

VANDERBILT ASSESSMENT *of* LEADERSHIP *in* EDUCATION™



VAL-ED Users' Guide: Implementation and Interpretation

Table of Contents

(linked for convenience)

Acknowledgements

About the Authors

Preface

Chapter 1. Introduction*	p. 6
Chapter 2. Conducting a VAL-ED Assessment	p. 13
Chapter 3. Interpreting and Using VAL-ED Results	p. 19
Chapter 4. Summary of the Technical Soundness and Utility of the VAL-ED	p. 32
References	p. 37
List of Tables	p. 42
List of Figures	
Appendix A. Sample Principal Report	p. 43

*Please note that sections of Chapter 1 are from two soon to be published articles by the author team in *The Elementary School Journal* and *Journal of Leadership and Policy in Schools*. Specifically, the references for these articles are:

Goldring, E., Cravens, X.C., Murphy, J., Elliott, S.N., Carson, B., & Porter, A.C. (in press). The evaluation of principals: What and how do states and districts assess leadership? *Elementary School Journal*.

Goldring, E., Porter, A., Murphy, J., Elliott, S.N., & Cravens, X. (2009). Assessing learning-centered leadership: Connections to research, professional standards, and current practices. *Journal of Leadership and Policy in Schools*, 8,1-36.

Acknowledgements

The authors wish to acknowledge the encouragement and financial support in the form of a grants from the Wallace Foundation and Institute of Education Sciences. Without this support, the VAL-ED would not exist. We also have benefited tremendously from the insights and assistance of a number of pre- and post-doctoral students: Xiu Cravens, Jennifer Frank, Morgan Polikoff, and Timothy Zeidner, in the design of the instrument. Henry May, along with others, provided expert guidance with validation and normative data analyses. Finally, we thank James O'Toole for his skillful and persistent efforts to recruit a national sample of principals, their teachers, and supervisors to provide us the normative basis for interpreting VAL-ED scores. As noted, it has taken real teamwork to design the VAL-ED and we are very thankful for each member's expertise and efforts over the past several years!

SNE
JM
EG
ACP

About the Authors of this User's Guide

Stephen N. Elliott is an educational psychologist with a specialization in school psychology and applied measurement. He is a Professor of Special Education and the Dunn Family Chair of Educational and Psychological Assessment in Peabody College at Vanderbilt University. Steve directs the Learning Sciences Institute at Vanderbilt and teaches courses in assessment and scale development. He has designed, developed, and conducted norming and validity studies, in collaboration with major test publishers, for the *Social Skills Rating System*, *Social Skills Improvement System*, *Academic Competence Evaluation Scales*, *Academic Intervention Monitoring System*, and the *Student Self-Concept Scales*. Steve has also been the lead developer on statewide alternate assessment systems for Hawaii, Idaho, and Mississippi departments of education. These alternate assessment systems involve multiple raters, multiple sources of evidence collected over time, criterion referenced scoring rubrics, and integrated proficiency reports.

Joseph Murphy is a Professor of Leadership, Policy, and Organizations at Vanderbilt University. He is a former administrator at the school, district, and state levels. He has published extensively on the topic of leadership in the leading journals in education as well as in the major practitioner outlets. His work on school administration focuses on the instructional dimensions of school leadership, beginning with his (and Philip Hallinger's) seminal work on principal and superintendent instructional leadership in the 1980's through his most recent volume (and related research) on teacher leadership ([Connecting Teacher Leadership and School Development](#), 2005). His work has had an important influence on the shape and texture of school leadership in the United States, perhaps most noticeably as the foundation for the ISLLC [Standards for School Leaders](#). Murphy has had considerable experience in developing processes and tools for leadership assessment. For example, he led the teams that created the specifications (with ETS) for the national licensure examination (School Leader Licensure Assessment) and for the statewide administrator evaluation system in Delaware.

Ellen Goldring is the Alexander Heard Distinguished Service Professor and Professor of Education Policy and Leadership at Peabody College, Vanderbilt University. She recently served as the coeditor of *Educational Evaluation and Policy Analysis*. She is also a member of the *Task Force on Developing Research to Improve Educational Leadership*, cosponsored by the American Education Research Association, The Laboratory of Student Success, and University Council for Educational Administration. Her work has been funded by the Spencer Foundation, The WT Grant Foundation, the Bill and Melinda Gates Foundation and the Institute of Educational Sciences. Goldring has widespread experience in the field of educational leadership. Most recently, she has designed and implemented a principal leadership academy in partnership with the Metropolitan Nashville Public Schools.

Andy Porter is Dean of the Graduate School of Education at the University of Pennsylvania and George and Diane Weiss Professor of Education. Trained as a statistician/psychometrician, he has been an active education researcher for more than 30 years and has led successful education research and development centers for 25 years. Porter currently chairs the assessment and accountability technical panels for the states of Ohio, Wisconsin and Missouri. His tools for measuring the content of instruction in K12 schools for the subjects of math, science, and English language arts are used in over 30 states. He is an elected member and vice president of the National Academy of Education, member of the National Assessment Governing Board, Lifetime National Associate of the National Academies, and past-President of the American Educational Research Association. Websites: andyporter.org/; gse.upenn.edu/.

Preface

The Vanderbilt Assessment of Leadership in Education (VAL-ED) is markedly different from other leadership evaluation and assessment frameworks employed by states and districts throughout the United States. First, the VAL-ED uses 360 degree feedback, from teachers, principals, and supervisors. Second, the content of the proposed assessment is *learning-centered* leadership behaviors, behaviors that are related to increases in student achievement. Third, the assessment is of leadership *behaviors*, not knowledge, dispositions, or personal characteristics of leaders. Fourth, the VAL-ED requires respondents to identify *evidence* on which they are basing their assessment of principal behaviors. Fifth, the psychometric properties are clearly documented and the authors have an ongoing program of research that will continue to yield evidence concerning the reliability and validity of VAL-ED scores. Information on norms, standards, and uses is available. In short, the VAL-ED is conceptually and theoretically grounded and the resulting scores are reliable and valid for purposes of evaluating learning-centered leadership.

When completed as designed by a representative and large portion of eligible respondents in a school, the VAL-ED can provide useful results for the purposes of evaluating the performance of a principal and also identifying leadership behaviors for improvement. The results from the VAL-ED address fundamental questions such as:

- ◆ *Who participated in the assessment of the principal's leadership behavior?*
- ◆ *What evidence did these respondents report using to make their effectiveness ratings of the principal?*
- ◆ *How effective is the principal's leadership behavior judged in comparison to a national sample of principals?*
- ◆ *How effective is the principal's leadership behavior in comparison to the VAL-ED proficiency standards?*
- ◆ *To what degree did the three respondent groups --teachers, supervisor, and principal-- agree with regard to effectiveness ratings?*
- ◆ *Which areas of leadership behavior represent areas of relative strength and which represent areas for possible improvement?*

When used two or more times across years, the VAL-ED can also address questions about behavior change and about the effects of programs designed to improve learning-centered leadership behaviors. This User's Guide is designed to provide the reader fundamental information for using and interpreting the results of the VAL-ED.



Chapter 1 Introduction

Overview of the VAL-ED

The Vanderbilt Assessment of Leadership in Education (VAL-ED) is an evidenced-based, multi-rater rating scale that assesses principals' behaviors known to directly influence teachers' performance, and in turn students' learning. The VAL-ED measures critical leadership behaviors for the purposes of diagnostic analyses, performance feedback, progress monitoring, professional development planning, and summative evaluation.

Principal leadership is an essential element of successful schools. To date, much of the work on developing educational leadership for school improvement has focused on licensure; program accreditation; and professional development, including coaching and mentoring. The identification and development of effective leadership, however, has been significantly hampered by the paucity of technically sound tools for assessing and monitoring the performance of school leaders. Until the publication of the VAL-ED, there has not been a school leadership assessment instrument that has undergone scientific, psychometric development.

With initial funding from the Wallace Foundation, we addressed this problem by working on the development of an instrument to assess the effectiveness of school leaders as evaluated by teachers, supervisors, and principals themselves. The resulting instrument is a paper or on-line assessment that utilizes a multi-rater, evidence-based approach to measure the effectiveness of leadership behaviors known to influence teacher performance and student learning. The VAL-ED is a 360 degree assessment; that is, key people surrounding the principal (i.e., teachers, the principal's supervisor, and the principal him or herself) respond to the behavior inventory. The VAL-ED measures *core components* and *key processes*. Core components refer to characteristics of schools that support the learning of students and enhance the ability of teachers to teach. Key processes refer to how leaders create and manage those core components. Effective learning-centered leadership is at the intersection of the two dimensions: core components created through key processes. The outcomes of the assessment include behavior profiles, interpretable from both norm-referenced and standards-referenced perspectives, and suggested clusters of behaviors for improvement.

The VAL-ED was constructed to (a) work well in a variety of settings and circumstances, (b) be unbiased, (c) construct valid, (d) reliable, (e) feasible for widespread use (both on-line and paper-and-pencil versions), (f) provide accurate and useful reporting of results, (g) yield a diagnostic profile for formative purposes, (h) used to measure progress over time in the development of leadership, and (i) predict important outcomes.

Learning-Centered Leadership Framework: The Blueprint for VAL-ED

Our leadership assessment instrument is part of a comprehensive model of a leadership assessment system that captures in broad strokes how education leadership might be assessed. The model (see Figure 1.1) shows that leadership knowledge and skills, personal characteristics, and values and beliefs inform the actual leadership behaviors exhibited by individuals or teams in performing their leadership responsibilities. These leadership behaviors (the constructs measured in our assessment instrument and reviewed in detail below) then lead to school

performance on core components such as providing a rigorous curriculum and high-quality instruction. These school performances, in turn, lead to student success. Student success is defined as value-added, for example, improvements in student achievement, student attendance, student graduation rates, and college enrollment.

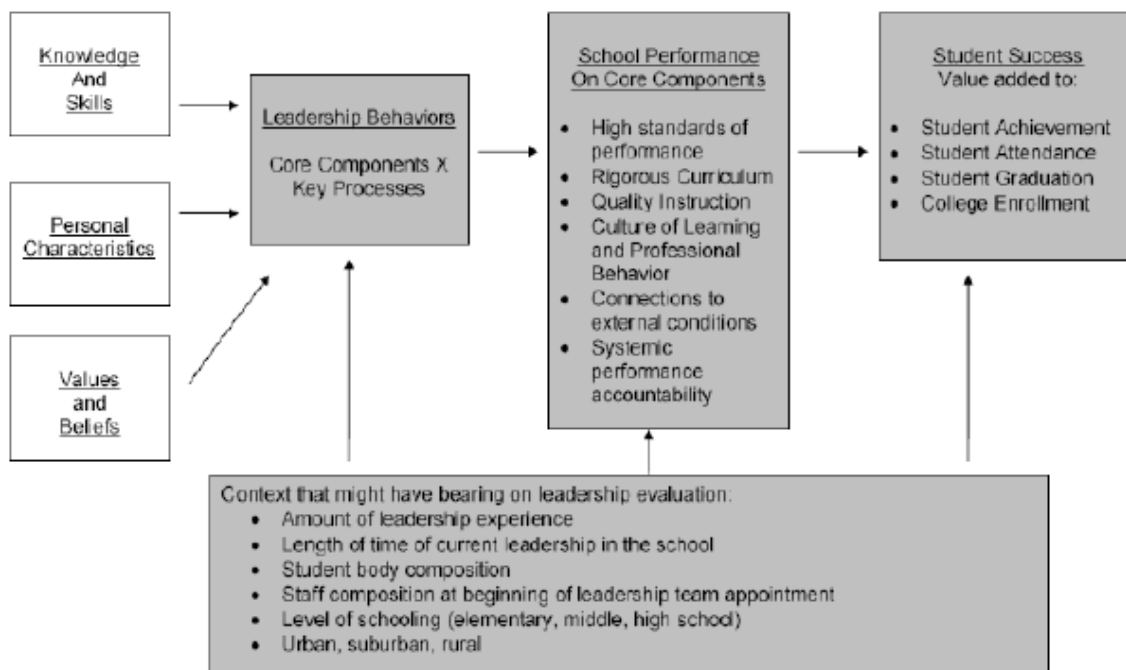


Figure 1.1. Concept framework for the Vanderbilt Assessment of Leadership in Education

Consistent with the empirical research (Hallinger & Heck, 1996; Heck & Hallinger, 1999; Leithwood et al., 2004), our assessment model does not envision direct effects of leadership behaviors on student success. Rather, the leadership behaviors lead to changes in school performance, which in turn lead to student success. Our leadership model also posits that there are aspects of the context within which leadership and schooling takes place that bear on leadership evaluation (Murphy & Meyers, 2008). Levels of experience, student body composition, staff composition, level of schooling, and geographic setting of the school can all have bearing on high-quality education leadership. The grey in Figure 1.1 indicates the focus of the VAL-ED assessment system, a piece of which is the VAL-ED assessment of leadership behaviors.

Inside this model, our assessment instrument of principals' leadership behaviors is defined by the intersection of 6 core components of school performance and 6 key processes which together make up our conception of principal leadership (See Figure 1.2).

Key processes						
Core components	Planning	Implementing	Supporting	Advocating	Communicating	Monitoring
High Standards for Student Learning						
Rigorous Curriculum (content)						
Quality Instruction (pedagogy)						
Culture of Learning & Professional Behavior						
Connections to External Communities						
Performance Accountability						

Figure 1.2. The VAL-ED Constructs of Core Components and Key Processes

The theory of action underlying our leadership assessment instrument focuses on two key dimensions of leadership behaviors: *core components* and *key processes*. Our framework states that school leadership assessment should include measures of the intersection of these dimensions. Does the leadership in the school support teachers to develop a culture of learning and professional behavior? Does the leadership implement programs to ensure there is a culture of learning and professional behavior? Does the leadership communicate effectively about the culture of learning?

The VAL-ED assesses the intersection of *what* principals must accomplish to improve academic and social learning for all students (the core components), and *how* they create those core components (the key processes). A substantial research base supports the constructs of the core components and key processes (See Knapp et al., 2003; Leithwood et al., 2004; Murphy et al., 2007; Goldring et al., 2007 for recent reviews). Core components refer to characteristics of schools that support the learning of students and enhance the ability of teachers to teach (Marks & Printy, 2003; Sebring & Bryk, 2000). Key processes are leadership behaviors, most notably aspects of transformational leadership traditionally associated with processes of leadership that raise organizational members' levels of commitment and shape organizational culture (Burns, 1978; Conley & Goldman, 1994; Leithwood, 1994).

