2017

Virginia Principals' Knowledge of Classroom Assessment and Support of Assessment for Learning Practices

Rachel Previs Ball

College of William and Mary, rachelprevisball@gmail.com

Follow this and additional works at: https://scholarworks.wm.edu/etd

Part of the Educational Leadership Commons

Recommended Citation


http://dx.doi.org/doi:10.21220/W4R36G

This Dissertation is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
VIRGINIA PRINCIPALS’ KNOWLEDGE OF CLASSROOM ASSESSMENT
AND SUPPORT OF ASSESSMENT FOR LEARNING PRACTICES

A Dissertation

Presented to

The Faculty of the School of Education

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

by

Rachel Previs Ball

July 2017
VIRGINIA PRINCIPALS’ KNOWLEDGE OF CLASSROOM ASSESSMENT
AND SUPPORT OF ASSESSMENT FOR LEARNING PRACTICES

by

Rachel Previs Ball

Approved July 2017 by

Leslie W. Grant, Ph.D.
Chairperson of Doctoral Committee

Michael F. DiPaola, Ed.D.

Christopher R. Gareis, Ph.D.
DEDICATION

I dedicate this dissertation to my family. To my parents, thank you for teaching me the meaning of grit. Through my life, I have watched as you modeled the values of hard work and perseverance. Any success I accomplish is attributable to your sacrifices and love. And a special thank you to my husband and best friend, Carter, for all of your love and support. It is without a doubt that you deserve a spot on this paper with your name on it. As proud as I am of this milestone, my greatest accomplishments are my faith and family. I am truly blessed.
# TABLE OF CONTENTS

Acknowledgments ........................................................................................................ix

List of Tables .................................................................................................................xi

List of Figures ................................................................................................................xiii

Abstract .........................................................................................................................xiv

Chapter 1: The Problem .................................................................................................2

   Conceptual Framework ..............................................................................................7
   Knowing What to Teach and How To Assess ..............................................................12
   How We Use Assessment as Instruction and Involve Students in the Process ..........12
   How We Monitor Our Practices .................................................................................12
   How We Communicate About Student Learning .......................................................13
   Application of Conceptual Framework ....................................................................13

Research Questions ......................................................................................................13

Significance of the Study ...............................................................................................14

Definition of Terms .......................................................................................................17

Organization of the Remainder of the Study .................................................................20

Chapter 2: Review of the Literature ..............................................................................21

   Virginia Accountability .............................................................................................21
   Virginia Standards of Learning ...............................................................................22
   Accreditation Ratings ...............................................................................................22
       Full Accreditation ................................................................................................23
       Partial Accreditation .........................................................................................23
History of Formative Assessment ................................................................. 58
Impact of Formative Assessment on Student Achievement ...................... 61
Balanced Assessment Approach .................................................................. 62
Summary ........................................................................................................ 63

Chapter 3: Methodology .............................................................................. 64
Research Design .......................................................................................... 65
  Research Questions .................................................................................... 67
  Participants .................................................................................................. 67
  Instrumentation ......................................................................................... 69
    Survey ...................................................................................................... 69
    Interview .................................................................................................. 74
  Procedure .................................................................................................. 77
    Survey ...................................................................................................... 77
    Interview .................................................................................................. 77
  Analysis ..................................................................................................... 78
    Survey ...................................................................................................... 78
    Interview .................................................................................................. 80
  Ethical Considerations .............................................................................. 81
  Delimitations ............................................................................................. 82
  Limitations ................................................................................................ 82
  Assumptions .............................................................................................. 83
  Summary .................................................................................................... 84

Chapter 4: Results ....................................................................................... 85
Unethical Practices a Nonissue .......................................................... 130
Training on Appropriate Administration of SOLS ....................... 130
SOL Practices to Minimize Testing Irregularities ....................... 132
Common Practices to Ensure Valid Results ................................. 133
Interpretation of Assessment Results ........................................... 135

Chapter 5: Discussion ........................................................................... 138
Summary and Discussion of Findings .............................................. 139
Assessment Literacy Across Levels ............................................... 144
Types of Training .............................................................................. 145
Support of Assessment for Learning ............................................. 146
Alignment ......................................................................................... 147
Professional Development ............................................................ 148
Balanced Assessment ..................................................................... 149
Ethical Considerations ................................................................. 150
Limitations ...................................................................................... 151
Recommendations for Practice ..................................................... 153
Directions for Further Research .................................................... 156

Appendix A: Survey Part I ................................................................. 159
Appendix B: Assessment Literacy Inventory .................................. 161
Appendix C: Interview Protocol ..................................................... 177
Appendix D: Researcher as Instrument Statement ....................... 181
Appendix D: Researcher Participation Informed Consent Form ....... 186
References ....................................................................................... 189
Vita
ACKNOWLEDGEMENTS

As a primary school principal, I am greeted by eager faces each morning. My students’ smiles and welcoming hugs are a constant reminder of why I became an educator. My students constantly challenge me to make an impact. Their quest for knowledge has paralleled my own as I continually seek ways to make their educational experiences powerful and meaningful.

I have been fortunate to have been supported through my educational journey through impactful teachers and professors of my own. I have had the opportunity to learn from Dr. Leslie Grant through my undergraduate program, master’s program and now doctoral coursework, and she has graciously chaired my dissertation committee as I completed this final step in my journey. She constantly provides meaningful and valuable feedback, something students such as me crave and value.

I have additionally had the opportunity to learn from so many impactful professors during my time at William and Mary. I have valued their experiences as practitioners and researchers. My former advisor and now State Superintendent, Dr. Steven Staples, helped me map out my course trajectory as I began my graduate coursework. I am grateful for professors such as my advisor, Dr. Mike DiPaola, who reminded me the value of common sense and prioritizing my role as an instructional leader. Dr. Chris Gareis inspired me to continue my quest to seek reform in curriculum, instruction and assessment to meet the needs of our students. I have admired Dr. James Stronge for his knowledge and innovation in the field. And finally, I have benefited from the leadership and mentorship of Dr. Jan Rozzelle through my participation in the SURN Principal Academy. In my own school division, I have been grateful for the support of
my former superintendent and mentor, Dr. Mark Jones. I was fortunate to learn from his
servant leadership during his time in King William County Public Schools.

And finally, this study would not have been possible without the gracious
participation of my study participants. The principalship is a demanding profession, and I
am grateful for participants’ gift of time.
LIST OF TABLES

2. Data Collection and Data Analysis by Research Question ........................................67
3. Alignment of Standards with Items on the ALI and to Chappuis (2004) Framework ....74
4. Interview Questions ........................................................................................................76
5. Participants’ Years of Classroom Teaching Experience .............................................87
6. Participants’ Years of Administrative Experience .......................................................87
7. Accreditation Status and Principals’ ALI Scores .........................................................88
8. Participants’ Mean Score on ALI ................................................................................88
9. Range and Frequency of ALI Scores ..........................................................................89
10. Principals’ Mean Scores by Standard ..........................................................................90
11. Percentage Correct by Question for Standard One: Choosing Assessment Methods ...91
12. Percentage Correct by Question for Standard Two: Developing Assessment Methods .92
13. Percentage Correct by Question for Standard Three: Administering, Scoring and Interpreting Results ........................................................................92
14. Percentage Correct by Question for Standard Four: Using Assessment Results When Making Decisions ...........................................................................93
15. Percentage Correct by Question for Standard Five: Developing Valid Pupil Grading Procedures .................................................................................94
16. Percentage Correct by Question for Standard Six: Communicating Assessment Results ........................................................................................................94
17. Percentage Correct by Question for Standard Seven: Ethical Practices ....................95
18. School Level Designations for Participants Who Denoted Other for Level Assignment

