equals *strongly disagree* and five equals *strongly agree*.

Many of the skills listed in Table 3 can be measured with existing self-assessments of personality based on the Five Factor Model (FFM) of personality (e.g., Goldberg, 1993). The FFM states that most people's personality can be characterized by five factors, which are labeled Extraversion, Agreeableness, Neuroticism, Conscientiousness, and Openness (see Kyllonen et al., 2009; Matthews, Zeidner, & Roberts, 2006). Several of these factors are represented in Table 3: specifically, Openness, Agreeableness, and Conscientiousness (work

ethic). Furthermore, we hypothesize that Neuroticism (e.g., emotional stability) and Extraversion (e.g., outgoing, gregarious) are related to teacher leadership, such that those low in Neuroticism will be better teacher leaders than those high in Neuroticism, and those high in Extraversion will be slightly better teacher leaders than those low in Extraversion.

Attitudes can be self-assessed with similar Likert-type items using the framework of the Theory of Planned Behavior (TpB) (Ajzen, 1991). The TpB states that intentions are the best predictor of behavior, and attitudes (an evaluation of a behavior), subjective norms (how important others feel about a behavior), and perceived control (self-efficacy regarding a behavior) determine intentions. By utilizing the TpB, in addition to measuring teacher attitudes we can also measure some of the cultural and structural conditions listed in Table 4 by assessing subjective norms. The Theory of Planned Behavior has been demonstrated to be useful in predicting several types of behaviors (e.g., Armitage & Conner, 2001; MacCann & Roberts, 2009), and we believe that behaviors associated with teacher leadership should not be an exception.

Although self-assessments tend to be easier to create and more cost effective than other types of assessments, there are a few issues researchers should keep in mind when using them. These include issues such as how many points a scale should have (Krosnick, Judd, & Wittenbrink, 2005), whether it is advisable to use negatively (reverse) keyed items (e.g., Barnette, 2000; DiStefano & Motl, 2006), and whether various groups of people tend to respond to these items in different ways (e.g., Austin, Deary, & Egan, 2006). However, perhaps the most pressing issue is the fact that self-assessments are easier to fake than are other types of assessments (e.g., Griffith, Chmielowski, & Yoshita, 2007; Ziegler, MacCann, & Roberts, 2010). Simply put, to create a positive impression, people can state that they possess positive characteristics they do not necessarily have. To date, a definitive solution to this problem has not been found (Converse, Oswald, Imus, Hedricks, Roy, & Butera, 2008).

Due to these problems, it may not be feasible to use self-assessments in a licensure or hiring situation. Rather, self-assessments may prove to have more value as a type of educational tool. For example, teachers who are interested in becoming teacher leaders can take self-assessments to inform them on how their knowledge, skills, and attitudes compare to those of experienced teacher leaders. This assessment can provide them with valuable information as to skills they should improve upon in order to become effective teacher leaders. As such, graduate

level teacher leader programs in universities may use self-assessments at the beginning of student training.

Other ratings. One way to get around the problem of faking is to have others rate teacher leaders on their personal skills, attitudes, and knowledge. For example, aspiring teacher leaders may have their principals, professors, fellow teachers, and possibly even students rate their personal qualities. The assumption is that faking is less likely with this method because others will not be as motivated to inflate scores as the aspiring teacher leaders will themselves, although it is possible that some raters may be biased to rate some people more leniently than others. In support of the contention that other raters are less likely to fake than are self-raters, others-ratings are often more predictive of outcomes than are self-ratings (e.g., Kenny, 1994; MacCann, Minsky, Ventura, & Roberts, 2009).

Situational Judgment Tests (SJT). SJTs consist of a set of descriptions of real-world (everyday) problems (scenarios); the examinee is asked "what is the best way to handle this problem?" To respond, the examinee writes a solution to the problem, or rates the quality of proposed solutions (verbally described or also presented in video format). This general format can be used to assess various personal skills, such as communication skills, intercultural sensitivity, and potentially, leadership. These SJTs can be delivered either through paper-and-pencil based scenarios or through video. Because a certain amount of knowledge is required to correctly answer SJTs, they should be less susceptible to the faking problem than are self-report methods.

There is a long history, especially within the field of industrial-organizational psychology, suggesting that SJTs are fairly simple, economical simulations of job tasks (Kyllonen & Lee, 2005). Motowidlo and his colleagues (Motowidlo & Tippins, 1993; Motowidlo & VanScotter, 1994) suggested the idea of a situational judgment test as an inexpensive simulator, in which situations would be described (in words) rather than acted out or imitated in high fidelity. Since then, there have been many SJTs developed in areas as diverse as army leadership (Krokos, Meade, Cantwell, Pond, & Wilson, 2004), college success (Lievens & Coetsier, 2002), teamwork (Wang, MacCann, Zhuang, Liu, & Roberts, 2009), emotional management (MacCann & Roberts, 2008), and supervisory performance (Hedge, Hanson, Borman, Bruskiewicz, & Logan, 1996), to name a relevant, select few.

SJTs can be used both in a selection context, as a way to check that teacher leader candidates possess the requisite knowledge to perform the role, and in an educational context. That is, the SJT can be used as a training tool both to inform the teacher leader candidates what skills they are deficient on, and to help them practice appropriate responses to situations. For example, SJTs that provide scenarios depicting conflicts that come with the job of advocating for one's school with many potential responses can help the aspiring teacher leader experiment with different responses and get feedback on the efficacy of each of those responses.

Authentic Assessment

Authentic assessment is an assessment requiring the respondent to "use the same competencies, or combinations of knowledge, skills, and attitudes, that they need to apply in the criterion situation in professional life. The level of authenticity of an assessment is thus defined by its degree of resemblance to the criterion situation" (Gulikers et al., 2004, p. 69). The basic idea is that the assessment should require teachers to demonstrate being a teacher leader.

How is authentic assessment done? According to Mueller (2009, see also Darling-Hammond & Snyder, 2000) authentic assessment can be done in a fairly sequenced fashion: setting standards, creating tasks, developing criteria, and developing a scoring rubric. In the passages that follow, we discuss each of these components as it might apply to the development of a teacher leadership assessment.