Core Components

Our framework includes 6 core components that represent the constructs of effective learning-centered leadership as grounded in the literature:

High Standards for Student Learning. We defined high standards for student learning as the extent to which leadership ensures there are individual, team, and school goals for rigorous student academic and social learning. There is considerable evidence that a key function of

effective school leadership concerns shaping the purpose of the school and articulating the school's mission (Hallinger & Heck, 2002; Knapp et al., 2003; Murphy et al., 2007). In our framework, we do not assess the mere presence of goals for student learning, but specifically emphasize the quality of the school goals, namely the extent to which there are high standards and rigorous learning goals. The research literature over the last quarter century has consistently supported the notion that high expectations for all, including clear and public standards, is one key to closing the achievement gap between advantaged and less advantaged students, and for raising the overall academic achievement of all students (Betts & Grogger, 2003; Brookover & Lezotte, 1977; Newmann, 1997; Purkey & Smith, 1983).

Rigorous Curriculum. We define a rigorous curriculum as the content of instruction, as opposed to the pedagogy of instruction, which is dealt with in the following section. Rigorous curriculum is defined as ambitious academic content provided to all students in core academic subjects. School leaders play a crucial role in setting high standards for student performance in their schools. These high standards, however, must be translated into ambitious academic content represented in the curriculum students experience. Murphy and colleagues (2007) argued that school leaders in productive schools are knowledgeable about and deeply involved in the school's curricular program (Marzano et al., 2005). These leaders work with colleagues to ensure that the school is defined by a rigorous curriculum program in general and that each student's program, in particular, is of high quality (Newmann, 1997; Ogden & Germinario, 1995). Learning-centered leaders ensure that each student has an adequate opportunity to learn rigorous content in all academic subjects (Boyer, 1983).

Quality Instruction. A rigorous curriculum (i.e., ambitious academic content) is insufficient to ensure substantial gains in student learning; quality instruction (i.e., effective pedagogy) is also required (Leithwood et al., 2004). Quality instruction is defined as effective instructional practices that maximize student academic and social learning. This component reflects research findings over the course of the past few decades about how people learn (National Research Council, 1999). That work makes clear that teachers' pedagogical practices must draw out and work with the pre-existing understanding that students bring to the classroom. Effective instructional leaders understand the properties of quality instruction and find ways to ensure that quality instruction is experienced by all students in their schools. They spend time on the instructional program, often through providing feedback to teachers and supporting teachers to improve their instruction (Wellisch et al., 1978; Marzano et al., 2005).

Culture of Learning and Professional Behavior. Another core component in our assessment framework is leadership that ensures there are integrated communities of professional practice in the service of student academic and social learning—that is, a healthy school environment in which student learning is the central focus. Research has demonstrated that schools organized as communities, rather than bureaucracies, are more likely to exhibit academic success (Bryk & Driscoll, 1988; Lee, Smith, & Croninger, 1995; Louis & Miles, 1990). Further research supports the notion that effective professional communities are deeply rooted in the academic and social learning goals of the schools (Little, 1982; Rosenholtz, 1989). Often termed teacher professional communities, these collaborative cultures are defined by elements such as shared goals and values, focus on student learning, shared work, deprivatized practice, and reflective dialogue (Louis, Marks, & Kruse, 1996). School leadership plays a central role in the extent to which a school exhibits a culture of learning and professional behavior and includes integrated professional communities (Bryk, Camburn, & Louis, 1999; Louis, Marks, & Kruse, 1996)

Connections to External Communities. Leading a school with high expectations and academic achievement for all students requires robust connections to the external community. There is a substantial research base that has reported positive relationships between family involvement and social and academic benefits for students (Henderson & Mapp, 2002). A study of standards-based reform practices, for instance, found that teacher outreach to parents of low-performing students was related to improved student achievement (Westat and Policy Studies Associates, 2001). Similarly, schools with well-defined parent partnership programs show achievement gains over schools with less robust partnerships (Shaver & Walls, 1998). Learning-centered leaders play a key role in both establishing and supporting parental involvement and community partnerships.

Systemic Performance Accountability. There is individual and collective responsibility among the leadership, faculty, students, and the community for achieving the rigorous student academic and social learning goals. Accountability stems from both external and internal accountability systems (Adams & Kirst, 1999). External accountability refers to performance expectations that emerge from outside the school and the local community. Simultaneously, schools and districts have internal accountability systems with local expectations and individual responsibilities. Internal goals comprise the practical steps that schools must take to reach those targets. Schools with higher levels of internal accountability are more successful within external accountability systems, and they are more skillful in areas such as making curricular decisions, addressing instructional issues, and responding to various performance measures (Bryk & Schneider, 2002; Elmore, 2005). Learning-centered leaders integrate internal and external accountability systems by holding their staffs accountable for implementing strategies that align teaching and learning with achievement goals and targets set by policy.

Key Processes

Our conceptual framework features 6 key process constructs. Following a systems view of organizations, we acknowledge the processes are interconnected, recursive, and reactive to one another, but for purposes of our assessment and descriptive analysis we review each individually.

Planning. An essential process of leadership is planning. We define planning as articulating shared direction and coherent policies, practices, and procedures for realizing high standards of student performance. Planning helps leadership focus resources, tasks, and people. Learning-centered leaders do not see planning as a ritual or as overly bureaucratic. They engage in planning as a mechanism to realize the core components of the school. Effective principals are highly skilled planners and in fact, they are proactive in their planning work (Leithwood & Montgomery, 1982). Planning is needed in each of the core components; it is an engine of school improvement that builds common purpose and shared culture (Goldring & Hausman, 2001; Teddlie, Stringfield, Wimpleberg, & Kirby, 1989).

Implementing. After planning, leaders implement; they put into practice the activities necessary to realize high standards for student performance. In a comprehensive review of the research on implementation of curriculum and instruction, Fullan and Pomfret (1977) concluded that “implementation is not simply an extension of planning... it is a phenomenon in its own right” (p. 336). Effective leaders take the initiative to implement and are proactive in pursuing their school goals (Manasse, 1985). Learning-centered leaders are directly involved in implementing policies and practices that further the core components in their schools (Knapp et al., 2003). For example, effective leaders implement joint planning time for teachers and other

structures as mechanisms to develop a culture of learning and professional behavior (Murphy, 2005). Similarly, they implement programs that build productive parent and community relations as a way to achieve connections to external communities (Leithwood & Jantzi, 2005).

Supporting. Leaders create enabling conditions; they secure and use the financial, political, technological, and human resources necessary to promote academic and social learning. Supporting is a key process that ensures the resources necessary to achieve the core components are available and used well. This notion is closely related to the transformational leadership behaviors associated with helping people be successful (Leithwood & Jantzi, 2005). The literature is clear that learning-centered leaders devote considerable time to supporting teachers in their efforts to strengthen the quality of instruction (Conley, 1991; Leithwood & Jantzi, 1990). This support takes varied forms. Leaders demonstrate personal interest in staff and make themselves available to them (Marzano et al., 2005). Leaders also provide support for high-quality instruction by ensuring that teachers have guidance as they work to integrate skills learned during professional development into their instructional behaviors (Murphy et al., 2007).

Advocating. Leaders promote the diverse needs of students within and beyond the school. Advocating for the best interests and needs of all children is a key process of learning-centered leadership (Murphy et al., 2007). Learning-centered leaders advocate for a rigorous instructional program for all students. They ensure that policies in the school do not prevent or create barriers for certain students to participate in classes that are deemed gateways to further learning, such as algebra. They ensure that special needs students receive content-rich instruction. Similarly, effective leadership ensures that all students are exposed to high-quality instruction; they manage the parental pressures that often create favoritism in placing students in particular classes. Both the instruction and content of the school's educational programs honor diversity (Ogden & Germinario, 1995; Roueche & Baker, 1986). Through advocacy, learning-centered leadership works with teachers and other professional staff to ensure that the school's culture both models and supports respect for diversity. (Butty, LaPoint, Thomas, & Thompson, 2001; Goldring & Hausman, 2001).

Communicating. Leaders develop, utilize, and maintain systems of exchange among members of the school and with its external communities. In studying school change, Loucks and colleagues (1982) found that "principals played major communication roles, both with and among school staff, and with others in the district and in the community" (p. 42). Learning-centered leaders communicate unambiguously to all the stakeholders and constituencies both in and outside the school about the high standards of student performance (Leithwood & Montgomery, 1982; Knapp et al., 2003). Leaders communicate regularly and through multiple channels with families and community members, including businesses, social service agencies, and faith-based organizations (Edmonds & Frederiksen, 1978; Garibaldi, 1993; Marzano et al., 2005). Through ongoing communication, schools and the community serve as resources for one another that inform, promote, and link key institutions in support of student academic and social learning.

Monitoring. Leaders systematically collecting and analyzing data to make judgments that guide decisions and actions for continuous improvement. This is monitoring. Early on, the effective schools literature identified monitoring school progress in terms of setting goals, assessing the curriculum, and evaluating instruction as a key role of instructional leadership (Hallinger & Murphy, 1985; Purkey & Smith, 1983). Learning-centered leaders monitor the school's curriculum, assuring alignment between rigorous academic standards and curriculum coverage (Eubanks & Levine, 1983). They monitor students' programs of study to ensure that all

students have adequate opportunity to learn rigorous content in all academic subjects (Boyer, 1983; Hallinger & Murphy, 1985). Learning-centered leadership also undertakes an array of activities to monitor the quality of instruction, such as ongoing classroom observations (Heck, 1992). Monitoring student achievement is central to maintaining systemic performance accountability.

Development of a Technically Sound Assessment of Leadership

To measure principals' behaviors and actions at the intersection of core components and key processes, we developed the VAL-ED, a multi-respondent (Principal, Teacher, and Supervisor) rating scale that requires respondents to make judgments about a principal's leadership behaviors that influence teachers' performance and students' learning. Our 36 cell conceptual model of leadership provided the framework for writing items that described leaders' behaviors represented by the cell. Each cluster of items in a cell thus served as indicators of the construct of leadership we desire to measure (see Figure 2).

The design of the VAL-ED is directly influenced by technical standards for high-quality assessments (*Standards for Educational and Psychological Testing*, AERA, APA, & NCME, 1999), principles of universal design (NCEO, September 2006), and time-tested practices of item and test development (Downing, 2006; Haladyna, Downing, & Rodriguez, 2002). Collectively, these professional documents and the published research on test development and high-quality assessment of human performance provide strong guidelines for designing a high-quality and successful assessment program for school leaders.

The development of a technically sound assessment is an ongoing process that begins with the conceptualization of the instrument and continues well after the instrument is published. The development of the VAL-ED was guided by a comprehensive plan that involved: (1) specifying the purposes of the assessment, (2) defining content assessed, (3) writing items, (4) developing test specifications for validity evidence plans, (5) designing instructions and response format, (6) piloting test forms, (7) designing scoring and interpretation frameworks for scores, (8) conducting studies that yield evidence for the reliability and validity of the scores, (9) refining items, format, and score interpretation procedures, (10) field-testing forms with a representative sample, (11) developing norms and standards to guide interpretation of results, and (12) writing a technical manual that summarizes technical characteristics and sound uses of the assessment.

Using classic test theory methodology, as well as item analyses tactics from Item Response Theory (IRT), we conducted a set of studies that provide substantial evidence for the reliability and validity of VAL-ED score inferences. These studies are described in detail in the VAL-ED Technical Manual (Porter, Murphy, Goldring, & Elliott, 2008). Collectively, the studies create the technical foundation for the VAL-ED scores and thus make it possible to use the instrument with confidence to evaluate principals' leadership behaviors. In the next two chapters, we focus on the administration and interpretation of results from the VAL-ED. We conclude this Guide with chapters on Applications of the VAL-ED and Frequently Asked Questions.



Chapter 2

Conducting a VAL-ED Assessment

The primary purpose of the VAL-ED is to evaluate the leadership behaviors of a school's principal. To achieve reliable and valid results from the VAL-ED requires that it be implemented with integrity with a robust, representative sample of educators who have had the opportunity to interact with the principal for at least 2 months, and ideally 6 months. The details of this implementation process are provided in this chapter.

Directions for Completing the VAL-ED

Respondents are asked how effective the principal is at specific actions that effect core components of learning-centered leadership. The effectiveness ratings range from 1 = Ineffective to 5 = Outstandingly Effective for each of 72 behaviors. These behaviors sample all 36 cells of our conceptual model of leadership equally and thus serve as indicators of the construct of leadership we desire to measure (see Figure 1.1). Respondents rate the extent to which the principal ensures behaviors and actions are taken in the school, thus acknowledging that principals do not necessarily perform the behavior themselves, but often designate and distribute these leadership practices and behaviors throughout the school. The specific directions given to Principal respondents are:

1. *Read each item describing a leadership behavior. In some cases, you may not have actually performed the behavior, but you have ensured that it was done by others in the school. Either way the behavior should be rated.*
2. *Check (✓) the key **Sources of Evidence** you use for the basis of your assessment. Note, at least one source of evidence must be checked for an item before you make an Effectiveness rating. If you check **No Evidence**, then **Ineffective** must be marked in the Effectiveness column.*
3. *If you check **any sources of evidence other than No Evidence**, always make an effectiveness rating. The number of **Sources of Evidence** checked is **not** necessarily indicative of the effectiveness rating.*
4. *Circle the **1 to 5 Effectiveness Rating** to indicate how effectively the behavior was performed. **Outstandingly effective** means you (or your designee) has carried out a particular behavior (e.g., providing necessary support) with a very strong, positive effect on the targeted area of school activity (e.g., rigorous curriculum). **Ineffective** means you (or your designee) has either not done the particular behavior (e.g., not provided necessary support) or has carried out the behavior with very low quality that does not have a positive effect on the targeted area of school activity (e.g., rigorous curriculum).*

A set of four items and the response format for the VAL-ED is illustrated in Figure 2.1. The higher the score, the more effective a principal is in exhibiting the desired leadership behaviors. The end goal of the interpretation is to be able to make a reliable and valid attribution about a principal based on the input from multiple respondents who have observed and interacted with him/her.