19. ALI Scores by School Level

20. Independent Samples t-test for School Level and ALI Score

21. Primary Method of Assessment Training and ALI Score

22. Analysis of Variance

23. Contextual Background for Interview Participants

24. Assessment Leadership Competencies and Corresponding Codes and Themes

25. Assessment Leadership Codes and Themes

26. Support of Assessment for Learning Principles

27. Alignment

28. Professional Development Themes

29. Balanced Assessment Themes

30. Ethical Considerations Themes
LIST OF FIGURES


   Assessment Leadership Competencies ..............................................................................51
The purpose of this study was to investigate the assessment literacy of Virginia principals and describe how principals with varying levels of assessment literacy integrate assessment leadership practices that support assessment for learning. This study investigated the differences in assessment literacy between elementary and secondary principals and across principals’ predominant method of training in assessment. Mertler and Campbell’s (2005) Assessment Literacy Inventory (ALI) was used to obtain measures of overall assessment literacy and determine relative strengths and weaknesses across the seven Standards for Teacher Competence in the Educational Assessment of Students. There were no significant differences in assessment literacy across levels or as a result of type of training in assessment. Participants scored highest in their ability to recognize unethical practices and their relative weakness was in developing assessment methods. Qualitative interviews were conducted with six principals with higher levels of assessment literacy and six principals with lower levels. Interviews were analyzed for assessment leadership practices related to: support of assessment for learning principals, alignment, professional development, balanced assessment, and ethical practices. Discrepancies between principals with higher and lower levels of assessment literacy were in the areas of professional development on learning targets and the alignment of instruction to learning targets. Principals with across levels of assessment literacy described using professional learning communities and instructional specialists to support grouping for instruction. Additionally, principals
described balanced assessment systems with multiple measures of formative and summative assessments. There were commonalities in ethical practices and considerations across principals.

RACHEL PREVIS BALL
PROGRAM IN EDUCATIONAL POLICY, PLANNING, AND LEADERSHIP
THE COLLEGE OF WILLIAM AND MARY IN VIRGINIA
VIRGINIA PRINCIPALS’ KNOWLEDGE OF CLASSROOM ASSESSMENT AND SUPPORT OF ASSESSMENT FOR LEARNING PRACTICES
CHAPTER 1: THE PROBLEM

Schools throughout the nation strive to meet elusive state achievement targets, and Virginia schools are not immune to these accountability woes. Standing out as one of the few states in the nation that resisted the adoption of the Common Core Standards, Virginia schools base achievement and performance according to their own curriculum and assessments derived from the Virginia Standards of Learning (SOLs).

Staggering levels of schools and districts failed to meet state accreditation benchmarks since the revision and subsequent adoption of more rigorous Standards of Learning assessments within the past decade, with only 68% of schools earning full accreditation status in 2014-2015 (Virginia Department of Education, 2014). Revised math standards were adopted in 2009, and reading ensued the following year. Substantial declines in pass rates followed and subsequently, school accreditation ratings. Declining scores adversely impacted Virginia schools’ accreditation ratings, as acknowledged by then State Superintendent, Dr. Patricia Wright. Wright addressed the impact of more rigorous standards in a news release, "Even with three-year averaging mitigating the impact of the new tests, we will see some schools slip from Fully Accredited to Accredited with Warning” (Virginia Department of Education, 2013, para. 16).

As predicted, changes to standards and assessments resulted in a decline in accreditation, plummeting to a mere 68% of schools meeting the classification of Fully Accredited according to the 2014-2015 state accreditation ratings, and only 22 of 132
school divisions achieved full accreditation in all their schools (Virginia Department of Education, 2014). Virginia schools faced a 25% decline in schools earning full accreditation since 2012-2013, during which 93% of schools were fully accredited (Virginia Department of Education, 2014).

Despite gains in school accreditation, the 2016-2017 accreditation ratings revealed that an alarming 19% of Virginia schools had still not met the requirements for full accreditation (Virginia Department of Education, 2016). Virginia schools’ failure to meet the requirements for full accreditation further highlights the need for instructional leadership to increase student achievement on demanding standardized assessments in the Commonwealth.

As accountability demands place increasingly greater emphasis on student outcomes and achievement in public schools, the role of the principal as instructional leader continues to have a prominent role in the literature (Hallinger, 2005). The principalship is a multifaceted role; however, the accountability movement and high stakes testing prioritize the need for administrators with a strong capacity for instructional leadership. Assessment leadership, as facet of instructional leadership, is necessary for enhancing student-learning outcomes because of the integral nature of instruction and assessment and the implications for student learning (Stiggins & Duke, 2008). Because of the impact of high-stakes testing, “there is a need to consider the role of assessment leadership as an expectation of contemporary instructional leaders” (Noonan & Renihan, 2006, p. 7). The role of the principal as assessment leader is highlighted by Stiggins and Duke (2008):
As a response to the amount of time spent assessing students and using assessment results, principals must ensure teachers employ sound classroom assessment practices. The typical teacher will spend a quarter to a third of her or his available professional time involved in assessment-related activities. If they do it well, both teachers and students gain access to evidence that can be used in making sound instructional decisions. If they do it poorly, learning will suffer. In spite of this, little of principals’ preparation time is spent learning about assessments. (p. 286)

Because of the time involved in student assessment and the opportunity for improved student achievement and outcomes, principals as assessment leaders must ensure that teachers are equipped with the knowledge and skills to be able to implement quality assessments for learning, but first, they must be assessment literate themselves.

Instructional leaders must be knowledgeable of sound assessment practices in order to address changes in schools and oversee the growth and development of these sound practices in teachers. The principal, as instructional leader, plays a significant role in the school improvement process by “focusing on learning, encouraging collaboration, using data to improve learning, providing support, and aligning curriculum, assessment, and instruction” (Lunenburg, 2013, p. 37). Student learning is directly influenced by curriculum, instructional practices, and assessment, and instructional leaders play an integral role in supporting teachers and ultimately improving student outcomes through this process of support (Glickman, 2002). Stiggins (1991) highlighted the need for assessment literate leaders in education:
No longer is it sufficient for purposes of ‘accreditation’ that educators simply build the proper facilities, buy the right textbooks, maintain the proper student/teacher ratios, and have enough books in the library. Today, such process variables are to be used in the right combinations to produce the desired product: measured student learning. (p. 534)

In order to be assessment leaders, Stiggins (1991) asserted that educators must develop their own assessment literacy. In order to do so, they must understand what comprises a quality assessment and recognize appropriate inferences that can be drawn from various assessments. Additionally, assessment literate educators anticipate positive or negative impacts of assessments and data. Assessment literate educators also recognize the impact of extraneous variables and sampling when interpreting assessment results (Stiggins, 2001, p. 535). Just as the principalship is a complex role, so is the role of assessment leader.

As the instructional leaders of a building, principals are responsible for developing their own requisite levels of assessment literacy and providing assistance to teachers in furthering their own levels of assessment literacy. McMillan (2003) recommended that teachers receive assistance as they find the appropriate “balance” between formative and summative assessment as well as the implications of each (p. 41).

“Deep understanding about assessment, for teachers, involves reflection, application, reasoning, and problem solving, just as it does for students. Effective assessment decision making involves the complexities involved with self-awareness of how their interpretations and judgments influence the assessment process” (McMillan, 2003, p. 39).
Principals must ensure teachers continue to develop and hone their own assessment literacy if they are to maximize the instructional impact of student assessment.