Setting standards. The first stage is to create statements that explicitly outline the abilities and level of performance expected from teacher leaders. Typically this is done through the formulation and codification of standards (see Kansas example in Appendix A). Mueller (2009) recommends a three-step process in which standard writers should reflect on what it is they would like the respondent to be able to do (e.g., Is it necessary for teachers to be excellent classroom teachers in order to be teacher leaders?), then review their thinking, and finally, write the standard. The standard should be written clearly and in such a way that it can be assessed, should not be too broad or too narrow, and should not mention the specific task a respondent will be required to perform.

Creating tasks. Once standards have been written, performance tasks that elicit evidence of proficiency at meeting these standards should be created (e.g., Darling-Hammond & Snyder, 2000; Mueller, 2009). An *authentic* task should mimic, simulate, or document activities carried out in professional practice (Gulikers et al., 2004). Thus, for example, videos of

potential teacher leaders mentoring junior teachers would be an authentic task demonstration of teacher leadership, whereas scores on a test of pedagogical knowledge might not. There are several types of tasks that can be performed in authentic assessments, some of which are described below.

One type of task would be to analyze a case (e.g., Mueller, 2009). A case can be a first-person account of one's teaching experience, or an analysis of a student, teacher, or situation written by a third person (Darling-Hammond & Snyder, 2000). Aspiring teacher leaders might read through these cases and analyze them to identify themes relevant to teacher leadership. In addition, they might write their own cases by researching teachers, schools, or students from one or more perspectives. Cases can be written about curriculum selection or analyzing teaching decisions in context (Darling-Hammond & Snyder, 2000), as well. Aspiring teacher leaders' analyses of cases can be used either as an assessment of their knowledge, or as a tool for reflecting on their practice. Some example case studies that can be written include case studies of curriculum selection and analyzing teaching decisions in context (Darling-Hammond & Snyder, 2000).

Another type of task would be an exhibition of one's performance (e.g., Darling-Hammond & Snyder, 2000; Mueller, 2009). There are several ways one might exhibit performance as a teacher leader. For example, aspiring teacher leaders could create videos demonstrating their work mentoring junior teachers. Consonant with standard 6 of the Kansas standards ("the teacher leader is able to identify and assess opportunities for educational improvement,"), a teacher could develop a needs assessment for identifying problems that the school or district might address. Teachers could document professional development plans related to fostering their own leadership abilities, or provide a written journal of their day-to-day leadership-related activities.

A third type of task is to create portfolios that provide evidence of requisite skills and proficiencies in teacher leadership (Darling-Hammond & Snyder, 2000). This task type requires collecting and documenting pre-existing demonstrations of one's performance. Ideally, this evidence would be gathered over a significant period of time and across many contexts. Such a portfolio might include lesson plans, assignments, videotapes, journals, peer assessments, and so forth.

Finally, a fourth type of task would be to conduct a problem-based inquiry (e.g., Darling-Hammond & Snyder, 2000; Mueller, 2009), in which potential teacher leaders attempt to answer a question by conducting research, similar to conducting a master's thesis or doctoral dissertation. For instance, a research question such as "do teachers learn new skills through learning in groups better than learning individually?" can be answered through archival, survey, quasi-experimental, or even experimental methods.

Defining criteria. A criterion is an indicator of what good performance on a task looks like. Characteristics of good criteria include that it should be clearly and concisely stated so that there is no ambiguity as to its meaning; it should be observable; and it should be a statement of behavior, as opposed to, for example, a statement of attitude. Thus, a criteria statement such as "the teacher leader makes positive statements when working with mentees," is preferred to "the teacher leader has a positive attitude toward mentees." Although multiple criteria are desirable, the number should be limited to measure only the most essential elements of a task. Piloting scoring criteria is also important to ensure that nothing important has been omitted.

Creating rubrics. After criteria are created, rubrics can be created with which the respondent's performance can be scored. There are two basic types of rubrics: analytic and holistic. On an analytic rubric, respondent performance is judged separately for each criterion. For example, the scorer might rate the respondent's performance on a criterion scale with one to three points, with descriptors defining the meaning of each point. This is repeated for each criterion. In a holistic rubric, the scorer judges the respondent's performance by considering how well they performed all the criteria together. All criteria are considered while the scorer makes a judgment of the respondent's overall performance. A discussion of the criteria that make an authentic assessment *authentic* may be found in Appendix B.

Knowledge Assessment

A third type of assessment is a measure of teacher leaders' knowledge. This type of measure assesses one's declarative knowledge as it relates to teacher leadership and can include both multiple choice and constructed response-type items. In fact, eight of the Kansas standards include a knowledge component, which could most efficiently be measured using this type of assessment.

Validation of Teacher Leadership Assessments

Validation is the process of evaluating whether the interpretations and uses of an assessment are warranted (American Psychological Association, National Council on Measurement in Education, & American Education Research Association, 1999; Kane, 2006). Validation typically involves presenting logical and analytical arguments and/or empirically-based evidence verifying the assessments' relation with other pertinent constructs. As an organizing framework in our discussion of the validation of teacher leader assessments, we follow Kane's (2006) argument-based approach, which is consistent with the unitary concept of validity as outlined in the *Standards for Educational and Psychological Testing* (American Psychological Association, National Council on Measurement in Education, & American Education Research Association, 1999).

In the paragraphs that follow, we briefly discuss Kane's (2006) framework and relate each piece of the framework to the validation of both authentic and traditional assessments of teacher leadership. The general framework is depicted in Figure 1. The framework depicted in the figure presumes one general assessment that measures a trait known as "teacher leadership," and outlines several steps that can be taken to validate this assessment. In reality, multiple assessments targeting different aspects of teacher leadership (e.g., work ethic and mentoring are also included in Figure 1) may be created, and each needs to go through a validation process, though the number of steps of the process required will vary.

Figure 1 is composed of two sections: (a) *Hypothesized empirical relationships*, and (b) the *interpretative argument* (Kane, 2006). The hypothesized empirical relationships section is composed of a *target domain*, which is the full range of possible observations associated with teacher leadership, and the underlying trait (teacher leadership) is represented by the oval at the top left of the figure. Observations of teacher leadership can be influenced by other traits, such as those listed in Table 3, by context (e.g., cultural and structural conditions listed in Table 4), and by the method of observing teacher leadership (e.g. videos of performance; self-report scale ratings). Because the boundary of the target domain is usually not specified very precisely, it is signified with a dashed line. A smaller set of the target domain that is more precisely defined and will actually be measured by the assessment is in the universe of generalization box.