Figure 2.1. Sample VAL-ED Items on Principal Form

Rigorous Curriculum (content)		Sources of Evidence						Effectiveness Rating				
		Check Key Sources of Evidence						Mark One Circle to Indicate How Effective				
		Reports from Others	Personal Observations	School Documents	School Projects or Activities	Other Sources	No Evidence	Ineffective	Minimally Effective	Satisfactorily Effective	Highly Effective	Outstandingly Effective
How effective am I at ensuring the school...												
Planning	13. develops a rigorous curriculum for all students.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	14. plans access to rigorous curricula for students with special needs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Implementing	15. creates rigorous sequences of learning experiences/courses.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	16. implements a rigorous curriculum in all classes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Respondents

The VAL-ED is designed to be completed by three types of respondents: *teachers*, *supervisors of principals*, and *principals themselves*. All individuals completing the instrument should know the principal and have worked with the principal in the same school for at least 2 months prior to the completion of the instrument. With regard to the teacher group, we encourage a broad definition so as to include faculty or staff who are actively engaged in an effort to improve the social and academic learning of students. This may include some individuals who typically are not considered teachers, but support teachers such as librarians, reading specialists, counselors, and teacher's aides. In some school districts, principals may have more than one supervisor. If so, all official supervisors are encouraged to participate in the evaluation of the principal. Finally, the principal being evaluated conducts a self-evaluation. When teachers, supervisors, and the principal, are included in rating the principal, what is commonly called a 360 assessment is achieved.

Teachers and supervisors complete the same version of the VAL-ED, which differs from the principal's version with regard to one aspect of evidence documentation. Specifically, teachers and supervisors can indicate they have "no evidence" upon which to base an effectiveness rating of a principal's behavior. Principals do not have the option of checking "no evidence." They simply rate themselves as "ineffective."

Desired Teacher Participation. The VAL-ED is designed to have all teachers in a school rate the effectiveness of a principal's behavior. The representativeness and size of the group of teachers participating in the assessment of a principal's leadership behaviors matters. The goal should be to have all teachers participate. We consider response rates of 75% or higher as high. When the response rate is below 50% of teachers, one must be concerned about the validity of

the resulting scores. Missing respondents from those eligible provokes a number of possible questions about the administration procedures and the relationships among respondents and a principal. Thus, it is important to take steps to ensure a high rate of participation.

Issues Influencing Response Rates. Three issues highly related to achieving a representative and large response rate to the VAL-ED are: (1) time and timing of the assessment, (2) guaranteeing anonymity of teachers' responses to the VAL-ED, and (3) appointing an evaluation coordinator who does not participate in the assessment and is seen as objective. The 72-item VAL-ED typically requires 20 to 25 minutes to complete once the directions have been read and the example items examined. Thus, a period of up to 45 minutes is recommended to ensure complete and thoughtful responses to all items from all participants. When using the paper version of the instrument, a good practice is large group administration, such as during a faculty meeting. If the respondents are completing the VAL-ED online or the paper administration is being conducted within an "assessment period" (e.g., 1 week), then the specified amount of time needed to complete the instrument may be less of a challenge.

Regardless of the amount of time to complete the VAL-ED or the timing in the school year when it is completed, the response rate from teachers is strongly affected by the guaranteed anonymity of responses. Such a guarantee begins with the appointment of an evaluation coordinator who provides an accurate statement about the purpose and use of the assessment, ensures there is no connection of respondents' names to response forms or individual score summaries, ensures the VAL-ED response forms or data files are secure from beginning to end of the assessment process, and limits the role of the principal in the administration and management process by allowing him/her to only provide positive encouragement to participate and possible evidence for consideration when completing effectiveness ratings. The principal should not be present when teachers are completing the VAL-ED. A principal's actions to encourage positive ratings or to suggest negative consequences for not completing a VAL-ED response form is unprofessional and likely will result in an invalid assessment.

Implementation Steps

Before assessing a school principal's leadership behavior using the VAL-ED it is important to (1) have a clear purpose for the assessment and (2) be sure that the leadership behaviors assessed by the VAL-ED are relevant. If these two conditions can be met, then the assessment of the principal using the VAL-ED can be implemented. To ensure the most valid assessment possible, we suggest the following 5 steps:

Step 1. Select the version of the VAL-ED to administer.

- A. Format: Decide whether to administer the Paper or Online version of the VAL-ED.
- B. Form: Forms A and Form C are parallel. Both forms measure the same Key Components and Processes using randomly equivalent sets of different items. The purpose of having two forms are a) the focus of interpretation should be on the sets of all possible behaviors of the intersection of a core component and a key process not the specific set of two behaviors on a particular form and b) parallel forms are useful in measuring change in leadership behavior over time. Teachers, Supervisors, and Principals must all use the same form for any one time of assessment. For re-assessment of the principal over time, for example as part of an evaluation of a professional improvement program, one form can be used for baseline

or pre-intervention measurement and the other form for post-intervention assessment. For an initial assessment, we recommend using Form A.

Step 2. Select a person to oversee the assessment process.

Individuals qualified to oversee the assessment process should be familiar with the VAL-ED and able to respond to any questions or technical problems that might arise during the assessment. The person(s) who oversees the process should not concurrently be assessing the principal.

Step 3. Decide *how* and *when* to conduct the assessment.

- A. Participants: Decide who should be asked to complete the VAL-ED.
- B. Timing: Decide when to conduct the assessment. Respondents need at least 2 months of interaction with a leader to have relevant and representative evidence for making valid behavior ratings. We recommend conducting the assessment in February or later in the school year.
- C. Duration: The assessment is designed to take no more than 20-25 minutes and ideally is completed in a single sitting.
- D. Time: Decide what time of day (before, after or during school) is best to schedule the administration. For the paper and pencil version, we recommend using a faculty meeting to ensure high response for teachers.

Step 4. Conduct the assessment.

- A. Discuss the importance of anonymity of ratings and assessment materials. Encourage respondents to work independently and not to discuss their ratings with others.
- B. Explain how to complete either the paper or online version of the VAL-ED as appropriate. Answer any questions that respondents might have at that time.
- C. If administering the paper version, be sure to collect and store all materials in a manner that respects anonymity and minimizes disruption. Do follow-ups as appropriate to ensure high response rates for each of the response groups. Follow-up can include reminders to complete the assessment and priority access to additional copies of the assessment when using the paper version.
- D. For the online version, provide respondents the VAL-ED website address, user name, and password for them to use to gain access to the appropriate assessment. Once online, each respondent will have support and a help directory.
- E. Document any irregularities you think might have influenced the assessment session.
- F. Set a deadline for completing the assessment.

Step 5. Conclude the assessment.

- A. Officially conclude the assessment by indicating the deadline for submitting assessments
has been reached and thank respondents for completing the VAL-ED.
- B. Organize materials/files for submission to Discovery Education Assessment (DEA) for scoring. Be sure to indicate who is to receive copies of the Principal's Report and store assessments in a secure location until they are shipped to Discovery.

- C. Indicate where copies of reports should be sent and who is to receive copies of the principal's final score report.
- D. Verify that all materials or data files have been received by DEA.

Behavior Ratings and Evidence for Ratings of Effectiveness of Leadership Behavior

As with any assessment tool, rating scales have limitations and should be used as part of a more comprehensive database to increase the likelihood that their resulting scores are reliable and valid. Perhaps an important limitation is that without tangible evidence about a principal's behavior, we do not know if self-reports and ratings by others are associated with what the principal actually does. Normative press and social desirability are also potentially problematic influences on the validity of self-report assessments. Most of the potential shortcomings of rating scales can be overcome with the use of evidenced-based, multi-rater systems that utilize objective scoring rubrics and repeated measures of performance. In addition, when items are phrased in behavioral terms and results reported in ways that lead to opportunities for improvement, research has found that rating scale methods can be highly efficient and valid measures of human behavior (Desimone, in press). Thus, the VAL-ED uses rating scale methodology and encourages respondents to recall, document, and evaluate evidence of leadership behaviors from their interactions with the principal. There is not a requirement to actually collect the evidence, although a respondent could do so or a principal could provide a portfolio of evidence he or she has developed as part of the evidential basis for effectiveness judgments. The actual evidence available to most respondents is substantial and takes the form of work products, observations, interviews, and school events. In summary, samples of tangible and reliable evidence documented by respondents play an important role in the assessment of a principal's leadership behavior. The VAL-ED assessment process is designed to be evidenced-based and honors the reports of respondents who have substantial interactions with the principal whom they are assessing.

Conclusions and Implementation Tips

The VAL-ED provides an evidence-based, multi-rater method for assessing learning-centered leadership behaviors of a school's principal. The VAL-ED is conceptually and theoretically grounded and its resulting scores, when implemented as described in this chapter, provide reliable and valid estimates of a principal's behaviors. The instrument privileges behaviors that influence the practices of teachers and staff, and in turn are related to increases in student achievement. The use of multiple respondents -- teachers, supervisors, and principals themselves -- that work in the same environment, effectively surrounding the principal, and the requirement for effectiveness ratings to be based on tangible observations or reports of behavior, both help to enhance the salience and the comprehensiveness of the resulting scores.

To achieve the VAL-ED's potential to inform principals and others interested in his/her leadership behaviors, it must be implemented with integrity. This means that the VAL-ED needs to be used for its intended purposes and be administered as designed. The following seven tips summarize key use and implementation guidance:

Tip #1. The teachers invited to complete the VAL-ED should result in a representative and reasonable large sample, ideally greater than 75% of those in the school. When all the appropriate supervisors are included along with the principal, the result is a 360 assessment.

Tip #2. The assessment should be coordinated by a person who is recognized as neutral or objective and not also completing the VAL-ED for the school's principal.

Tip #3. The assessment should only be undertaken, at the earliest, during the end of the second month of a school year. This increases the likelihood that respondents have had a reasonable opportunity to interact with the principal they are assessing. *We suggest waiting until February.*

Tip #4. Teachers must be guaranteed that their responses are anonymous.

Tip #5. Respondents must be given adequate time to read, reflect on evidence, and rate the behaviors of the principals. The instrument is designed to take 20 to 25 minutes, but 45 minutes should be allocated to the task to ensure that all respondents have time to complete the task.

Tip #6. The completed VAL-ED paper response forms must be secured in a manner to ensure confidentiality of responses and the objectivity of the assessment. Without confidentiality, the creditability of the results can be compromised.



Chapter 3

Interpreting and Using VAL-ED Results

When completed as designed by a representative and large portion of eligible respondents in a school, the VAL-ED can provide useful results for the purposes of evaluating the performance of a principal and also identifying leadership behaviors for improvement. The results from the VAL-ED address six fundamental questions. These are:

- ◆ *Who participated in the assessment of the principal's leadership behavior?*
- ◆ *What evidence did these respondents report using to make their effectiveness ratings of the principal?*
- ◆ *How effective is the principal's leadership behavior judged in comparison to a national sample of principals?*
- ◆ *How effective is the principal's leadership behavior judged in comparison to the VAL-ED proficiency standards?*
- ◆ *To what degree did the three respondent groups --teachers, supervisor, and principal-- agree with regard to these effectiveness ratings?*
- ◆ *Which areas of leadership behavior represent areas of relative strength and which represent areas for possible improvement?*

When used two or more times across years, the VAL-ED can also be used to address questions about behavior change and possible questions about the effects of programs designed to improve learning-centered leadership behaviors. In the remainder of this chapter, we focus on understanding and using the results from the VAL-ED. To accomplish these goals, however, several important technical aspects fundamental to the VAL-ED scores and scoring must be examined.

VAL-ED Scores

The VAL-ED behavior inventory provides information on a total score, six subscales for core components and six subscales for key processes separately for each respondent group and overall averaged across respondent groups. The core components and key processes are based on the same information, so while their information is redundant, the two separate profiles offer diagnostic information as to how a principal's behaviors might be improved leading to a more effective school and, in turn, improved student achievement.

Effectiveness Scores

The most fundamental score resulting from the VAL-ED is the Principal's Overall Total Effectiveness score. This score is based on the average ratings of all respondents where each respondent group is equally weighted and is reported in the 5-point effectiveness metric used to rate each of the 72 items on the instrument. Thus, the Principal's Overall Total Effectiveness

score and the Core Component and Key Process subscale scores are all reported on a continuous scale from a low of 1.0 (Ineffective) to a high of 5.0 (Outstandingly Effective). Table 3.1 provides data on the mean effectiveness scores (averaged across Forms A and C) for our national field trial sample of principals from nearly 300 schools. As you can observe from this data, the mean overall effectiveness rating for these principals was 3.61, which translated functionally to a performance significantly above "Satisfactorily Effective" but below "Highly Effective." The various subscales ratings of effectiveness for both Core Components and Key Process behaviors all fall within this same range of functioning.

Table 3.1 VAL-ED Mean and Median Effectiveness Scores

VAL-ED Scale or Subscale	Mean (SD)
Overall Effectiveness	3.61 (.35)
<i>Core Components</i>	
High Standards	3.68 (.37)
Rigorous Curriculum	3.58 (.37)
Quality Instruction	3.70 (.37)
Culture of Learning	3.76 (.39)
External Communities	3.45 (.39)
Performance Accountability	3.48 (.40)
<i>Key Processes</i>	
Planning	3.59 (.36)
Implementing	3.60 (.38)
Supporting	3.73 (.36)
Advocating	3.55 (.36)
Communicating	3.64 (.37)
Monitoring	3.56 (.39)

Note. 1.00 = Ineffective; 2.00 = Minimally Effective; 3.00 = Satisfactorily Effective; 4.00 = Highly Effective; 5.00 = Outstandingly Effective

Validity Indicators for VAL-ED Scores

Before interpreting scores from the VAL-ED, or any test, one must be confident that they are reliable and valid. In other words, for the VAL-ED scores to be meaningful and helpful to those using the instrument, they must have been proven to meet high standards for technical soundness and the instrument must have been completed as designed by a representative sample of respondents from a principal's school. A companion manual, the *VAL-ED Technical Manual*, provides comprehensive documentation and synthesis of the reliability and validity evidence of the VAL-ED prior to its publication. The VAL-ED website (www.valed.com) provides technical reports and research studies by the authors and other researchers that extend the empirical evidence supporting the interpretation of VAL-ED scores as reliable and validity indicators of the effectiveness of principals' leadership behaviors.