Principals share a role in developing teachers’ capacity to employ appropriate assessment practices within the classroom that support student learning. Teacher preparation programs might include some assessment literacy preparation; however, not all teachers benefited from or participated in this type of formal training; therefore, professional development is necessary (Popham, 2009). Assessment literacy is needed for teachers in order to maximize the instructional utility of assessment:

Thus, it seems that assessment literacy is a commodity needed by teachers for their own long-term well-being, and for the educational well-being of their students. For the foreseeable future, teachers are likely to exist in an environment where test-elicited evidence plays a prominent instructional and evaluative role. (Popham, 2009, p. 11)

Because of the significant role assessment plays in informing classroom instruction, the principal shares responsibility in ensuring that teachers’ have requisite levels of assessment literacy.

The need for assessment for learning in classrooms is of pivotal importance because of its positive impact on student learning. “There is a body of firm evidence that supports formative assessment as an essential component of classroom work and that its development can raise standards of achievement” (Black & Wiliam, 2010, p. 90). Formative assessment practices yielded higher gains than a host of educational interventions, with promising effect sizes ranging from 0.4 to 0.7. The need for schools and classroom teachers to employ assessment for learning strategies is critical, and it is
the responsibility of principals to ensure that these practices are employed with fidelity in order to raise student achievement.

As a result of the need for assessment leadership to address the challenging Virginia standards and assessments and improve student achievement and accreditation ratings across schools in the Commonwealth, this study examined principals’ knowledge of classroom assessment practices as well as the ways in which principals support assessment for learning practices within their settings. It is essential to first gauge principals’ levels of assessment literacy in a post-No Child Left Behind context. This study examined principals’ support of assessment for learning strategies. The link between assessment for learning strategies and student achievement highlights the need for strong assessment leadership that supports quality use and integration of student assessments.

**Conceptual Framework**

To further understand the role of assessment leadership, Chappuis, Stiggins, Arter, and Chappuis (2004) presented 10 competencies that reflect leadership skills and understandings that support assessment for learning. These competencies encompass Black and Wiliam’s (1998) assessment for learning research as well as Stiggins’ (2002) recommendations for a balanced assessment approach (as cited in Chappuis et al., 2004). These competencies involve the creation of quality assessments that are implemented within the classroom to facilitate and inform instruction. When addressed with fidelity, these competencies provide a framework for leaders to improve student achievement. Below are the 10 competencies that comprise this framework:
1. The leader understands the standards of quality for student assessment and how to ensure that these standards are met in all assessments.

2. The leader understands the principles of assessment for learning and works with staff to integrate them into classroom instruction.

3. The leader understands the necessity of clear academic achievement targets, aligned classroom-level achievement targets, and their relationship to the development of accurate assessments.

4. The leader knows and can evaluate the teacher’s classroom assessment competencies and helps teachers learn to assess accurately and use the results productively.

5. The leader can plan, present, and/or secure professional development activities that contribute to the use of sound assessment practices.

6. The leader accurately analyzes student assessment information, uses the information to improve curriculum and instruction, and assists teachers in doing the same.

7. The leader develops and implements sound assessment and assessment-related policies.

8. The leader creates the conditions necessary for the appropriate use and reporting of student achievement information, and can communicate effectively with all members of the school community about student assessment results and their relationship to improving curriculum and instruction.

9. The leader understands the attributes of a sound and balanced assessment system.
10. The leader understands the issues related to the unethical and inappropriate use of student assessment and protects students and staff from such misuse. (Chappuis et al., 2004, p. 125)

These 10 competencies have been categorized into four overarching areas: “knowing why something is important, knowing what we need to do, knowing how to do it, and knowing when we do it” (Chappuis, 2004, p. 20). These four overarching areas were based on Waters, Marzano, and McNulty’s (2003) knowledge taxonomy referred to as the balanced leadership framework. Using these overarching categories, Chappuis (2004) developed four domains of assessment leadership, as referred to in Figure 1.

Figure 1. Chappuis (2004) domains for assessment leadership. This figure was developed for the purpose of this study to provide a pictorial representation of the domains of assessment leadership. These four domains encompass Chappuis et al. (2004)’s assessment leadership competencies. Adapted from “Leading Assessment for Learning: Using Classroom Assessment in School Improvement. Texas Association of School Administrators Professional Journal Insight, 18(3), 18-22. Retrieved from http://downloads.pearsonassessments.com/ati/downloads/insightnograph.pdf

To conceptualize the framework for the purpose of this study, the Chappuis et al. (2004) 10 competencies were subsequently categorized into one of the four previously mentioned domains, as referred to in Table 1. Chappuis (2004) specifically categorized some of these competencies within this framework; however, competencies that were not
specifically categorized were further delineated into one of the four domains by the researcher for the purpose of this study.
### Table 1

**Chappuis (2004) Framework for Assessment Leadership and Alignment to Assessment Leadership Competencies**

<table>
<thead>
<tr>
<th>Assessment leadership domain</th>
<th>Competency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knowing What to Teach and How To Assess</strong></td>
<td>1. The leader understands the standards of quality for student assessment and how to ensure that these standards are met in all assessments.</td>
</tr>
<tr>
<td></td>
<td>3. The leader understands the necessity of clear academic achievement targets, aligned classroom-level achievement targets, and their relationship to the development of accurate assessments.</td>
</tr>
<tr>
<td></td>
<td>5. The leader can plan, present, and/or secure professional development activities that contribute to the use of sound assessment practices.</td>
</tr>
<tr>
<td></td>
<td>9. The leader understands the attributes of a sound and balanced assessment system.</td>
</tr>
<tr>
<td></td>
<td>10. The leader understands the issues related to the unethical and inappropriate use of student assessment and protects students and staff from such misuse.</td>
</tr>
<tr>
<td><strong>How We Use Assessment as Instruction and Involve Students in the Process</strong></td>
<td>2. The leader understands the principles of assessment for learning and works with staff to integrate them into classroom instruction.</td>
</tr>
<tr>
<td><strong>How we Monitor Our Practices</strong></td>
<td>4. The leader knows and can evaluate the teacher’s classroom assessment competencies and helps teachers learn to assess accurately and use the results productively.</td>
</tr>
<tr>
<td></td>
<td>6. The leader accurately analyzes student assessment information, uses the information to improve curriculum and instruction, and assists teachers in doing the same.</td>
</tr>
<tr>
<td></td>
<td>7. The leader develops and implements sound assessment and assessment-related policies.</td>
</tr>
<tr>
<td><strong>How We Communicate About Student Learning</strong></td>
<td>8. The leader creates the conditions necessary for the appropriate use and reporting of student achievement information, and can communicate effectively with all members of the school community about student assessment results and their relationship to improving curriculum and instruction.</td>
</tr>
</tbody>
</table>
Knowing What to Teach and How to Assess

This domain relates primarily to a leader’s knowledge of quality assessment creation and use, alignment of assessments to learning intentions, understanding acceptable uses and interpretations of assessment, and finally, a leader’s ability to provide professional development for staff regarding the creation and use of assessments. This domain encompasses competencies 1, 3, 5, 9, and 10 (Chappuis, 2004; Chappuis et al., 2004).

How We Use Assessment as Instruction and Involve Students in the Process

This domain specifically addresses the integral nature of curriculum, instruction, and assessment. Competency two is addressed within this domain. The leader is responsible for understanding assessment for learning and assisting staff with classroom implementation. Black and Wiliam (2010) highlighted the positive relationship between formative assessment and student achievement (Chappuis, 2004; Chappuis et al., 2004;).  

How We Monitor Our Practices

This domain addresses the need for assessment leaders to monitor teachers’ use and application of assessments and teachers’ effectiveness in this process. Principals must ensure teachers can articulate the reason for each assessment. Once students have been assessed, principals must monitor the ways in which teachers interpret, communicate, and use the results. This competency also addresses the assessment policies that a leader creates that will support or hinder assessment for learning. Because assessment for learning focuses on the reciprocal nature of instruction and assessment, emphasis must be placed on the feedback provided to students as a result of the assessment (Chappuis, 2004; Chappuis et al., 2004).
How We Communicate About Student Learning

The fourth domain aligns with competency eight. The principal is responsible for ensuring that stakeholders are informed about how and why content is assessed, how items will be assessed and scored, and how to understand the results. This should not be limited to standardized assessment results or final grades on report cards (Chappuis, 2004; Chappuis et al., 2004).