The process of validating what is measured begins in the box below the universe of generalization box and follows the arrows. Validation is comprised of two types of arguments. The *interpretive argument*, "specifies the proposed interpretations and uses of test results by laying out the network of inferences and assumptions leading from the observed performances to the conclusions and decisions based on the performances" (Kane, 2006, p. 23). The interpretive argument for Figure 1 is detailed in Table 5. Basically, the interpretive argument makes the inherent inferences and assumptions in the assessment explicit. The *validity argument* is an evaluation of the interpretive argument (Kane, 2006). For each stage of the interpretive argument, possible validity argument evidence is listed in Figure 1.

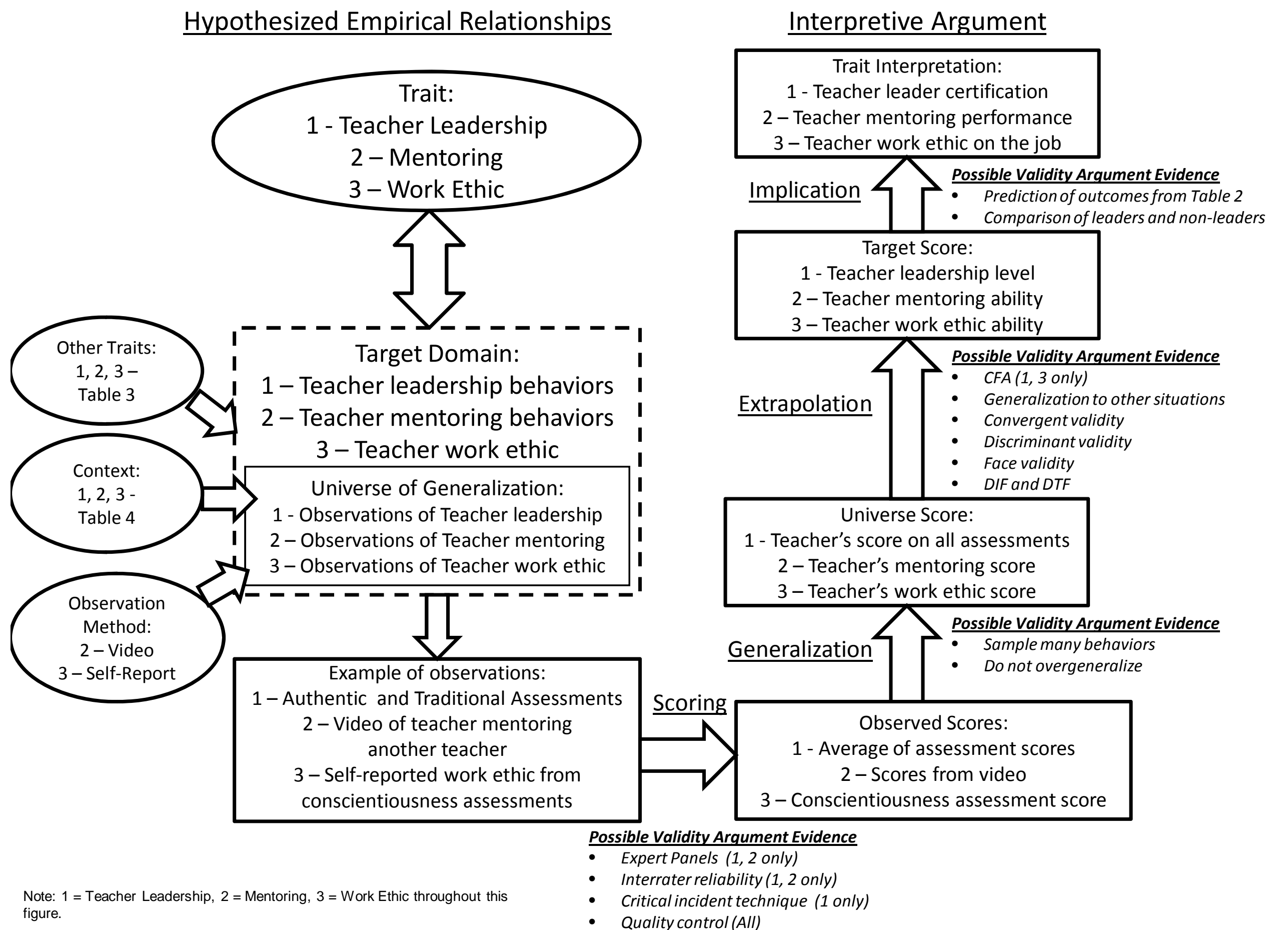

Figure 1. **Hypothesized empirical relationships, interpretive argument, and possible validity evidence for teacher leadership in general, authentic assessment of teacher leadership, and traditional assessment of teacher leadership (adapted from Kane, 2006).**

The interpretive argument is composed of four stages: *scoring*, *generalization*, *extrapolation*, and *implication*. In the following pages we describe each stage, how it may differ in traditional versus authentic assessments, and provide examples of validation studies that apply to the types of teacher leadership assessments we have described.

Scoring. When observations of teacher leadership are made they must be scored. For example, scoring rules must be created for both an authentic assessment consisting of a video demonstration of a potential teacher leader's mentoring ability and a SJT assessment of an aspiring teacher leader's teamwork ability. The interpretive argument in Table 5 includes four assumptions or inferences that are often made when observations are scored. Each of these assumptions or inferences can be assessed in the validation process.

Table 5

Interpretive argument for teacher leadership (adapted from Kane, 2006)

1. Scoring
    - 1.1 The scoring rule is appropriate
    - 1.2 The scoring rule is applied as specified
    - 1.3 The scoring is free of bias
    - 1.4 The data fit the scoring model
2. Generalization
    - 2.1 The sample of observations adequately represents teacher leadership behaviors
    - 2.2 The sample of observations is large enough to control random error
3. Extrapolation
    - 3.1 The score on the teacher leader assessment adequately represents a teacher's teacher leadership ability
    - 3.2 There are not systematic errors that undermine extrapolation
4. Implication
    - 4.1 The implications associated with the teacher leader ability score are appropriate
    - 4.2 The properties associated with the scores on the teacher leader assessments support the implications associated with the teacher leader ability score

The scoring procedures of an authentic assessment, such as a portfolio or video of one's performance, can be validated in several ways. For instance, expert panels can be formed to create scoring rules, review scoring criteria, ensure that scoring is free of bias, etc. These expert panels can consist of a diverse set of people, such as experienced teachers, administrators, or researchers in education. When humans are called on to rate performance, inter-rater reliability is typically calculated to assess the consistency of scoring.