In addition to the research-base used to develop the VAL-ED, each Principal Report also provides individual principals and other users with information about the response rates from those persons invited to complete the instrument, the evidence they reported to use as the basis of their ratings, and the percentage of their "Don't Know" responses. As noted on the Principal

Report form, a response rate of greater than or equal to 75% is high, 50% to 74% is moderate, and below 50% is low. When response rates are low, the resulting scores should be interpreted with caution.

The VAL-ED instrument asks each respondent to think about the item describing a principal's behavior and before rating the principal on effectiveness as to that item, first identify what sources of evidence they have for making their effectiveness rating. Alternatives from which they can select are: (a) reports from others, (b) personal observations, (c) school documents, (d) school projects or activities, (e) other sources, and (f) no evidence. When "No evidence" was checked by a supervisor or teacher, the principal's effectiveness rating had to be "ineffective" or the respondent indicated "Don't know." Principals were not, however, given the option of "Don't know." If a principal checked "No evidence," then he/she was required to rate "ineffective" for that item. For the online version, the respondent is forced to conform to these rules. For the paper and pencil form, however, an item is dropped when a respondent does not conform to these rules. If a significant number of such violations of the item completion rules occur, the resulting scores become questionable. A principal's effectiveness then is the mean item responses for the reduced set of items. Similarly, when no effectiveness rating is given and instead "Don't know" is checked the item does not count toward the principal's effectiveness rating.

Table 3.2 reports as an average across the 72 items on both VAL-ED Forms the percent of times respondents indicated each type of source of evidence or no evidence. Since there were no restrictions on the number of sources of evidence that could be checked for any individual item, the percentages do not add to 100% across types of evidence within a respondent group. No evidence was checked 2.6% of the time by principals, 7.5% of the time by supervisors, and 10.3% of the time by teachers. These are acceptable percentages for the "No evidence" response. As indicated in Table 3.2, the most common source of evidence used by all respondents was personal observation. In contrast to teachers, however, supervisors and principals indicated school documents as a source of evidence nearly as frequently as personal observations, while teachers were less than half as likely to select school documents as they were personal observations. All of the sources of evidence options were selected frequently by each respondent group. Such distributed "portfolios" of evidence are generally seen as positive and likely to be revealing of behaviors in different settings or conditions, thus, adding to the generalizability of the ratings.

Table 3.2 Sources of Evidence Used by Respondent Group

	Reports from Others	Personal Observations	School Documents	School Projects or Activities	Other Sources	No Evidence
Principals	33.27%	61.91%	56.84%	36.19%	26.31%	2.60%
Supervisors	36.78%	58.01%	53.26%	27.18%	19.96%	7.52%
Teachers	24.24%	65.04%	29.27%	20.54%	10.04%	10.25%

Table 3.3 documents the percent of respondents checking “Don’t know” averaged across items for total score and each of the core components and key processes by respondents in our national field trial. The highest percentage of “Don’t knows” was checked for External Communities and Performance Accountability by both supervisors and teachers. Total response rates of "Don't know" that exceed 25% should elicit caution when interpreting a principal's scores.

Table 3.3 Percent of "Don't Know" Responses by Subscale

	Supervisors: Percent Don't Know	Teachers: Percent Don't Know
Total Score	9.31%	15.05%
<i>Core Components</i>		
High Standards	4.07%	7.99%
Rigorous Curriculum	7.64%	13.32%
Quality Instruction	7.64%	10.86%
Culture of Learning	5.94%	8.49%
External Community	18.96%	28.97%
Performance Accountability	11.58%	20.64%
<i>Key Processes</i>		
Planning	7.28%	14.41%
Implementing	9.48%	12.82%
Supporting	5.91%	8.37%
Advocating	11.58%	19.34%
Communicating	8.30%	12.16%
Monitoring	13.29%	23.16%

Approaches to Interpreting Assessment Results

Norm-Referenced Scores: Comparisons to National Sample of Principals

To facilitate interpretation of the total and subscale scores we translate total mean scores for each respondent's form of the VAL-ED to a percentile rank score based on our 2008 national field trial. These original norms are created separately for teacher responses, principal responses, supervisor responses, and for the combination of all three of these respondent groups on both Forms A and C of the VAL-ED. The resulting normative information allows for the comparison of a principal's behaviors on total score and subscale profiles to a national sample of principals who lead elementary, middle, and high schools representative of urban, suburban, and rural locales in all regions of the country. The norms for the VAL-ED will be updated periodically to ensure sensitivity to changes in leadership behavior and schools nationally.

Percentile ranks indicate the percentage of individuals in the norm group who scored at or below a given raw score. Percentiles range from 1 to 99. A percentile rank of 50 indicates that the individual's score is average in his or her group. Percentiles are widely used because they can be described in a simple manner and are understood by most educators. It is important, however,

to recognize the limitations of percentiles. Unlike standard scores, which are equal interval measures, percentiles represent a rank-order type of measure. Score units are unequal; thus, percentiles cannot be arithmetically manipulated like standard scores.

Table 3.4 provides a section of the norms table for the combined leadership behavior ratings on VAL-ED forms by our national field trial sample of principals, teachers, and supervisors. Many people refer to such a table as a Look-Up Norms table because once a raw score for both the total scale or for subscales is determined it can be looked up and translated to a percentile rank in the distribution of all scores from a group of respondents in a defined sample.

**Table 3.4 Section of Norms Table for VAL-ED Overall Effectiveness
Based on Combined 360 Ratings of Principals**

Overall Effective Score	Percentile Rank	Proficiency Level
4.49	98.9	Distinguished
4.11	89.9	Distinguished
3.95	79.9	Proficient
3.77	69.9	Proficient
3.67	59.9	Proficient
3.55	49.9	Basic
3.49	39.9	Basic
3.40	29.9	Basic
3.28	19.9	Below Basic
3.09	9.9	Below Basic
2.74	0.9	Below Basic

Examine Table 3.4 and observe that the Mean Total Effectiveness scores are provided from a high of 4.49 to a low of 2.74. This lowest score has a percentile rank of .9 while the highest score has a percentile rank of 98.9. The Mean Total Effectiveness score for Form A at the 50%tile is 3.55. This same basic approach to providing normative information is used for each of the Core Component subscale scores and the Key Process subscale scores. Given that all VAL-ED forms are computer scored by Discovery Education Assessment users will not actually look up their own scores.

You will also notice in Table 3.4 that for each Mean Total Effectiveness scores there is also a Proficiency Level identified with the lowest level called Below Basic and the highest level called Distinguished. These Proficiency Levels represent another interpretive approach for the VAL-ED and are discussed next.

Criterion-Referenced Scores: Comparison to Proficiency Standards with Cut-Scores

For our criterion-referenced interpretive framework, we defined proficiency levels and used a national panel of educational leaders to establish the scores (i.e., cut scores) that were needed to marginally meet each level. These proficiency cut scores were set on total score mean item response, using an equal-weight average across the three respondent groups. We used a modified Bookmark approach to setting proficiency standards (Cizek, 2001) for the following four proficiency levels:

- *Below Basic* – A leader at the below basic level of proficiency exhibits leadership behaviors of core components and key processes at levels of effectiveness that over time are unlikely to influence teachers to bring the school to a point that results in acceptable value-added to student achievement and social learning for students.
- *Basic* – A leader at the basic level of proficiency exhibits leadership behaviors of core components and key processes at levels of effectiveness that over time are likely to influence teachers to bring the school to a point that results in acceptable value-added to student achievement and social learning for some sub-groups of students, but not all.
- *Proficient* – A proficient leader exhibits leadership behaviors of core components and key processes at levels of effectiveness that over time are likely to influence teachers to bring the school to a point that results in acceptable value-added to student achievement and social learning for all students.
- *Distinguished* – A distinguished leader exhibits leadership behaviors of core components and key processes at levels of effectiveness that over time are virtually certain to influence teachers to bring the school to a point that results in strong value-added to student achievement and social learning for all students.

The three cut scores used to differentiate these four levels of leadership proficiency are: **3.29** between Basic and Below Basic; **3.60** between Basic and Proficient; and **4.00** between Proficient and Distinguished. The result of these cut scores is that principals who earn a mean item response score averaged across all respondent groups in the range of 1.0 to 3.28 will be described as behaving at the Below Basic level. Principals who earn a mean item response score averaged across all respondent groups in the range of 3.29 to 3.59 will be described as behaving at the Basic level. Principals who earn a mean item response score averaged across all respondent groups in the range of 3.60 to 3.99 will be described as behaving at the Proficient level. And finally, principals who earn a mean item response score averaged across all respondent groups in the range of 4.00 to 5.00 will be described as behaving at the Distinguished level. Based on our national field trial with 300 principals, these cut scores resulted in 17% of principals at the Below Basic level, 33% at the Basic level, 36% at the Proficient level, and 14% at the Distinguished level of proficiency.

The Principal Report

It is now time to apply all the information that one gains from completing the VAL-ED to advance a better understanding of the effectiveness of a principal's learning-centered leadership behaviors. As noted at the outset of this chapter, there are six fundamental questions about the assessment of a principal's behavior that can be meaningfully answered with the VAL-ED when it is completed by a representative group of teachers, the principal, and his/her supervisor(s). The questions are:

1. *Who participated in the assessment of the principal's leadership behavior?*
2. *What evidence did these respondents report using to make their effectiveness ratings of the principal?*

3. *How effective is the principal's leadership behavior judged in comparison to a national sample of principals?*
4. *How effective is the principal's leadership behavior judged in comparison to the VAL-ED proficiency standards?*
5. *To what degree did the three respondent groups --teachers, supervisor, and principal-- agree with regard to these effectiveness ratings?*
6. *Which areas of leadership behavior represent areas of relative strength and which represent areas for possible improvement?*

The answers to these questions are organized sequentially in the 9-page VAL-ED Principal Report (see Appendix A for a complete sample Principal Report). For purposes of connecting the content of the Principal Report with each of these six questions, we have excerpted sections from the report and highlight key explanatory information with callouts in Figures 3.1 through 3.6.

Figure 3.1. Respondent information from the Principal Report

	Possible Respondents	Actual Respondents	Percent (%) Responding
Principal	1	1	100 %
Teachers	22	22	100 %
Supervisor	1	1	100 %

Who participated in the assessment?

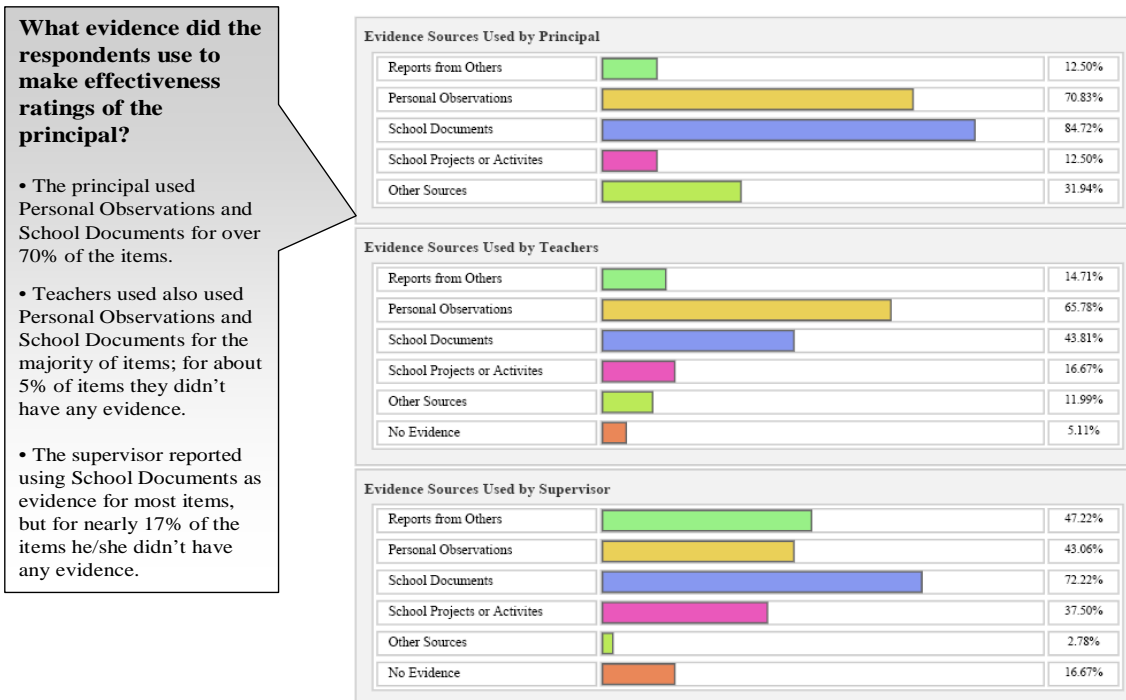
- 1 principal, 22 teachers, & 1 supervisor completed the VAL-ED.
- This results in an excellent response rate.

A response rate of greater than or equal to 75% is high, 50% to 74% is moderate, and below 50% is low. When response rates are low, resulting scores should be interpreted with caution.