Application of Conceptual Framework

This study addressed two domains of the conceptual framework, specifically Knowing What to Teach and How to Assess and How We Use Assessment as Instruction and Involve Students in the Process. This study encompassed principals’ knowledge of and support of assessment for learning practices. The focus is on these two domains because they relate most directly to principals’ knowledge of classroom assessment and its relationship to instruction. Because of the positive relationship between student achievement and formative assessment (Black & Wiliam, 2010), it is critical to focus on principals’ knowledge of the appropriate use of assessment for learning so they may support teachers in this process. This conceptual framework served as the basis for the research questions guiding this study.

Research Questions

Question 1: To what degree are Virginia principals knowledgeable of classroom assessment practices as measured by the Assessment Literacy Inventory?

Question 2: What are the differences, if any, among principals’ assessment literacy related to level assignment (elementary versus secondary) and type of assessment training?
Question 3: What is the relationship between principals’ knowledge of classroom assessment practices and leadership practices that support assessment for learning?

**Significance of the Study**

The need for assessment leadership is highlighted by accreditation ratings across Virginia. The Commonwealth reported an alarming 19% of Virginia schools did not meet the requirements for full accreditation (Virginia Department of Education, 2016). The need to improve student achievement to address more rigorous standards is pervasive across schools and divisions throughout the commonwealth. Accountability demands plague the principalship because he or she is responsible for leading a school and improving student outcomes and accreditation ratings. As the instructional leader of the building, principals are responsible for ensuring quality instruction and student learning. Leithwood and Riehl (2003) acknowledged the impact of accountability on stakeholders and school leadership:

Local, state and federal achievement standards for ambitious learning for all children have changed the landscape of educational accountability. Pressure is on actors at all levels, from students themselves to teachers, principals, and superintendents. In these times of heightened concern for student learning, school leaders are being held accountable for how well teachers teach and how much students learn. (p. 2)

Scholars in educational assessment assert that educational leaders must be assessment literate in order to address accountability challenges. Popham (2004) asserted, “educational accountability and assessment literacy are almost joined at the hip—or should be” (p. 82). He further highlighted that accountability systems are driven by
student achievement on standardized assessments. Despite the imperative role of assessments and data, many educators lack knowledge in sound assessment practices. Popham (2004) warned, “such assessment illiteracy is surely a prescription for professional suicide” (p. 82).

Assessment literacy is defined as the ability to understand the characteristics of high and low quality assessments and the capacity to relate this understanding to student outcomes (Stiggins, 1991, p. 535). Educators who are assessment literate can determine the impact of the assessment on student achievement and can use the results to understand which outcomes are significant. Assessment illiteracy will result in negative consequences for students (Stiggins, 1991).

Since accreditation is determined at the building level, the need for assessment leadership within schools often rests on the shoulders of the principal. The principal’s role in classroom and state assessment is all encompassing and requires that he or she be knowledgeable of assessment practices. He or she must be comfortable with interpreting assessments and assisting internal and external stakeholders in understanding what the results mean. Additionally, he or she is responsible for ensuring that quality assessments are used because of the critical role assessments play in decision-making (Stiggins, 2001).

In addition to ensuring they are personally assessment literate, principals must also extend this understanding to other educators within their building. Principals must guarantee each teacher “is a competent, confident master of the achievement targets that students are to hit. That mastery represents an essential foundation of accurate classroom assessment” (Stiggins, 2001, p. 16). Principals must craft professional development that
promotes assessment literacy and support teachers’ development of assessment literacy within the school community (Stiggins, 2001).

As a result of the link between assessment for learning and student achievement, this study examined principals’ assessment literacy and the practices they employed to promote and facilitate teachers’ development of assessment literacy in a post-No Child Left Behind (NCLB) context. Impara, Plake, and Fager (1993) first compared the assessment literacy of administrators and teachers using a national sample. Impara and Plake (1995) further analyzed differences in assessment literacy in administrators, teachers, and counselors in a Virginia sample of educators. To date, studies of principals’ assessment literacy in Virginia have not been examined in a post-NCLB context. Additionally, the literature has not explored the relationship between principals’ levels of assessment literacy and their support of assessment for learning practices. This study served to inform the professional development needs of administrators across the state. Additionally, it examined administrator’s application of assessment for learning strategies that help or hinder student achievement across the Commonwealth. Because of the power of formative assessment to improve student achievement, it is imperative that principals create school structures that support the utilization of assessment for learning.
Definition of Terms

- Accreditation Denied: a status designated to Virginia schools that do not meet full or provisional accreditation ratings over four consecutive years (Virginia Department of Education, 2016, p. 2)

- Assessment: “Any systematic method of obtaining information, used to draw inferences about characteristics of people, objects, or programs; a systematic process to measure or evaluate the characteristics of performance of individuals, programs, or other entities, for purposes of drawing inferences; sometimes used synonymously with test” (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014, p. 216)

- Assessment for Learning: “teachers use classroom assessment and the continuous flow of information about student achievement that it provides to advance, not merely check on, student learning” (Chappuis et al., 2004, p. 35)

- Assessment Leadership: instructional leadership that facilitates teachers’ integration of quality assessment practices that support the integral nature of teaching and learning (Stiggins & Duke, 2008, p. 286)

- Assessment Literacy: “knowledge about testing that supports valid interpretations of test scores for their intended purposes, such as knowledge about test development practices, test score interpretations, threats to valid score interpretations, score reliability and precision, test administration, and use” (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014, p. 216)
Balanced Assessment: the use of multiple data points and multiple assessment formats from formative and summative assessments to guide instructional decision making

Conditionally Accredited: a status designated to Virginia public schools within one year of opening (Virginia Department of Education, 2016, p. 3)


Fully Accredited: a status designated to Virginia schools whose overall adjusted pass rates meet a 75% benchmark in English and 70% in mathematics, science, and history. High schools are fully accredited if they have an 85 or higher on the Graduation Completion Index (Virginia Department of Education, 2016, p. 1)

Graduation & Completion Index: a calculation that measures on-time graduation and student completion outcomes (Virginia Department of Education, 2016, p. 2)

Instructional Leadership: leadership focused on enhancing classroom instruction and student learning (Leithwood, Louis, Anderson, & Wahlstrom, 2004).