In most traditional assessments, scoring rules are less involved, typically being comprised of summed scores or mean scores from responses to items that closely match the sample of behavior under consideration. Methods like SJTs, however, may benefit from scoring procedures more analogous to authentic assessments; i.e., expert panels (see Flanagan, 1954 for a review *critical incident technique* as applied to SJTs) to create scoring rules, review scoring criteria, check for bias, etc. Multiple raters may be used, requiring the computing of inter-rater reliability estimates.

Generalization. The next assumption of the interpretive argument is that the teacher leader behaviors assessed adequately represent the universe of behaviors we are interested in. Do observed scores generalize to a broader universe score? For example, do the behaviors measured by a video demonstrating an aspiring teacher leader's mentoring ability adequately represent that person's ability to mentor? Does a self-report assessment of work ethic adequately represent an aspiring teacher leader's propensity to work hard?

Generalization is often a problem for authentic assessment: "[T]he proponents of authentic assessments have tended to emphasize the extent to which the performances observed in testing matched the target performance while taking the generalization of observed scores over tasks, occasions, and conditions of observation for granted, even though empirical research consistently indicates that generalizability over performance tasks cannot be taken for granted" (Kane, 2006, p. 57). To counter this problem, creators of authentic assessments are advised to attempt to sample as many behaviors as possible in order to gather observations that are representative of the behavior in question. An assessment that requires aspiring teacher leaders to create two videos demonstrating their mentoring ability in two different situations is preferable to an assessment that requires just one video, for example. It is also important not to overgeneralize or overstate what an assessment is measuring. One should be careful in claiming that an assessment measures overall mentoring ability when it provides only one demonstration of mentoring.

Traditional assessments, such as self-report personality and attitude assessments have the ability to sample many behaviors in a relatively short time. Thus, generalization may not be as problematic for traditional assessments as for authentic assessments. It remains crucial, however, not to overstate what one is measuring here either. One should be careful in claiming, for example, that an assessment of extraversion is an assessment of teacher leadership ability,

because teacher leadership likely requires many more personality characteristics than extraversion alone. Thus, it is important both to sample many behaviors and not to overgeneralize when one is creating traditional assessments as well.

Extrapolation. The assumption in the interpretive argument is that the score on the teacher leader assessment is an accurate representation of one's ability as a teacher leader. This evidence can be analytic or empirical. Analytic evidence "relies on conceptual analyses and on judgments about the relationship between the universe of generalization and target domain," while empirical evidence, "examines relationships between observed scores and other scores associated with the target domain" (Kane, 2006, p. 35).

Because authentic assessments focus on the performance of tasks conducted in the real world, people not trained in psychological or educational assessment may be more likely to be able to intuitively make the connection between the assessment and real-world performance than they would with traditional assessment – what is traditionally called *face* validity. In a court of law, such face validity can be seen as a fair assessment of one's ability, giving its strong *legal defensibility*. A high level of face validity can also lead to greater buy-in from respondents themselves, leading to greater engagement in the task. Face validity can be evaluated using simple interview studies that ask teachers and administrators how well they believe each assessment represents teacher leadership.

Another analytic validity argument is to assess the differential item and differential test functioning (DIF & DTF) of assessments (e.g. Holland & Wainer, 1993). That is, scores on individual items within assessments (DIF) and whole tests (DTF) should vary only on construct-relevant dimensions. For example, scores on videos of aspiring teacher leaders' mentoring ability should vary based only on differences in mentoring ability. They should not, however, vary based on the quality of the video camera used to record the interaction. One might predict, for example, that videos taken with high quality video cameras might be rated higher than videos taken with low quality video cameras due to the greater perceptual fluency afforded by the clarity of the picture taken with higher quality cameras. In this case, ratings will vary based on the construct-irrelevant variable of camera quality and call the validity of the assessment into question.

Furthermore, studies should be conducted that assess convergent (how the assessment correlates with different measures of the same ability) and discriminant (how the assessment

correlates with measures of different abilities) validity (Campbell & Fiske, 1959). For example, scores on an authentic assessment requiring an aspiring teacher leader to demonstrate his or her ability to work well with others should theoretically correlate more highly with existing measures of teamwork (convergent validity) than with existing measures of, for example, self-esteem (discriminant validity).

Another way to provide evidence for the validity of the extrapolation assumption of authentic assessment might be to conduct *validity generalization* studies (Kane, 2006). Validity generalization would involve applying the results of one authentic assessment to the prediction of the results of other, related assessments. For example, two authentic assessments of an aspiring teacher leader's mentoring ability can be created and scores on one assessment could be used to predict scores on the other assessment. Evidence for the validity of the assessments would be gathered if the assessments significantly predicted each other.

Validity arguments for traditional assessments are identical to the arguments for authentic assessments. For example, one can conduct interview studies to assess face validity, although one may find lower face validity because it will not be immediately clear to some individuals how items such as one might find on a self-report personality test will apply to teacher leadership. As noted, low face validity makes such tests more susceptible to legal challenge in high stakes situations. A DIF study, for example, may find that self-reported extraversion varies as a function of cultural norms for reporting one's behavior rather than actual personality. People from some cultures may not report extraverted behaviors even though they actually engage in them because reporting such behaviors may be seen as a lack of humility in certain cultures. In this case, the validity of the extraversion assessment may be called into question because the assessment is varying based on cultural norms for reporting one's behavior rather than what the assessment creator was interested in (i.e., extraversion).

An example of a convergent and discriminant validity study might be to examine whether an assessments of teaching attitudes correlates highly with other assessments of teaching attitudes, but not highly with assessments of religious attitudes. An example of a validity generalization study would be to give assessments to several populations of individuals and over several different situations. A replication of results across diverse populations and situations would provide support for the validity of the assessment.

Finally, *Confirmatory Factor Analysis (CFA)* can be conducted for traditional assessments to ensure that the structural properties of the assessment are as predicted. For example, if a personality assessment, such as one designed to measure extraversion, is hypothesized to be unidimensional (i.e., only measures extraversion), a CFA can verify that the assessment indeed measures only extraversion, and not some combination of extraversion and agreeableness.

Implication. The final assumption of the interpretive argument is that the implications associated with the teacher leadership score given are appropriate. Do the teacher leadership scores predict any other variables in any way? Do the teacher leadership scores predict differences in teachers in any way?