Figure 3.1 is excerpted from page 2 of the Principal Report and provides a clear record of number of possible respondents for each respondent groups in a given school and then actual number and percentages of respondents who completed and submitted VAL-ED response forms. As indicated in the figure, a response rate of 75% or higher is consider high and thus desirable

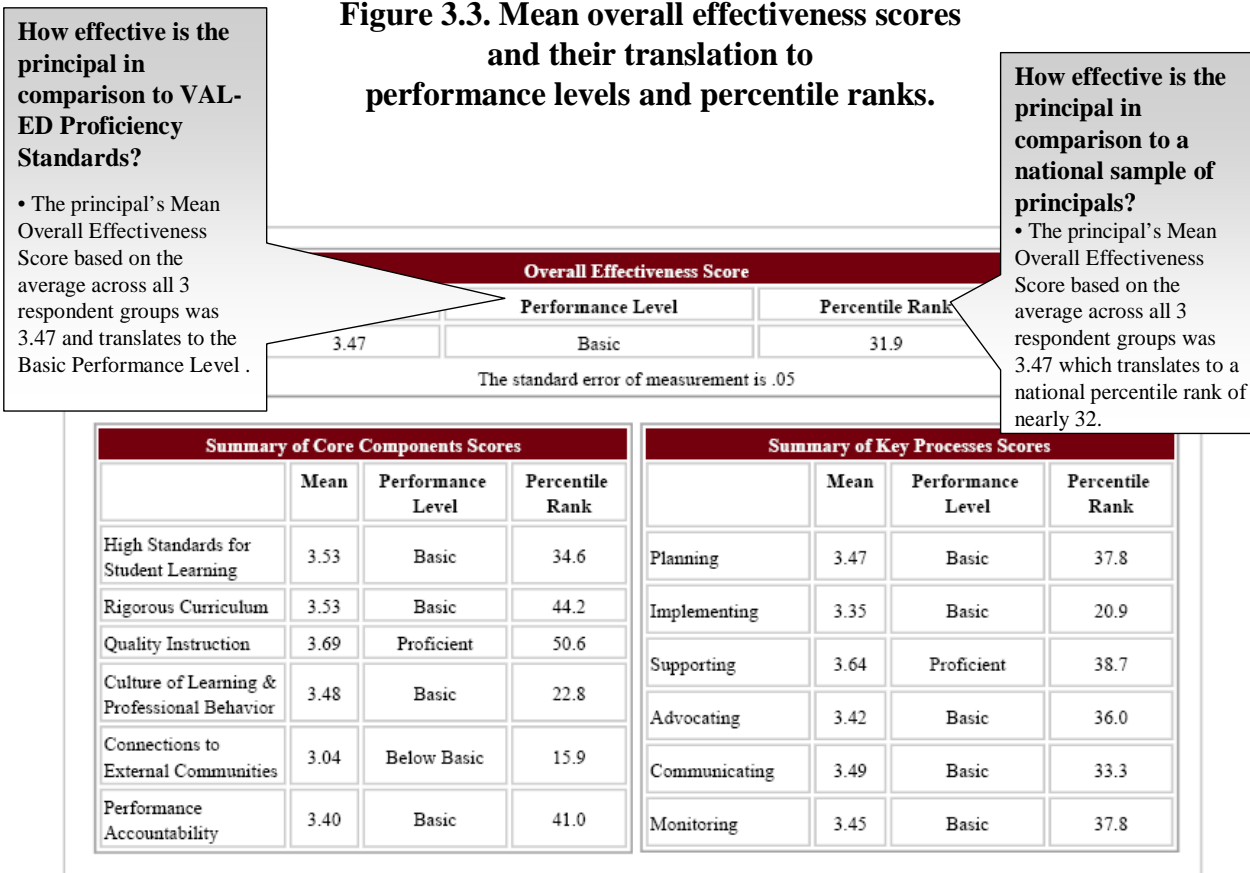
because it increases the likelihood that the resulting assessment data is representative of the respondents who interact with the principal.

Figure 3.2 Summary of evidence sources used by respondents for VAL-ED items.



Information about the evidence respondents reported using as the basis of their effectiveness ratings is portrayed in the graphic featured as Figure 3.2. This figure is an excerpt from page 2 of the Principal Report and provides data on the percentage of items for which each of several different types of evidence were used. Recall that multiple types of evidence can be used for a given item and also that Principals are not given the choice of "No Evidence." In the report illustrated in Figure 3.2, the respondents tended to use a good mixture of types of evidence. It should also be noted that the Supervisor reported on 12 (or 16.67%) items that he/she did not have any evidence to make a rating. This number of items without evidence for a particular respondent group is higher than the average established for the national field trial (see Table 3.2). With this information, the principal might, for future assessments, work to ensure evidence is available or more obvious for all respondent groups.

Figure 3.3. Mean overall effectiveness scores and their translation to performance levels and percentile ranks.



The most important quantitative results from the VAL-ED for most users are summarized on page 3 of the Principal Report. As highlighted in Figure 3.3 two key questions about a principal's leadership behaviors are answered with the test results aggregated across all three respondent groups. Specifically, the Overall Effectiveness Score and the related summary scores on the six Core Components and six Key Processes are provided as mean raw scores using the 1 (Ineffective) to 5 (Outstanding) Effectiveness scale. These mean effectiveness scores are then transformed to a criterion-referenced Performance Level (i.e., Below Basic, Basic, Proficient, or Distinguished) and to a norm-referenced Percentile Rank (i.e., 1st to 99th percentile). A reminder is also provided that the mean score, like any test score, is an observed score and is likely to have some error associated with it. The amount of error on the VAL-ED is very low and is best characterized by a range of $\pm .05$ around the mean score. So, for the principal for whom the present report was generated, her Mean Overall Effectiveness Score, if measured multiple times, is likely to be within the score band of 3.42 to 3.52.

Figure 3.4. Comparison of respondent groups' ratings of principal's behavior.

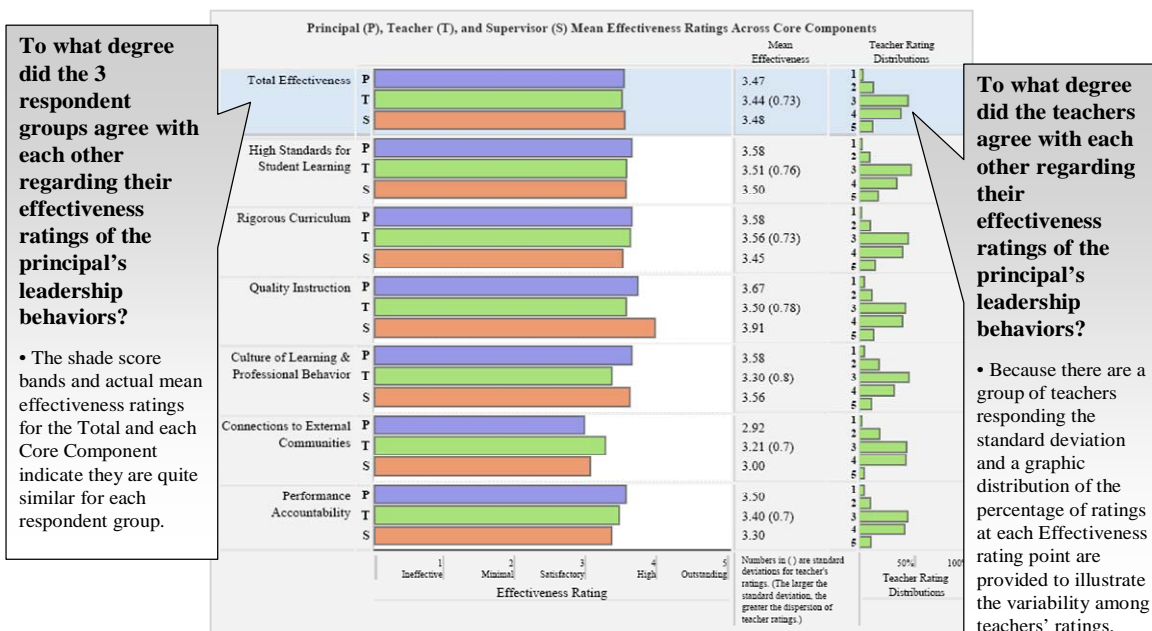


Figure 3.4 (from page 4 of the Principal Report) provides a comparison of each of the three respondent groups' ratings of the principal's behavior for each of the core components. This graphic information "unpacks" the overall mean ratings highlighted in Figure 3.3 and allows users to see how the different respondent groups evaluated the principal. A similar figure featuring a comparison of the respondent groups' ratings on the Key Processes (although not illustrated in this chapter) is provided on page 5 of the Principal Report. As highlighted by the callout on the right-hand side of the figure, detailed information about the distribution of teachers' effectiveness ratings is provided for the Total Effectiveness Scale and each of the Core Component Subscales.

Figure 3.5. Summary matrix identifying relative strengths and areas for improvement.

- Cells that are green represent areas of behavior that are 'proficient' or 'distinguished' (P).
- Cells that are yellow represent areas of behavior that are 'basic' (B).
- Cells that are red represent areas of behavior that are 'below basic' (BB).

Core Components	Key Processes					
	Planning	Implementing	Supporting	Advocating	Communicating	Monitoring
High Standards for Student Learning	B	B	P	BB	P	B
Rigorous Curriculum	B	BB	P	P	BB	B
Quality Instruction	B	P	P	B	P	P
Culture of Learning & Professional Behavior	P	B	B	P	B	BB
Connections to External Communities	BB	BB	BB	BB	BB	BB
Performance Accountability	B	BB	P	BB	B	B

Which areas of leadership behavior represent areas of relative strength and areas for possible improvement?

• This matrix provides an integrated summary of the principal's relative strengths and areas for possible growth.

• Cells with BB indicate skill areas in the greatest need of attention. These cells represent leadership behaviors where the mean item scores for the intersection of the Core Components and Key Processes across all three respondent groups fell in the Below Basic Performance Level.

For leadership improvement, perhaps the most useful outcome of the VAL-ED is information about their strengths and areas for improvement. Using the VAL-ED's Core Components and Key Processes Framework and the cut-scores based on the Performance Level descriptions, a matrix like that illustrated in Figure 3.5 is generated for each principal using the mean scores for the combined respondent groups. As indicated by the matrix above from page 6 of the Principal Report, cells labeled **P** indicate the leadership behaviors have Effectiveness ratings in a particular area that meet or exceed the Proficient level of performance. Examples of leadership behavior areas that represent this level of performance are: Supporting and Communicating High Standards for Student Learning and Implementing, Supporting, Communicating, and Monitoring Quality Instruction. Cells labeled **B** indicate behaviors at the Basic level of performance. In the matrix represented in Figure 3.5, there are 13 cells that are characterized as being at the Basic level of performance. Examples of leadership behavior areas that are at this level of performance included Implementing, Supporting, and Communicating about a Culture of Learning & Professional Behavior. Finally, cells labeled **BB** indicate areas where the effectiveness of the leadership behaviors is characterized as Below Basic. In the example portrayed, there are 12 cells that are so designated. All key processes associated with the Connections to External Communicates are examples of leadership behavior areas that meet this level of performance. Cells labeled BB are behavioral areas in need of the greatest improvement.

Figure 3.6. The lowest rated leadership behavior clusters for possible improvement.

Leadership Behaviors for Possible Improvement

The leadership behaviors listed in each cluster on the following pages are representative of the lowest rated core component by key process areas of behavior. If fewer than six core component by key process areas of behavior are listed, that is because the principal had fewer than six that were below distinguished. If no behavior clusters are provided it indicates the principal's current learning-centered leadership behaviors are considered acceptable.

The behaviors on each page that are **boldface** type are those that were actually assessed in the evaluation. The other behaviors represent the entire pool of VAL-ED behaviors for each core component by key process. All of these behaviors are relevant targets for improvement.

For a list of all the leadership behaviors associated with each core component area, consult the VAL-ED Users' Guide.

Connections to External Communities X Advocating
<ul style="list-style-type: none"> • Advocates for social services needed by students and families. • Advocates for students in need of special services with the external community. • Challenges teachers to work with community agencies to support students with low achievement. • Challenges barriers from outside the school that can inhibit learning • Advocates to district decision makers to promote the needs of all students. • Promotes mechanisms for reaching families who are least comfortable at school.

The final excerpt from the VAL-ED Principal Report is presented in Figure 3.6. This figure is from the section of the report titled "Leadership Behaviors for Possible Improvement." In this section, nearly all principals receive a list of the six lowest rated Core Component-Key Process clusters with six items from the pool of VAL-ED leadership behaviors. If a principal receives fewer than six clusters it is because he/she had fewer than six clusters below Distinguished. Boldfaced items in each cluster are those that were actually rated for the principal; the remaining four items represent other examples of leadership behaviors representative of the same cluster. The figure above illustrates that the Connections to External Communities X Advocating behavior cluster was one of the lowest rated for the selected principal. We know from Figure 3.5 that this cell cluster was labeled BB for Below Basic. Based on this finding and the boldfaced items in the cluster, it is suggested that the principal target leadership behaviors in this cell.

With the completion of this examination of the Principal Report, we trust you now have a clear understanding of the comprehensive and detailed information provided by the VAL-ED. Individuals who want to improve their learning-centered leadership behavior must first accurately assess it and then have a method for meaningfully interpreting the results. The VAL-ED accomplishes these assessment and communication goals and lays the foundation for an improvement plan!

Conclusions and Usage Tips

When used as intended with a robust sample that have interacted with a principal for several months, the VAL-ED yields reliable and valid scores about key learning-focused leadership behaviors and provides user-friendly interpretations of these scores. As you use and

communicate the results of a VAL-ED assessment of a principal, we encourage you to keep the following tips in mind:

Tip #1. The VAL-ED measures leadership behaviors, not attitude or personal characteristics.

Tip #2. Scores based on respondents that represent fewer than 50% of those invited to assess the principal may be unrepresentative and should be used with caution.

Tip #3. Scores based on respondents who report high rates of No Evidence or Don't Know responses may be unrepresentative and should be used with caution.

Tip #4. The results of the VAL-ED provide technically sound scores, but they should be used along with other information when making important personnel decisions or they should be repeated to confirm original conclusions.



Chapter 4

Summary of the Technical Soundness and Utility of the VAL-ED

The VAL-ED is a 360 degree assessment of a principal's leadership behaviors. It facilitates key people surrounding the principal (i.e., teachers, the principal's supervisor, and the principal him or herself) to rate important learning-centered leadership behaviors of the principal. The VAL-ED measures *core components* and *key processes*. Core components refer to characteristics of schools that support the learning of students and enhance the ability of teachers to teach. Key processes refer to how leaders create and manage those core components. Effective learning-centered leadership is at the intersection of the two dimensions: core components created through key processes. The outcomes of the assessment are profiles, interpretable from both norm-referenced and standards-referenced perspectives, and suggested clusters of behaviors for improvement.

A series of psychometric studies has been completed to support our claims that the VAL-ED (a) works well in a variety of settings and circumstances, (b) is unbiased, (c) yields reliable and valid scores, (d) is feasible for widespread use (both on-line and paper-and-pencil versions), (e) provides accurate and useful reports of assessment results for summative purposes, and also (f) yields a diagnostic profile for formative purposes. The details of the VAL-ED development process and the studies that provide the support for validity and usage claims are provided in the *VAL-ED Technical Manual*. In the remainder of this chapter, we provide information on the technical aspects of rating scales in general and also a summary of the particular evidence we have established that supports the use of the VAL-ED for the purposes of assessing principals' leadership behaviors.

Reliability and Validity Evidence Summary

The conceptual framework that drives the development, reporting, and interpretation of the VAL-ED results is six core components by six key processes. The core components are features of effective schools. The key processes are leadership behaviors that principals can employ to lead their schools to a point in time when they are strong on each of the core components. Thus, for each core component, there are six key processes of leadership behaviors. The six-by-six matrix identifies 36 cells with two items in each cell. Forms A and C were created by randomly selected two items from a set of items written for each of the 36 cells.