Partially Accredited: Approaching Benchmark-Graduation and Completion Index: a status designated to Virginia public schools who meet adjusted pass rates and miss the Graduation and Completion Index by one point (Virginia Department of Education, 2016, p. 1)
• Partially Accredited: Approaching Benchmark-Pass Rate: a status designated to Virginia public schools that score two points below the SOL pass rates for full accreditation (Virginia Department of Education, 2016, p. 1)

• Partially Accredited: Improving School-GCI: a status designated to Virginia public high schools that meet adjusted pass rates and have improved their GCI; however, do not meet the full or narrow margin (Virginia Department of Education, 2016, p. 2)

• Partially Accredited: Improving School-Pass Rate: a status designated to Virginia public schools that do not meet the requirements of Full Accreditation or Partially Accredited: Approaching Benchmark-Pass Rate but demonstrate appropriate progress (Virginia Department of Education, 2016, p. 2)

• Partially Accredited: Reconstituted School: a status designated to Virginia public schools that do not meet the requirements for full accreditation for four consecutive years and are approved by the Virginia Department of Education to reconstitute (Virginia Department of Education, 2016, p. 2)

• Partially Accredited: Warned School-Pass Rate: a status designated to Virginia public schools whose SOL adjusted pass rates have not made acceptable progress and do not fall within the margin for full accreditation (Virginia Department of Education, 2016, p. 2)

• Partially Accredited: Warned School-GCI: a status designated to Virginia public schools who reached the benchmark for full accreditation using adjusted SOL pass rates but have not met the GCI requirements, nor fall within the narrow margin or have not made acceptable progress towards meeting the GCI (Virginia Department of Education, 2016, p. 2)
• Standards of Learning (SOL): Virginia state standards and corresponding curriculum with corresponding assessments that influence state accreditation ratings (Virginia Department of Education, 2016, p. 1)

• State Accreditation: a system employed by the Virginia Board of Education, with updated 2015-2016 ratings that provides information to stakeholders about a school’s performance and progress towards meeting state benchmarks (Virginia Department of Education, 2016, p. 1)

• Summative Assessment: “The assessment of a test taker’s knowledge and skills typically carried out at the completion of a program of learning, such as the end of an instructional unit” (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014, p. 224)

**Organization of the Remainder of the Study**

Chapter 2 is inclusive of a review of the literature on assessment leadership, assessment literacy, and assessment for learning practices. The literature review consists of research that supports the Chappuis (2004) conceptual framework for assessment leadership. Chapter 3 describes the methodology used for conducting research. Chapter 4 is inclusive of major quantitative and qualitative analyses and findings related to the research questions. Finally, Chapter 5 presents an overall summary of the results as well as future recommendations and implications for research.
CHAPTER TWO: REVIEW OF THE LITERATURE

To conceptualize this study, an overview of Virginia’s accountability system is first described and explained using the Virginia Standards of Learning (SOLs). In order to meet the rigorous standards, the literature describes the role and significance of instructional leadership. Assessment leadership, as a facet of instructional leadership, is further described in the review. The Chappuis (2004) conceptual framework for assessment leadership is depicted and described to further explain the complexities of assessment leadership. The review subsequently describes the literature that supports this conceptual framework, with a review of classroom assessment standards. An overview of the use and evolution of formative assessments is explained and how formative assessment supports student achievement. Finally, the role of formative assessment in a balanced assessment system is also explained through the literature. Because of the contextual nature of this study, Virginia’s accountability system is described in the section below.

**Virginia Accountability**

The passage of the Every Student Succeeds Act (ESSA) promulgated the nationwide accountability movement originally prompted by No Child Left Behind, ensuring that all students will meet rigorous standards. Major changes of the ESSA provided greater latitude to states and local school divisions in ensuring school improvement measures. Additionally, the ESSA called for a balanced approach to
assessment to protect instructional time while still ensuring progress is appropriately monitored and communicated to stakeholders (The White House Office of the Press Secretary, 2015). Virginia has developed its own accountability system using the Virginia Standards of Learning that guide the curriculum, and assessment in K-12 public schools in Virginia. Under ESSA, states must assess reading and mathematics and provide disaggregated information about subgroup performance (Virginia Department of Education, 2016).

**Virginia Standards of Learning**

Virginia’s accountability system is based on a series of standards referred to as Standards of Learning (SOL). The intention of SOL tests is to demonstrate the degree to which students have met Virginia standards in the areas of English, mathematics, science, and history. Additionally, the assessments are designed to determine subgroup proficiency. These assessments serve to “identify schools in need of assistance and to inform parents and the public about the progress of schools through the awarding of annual accreditation ratings” (Virginia Department of Education, 2016, p. 1)

**Accreditation Ratings**

The Virginia Board of Education designates accreditation ratings as a means of communicating to stakeholders about a school’s overall performance. Schools may be recognized for meeting all benchmarks by being *Fully Accredited, Partially Accredited,* or have their *Accreditation Denied.* The revised system now acknowledges improvement and progress in schools that have not yet met the requirements for full accreditation (Virginia Department of Education, 2016).
**Full accreditation.** To achieve full accreditation, schools must have a minimum of 75% in the content areas of English, and 70% or higher in the areas of math, science, and history. Additionally, high schools must also have a Graduation Completion Index (GCI) of 85 (Virginia Department of Education, 2016).

**Partial accreditation.** Schools may be classified as *Partially Accredited:* Approaching Benchmark-Pass Rate if they miss the adjusted SOL pass rates for full accreditation by two points. A school that misses the GCI by one point is deemed *Partially Accredited: Approaching Benchmark-Graduation and Completion Index.* Schools may also be recognized as *Partially Accredited: Improving School-Pass Rate* if they do not meet categories previously stated but if they demonstrate progress or if they demonstrate improvement in low performing subgroups, as quantified by the Virginia Department of Education. Similarly, a school is deemed *Partially Accredited: Improving School-GCI* if it has met the adjusted pass rates, demonstrating at least one point improvement towards the GCI but not met the GCI requirements. A school is designated as *Partially Accredited: Warned School-Pass Rate* if it is not near the pass rate and not making adequate progress towards reaching the established pass rates. A *Partially Accredited: Warned School-GCI* has met adjusted pass rates but not made appropriate progress towards the GCI. A *Partially Accredited: Reconstituted School* demonstrates that they have not met requirements for four years and have reconstituted upon approval by the Board of Education. If it does not meet accreditation requirements within an established window or does not renew its status, it may be reclassified as *Accreditation Denied* status (Virginia Department of Education, 2016).
Accreditation denied. A school receives *Accreditation Denied* status when it has not met accreditation requirements for four years in a row, which will result in a memorandum of understanding between the local governing bodies and the Board of Education. Schools must provide parents notice of this rating and have a time bound corrective action plan (Virginia Department of Education, 2016).

Conditional accreditation. *Conditionally Accredited* is a status that can be awarded for one year for new schools if students previously attended another school (Virginia Department of Education, 2016).

System of accountability. Since accreditation ratings are contingent on student achievement on the SOL assessments, schools must rise to meet state standards and demonstrate appropriate levels of student achievement. Systems of accountability, driven by high-stakes testing, have large standing implications for schools. “The impact of testing on curriculum, teaching, school systems, pupil motivation and teachers’ practice should leave us in no doubt as to the power of testing, particularly high-stakes testing, to affect teaching and learning” (Gipps, 1994, p. 57). The principal, as instructional leader, is responsible for the accreditation status of his or her school. Because of this accountability, the role of instructional leadership merits further exploration in the sections below.

**Instructional Leadership**

The principalship is a complex role associated with competing demands and responsibilities. A principal is expected to raise student achievement, facilitate change, manage staff, oversee organizational functions, and provide the instructional vision and direction for the school. Despite a need to improve student outcomes, many principals
fail to prioritize their role as instructional leader by allowing other responsibilities to overshadow this instructional priority (Cotton, 2003; Fink & Resnick, 2001). Heightened attention and focus on school accountability reframed the principalship such that student learning evolved as the primary focus. Schools are challenged with accountability demands, and “in order to meet the challenges associated with national and state expectations, principals must focus on teaching and learning” (Stronge, Richard, & Catano, 2008, p. 4). A focus on student outcomes and subsequently the need to improve quality instruction is central to the role of principals as instructional leaders; however, the literature has broadly defined this role.

An instructional leader is defined as someone who “encourages a focus on improving the classroom practices of teachers as the direction for the school” (Leithwood et al., 2004, p. 6). Using this operational definition of instructional leadership, principals are responsible for overseeing the instructional direction of the school through creation and support of a clear vision. This can be achieved through creating and monitoring school improvement plans, providing appropriate professional development for staff, building the capacity for leadership in others, and conducting staff evaluations. Instructional leadership also involves data driven decision-making. This review of the literature emphasized that student performance in these high-stakes contexts is contingent on a principal’s instructional leadership (Leithwood et al., 2004; Stronge et al., 2008).