Within this context, validity arguments focus on the implications of an assessment stating that a teacher is skilled enough to be named a certified teacher leader. Some implications of teacher leadership are listed in column one of Table 2. Treating these as outcome variables creates a framework of validity studies predicting these outcomes from assessments in using regression analysis. For example, one can operationalize "student benefits" to mean an increase in GPA or test scores as an outcome variable and then predict it by teachers' scores on assessments. Another example would be to use test scores to predict an outcome variable based on asking teachers and principals to identify those teachers in their school that they would consider teacher leaders. Supporting evidence of valid implications of assessments scores would be high agreement or correlations between those who were identified as teacher leaders by their peers and principals and higher scores by those teachers on the assessments. Similar methods apply to traditional assessments, however, traditional measures also lend themselves well to logistic regression analysis, which has the advantage of determining the predictive power of several assessments simultaneously while controlling for all other assessments.

Validation Examples

Below are examples of how we see the validation process of authentic and traditional assessments. Although several validity arguments could be made for each assessment type at each level of the interpretive argument, for the sake of brevity we focus on only a few possibilities per level. These are listed in Table 6. Because validation is a "process of evaluating the plausibility of proposed interpretations and uses" (Kane, 2006, p. 17), none of these

arguments alone are sufficient to establish validity. Rather, several logical, analytical, and empirical arguments must be made and these arguments must be continually evaluated.

Authentic Assessment Validation Example: Video Demonstration of Mentoring Skills

For the first example, imagine that an aspiring teacher leader is asked to demonstrate his or her mentoring skills by providing a video that shows him or her mentoring a junior colleague. The video shows the junior teacher teaching a class, followed by a one-on-one meeting with the senior teacher giving advice to the junior teacher on improving his or her teaching style. The video can be scored on a number of dimensions, such as the quality of the advice given and the mentor's interpersonal style. See Figure 1 for the full validation model for this example.

In the *hypothesized empirical relationships* section, the trait is mentoring and the target domain consists of all possible mentoring behaviors. The universe of generalization consists of the mentoring behaviors that will be observed. Other considerations that should be made (listed in the ovals to the left) include a variety of other traits that might influence the behaviors observed, the context in which the behaviors are observed, and the observation method (video in this example).

Table 6

Validity evidence for teacher leadership example assessments.

	Example Assessments	
Argument Component	**Authentic Assessment: Video demonstration of mentoring skills**	**Traditional Assessment: Self-report of work ethic**
Scoring	• Expert panel creates scoring procedures • Videos scored by two raters • Scoring and data entry checked for quality	• Data checked to ensure it is entered and coded correctly
Generalization	• Teachers videotape multiple examples of mentoring • Generalizability claims limited to similar circumstances (e.g., mentoring of teaching of algebra should not be generalized to mentoring of teaching English)	• Several behaviors related to work ethic assessed • Avoid overgeneralizing claims beyond those related to work ethic (e.g. should not claim that a work ethic assessment also assesses creativity)

Extrapolation	• Examine how well scores on videos correlate with other mentoring assessments • DIF analyses conducted (e.g., do scores vary as a function of video quality rather than mentoring)	• CFA to confirm hypothesized factor structure of work ethic assessment. • Assess convergent validity: assessment should relate to other assessments of work ethic • Assess discriminant validity: assessment should not be highly related to other assessments of other constructs (e.g., self-esteem) • DIF analyses (e.g., do scores vary as a function of opportunities in school rather than work ethic?)
Implication	• Video scores used to predict performance of mentee's students • Examine whether video scores differentiate those who are considered leaders from those who are not	• Predict important outcomes from work ethic assessment (e.g. student test scores) • Examine whether work ethic scores differentiate those who are considered leaders from those who are not

Traditional Assessment Validation Example: Self-report of Work Ethic

For the second example, imagine that an aspiring teacher leader is asked to demonstrate his or her work ethic by completing a questionnaire that asks about his or her behaviors as they relate to achievement orientation, commitment, persistence, resourcefulness, and responsibility. An example persistence item might be, "I continue working until the job is finished," answered on a 7-point scale ranging from 1 = *strongly disagree* to 7 = *strongly agree*.

In the hypothesized empirical relationships section, the trait is work ethic and the target domain consists of all possible behaviors related to work ethic. The universe of generalization consists of the work ethic-related behaviors that will be observed. Other considerations that should be made include other traits (e.g., teamwork) that might influence the behaviors observed, the context in which the behaviors are observed, and the observation method (self-report questionnaire in this example). See Figure 1 for the full validation model for this example.

Other Validation Considerations

Several other considerations should be made in the process of validating assessments of teacher leadership. One issue to consider is the formation of standards for teacher leadership (see Appendix A for an example from the State of Kansas). Once standards are formulated, they can be validated and can also be used for the purpose of validating other assessments. The appropriateness of the interpretations and uses of standards can be assessed in the same way as has been described above. Furthermore, because many states will define a "teacher leader" as one who meets established standards, the standards themselves can be used as criteria to help validate other assessments of teacher leadership-related constructs. These validation processes may prove to be of value for several reasons, including determining the appropriateness of the use of individual standards and the evaluation of the efficacy of teacher leaders in improving education in states and districts.

The potential social consequences of creating a position called "teacher leader" should also be considered (Kane, 2006). That is, in addition to its use in the prediction of future performance, the development and use of an assessment at times has consequences for society as a whole. The fact that an important organization is aspiring to measure something can serve as a signal to society that what is being measured is indeed of value. As an example, consider the consequential validity of the College Board's addition of a writing test to the SAT. Although The College Board has been criticized because the writing test increases demands on examinees without a substantial increase in predictive validity, high school teachers and principals report that writing has become a larger priority in their school or district since the inception of the SAT writing test (Noeth & Kobrin, 2007). In the case of authentic assessment, the development of standards of performance for teacher leaders can send a clear signal to the educational community that it is important for teacher leaders to acquire the requisite skills with which to meet these performance standards. There could, however, be unintended negative consequences to the creation of a *teacher leader* position. Designating some teachers as "leaders," for instance, could cause resentment among those teachers not designated as such. The resulting decrement in school morale could then have a negative impact on student achievement. Although not all agree, some state that the presence of negative consequences such as these would detract from an assessment's validity (Kane, 2006).