Results from the VAL-ED are reported in terms of mean item effectiveness on a 1-5 effectiveness rating scale. Percentile ranks are based on a national field trial and performance standards were set by a 22-member panel of experts. There is an overall total effectiveness score, which is a function of the responses to all 72 items across the supervisor, the principal, and the teachers, where supervisor, principal, and teachers are weighted equally. Results are also reported separately for each core component and each key process. Finally, the total score as well as subscale scores for the six core components and the six key processes are reported separately for each respondent group.

The first and most important validity evidence for the content of the VAL-ED is that the items were written against the six-by-six conceptual framework. The conceptual framework is based on the literature on school leadership effects on student achievement. The development of the VAL-ED was embedded in a research paradigm. First, items were written to fit each of the 36 cells in the core components (6) by key processes (6) conceptual framework. As many items were written as leadership behaviors could be identified. For some cells more than 10 behaviors were identified and items written.

A sample of principals were asked to sort items into cells as a second check on the content validity of the items; sorting accuracy was good but some items were dropped and some re-written for clarity of target cell. Two rounds of cognitive interviews were conducted in three districts each. In each case, principals, teachers, and principals' supervisors from elementary, middle, and high schools participated in the cognitive interviews. Based on the results, the instrument was revised and tried again. When the instrument was judged to be ready, a nine-school pilot test was conducted in one district involving elementary, middle, and high schools. Based on the pilot, the instrument was seen to have good internal consistency reliability, good construct validity, good face validity, but the instrument was too long and the effectiveness scale was not being used across its full range. The instrument was revised to be shorter and have different benchmarks for the effectiveness scale. Cognitive interviews were conducted on the online instrument (the paper and pencil and online instruments are virtually identical). The positive results led to piloting the instrument once again, this time in eleven schools, again across elementary, middle, and high school. The results for this second pilot were encouraging. The reliability remained high. More of the range of the response scale was used. Completion time was seen as less of a problem and confirmatory factor analysis on the teacher data supported the conceptual framework against which the items were written.

A fairness review of the VAL-ED instructions and items was conducted to identify and remove aspects of test items or directions that might hinder respondents from completing the instrument. The fairness review was based on the fairness guidelines published and used by ETS. A panel of nine individuals completed the fairness review. Eight worked in public schools as either a teacher behavior specialist or administrator. The results indicated no fairness concerns in instructions or introductory content. Thirteen items on Form A and fourteen items on Form C were identified as requiring some edits to improve their fairness characteristics and these edits were made.

With Form A and Form C of the VAL-ED assessment of school leadership in final form, a national field trial was undertaken to establish the psychometric properties of the assessment, to establish percentile ranks for reporting, to build an item-ordered booklet for the Bookmark performance standard setting, to further investigate the perceived feasibility of the instrument, and to investigate design effects, including level of schooling, locale, and the parallelness of Forms A and C. The obtained sample had principal data on 235 schools, supervisor data on 253, and teacher data on 245. For 218 schools, there was data from all three respondent groups. Based on the national field trial, supervisors were seen as slightly more positive on principal effectiveness than were principals with teachers in between. Controlling for other factors, high school principals were seen as slightly less effective than elementary or middle school principals, but the difference was small. Suburban school principals were seen as more effective than rural school principals and no significant differences were found between Forms A and C.

Confirmatory factor analysis, exploratory factor analysis with oblique factor rotations, and investigations of mean differences among core components, key processes, and their

interactions were conducted to investigate the construct validity of the instrument. Confirmatory factor analysis supported both core components and key processes. Exploratory factor analysis identified factors for Performance Accountability, Connections to External Community, and Culture of Learning and Professional Behavior, as well as the key processes of Supporting and Advocating. There were significant differences between the means of the core components, with the exception that Rigorous Curriculum and Performance Accountability were not significantly different one the other, nor were Quality of Instruction and Culture of Learning and Professional Behavior. For key processes, Supporting was significantly different from all other key processes, but the other key processes were not significantly different one from another. Of the 630 pairs of contrasts among the 36 cells, for Form A 44% were significant and for Form C, 47%. Finally, generalizability theory was used to investigate the reliability of contrasts between core components and between key processes. The reliabilities of the differences were surprisingly good given the notoriously low reliability of difference scores. The results were similar to the results from the exploratory factor analysis. The reliabilities contrasting Culture of Learning, External Communities, Performance Accountability, and Rigorous Curriculum were all strong. For key processes, the reliabilities contrasting Supporting and Advocating were both strong.

The relationship between responses from principals, supervisors, and teachers was investigated. All three respondent groups were positively intercorrelated with the highest intercorrelation between principals and teachers (approximately .25), followed by the correlation between supervisors and teachers and the lowest correlation between principals and supervisors. Clearly, the information from one respondent group was not redundant with the information from the other respondent groups, a finding that supports the 360 degree approach to assessment.

Investigations of the use of sources of evidence, the “Don’t know” option on the effectiveness scale, and a variety of possible errors that could be made in filling out the instrument revealed no problems.

Nine questions were asked of respondents after they completed the assessment. All three respondent groups indicated that they found the items (a) focusing on important leadership behaviors, (b) understandable, (c) not biased, (d) appropriate for elementary, middle, and high school levels, and (e) the instrument was easy to use. Similarly, all respondent groups agreed that teachers should have input into the assessment of principal leadership, but they neither agreed nor disagreed that the VAL-ED should be used to hold principals accountable in their district. Perhaps had the item asked about formative as well as summative evaluations, the responses would have been more enthusiastic about use of the instrument.

A Bookmark method was used to set performance standards for distinguished, proficient, basic, and below basic. A national panel of 22 experts participated. Ultimately, the standards were set yielding 17% of the national field trial principals below basic, 50% below proficient, and 86% below distinguished.

With the completion of the national field trial, the VAL-ED assessment of school leadership has been documented to have excellent reliability, strong validity, initial national norms for reporting percentile ranks, and performance standards to identify distinguished, proficient, basic, and below basic principals. The norms and the proficiency levels apply to both Form A and C, which can be used interchangeably and for each form, a paper and pencil as well as an online version of the assessment.

Conclusions about the Technical Aspects and Utility of the VAL-ED

Many test developers use a set of basic questions about the technical soundness of assessment to guide their research and to organize their conclusions about the resulting test. These questions for the VAL-ED are as follows:

The *Content Question*. How well does the sample of items represent the domain of behaviors to be measured?

The *Consistency Question*. How consistent are the results of an assessment taken at different points in time or on alternate forms of the assessment?

The *Test-Criterion Relationship Question*. How well do principals' performances on the assessment predict future performances or estimate current performances on some valued measure of skills other than the test itself?

The *Construct Question*. How well can consumers interpret performance on the assessment as a meaningful measure of the behaviors the assessment is intended to measure?

The *Change Question*. How well does the assessment measure change in measured behaviors?

The *Consequences Question*. How well does the use of the VAL-ED accomplish the intended purposes of the assessment and avoid unintended effects?

With regard to validity question 1, "how well does the sample of items represent the domain of behaviors to be measured?" the strongest evidence comes from the procedures followed to build the instrument. Additional evidence came from results from the sorting study and the cognitive labs. The instrument development phase was specifically focused on writing an exhaustive set of appropriately-grained, non-redundant items to fit into the 36-cell conceptual framework. If no other validity work had been done, instrument development would have created a set of items that represented the domain of possible items well. A sorting study was conducted, and results revealed that principals were able to accurately place items correctly into the exact cell with great frequency. Where principals were not able to sort properly, improvements to item wording were made to ensure item fit. Finally, respondents to the cognitive labs indicated that no important items were missing from the instruments. These three pieces of evidence provide support for the claim that the items on the VAL-ED represent the domain of important principal leadership behaviors well.

With regard to the second question, "How consistent are the results of an assessment taken at different points in time or on alternate forms of the assessment?" The evidence from the VAL-ED national field trial indicated that Forms A and C were vitually equivalent and forms exhibited high levels of internal consistency. Thus, this evidence indicates the resulting scores can be highly reliable.

With regard to the third question, "How well do principal's performances on the assessment estimate current performances on some valued measure of skills (other than the test itself)?" evidence comes from high correlations among the effectiveness ratings across respondent groups. More research is underway to provide evidence for the predictive relationship between VAL-ED scores and other important performances concerning teachers' and students' behaviors.

With regard to the fourth question, “How well can consumers interpret performance on the assessment as a meaningful measure of the behaviors the assessment purports to measure,” Results from the analyses of the performance level descriptors provided evidence that respondents understood the PLDs, and that they appropriately differentiated among levels of effectiveness. However, suggestions were made to improve the PLDs, and the PLDs were changed to improve clarity. Additional work to address construct validity is ongoing.

With regard to the fifth question, “How well does the assessment measure change or achievement progress on content that is valued by educators?,” additional evidence as to the VAL-ED’s ability to measure change is to be gathered in future work.

Finally, with regard to the sixth question “How well does the use of the VAL-ED results accomplish the intended purposes of the assessment and avoid unintended effects?” evidence suggests that the use of the VAL-ED accomplishes the intended purposes of the assessment. Cognitive lab results indicate that respondents can successfully navigate the online interface of the VAL-ED. The study of scale use suggests that respondents are using the full effectiveness scale after the modifications that have been made. A planned consequences study will provide further evidence to address this question.

As described in the VAL-ED Technical Manual, future studies will be conducted by the authors as a result of a grant from the Institute of Education Sciences (IES), as well as by independent researchers interested in learning-centered school leadership. As additional evidence becomes available, this technical manual will be updated with new evidence so stay tuned and connect to www.valed.com for more evidence about the valid use of VAL-ED. Still, the results to date support that the instrument yields scores that reliably and validly measure principal leadership effectiveness and can be used for the purposes of determining principals’ norm-referenced and standard-referenced standing among fellow professionals, as well as providing them guidance regarding areas for improvement.



References

- Adams, J. E., & Kirst, M. W. (1999). New demands and concepts for educational accountability: Striving for results in an era of accountability. In J. Murphy & K. Louis (Eds.), *Handbook of research on educational administration, 2nd edition* (pp. 463-489). San Francisco: Jossey-Bass.
- American Educational Research Association (1999). *Standards for educational and psychological testing*. Washington, D.C.: Author.
- Betts, J. R., & Grogger, J. (2003). The impact of grading standards on student achievement, educational attainment, and entry-level earnings. *Economics of Education Review, 22*, 343-352.
- Boyer, E.L. (1983). *High school: A report on secondary education in America*. New York: Harper & Row.
- Brookover, W. B., & Lezotte, L. W. (1977). *Changes in school characteristics coincident with changes in student achievement*. East Lansing: College of Urban Development, Michigan State University
- Bryk, A. S., & Driscoll, M. E. (1988). *The high school as community: Contextual influences and consequences for students and teachers*. Madison, WI: University of Wisconsin-Madison, National Center on Effective Secondary Schools.
- Bryk, A., Camburn, E., & Louis, K. S. (1999). Professional community in Chicago elementary schools. Facilitating factors and organizational consequences. *Educational Administration Quarterly, 35*, 751-781.
- Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage.
- Burns, J. M. (1978). *Leadership*. New York: Harper & Row.
- Butty, J., LaPoint, V., Thomas, V., & Thompson, D. (2001). The changing face of after school programs: Advocating talent development for urban middle and high school students. *NASSP Bulletin, 58*(262), 22-34.
- Conley, D. T. (1991). Lessons from laboratories in school restructuring and site-based decision making. *Oregon School Study Council Bulletin, 34*(7), 1-61.

- Conley, D. T., & Goldman, P. (1994). Ten propositions for facilitative leadership. In J. Murphy & K. S. Louis (Eds.), *Reshaping the principalship: Insights from transformational reform efforts*. Thousand Oaks, CA: Corwin Press.
- Council of Chief State School Officers (CCSSO). (1996). *Interstate School Leaders Licensure Consortium Standards for School Leaders*. Washington, DC: Author.
- Downing, S.M., & Haladyna, T.M. (Eds.) (2006). *Handbook of test development*. Mahwah, NJ: LEA.
- Edmonds, R., & Frederiksen, J. R. (1978). *Search for effective schools: The identification and analysis of city schools that are instructionally effective for poor children*. Cambridge, MA: Harvard University, Center for Urban Studies.
- Elmore, R. F. (2005). Accountable leadership. *The Educational Forum*, 69, 134-142.
- Eubanks, E. E., & Levine, D. U. (1983, June). A first look at effective schools projects in New York City and Milwaukee. *Phi Delta Kappan*, 64(10), 697-702.
- Fullan, M., & Pomfret, A. (1977). Research on curriculum and instructional implementation. *Review of Educational Research*, 47, 335-397.
- Garibaldi, A. M. (1993). *Improving urban schools in inner-city communities* (Occasional Paper No. 3). Cleveland, OH: Cleveland State University, Levine College of Urban Affairs, Urban Child Research Center.
- Goldring, E., Cravens, X., Porter, A., Murphy, J., & Elliott, S. (2007). *The State of Educational Leadership Evaluation: What do States and Districts Assess?*
- Goldring, E., & Hausman, C. (2001). Civic capacity and school principals: The missing link in community development. In R. Crowson & B. Boyd (Eds.), *Community development and school reform*. Greenwich, CT: JAI Press.
- Goldring, E., Porter, A., Murphy, J., Elliott, S. N., & Cravens, X. (2007). *Assessing Learning-Centered Leadership: Connections to Research, Professional Standards, and Current Practices*. New York, NY: Wallace Foundation.
- Haladyna, T.M., Downing, S.M., & Rodriguez, M.C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15 (3), 309-334.
- Hallinger, P., & Heck, R. H. (1996). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly*, 32(1), 5-44.