Through a synthesis of 25 years of research findings, Hallinger (2005) highlighted that the principal as an instructional leader is responsible for the following actions:

- Creating a shared sense of purpose in the school, including clear goals focused on student learning;
• Fostering the continuous improvement of the school through cyclical school development planning that involves a wide range of stakeholders;
• Developing a climate of high expectations and a school culture aimed at innovation and improvement of teaching and learning;
• Coordinating the curriculum and monitoring student learning outcomes;
• Shaping the reward structure of the school to reflect the school’s mission;
• Organizing and monitoring a wide range of activities aimed at the continuous development of staff; and
• Being a visible presence in the school, modeling the desired values of the school’s culture. (p. 233)

Less conventional notions of instructional leadership describe the role of principals in “organizational management” (Horng & Loeb, 2010, p. 66). In this regard, principals are responsible for hiring and retaining effective teachers as well as providing the necessary resources and professional development to ensure quality outcomes.

As the literature continues to grapple with how to define instructional leadership and the various responsibilities of principals as instructional leaders, a host of research examines the impact of instructional leadership on student outcomes and on teaching practices, as summarized in the section below.

**Impact of Instructional Leadership**

A mirage of factors can potentially impact student achievement, positively or negatively; however, leadership ranks high within school factors for its influence. In a comparison with various school-related factors for their impact on student achievement, leadership is only surpassed by instruction and teaching (Leithwood et al., 2004;
Leithwood & Riehl, 2003). This finding stresses the importance of leadership, and it underscores the need for quality instructional leadership across K-12 contexts. Additionally, one of the major distinctions between a principals’ effectiveness is their degree of envelopment with the instructional program of the school (Cotton, 2003). While these studies identified the potential impact of leadership, the literature dives deeper into the aspects of leadership that are associated with raising student achievement.

When analyzing the effectiveness of leadership through the lens of student outcomes, instructional leadership has a greater impact on overall student achievement as opposed to more traditional conceptions of leadership such as transformational leadership within a school setting (Hattie, 2009; Robinson, Lloyd, & Rowe, 2008). Instructional leadership yielded higher effect sizes on student achievement than the often more glamorized role of transformational leader, known for its ability to be a catalyst for change within an organization. In a study of types of leadership and student achievement, researchers found that “the impact of instructional leadership on student outcomes is three to four times greater than that of transformational leadership” (Robinson et al., 2008, p. 655). These differences can be attributed to the different foci of transformational and instructional leadership. Instructional leadership mobilizes staff towards an instructional vision and developing teacher pedagogy; conversely, transformational leadership is driven by the relationship between leadership and staff (Robinson et al., 2008). At the core of their work, principals as instructional leaders maintain a school wide focus on student learning (Hattie, 2009, p. 83).

Brown’s (2001) meta-analysis on the influence of school leadership on student achievement also supported the role of instructional leadership compared to other
approaches to leadership. Instructional leadership yielded higher effect sizes in elementary contexts \((d = 0.75)\) versus secondary contexts \((d = 0.44)\). This meta-analysis revealed “that leadership does influence school effectiveness, and the instructional approach to leadership assumes preeminence over other approaches” (p. 113).

Marzano, Waters, and McNulty (2005) also examined the impact of school leadership on student achievement. A correlation coefficient of 0.25 was revealed between a principal’s leadership and student achievement across 69 studies and 1.4 million students. This finding suggests that effective leadership was positively correlated to student achievement. This correlation implies that “a highly effective school leader can have a dramatic influence on the overall academic achievement of students” (p. 10). While the findings of these meta-analyses specifically addressed correlations between a principal and student achievement, other studies examined the indirect impact of instructional leadership.

**Indirect Impact of Instructional Leadership**

While accountability movements greatly emphasize the role of a leader in improving student outcomes, multiple studies demonstrated that the leader’s influence on student achievement is indirect. This indirect influence is a result of a leader’s influence on teaching practices as well as school vision and goals (Hallinger, 2005; Leithwood & Riehl, 2003; Leithwood et al, 2004). Additionally, through alignment of practices, instructional leaders ensure that the actions of the organization support the school’s mission (Hallinger, 2005).

Although their impact on student achievement may be indirect, principals do have the potential to directly influence teachers and the overall organization. Principals as
instructional leaders have the greatest impact “on teachers’ motivation and working conditions; their influence on teachers’ knowledge and skills produces less impact on student achievement” (Louis, Leithwood, Wahlstrom, & Anderson, 2010, p. 19). This review further revealed that teachers and principals demonstrated that instructional leadership that emphasizes goals, monitoring professional development, and providing opportunities for collaboration are pivotal instructional leadership strategies.

Through a substantive review of the literature to discern if leadership has a direct or indirect influence on student achievement, Leithwood, Harris, and Hopkins (2008) concluded “there is not a single documented case of a school successfully turning its pupil achievement trajectory in the absence of talented leadership” (p. 29). This highlights the need for strong, instructional leadership in schools that are not meeting achievement benchmarks because of their potential to act as a “catalyst” within their building (p. 29).

Earlier research on instructional leadership investigated the influence of leadership on teaching practices without directly examining the impact on student achievement. In a survey of 809 teachers, Blase and Blase (1999) found that instructional leadership can “have strong enhancing effects on teachers emotionally, cognitively, and behaviorally” (p. 367). The “impact achieved by principals on school outcomes (i.e., student achievement) derives, in part, from the principals’ interaction with and influence on teachers” (p. 368). Specifically, focusing on collegial conversations, supporting professional development, and encouraging reflective practices were distinguished as leadership practices that led to changes in teaching practices. This study provides further support for the role of professional development as an important facet of instructional
leadership; however, it does not address the impact these practices have on student achievement. Central to a principal’s role is assessing the needs of the school and determining the best manner with which to deploy the resources to support the professional team (Portin, Schneider, DeArmond, & Gundlach, 2003).

Assessment Leadership

The aforementioned research highlights the critical need for instructional leadership to improve student outcomes. A major responsibility of the principal is to fulfill the instructional vision of the school, and the use of assessment data to drive instruction is a vital aspect of the principal’s role. Lunenburg (2013) highlighted the integral role of assessments within the role of the building leader:

The instructional leadership of the principal is a critical factor in the success of a school's improvement initiatives and the overall effectiveness of the school. The primary responsibility of the principal is to promote the learning and success of all students. School principals can accomplish this goal by focusing on learning, encouraging collaboration, using data to improve learning, providing support, and aligning curriculum, assessment, and instruction. (p. 37)

The literature recognized assessment leadership as a form of instructional leadership that is needed to improve assessment practices and enhance the teaching and learning practices within schools. Stiggins and Duke (2008) made the case that effective instructional leadership needs strong assessment leadership. “Instructional leadership also requires an understanding of the role of sound assessment in efforts to improve teaching and learning. The well-prepared principal is ready to ensure that assessments are of high quality and used effectively” (Stiggins & Duke, 2008, p. 286).
The principal plays a significant role in the utilization of effective assessment practices within buildings. Principals who clearly understood the role of formative assessment and concentrate on students within classroom observations demonstrate the ability to effectively support the integration of formative assessment practices within their buildings (Moss, Brookhart, & Long, 2013). Principals play a pivotal role in this process:

The formative assessment process creates an evidence-based culture focused on student learning and achievement rather than on the instructional activities of the adults in the school. As a result, formative assessment promotes a cultural shift from teacher-centered to student-centered evaluative beliefs and normative practices. Administrative leadership is both the catalyst and driving force for the kind of cultural change in classrooms that formative assessment requires. When administrators see formative leadership as their target, they also see themselves as the leading learners in their schools, view teachers as learners, and enter into meaningful learning partnerships with teachers and students. (Moss, Brookhart, & Long, 2013, p. 213)

Assessment leadership encompasses an array of competencies and actions to support assessment for learning. Stiggins (2001) asserted that assessment leaders must demonstrate that they understand assessment for learning principles and they must first possess their own requisite levels of assessment literacy. Additionally, assessment leaders know how to support teachers in the integration of assessment for learning practices and in their understanding of the role of assessment data. Finally, assessment leaders recognize the interplay between assessment and learning. Stiggins (2001) articulated that
schools must have “clear and appropriate achievement targets” and “an assessment
literate faculty” in order to improve the teaching and learning process within schools (pp.
18-19). It is the role of the assessment leader to ensure that structures and supports are
established to support assessment for learning practices.