Conclusion

If teacher leadership is to become formalized into another level of the teaching profession, valid and reliable assessments of teacher leadership are needed. The current paper provides some ideas for initial steps toward providing such assessments. However, much more work is needed. Research on teacher leadership requires much more definitional clarity with which researchers can create precise operations and measurements. Quality assessments will enhance the potential of research to answer important questions, such as, "Do students of teacher leaders produce better learning outcomes in their students?" "How does quality teaching relate to teacher leadership?" and "How do beginning teachers benefit from teacher leaders?" The knowledge gained from this research has the potential to help guide strategies that improve the education system and teacher retention rates.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179–211.

American Psychological Association, National Council on Measurement in Education, & American Educational Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Armitage, C., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*, 471–499.

Austin, E. J., Deary, I. J., & Egan, V. (2006). Individual differences in response scale use: Mixed Rasch modelling of responses to NEO-FFI items. *Personality and Individual Differences, 40*, 1235–1245.

Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement, 60*, 361–370.

Barth, R. S. (2001). Teacher leader. *Phi Delta Kappan, 82,* 443–449.

Bascia, N. (1996). Teacher leadership: Contending with adversity. *Canadian Journal of Education, 21*, 155–169.

Behrstock, E., & Clifford, M. (2009). *Leading gen Y teachers: Emerging strategies for school leaders.* Retrieved from http://www.tqsource.org/publications/February2009Brief.pdf.

Billingsley, B. S. (2007). Recognizing and supporting the critical roles of teachers in special education leadership. *Exceptionality, 15,* 163–176.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Childs-Bowen, D., Moller, G., & Scrivner, J. (2000). Principals: Leaders of leaders. *NASSP Bulletin, 84*, 27–34.

Converse, P. D., Oswald, F. L., Imus, A., Hedricks, C., Roy, R., & Butera, H. (2008). Comparing personality test formats and warnings: Effects on criterion-related validity and test-taker reactions. *International Journal of Selection and Assessment, 16,* 155–169.

Crowther, F., Kaagen, S. S., Ferguson, M., & Hann, L. (2002). *Developing teacher leaders: How teacher leadership enhances school success*. Thousand Oaks, CA: Corwin Press.

Danielson, C. (2006). *Teacher leadership that strengthens professional practice.* Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L. (1988). Policy and professionalism. In A. Lieberman (Ed.), *Building a professional culture in schools* (pp. 55–77). New York, NY: Teachers College Press.

Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment in teaching in context. *Teaching and Teacher Education, 16,* 523–545.

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling: A Multidisciplinary Journal*, *13*, 440–484.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51,* 327–359.

Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology*, *2*, 300–319.

Fullan, M. G. (1994). Teacher leadership: A failure to conceptualize. In D. R. Walling (Ed.), *Teachers as leaders* (pp. 241–253). Bloomington, IN: Phi Delta Kappa Educational Foundation.

Gillespie, R. (1991). *Manufacturing knowledge: A history of the Hawthorne experiments*. Cambridge, England: Cambridge University Press.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*, 26–34.

Griffith, R. L., Chmielowski, T., & Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. *Personnel Review, 36*, 341–357.

Grissmer, D., & Kirby, S. (1997). Teacher turnover and teacher quality. *Teachers College Record*, *99*, 45–56.

Gulikers, J. T. M., Bastiaens, T. J., & Krischner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology, Research and Development, 52,* 67–86.

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

Hedge, J. W., Hanson, M. A., Borman, W. C., Bruskiewicz, K. T., & Logan, K. K. (1996, May). *Predicting the crew resource management skills of Air Force pilots* (Institute Report No. 283). Tampa, FL: Personnel Decisions Research Institutes, Inc.

Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum.

Ingersoll, R., & Smith, T. (2003.) The wrong solution to the teacher shortage. *Educational Leadership, 60*, 30–33.

Institute for Educational Leadership. (2001). *Leadership for student learning: Redefining the teacher as leader*. Retrieved from http://www.iel.org/programs/21st/reports/teachlearn.pdf.

Institute for Educational Leadership. (2008). *Teacher leadership in high schools: How principals encourage it. How teachers practice it*. Retrieved from www.iel.org/pubs/metlife_teacher_report.pdf.

Interstate School Leadership Licensure Consortium. (1996). Standards for school leaders. Washington, DC: Council of Chief School Officers.

Joyce, B., & Showers, B. (1995). *Student achievement through staff development: Fundamentals of school renewal.* White Plains, NY: Longman.

Kane, M. T. (2006) Validation. In R. L. Brennan (Ed.), *Educational Measurement* (pp. 17–64). New York, NY: Praeger.

Katzenmeyer, M., & Moller, G. (2001). *Awakening the sleeping giant: Helping teachers develop as leaders* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, NY: Guilford Press.

Killion, J., & Harrison, C. (2006). *Taking the lead: New roles for teachers and school-based coaches.* Oxford, OH: National Staff Development Council.

Krokos, K. J., Meade, A. W., Cantwell, A. R., Pond, S. B., & Wilson, M. A. (2004, April). *Empirical Keying of Situational Judgment Tests: Rationale and Some Examples*. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Chicago, IL.

Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). Attitude measurement. In D. Albarracin, B. T. Johnson, & M. P. Zanna (Eds.), *Handbook of attitudes and attitude change.* Mahwah, NJ: Erlbaum.

Kyllonen, P. C., & Lee, S. (2005). Assessing problem solving in context. In O. Wilhelm & R. W. Engle (Eds.) *Handbook of understanding and measuring intelligence.* (pp. 11–25). Thousand Oaks, CA: Sage.

Kyllonen, P. C., Lipnevich, A. A., Burrus, J., & Roberts, R. D. (in press). *Personality, motivation, and college readiness: A prospectus for assessment and development*. Princeton, NJ: ETS.

LeBlanc, P. R., & Shelton, M. M. (1997). Teacher leadership: The needs of teachers. *Action in Teacher Education, 19*, 32–48.

Lievens, F., & Coetsier, P. (2002). Situational tests in student selection: An examination of predictive validity, adverse impact, and construct validity. *International Journal of Selection and Assessment, 10,* 245–257.

Little, J. W. (2003). Constructions of teacher leadership in three periods of policy and reform activism. *School Leadership and Management, 23,* 201–419.

MacCann, C., Minsky, J., Ventura, M., & Roberts, R. D. (2009). *Mother (and father) know best: Comparing self- and parent-reported personality in predicting academic achievement*. Manuscript submitted for publication.