- Hallinger, P., & Heck, R. (2002). What do you call people with visions? The role of vision, mission, and goals in school improvement. In K. Leithwood, P. Hallinger, G. Furman, J. MacBeath, B. Mulford, & K. Riley (Eds.), *The second international handbook of educational leadership and administration*. Dordrecht, The Netherlands: Kluwer.
- Hallinger, P., & Murphy, J. (1987). Instructional leadership in the school context. In W. Greenfield (Ed.), *Instructional leadership: concepts, issues, and controversies*. Boston: Allyn & Bacon.
- Heck, R. (1992). Principals' instructional leadership and school performance: Implications for policy development. *Educational Evaluation and Policy Analysis*, 14(1), 21-34.
- Heck, R. H., & Hallinger, P. (1999). Next generation methods in the study of leadership and school improvement. In J. Murphy, & K. S. Louis (Eds.). Handbook of Research on Educational Administration (2nd ed., pp. 141-162). San Francisco: Jossey-Bass.
- Henderson, A. T., & Mapp, K. L. (2002). *A new wave of evidence: The impact of school, family, and community connections on student achievement*. Austin, TX: Southwest Educational Development laboratory.
- Knapp, M. S., Copland, M. A., & Talbert, J. F. (2003, February). Leading for learning: Reflective tools for school and district leaders. Seattle: Center for the Study of Teaching and Policy, University of Washington.
- Lee, V. E., Smith, J. B., & Croninger, R. G. (1995). *Another look at high school restructuring. More evidence that it improves student achievement and more insights into why*. Madison, WI: Center on Organization and Restructuring of Schools.
- Leithwood, K. (1994). Leadership for school restructuring. *Educational Administration Quarterly*, 30, 498-518.
- Leithwood, K. (2001). School leadership in the context of accountability policies. *International Journal of Leadership In Education*, 4(3), 217-236.
- Leithwood, K., & Jantzi, D. (1990). *Transformational leadership: How principals can help reform school cultures*. Paper presented at the annual meeting of the Canadian Association for Curriculum Studies, Victoria, B.C.
- Leithwood, K., & Jantzi, D. (2000). The effects of transformational leadership on organizational conditions and student engagement. *Journal of Educational Administration*, 38(2), 112-129.
- Leithwood, K., & Jantzi, D. (2005). A review of transformational school leadership research, 1996-2005. *Leadership and Policy in Schools*, 4, 177-199.

- Leithwood, K., & Montgomery, D. J. (1982). The role of the elementary school principal in program improvement. *Review of Educational Research*, 52(3), 309-339.
- Leithwood, K. A., & Montgomery, D. J. (1986). *Improving principal effectiveness : the principal profile*. Toronto, Ontario: Ontario Institute for Studies in Education.
- Little, J. W. (1982, Fall). Norm of collegiality and experimentation: Work-place conditions of school success. *American Educational Research Journal*, 19(3), 325-340.
- Louis, K. S., Marks, H., & Kruse, S. (1996). Teachers' professional community in restructuring schools. *American Educational Research Journal*, 33(4), 757-798.
- Louis, K. S., & Miles, M. B. (1990). *Improving the urban high school: What works and why*. New York: Teachers College Press.
- Marks, H., & Printy, S. (2003). Principal leadership and school performance: An integration of transformational and instructional leadership. *Educational Administration Quarterly*, 39, 370-397.
- Marzano, R. J., Waters, T., & McNulty, B. A. (2005). *School leadership that works: From research to results*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Murphy, J. (2005). Unpacking the foundations of ISLLC Standards and addressing concerns in the academic community. *Educational Administration Quarterly*, 41, 154-191.
- Murphy, J., Goldring, E., Elliott, S. N., & Porter, A. (2006). *Learning-centered leadership: a conceptual foundation* Wallace Foundation
- Murphy, J. & Meyers, C. V. (2008). *Turning around failing schools: Leadership lessons from the organizational sciences*. Thousand Oaks, CA: Corwin.
- National Research Council. (1999). In M. S. Donovan, J. D. Bransford, & J. W. Pellegrino (Eds.), *How people learn: Bridging research and practice*. Washington, DC: National Academy Press.
- Newmann, F. M. (1997). How secondary schools contribute to academic success. In K. Borman & B. Schneider (Eds.), *Youth experiences and development: Social influences and educational challenges*. Berkeley, CA: McCutchan.
- Ogden, E. H., & Germinario, V. (1995). *The nation's best schools: Blueprints for excellence*. Lancaster, PA: Technomic.
- Porter, A.C., Goldring, E., Murphy, J., Elliott, S. N., & Cravens, X. (2006). *A Conceptual Framework for the Assessment of Principal and Team School Leadership*: Wallace Foundation

- Porter, A.C., Murphy, J., Goldring, E., Elliott, S.N., Polikoff, M.S., & May, H. (2008). *Vanderbilt Assessment of Leadership in Education (VAL-ED): Technical manual*. Nashville, TN: Discovery Education Assessment.
- Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *The Elementary School Journal*, 83(4), 426-452.
- Rosenholtz, S. J. (1989). *Teachers' workplace: The social organization of schools*. New York: Longman.
- Roueche, J. E., Baker, G. A., Mullin, P. L., & Boy, N. H. O. (1986). *Profiling excellence in America's schools*. Arlington, VA: American Association of School Administrators.
- Sebring, P., & Bryk, A. (2000). School leadership and the bottom line in Chicago. *Phi Delta Kappan*, 81, 440-443.
- Shaver, A. V., & Walls, R. T. (1998). Effect of Title I parent involvement on student reading and mathematics achievement. *Journal of Research and Development in Education*, 31(2), 90-97.
- Teddlie, C., & Stringfield, S. (1993). *Schools do make a difference: Lessons learned from a 10-year study of school effects*. New York: Teachers College Press.
- Teddlie, C., Stringfield, S., Wimpelberg, R., & Kirby, P. (1989). Contextual differences in models for effective schooling in the United States. In B. Creemers, T. Peters, & D. Reynolds (Eds.), *School effectiveness and school improvement: Selected proceedings from the Second Inferential Congress* (pp 117-130). Amsterdam: Swets and Zeitlinger.
- Wellisch, J. B., MacQueen, A. H., Carriere, R. A., & Duck, G. A. (1978, July). School management and organization in successful schools. *Sociology of Education*, 51, 211-226.
- Westat and Policy Studies Associates. (2001). *The longitudinal evaluation of change and performance in Title I schools*. Washington, DC: U.S. Department of Education, Office of the Deputy Secretary, Planning and Evaluation Service.

List of Tables

(linked for convenience)

Table 3.1	VAL-ED Mean and Median Effectiveness Scores.	p. 20
Table 3.2	Sources of Evidence Used by Respondent Group.	p. 21
Table 3.3	Percent of "Don't Know" Responses by Subscale.	p. 22
Table 3.4	Section of Norms Table for VAL-ED Overall Effectiveness based on Combined 360 Ratings of Principals.	p. 23

List of Figures

(linked for convenience)

Figure 1.1.	Concept Framework for the Vanderbilt Assessment of Leadership in Education.	p. 7
Figure 1.2.	The VAL-ED Constructs of Core Components and Key Processes.	p. 8
Figure 2.1.	Sample VAL-ED Items of Principal Form.	p. 14
Figure 3.1.	Respondent Information from the Principal Report.	p. 25
Figure 3.2.	Summary of Evidence Sources Used by Respondents for VAL-ED Items.	p. 26
Figure 3.3.	Mean Overall Effectiveness Scores and Their Translation to Performance Levels and Percentile Rank.	p. 27
Figure 3.4.	Comparison for Respondent Groups' Ratings of Principal's Behavior.	p. 28
Figure 3.5.	Summary Matrix Identifying Relative Strengths and Areas for Improvement.	p. 29
Figure 3.6.	The Lowest Rated Leadership Behavior Clusters for Possible Improvement.	p. 30

Appendix A
Sample VAL-ED Principal Report



VANDERBILT ASSESSMENT of LEADERSHIP in EDUCATION™

Principal Report

Survey ID:	1141551	Date of Report:	December 08, 2008
School District:	Field Test District	Date of Evaluation:	N/A
School:	Field Test School	VAL-ED Form:	A

Purpose of the Assessment

The Vanderbilt Assessment of Leadership in Education or VAL-ED is designed to provide a summary of effectiveness of a principal's learning-centered leadership behaviors during the current school year. A comprehensive picture of the principal has emerged and is reported with input from teachers, the principal's supervisor and his or her own self-report.

The VAL-ED focuses on leadership behaviors defined by six core components and six key processes known to influence student achievement:

Core Components

- High Standards for Student Learning
- Rigorous Curriculum
- Quality Instruction
- Culture of Learning & Professional Behavior
- Connections to External Communities
- Performance Accountability

Key Processes

- Planning
- Implementing
- Supporting
- Advocating
- Communicating
- Monitoring

Respondents to the VAL-ED were asked: How effective the principal is at ensuring the school carries out specific actions that affect core components of learning-centered leadership. The effectiveness ratings, based on evidence, range from 1 (ineffective) to 5 (outstandingly effective) for each of 72 leadership behaviors.

This VAL-ED report addresses the questions of:

- (1) who responded?
- (2) what evidence was used to evaluate the principal?
- (3) what do the results say about the principal's current leadership behaviors?

The results are interpreted against both norm-referenced and standards-referenced criteria that highlight areas of strength and possible areas for improvement. A leadership development plan can be developed based on these results.

The VAL-ED provides technically sound scores when used as designed, however, it is recommended that it be used along with other information when making important evaluative decisions.

For more information about the VAL-ED, please visit our website: <http://www.valed.com>.

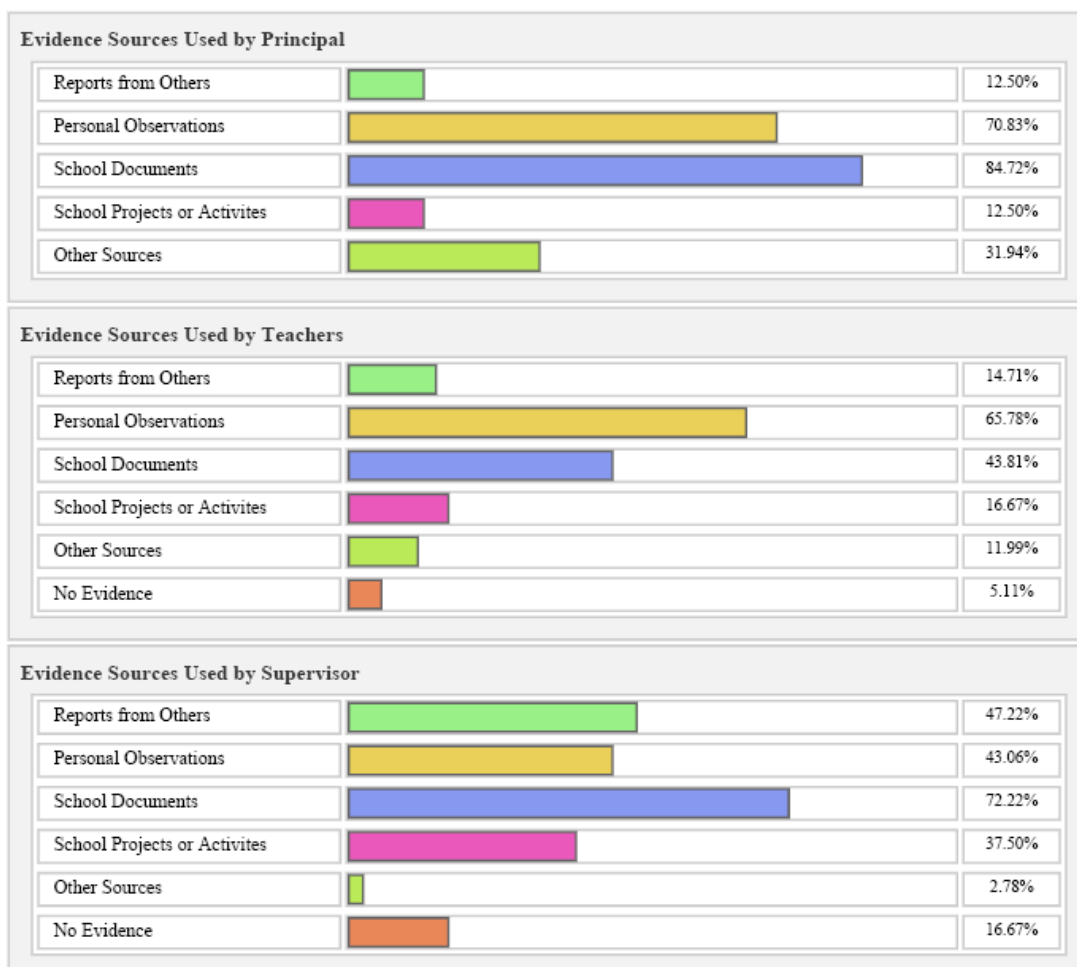
Who Responded and What Evidence Did They Use?

	Possible Respondents	Actual Respondents	Percent (%) Responding
Principal	1	1	100 %
Teachers	22	22	100 %
Supervisor	1	1	100 %

A response rate of greater than or equal to 75% is high, 50% to 74% is moderate, and below 50% is low. When response rates are low, resulting scores should be interpreted with caution.

Sources of Evidence

Ratings of a principal's behaviors should be based on evidence that is recent, relevant and representative. Evidence comes in many forms (e.g., observations of behavior, review of documents that record leadership actions and communications with people who have directly observed the principal's behavior). After reflecting on a sample of evidence, respondents effectiveness ratings of leadership behaviors are behaviorally-anchored and more accurate. The graphs below summarize each type of evidence used as a basis for effectiveness ratings of the leadership behaviors. The bars display the sources of evidence for each item used by the principal, all teacher and supervisor respondents in the school. Percentages are based on number of items for which a source of evidence was checked; these percentages need not sum to 100 across sources.



What are the Results of the Assessment?

VAL-ED provides a total score across all respondents as well as separately by respondent group. The scores from the teachers are based on the average across all teacher respondents. The total score, core component, and key process effectiveness ratings are interpreted against a national representative sample that included principals, supervisors, and teachers, providing a percentile rank. The results are also interpreted against a set of performance standards ranging from **Below Basic** to **Distinguished**. The scores associated with performance levels were determined by a national panel of principals, supervisors and teachers.

Below Basic	Basic	Proficient	Distinguished
A leader at the <u>below basic</u> level of proficiency exhibits learning-centered leadership behaviors at levels of effectiveness that are unlikely to influence teachers positively nor result in acceptable value-added to student achievement and social learning for students.	A leader at the <u>basic</u> level of proficiency exhibits learning-centered leadership behaviors at levels of effectiveness that are likely to influence teachers positively and that result in acceptable value-added to student achievement and social learning for some sub-groups of students, but not all.	A <u>proficient</u> leader exhibits learning-centered leadership behaviors at levels of effectiveness that are likely to influence teachers positively and result in acceptable value-added to student achievement and social learning for all students.	A <u>distinguished</u> leader exhibits learning-centered leadership behaviors at levels of effectiveness that are virtually certain to influence teachers positively and result in strong value-added to student achievement and social learning for all students.