Impact of Assessment Leadership

Principals’ support of assessment practices has been examined in relationship to
improved outcomes for schools. In an analysis of the subdomains related to assessment
leadership, a principal’s involvement with curriculum, instruction, and assessment has a
positive correlation ($r=0.20$) with student achievement. A principal’s knowledge of
curriculum, instruction, and assessment also has a positive relationship ($r=0.25$) with
student achievement. And finally, the degree to which a principal monitors student
achievement is also positively associated with student achievement ($r=0.27$). Each of
these correlations demonstrates the relationship between an aspect of assessment
leadership and student achievement (Marzano et al., 2005).

Other studies examined the impact of leadership in contexts in which instructional
improvements were warranted and made as a result of assessment leadership. Connell
(1996) examined school practices in schools that were able to have their names removed
from the state school improvement list. Schools shared a common focus in which they all
addressed the low academic achievement. Strategies to address these shortcomings
included curriculum alignment, enhancements in classroom instruction, progress
monitoring, a positive school culture for students and family, partnerships with external
stakeholders, the creation of an arts program, and changes in personnel. Duke (2004)
asserted that school improvements are unlikely in the absence of assessment leadership,
which requires a focus on teaching and learning and the integration of data to examine student performance.

**Principals’ Application of Assessment Leadership**

In addition to the research that examines the role of assessment leadership in student performance, the assessment practices of principals has been examined within the context of Connecticut Principals (Ulmer, 2002) and again with principals in Saskatchewan, Canada (Hellsten, Noonan, Preston, & Prytula, 2013). In addition to geographic differences between the samples, the Canadian principals were affected differently by accountability policies, specifically, because Hellsten et al. (2013) were not influenced by NCLB policies, unlike their Connecticut counterparts in Ulmer (2002). Canadian principals within this sample demonstrated that they were most likely to use student assessment and data to contribute to foster a constructive culture surrounding assessment within the school context, a finding that was paralleled in the Ulmer (2002) sample. Additionally, the Canadian principals demonstrated that they were likely to employ assessment data as a progress-monitoring tool to measure fulfillment of school wide goals. And finally, these principals demonstrated that they urged staff to rely on student assessment to identify individual student performance (Hellsten et al., 2013).

The research on assessment leadership and instructional leadership support the overarching need for principals with a clear vision of how to integrate assessments within the teaching and learning process; however, challenges are associated with the role of assessment leadership. Webber, Scott, Aitken, Lupart, and Scott (2013), highlighted these challenges:

In summary, leading assessment is complex and difficult. It requires the capacity
to go beyond traditional conceptualizations of leadership to build teacher capacity for assessment innovation. Leaders who draw upon the interplay among values, theoretical and procedural knowledge, professional skills and personal qualities to shape their leadership vision are more likely to achieve positive organizational change and enhanced professional cultures. Additional benefits include positive student outcomes, enhanced instructional practice, enriched partnerships with parents and community, increased assessment literacy, productive cultures and more effective monitoring and reporting practices. (p. 252)

**Principal’s Role in Developing Teachers’ Assessment Literacy**

A critical aspect of the principal’s role as instructional leader is to support teachers’ development, use, and understanding of quality assessments for learning. The use of formative assessment to improve student achievement requires support and training for teachers (Heritage, 2007; Moss et al., 2013; Renihan & Noonan, 2012; Stiggins, 2001; Webber et al., 2013). “Principals can be pivotal in the improvement of student learning by helping teachers develop and use sound classroom assessment that strengthens instruction and student learning” (Stiggins & Duke, 2008, p. 286). Staff development is critical when teachers do not possess the requisite skills to fully implement and utilize assessment to gauge and adjust instruction (Duke, 2004). A synthesis of research is described below regarding how principals can provide opportunities for teachers to expand their use and understanding of assessment practices within their classrooms.
Personal levels of assessment literacy are a prerequisite before a principal can work with teachers to develop their own assessment literacy. Moss et al. (2013) specified the significance of this understanding:

Until administrators have a deep understanding of formative assessment and develop specific descriptive language to name exactly what they see, they will not be able to recognize and understand formative assessment practices when they occur in the classroom. Until a principal or supervisor deeply understands formative assessment, classroom observations remain at the level of the principal telling the teacher what she did right instead of the principal partnering with the teacher to learn something about student achievement. (pp. 215-216)

The principals’ knowledge of formative assessment combined with their observation of formative assessment in action is essential to school wide practices that use assessment to improve the teaching and learning process in the classroom. “On a related note, we consider the appreciation of the integral relationship among teaching, learning and assessment to be an important antecedent to effective assessment leadership” (Noonan & Renihan, 2006, p. 11).

Staff development does not necessarily mean assessment should be examined and studied in isolation. Instead, rather, professional development should involve “continuous improvement with the objectives of aligning school expectations, providing regular feedback on student learning, and promoting thinking about classroom strategies for enhancing learning” (Noonan & Renihan, 2006, p. 8). This supports the overarching goal of linking learning and assessment as a cyclical process.

The role of the principal in supporting teachers in their use of formative
assessment practices is paramount. “Administrators need to both be part of and provide leadership for the intentional lesson-by-lesson focus on what students are actually doing to develop and produce evidence of their understanding of essential learning targets” (Moss et al., 2013, p. 217). Through this leadership, principals have the capacity to create schools with an orientation to student learning.

This focus on student learning is essential to the school improvement process as the principal monitors the effectiveness of teaching practices. “Effective assessment leaders will monitor improvements in assessment practices and celebrate professional growth as well as celebrate improvements in student achievement” (Webber et al., 2013, p. 249). Assessment leaders understand the role of each assessment used and how it supports the overarching goals (Cizek, 1995). As part of this process, assessment leaders understand that assessments yield data that addresses various interests because “every assessment poses different questions because the individuals who use the resulting information have different needs” (Stiggins & Duke, 2008, p. 288). Ultimately, the principal must in turn interpret and use the data to improve student achievement (Duke, 2004; Stiggins & Duke, 2008).

This school wide improvement and development of assessment-related skills is further enhanced when teachers are provided opportunities to learn from their colleagues and observe tangible examples of effective assessment practices within the classroom (Black & Wiliam, 2010). Additionally, the use of professional learning communities that support “learning and collaboration” is appropriate to support teachers in the development of their professional skillsets (Hollingworth, 2012, p. 377). It is through these supports and structures that principals will support the utilization of assessment for
learning strategies within their contexts and help address the needed levels of assessment literacy of their staff. Assessment leadership presents an opportunity for principals to improve the teaching and learning process; however, it is not without its share of obstacles in implementation.

As an example of such support, Hollingworth (2012) explored the role of leadership within a school setting as teachers developed knowledge and use of formative assessment. In this qualitative case study, the principals were “catalysts for innovation in instruction and classroom assessment” (p. 365). By providing time for implementation as well as time for professional collaboration, the principal was able to support the efforts of classroom teachers. Additionally, principals showed support by attending professional development in an effort to develop a greater understanding of the instructional impact of formative assessment.