MacCann, C., & Roberts, R. D. (in press). Development of a student health assessment system: Health knowledge, attitudes, and behaviors in middle-schoolers. Princeton, NJ: ETS.

MacCann, C., & Roberts, R. D. (2008). Assessing emotional intelligence with situational judgment test paradigms: Theory and data. *Emotion*, *8*, 540–551.

Matthews, G., Zeidner, M., & Roberts, R. D. (2006). Personality, affect, and emotional development. In P. A. Alexander & P. H. Winne (Eds.), *Handbook in educational psychology* (2nd ed., pp. 163–186). Mahwah, NJ: Lawrence Erlbaum.

Mizell, H. (2004). *Pioneers for professional learning.* Speech given to National Staff Development Council's School Based Staff Developer Learning Community, August 5, 2004, Dallas, TX.

Motowidlo, S. J., & Tippins, N. (1993). Further studies of the low-fidelity simulation in the form of a situational inventory. *Journal of Occupational and Organizational Psychology, 66,* 337–344.

Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology,79*, 475–480.

Mueller, J. (2009). *Authentic Assessment Toolbox.* Retrieved from http://jonathan.mueller.faculty.noctrl.edu/toolbox/index.htm.

Murphy, J. (2005). *Connecting teacher-leadership to school improvement.* Thousand Oaks, CA: Corwin Press.

National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for America's future.* New York, NY: Author.

Neufeld, B., & Roper, D. (2003). *Coaching: A strategy for developing instructional capacity – Promises and practicalities.* Washington, DC: Aspen Institute Program on Education and Providence, RI: Annenberg Institute for School Reform.

Noeth, R. J., & Kobrin, J. L. (2007). *Writing changes in the nation's K-12 educational system.* (College Board Research Note No. 34). New York, NY: The College Board.

Ovando, M. N. (1996). Teacher leadership: Opportunities and challenges. *Planning and Changing, 27,* 30–44.

Porter, A. C. (1986). Teacher collaboration: New partnership to attack old problems. *Phi Delta Kappan, 69,* 147–152.

Richardson, J. (2009). *Learning teams: When teachers work together, knowledge and rapport grows.* Retrieved from http://www.nsdc.org/news/tools/tools8-01rich.cfm.

Ryan, S. (1999, April). *Principals and teachers leading together*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Sherrill, J. A. (1999). Preparing teachers for leadership roles in the 21st century. *Theory Into Practice, 38,* 56–61.

Smylie, M. A. (1994). Redesigning teachers' work: Connections to the classroom. In L. Darling-Hammond (Ed.), *Review of research in education* (Vol. 20, pp. 129–177). Washington, D.C.: American Educational Research Association.

Smylie, M. A. (1995). New perspectives on teacher leadership. *The Elementary School Journal, 96,* 3–7.

Steffy, B. E., Wolfe, M. P., Pasch, S. H., & Enz, B. J. (2000). *Life cycle of the career teacher.* Thousand Oaks, CA: Corwin Press.

Wang, L., MacCann, C., Zhuang, X., Liu, L., & Roberts, R. D. (2009). Assessing teamwork and collaboration in high school students: A multimethod approach. *Canadian Journal of School Psychology, 24,* 108–124.

Wasley, P. A. (1991). *Teachers who lead: The rhetoric of reform and realities of practice*. New York, NY: Teachers College Press.

Yarger, S. J., & Lee, O. (1994). The development and sustenance of instructional leadership. In D. R. Walling (Ed.), *Teachers as leaders: Perspectives on the professional development of teachers* (pp. 223–237). Bloomington, IN: Phi Delta Kappa Educational Foundation.

York-Barr, J., & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research, 74,* 255–316.

Ziegler, M., MacCann, C., & Roberts, R. D. (Eds.) (in press). *New perspectives on faking in personality assessment*. New York, NY: Oxford University Press.

Appendix A

Kansas Teacher Leader Standards

Standard 1: The teacher leader is able to apply strategies of adult learning across teacher leadership activities.

Knowledge: The teacher leader demonstrates knowledge of. . .

- The differences in knowledge acquisition and transfer for children and adults
- Stages of career development and learning for colleagues
- Effective use of individual and group interactions such as collaboration, networking, facilitation, team building, and conflict resolution
- Effective listening, oral communication, presentation skills, and expression in written communication
- Research and exemplary practice on "organizational change and innovation"

Performance: The teacher leader. . .

- Demonstrates knowledge and skills for high quality professional learning for individuals as well as groups
- Assesses teachers' content knowledge and skills throughout professional learning
- Fosters mutually respectful and productive relationships among colleagues
- Uses effective communication skills and processes
- Demonstrates the ability to adapt to the contextual situation and make effective decisions
- Demonstrates knowledge of the role of creativity, innovation, and flexibility in the change process
- Improves colleagues' acquisition and application of knowledge and skills

Standard 2: The teacher leader is able to advance the professional skills of colleagues by demonstrating and applying expertise in observational skills and in providing quality feedback in order to support reflective practice focused on improving curriculum, instruction, and assessment.

Knowledge: The teacher leader demonstrates knowledge of. . .

- Research-based curriculum, instruction, and assessment and their alignment with desired outcomes
- Models and protocols for effective observation and feedback
- Role and use of critical reflection in improving professional practice

Performance: The teacher leader. . .

- Recognizes, analyzes, and improves quality of colleagues' professional and instructional practices
- Uses effective observation techniques to identify opportunities to improve curriculum, instruction, and assessment
- Provides observational feedback that improves curriculum, instruction, and assessment

- Develops, leads and promotes a culture of self-reflection and reflective dialogue
- Applies mentoring as well as coaching practices to support colleagues' individual and group professional improvement and career development

Standard 3: The teacher leader is able to improve the quality of colleagues' collaboration and interaction with families and other stakeholders.