Overview of Assessment Results

The Principal's Overall Total Effectiveness score based on the averaged ratings of all respondents is 3.47. Remember, this score is based on a 5-point effectiveness scale where 1=Ineffective; 2=Minimally Effective; 3=Satisfactorily Effective; 4=Highly Effective; 5=Outstandingly Effective. The Performance Level and national Percentile Rank for this score are documented in the table below.

Overall Effectiveness Score			
Mean Score	Performance Level	Percentile Rank	
3.47	Basic	31.9	
The standard error of measurement is .05			

Summary of Core Components Scores				Summary of Key Processes Scores			
	Mean	Performance Level	Percentile Rank		Mean	Performance Level	Percentile Rank
High Standards for Student Learning	3.53	Basic	34.6	Planning	3.47	Basic	37.8
Rigorous Curriculum	3.53	Basic	44.2	Implementing	3.35	Basic	20.9
Quality Instruction	3.69	Proficient	50.6	Supporting	3.64	Proficient	38.7
Culture of Learning & Professional Behavior	3.48	Basic	22.8	Advocating	3.42	Basic	36.0
Connections to External Communities	3.04	Below Basic	15.9	Communicating	3.49	Basic	33.3
Performance Accountability	3.40	Basic	41.0	Monitoring	3.45	Basic	37.8

An examination of the principal's Core Components mean item ratings ranged from a low of 3.04 for Connections to External Communities to a high of 3.69 for Quality Instruction. Similarly the principal's Key Processes mean item ratings indicate they ranged from a low of 3.35 for Implementing to a high of 3.64 for Supporting.

Assessment Profile and Respondent Comparisons

The principal's relative strengths and areas for development can be determined by comparing scores for each of the 6 Core Components and 6 Key Processes across different respondent groups. The next two graphs present an integrated visual summary of the results. They show the **Mean Effectiveness** associated with each Core Component and Key Process.

First, examine the profiles as recorded by each of the three respondent groups. These scores can be interpreted by

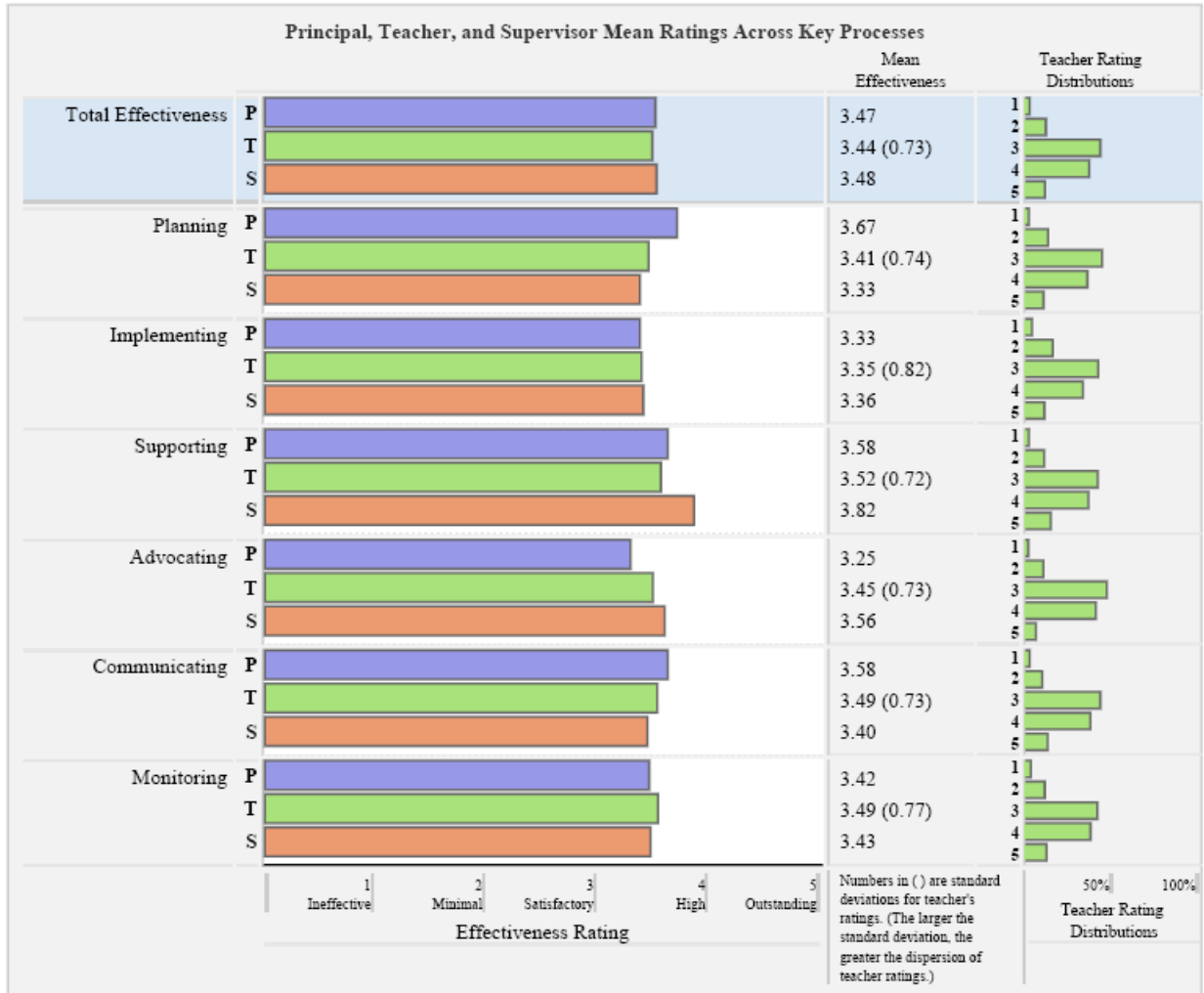
- (a) Comparisons among Core Components and Key Processes
- (b) Examination of scores among respondent groups
- (c) Comparisons to the mean effectiveness scale
- (d) Distribution of ratings among teachers



For each of the six Core Components in the graph, examine the effectiveness ratings. The ratings for a core component are based on twelve items. The higher the ratings, the more effective the leadership behaviors of the principal. When there are large differences between respondent groups, the focus should be on the results for each respondent group rather than the overall effectiveness score.

Assessment Profile and Respondent Comparisons (Cont'd.)

The ratings of the six Key Processes are based on 12 items that focus on a given Key Process. Again, the higher the score, the more effective the leadership behaviors of the principal. For more details about the technical aspects of the VAL-ED scores and tips on interpreting scores, visit the VAL-ED website <http://www.valed.com>.



Using Results to Plan for Professional Growth

The matrix below provides an integrated summary of the principal's relative strengths and areas for growth based on the mean item scores for the intersection of Core Components by Key Processes across the three respondent groups.

- Cells that are green represent areas of behavior that are 'proficient' or 'distinguished' (P).
- Cells that are yellow represent areas of behavior that are 'basic' (B).
- Cells that are red represent areas of behavior that are 'below basic' (BB).

Core Components	Key Processes					
	Planning	Implementing	Supporting	Advocating	Communicating	Monitoring
High Standards for Student Learning	B	B	P	BB	P	B
Rigorous Curriculum	B	BB	P	P	BB	B
Quality Instruction	B	P	P	B	P	P
Culture of Learning & Professional Behavior	P	B	B	P	B	BB
Connections to External Communities	BB	BB	BB	BB	BB	BB
Performance Accountability	B	BB	P	BB	B	B

Leadership Behaviors for Possible Improvement

The leadership behaviors listed in each cluster on the following pages are representative of the lowest rated core component by key process areas of behavior. If fewer than six core component by key process areas of behavior are listed, that is because the principal had fewer than six that were below distinguished. If no behavior clusters are provided it indicates the principal's current learning-centered leadership behaviors are considered acceptable.

The behaviors on each page that are **boldface** type are those that were actually assessed in the evaluation. The other behaviors represent the entire pool of VAL-ED behaviors for each core component by key process. All of these behaviors are relevant targets for improvement.

For a list of all the leadership behaviors associated with each core component area, consult the VAL-ED Users' Guide.

Leadership Behaviors for Possible Improvement

Connections to External Communities X Advocating

- Advocates for social services needed by students and families.
- Advocates for students in need of special services with the external community.
- **Challenges teachers to work with community agencies to support students with low achievement.**
- Challenges barriers from outside the school that can inhibit learning
- Advocates to district decision makers to promote the needs of all students.
- **Promotes mechanisms for reaching families who are least comfortable at school.**

Culture of Learning & Professional Behavior X Monitoring

- Evaluates teachers behaviors when monitoring the culture of learning.
- Evaluates students behaviors when monitoring the learning environment.
- Monitors the school culture.
- **Monitors the participation of every student in social and academic activities.**
- **Assesses the culture of the school from students perspectives.**
- Analyzes data regarding a safe and orderly school environment.

Performance Accountability X Implementing

- Implements programs and practices to hold faculty accountable to reach the highest levels of performance.
- Implements programs and practices that hold the school accountable to families for the learning of their children.
- Builds behavioral and academic accountability measures with input from faculty.
- Implements practices that hold leaders accountable for student learning.
- **Implements social and academic accountability equitably for all students.**
- **Uses faculty input to create methods to hold faculty accountable.**

Leadership Behaviors for Possible Improvement

Connections to External Communities X Implementing

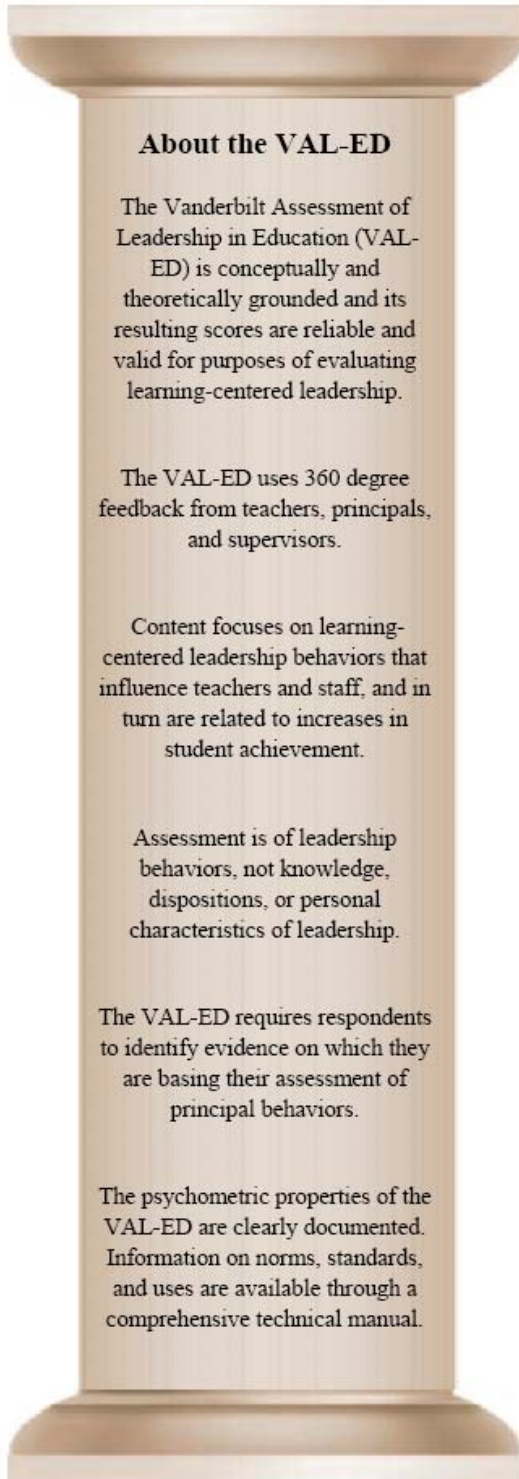
- **Builds business partnerships to support social and academic learning.**
- Implements programs to involve families in the educational mission.
- **Implements programs to help address community needs.**
- Builds a positive, open relationship with the community.
- Coordinates access to social service agencies to support students.
- Implements programs to help parents assist their children to be successful in school.

Connections to External Communities X Supporting

- Supports teachers to work with community agencies on behalf of students.
- **Secures additional resources through partnering with external agencies to enhance teaching and learning.**
- Secures technology from the district and/or the community to enhance teaching and learning.
- Secures resources to support school-community relationships.
- **Allocates resources that build family and community partnerships to advance student learning.**
- Motivates teachers to be responsive to all families.

Connections to External Communities X Planning

- Plans family education programs consistent with instructional goals.
- Plans for the use of external community resources to promote academic and social learning goals.
- **Develops a plan for community outreach programs consistent with instructional goals.**
- Plans activities with volunteers to advance social and academic goals.
- Plans activities to engage families in student learning.
- **Develops a plan for school/community relations that revolves around the academic mission.**



About the VAL-ED

The Vanderbilt Assessment of Leadership in Education (VAL-ED) is conceptually and theoretically grounded and its resulting scores are reliable and valid for purposes of evaluating learning-centered leadership.

The VAL-ED uses 360 degree feedback from teachers, principals, and supervisors.

Content focuses on learning-centered leadership behaviors that influence teachers and staff, and in turn are related to increases in student achievement.

Assessment is of leadership behaviors, not knowledge, dispositions, or personal characteristics of leadership.

The VAL-ED requires respondents to identify evidence on which they are basing their assessment of principal behaviors.

The psychometric properties of the VAL-ED are clearly documented. Information on norms, standards, and uses are available through a comprehensive technical manual.



"Leadership is a central ingredient - often the keystone element in school and district success as defined in terms of student achievement."

- **Joseph Murphy**
Vanderbilt University

"Assessments that provide ongoing performance feedback to school leaders about their learning-centered leadership behaviors can substantially help school leaders develop effective leadership for school improvement."

- **Ellen Goldring**
Vanderbilt University

Visit

<http://www.valed.com>

For more information and periodic updates on research and related articles on the use of VAL-ED

VAL-ED Authors

Andrew Porter, Joseph Murphy,
Ellen Goldring, & Stephen N. Elliott