Knowledge: The teacher leader demonstrates knowledge of. . .
- Child development and conditions in the home, culture and community and their influence on educational processes
- Contextual considerations of the family, school, and community and their interaction with educational processes
- Effective strategies for involvement of families and other stakeholders as part of a responsive culture

Performance: The teacher leader. . .
- Develops colleagues' abilities to form effective relationships with families and other stakeholders
- Recognizes, responds and adapts to contextual considerations to create effective interactions among families, communities, and schools
- Improves educational outcomes by promoting effective interaction and involvement of teachers, families, and stakeholders in the educational process

Standard 4: The teacher leader is able to initiate and facilitate colleagues' design and implementation of action research and analysis of data for individual and group decision-making

Knowledge: The teacher leader demonstrates knowledge of. . .
- Action research methodology
- Analysis of research data and development of a data-driven action plan that reflects relevance and rigor
- Implementation strategies for research-based change and for dissemination of findings for programmatic changes

Performance: The teacher leader. . .
- Models and facilitates relevant and targeted action research
- Models and facilitates analysis and application of research findings to improve educational outcomes
- Engages colleagues in identifying research questions as well as designing and conducting action research to improve educational outcomes
- Facilitates the analysis of data for informed decision making to improve educational results with a focus on increased productivity, effectiveness and accountability
- Assists with application and supports dissemination of action research findings to improve educational outcomes

Standard 5: The teacher leader is able to develop and support collaborative teams and promote collegial interactions that improve the effectiveness of practice.

Knowledge: The teacher leader demonstrates knowledge of. . .

- Collaboration, facilitation, team building, and conflict resolution techniques
- Influence of individual characteristics on group interactions
- Structures and processes for collaborative work
- The process of development of group goals and objectives

Performance: The teacher leader. . .

- Facilitates development of a responsive culture with shared vision, values, and responsibility
- Applies understanding of team members' characteristics to develop collaborative teams
- Guides purposeful collaborative interactions, inclusive of team members' ideas and perspectives
- Promotes team-based responsibility for assessing and advancing the effectiveness of practice
- Creates structures and processes for collaborative teams that promote collegiality and result in improved practice

Standard 6: The teacher leader is able to identify and assess opportunities for educational improvement, and advocate effectively for them within and beyond the school community.

Knowledge: The teacher leader demonstrates knowledge of. . .

- Effective identification and interpretation of data, research findings, and exemplary practices
- Alignment of opportunities with identified needs
- Synthesis of information to support a proposal for educational improvement
- Local, state and national policy decisions and their influence on instruction
- The process to impact policy and to advocate on behalf of students and the community

Performance: The teacher leader. . .

- Identifies and evaluates needs and opportunities
- Generates ideas to effectively address solutions/needs
- Analyzes feasibility of potential solutions and relevant policy context
- Advocates effectively and responsibly to relevant audiences for realization of opportunities

Standard 7: The teacher leader is able to inform and facilitate colleagues' selection or design, use, and interpretation of multiple assessments, along with other available data, to make informed decisions that improve the quality of instruction and student learning.

Knowledge: The teacher leader demonstrates knowledge of. . .

- Design and selection of suitable evaluation instruments and effective assessment practices for a range of purposes
- Analysis and interpretation of data from multiple sources
- Use of formative and summative data to inform the continuous improvement process

Performance: The teacher leader. . .

- Informs and facilitates colleagues' selection or design of suitable evaluation instruments to generate data that will inform instructional improvement
- Informs and facilitates colleagues' interpretation of data and application of findings from multiple sources (e.g., standardized assessments, demographics and other relevant sources) to guide instructional decisions and improve educational practice

Standard 8: The teacher leader is able to inform and facilitate the design and implementation of coherent, integrated and differentiated professional development based on assessed student and teacher needs.

Knowledge: The teacher leader demonstrates knowledge of. . .

- Selection and evaluation of career professional development resources appropriate to the identified need(s)
- The standards of high quality professional development and their relevance to improved learning
- Application of the concepts of adult learning to the design and implementation of professional development
- Effective use of professional development needs assessment, designs, protocols, and evaluation tools
- The role of 21st century skills and technologies in educational practice
- The role of shifting cultural demographics in educational practice

Performance: The teacher leader. . .

- Accurately identifies the professional development needs and opportunities for colleagues in the service of improving education
- Works with staff and staff developers to design and implement ongoing professional learning based on assessed teacher and student needs
- Involves colleagues in development and implementation of a coherent, systemic, and integrated approach to professional development aligned with school improvement goals
- Utilizes and facilitates the use of technology and media literacy as appropriate
- Continually assesses the effectiveness of professional development activities and adjusts appropriately

Appendix B

What Makes an Assessment *Authentic*?

One of the most important aspects of authentic assessment is that the assessment is indeed, authentic. One might assume that conducting an authentic assessment is relatively straightforward; after all, authentic assessment simply means demonstrating one's skills in real-life settings. However, a closer inspection of many of the steps described in this paper reveals that it is sometimes debatable whether an assessment is an authentic one. For example, a task which requires creating a video demonstrating one's ability to mentor junior teachers may be seen as an authentic task by some; however, the mere fact that the mentorship is being videotaped takes away from the authenticity of the task. After all, most mentorship is not conducted in front of a camera, and the knowledge that one's every actions are being recorded might influence the interactions that occur, thus creating a less-than-authentic task. In fact, the idea that people tend to change their behavior when they are being observed (the Hawthorne effect) is one of the most well-known findings from the field of psychology (Gillespie, 1991). Recently, Gulikers et al. (2004) have attempted to define authenticity by outlining five dimensions of authentic assessment. They are

- **Tasks**. An authentic task is a task a teacher leader performs that also occurs in professional life. It should mimic the complexity of the real-life task, respondents should feel ownership of the task, and respondents should feel the task is relevant.
- **Physical context**. The physical context of the assessment should closely simulate reality. Respondents should have similar resources and time available that they would in real life.
- **Social context**. A situation that normally requires collaboration should be assessed with a collaborative effort, and a situation that normally is handled by individuals should be assessed with an individual effort.
- **Assessment results or form**. An authentic result or form: a) should be a product that respondents can produce in real life; b) permits making valid inferences about the underlying competency; c) contains multiple indicators of skills,

knowledge, or abilities; and d) requires respondents to present and defend their work.

- **Criteria and standards**. Criteria and standards should be made explicit and transparent to respondents, as they would be in real life.

Although perhaps an incomplete list, an assessment that nonetheless follows these basic dimensions should be able to come closer to authenticity than would an assessment that does not follow this basic framework. Other dimensions of authenticity that could be added are numerous. For example, no mention is made of the psychological context of the assessment. That is, do all participants think, feel, and behave as they would in a real-life situation? The presence of task, physical context, and social context authenticity does not necessarily guarantee psychological authenticity. In fact, these three contexts may interact in several ways to unpredictably influence the psychological context of the situation.