Principals as assessment leaders must use assessment in an evaluative capacity but also to better understand what is occurring in the classroom. Principals must weigh the various roles of assessment, which involves “maintaining a fairly delicate balance between ensuring accountability and quality control, on one hand, and nurturing professional empowerment among teachers, on the other” (Renihan & Noonan, 2012, p. 4). This balance between accountability and teacher empowerment represents one of the challenges of assessment leadership as the users juggle the competing roles and demands of assessment.

**Obstacles to Assessment Leadership**

Various obstacles may impede the development of effective assessment leadership within schools. One of which may be the belief that principals have been
adequately trained to oversee and use assessments (Stiggins & Duke, 2008). Despite this obstacle, a major role of assessment leaders is to develop their own knowledge base of effective assessment principles and practices. Additionally, “principals must remove all barriers to the development of teachers’ assessment literacy. These include personal, institutional, and community barriers” (Stiggins, 2001, p. 24).

Cizek (1995) outlined the steps principals should take to lead an assessment system within a building. The first recommended step was to develop personal levels of assessment literacy. Technical experts should not be solely responsible for obtaining this knowledge; they should solicit the support of others involved with assessment to develop a broader understanding of the roles of various forms of assessment. As assessment leaders, principals:

must become assessment literate themselves. Without this basis of professional expertise, principals will remain unable to bring the issue of effective assessment to the fore as a school priority or provide the support teachers need to develop and use assessments productively in the classroom. (Stiggins, 2001, p. 24)

**Assessment Literacy of Principals**

Assessment literacy describes an individual’s level of knowledge related to quality assessment use and creation and his or her ability to understand and apply standards for quality assessment use and creation. Individuals with high levels of assessment literacy reduce threats to reliability and validity in an effort to ensure appropriate inferences about student achievement and learning are drawn. Assessment literacy requires that the user have a clear understanding of which assessment methods to
employ based on the learning outcomes and how to appropriately communicate and interpret the results of an assessment (Stiggins, 1991, 1995).

Assessment literacy does not necessarily require an in-depth understanding of psychometric principles; however, it does encompass an individual’s ability to make appropriate inferences regarding student learning based on assessment-related information (Popham, 2006). Those with high levels of assessment literacy are guided by two questions, “What does this assessment tell students about the achievement outcomes we value? And what is likely to be the effect of this assessment on students?” (Stiggins, 1991, p. 535). Although there is an implicit need for assessment literate educators to ensure the appropriate use and interpretation of assessments, the route to developing assessment literacy is plagued with multiple obstacles. The following section describes some of the obstacles educators face in their development of assessment literacy.

**Obstacles to Assessment Literacy**

One major obstacle in the development of assessment literate educators is educators’ apprehension and discomfort with assessment and evaluation (Stiggins, 1995, 2001). Many teachers lack an understanding of the connection between classroom instruction and standardized assessment. Discomfort with assessment may inhibit a teacher’s willingness to pursue occasions to develop their current knowledge of assessment practices in order to maximize their instructional utility (Stiggins, 1995). Despite the comprehensive understanding of assessment needed for educators, formal training in assessment is often limited to knowledge of standardized testing (Stiggins, 1991).
Another challenge is the limited time available to implement quality professional development on assessment. Extensive time in professional development is needed to develop assessment literacy, and furthermore, time is required to implement what is learned into the classroom (Stiggins, 1995). Even if sufficient time is allocated to training educators in assessment principles, incorporation of quality assessment into the classroom requires time (Stiggins, 1995). Assessment consumes a significant volume of teacher’s time. Teachers spend an estimated one quarter to a third of their time engaged in assessment activities; however, many teachers lack adequate preparation (Stiggins, 2014, p. 68). Again, proper training in assessment facilitates teachers’ ability to seamlessly and efficiently integrate assessment into classroom instruction to maximize its effectiveness in spite of time constraints. The principal is partially responsible for ensuring teachers are assessment literate.

Outside influences also represent potential hurdles in educators’ development of assessment literacy. Teachers and principals must grapple with the public perception that all teachers and administrators already have requisite levels of assessment literacy necessary to ensure quality teaching and learning are occurring (Stiggins, 1995). It may not be widely recognized that there is a need for assessment training for educators. Additionally, stakeholders may perceive that standardized assessments and report-card grades suffice as measures of student mastery of content. As a result of parents’ limited experience with assessment, they may lack an understanding of the importance of assessment and its implications (Stiggins, 2001). What external stakeholders understand and discuss about assessment is often narrowly defined by standardized assessment information.
Although standardized test scores are often used as the litmus test for student achievement, assessment literacy is more comprehensive than simply understanding the implications of standardized tests. One unintended consequence of standardized testing is that a “societal blind spot has been created by the common belief that standardized test results are the only truly acceptable evidence of student achievement” (Stiggins, 2014, p. 68). Additionally, if schools only rely on standardized assessment scores as a means to judge teacher quality or student achievement, they are discounting other influences on student achievement. Principals should support “balanced development and use of assessments. Both standardized tests and classroom assessments must be of high quality. Both must be effectively used for schools to improve, because both inform critical important decisions” (Stiggins, 2001, p. 15). In order to combat some of these issues associated with assessment literacy, the section below describes standards for developing requisite levels of assessment literacy.

**Standards for Developing Assessment Literacy**

In an effort to conceptualize a framework for improving assessment literacy, Stiggins (1995) has delineated five standards for quality assessment and has also specified how these standards relate to assessment literacy. Each standard is explained in the following section.

The first standard involves “starting with clear purposes” (Stiggins, 1995, p. 240). This involves the recognition that assessments are designed to fulfill various roles and provide different information. Assessments serve a variety of functions, including classroom-level assessments designed to inform instruction and larger scale assessment designed to analyze student achievement across a defined population. Specifically, for
building leaders, this involves providing the necessary resources and using assessments to
gauge the effectiveness of programs and staff.

The second standard involves “focusing on achievement targets” (Stiggins, 1995, p. 240). Assessment literate leaders recognize that there can be a wide array of desired outcomes. Some targets may pertain to acquisition of specific content or skills whereas others may involve the creation of products.

The third standard specifies that assessment literate educators should be skilled in “selecting appropriate assessment methods” (Stiggins, 1995, p. 241). At the building level, principals recognize that there are multitudes of assessment formats that serve different purposes. Assessment literate educators demonstrate the capacity to choose an assessment method that corresponds to the achievement outcomes and is appropriate for the students and tasks. This may include, but is not limited to assessment forms such as a selected-response format, essay assessment, performance assessment, or personal communication.

The fourth standard for assessment literacy emphasizes the importance of “sampling student achievement” (Stiggins, 1995, p. 242). When creating assessments, it is not feasible to include all possible questions that is inclusive of all related skills or content. As a result, the assessment creator must determine a representative sample of items that provides a good indicator of student mastery.

The fifth standard involves “avoiding bias and distortion” (Stiggins, 1995, p. 243). When assessing students, educators must be aware of the external threats to validity that can compromise the inferences drawn from the assessment. Additionally, the internal
threats to validity within the assessment may also have a negative impact. Knowledge of these standards will ultimately affect principals’ assessment literacy.

**Principals’ Levels of Assessment Literacy**

Using the aforementioned standards as a guide, the route to an assessment literate faculty should first be addressed by acknowledging and elevating the importance of principals developing their own assessment literacy (Cizek, 1995; Stiggins, 2001). Assessment literacy of principals has been analyzed and compared to teachers on a national scale within a pre-NCLB context (Impara et al., 1993). Other contextualized studies of principals’ assessment literacy have also been examined and compared to teachers and/or counselors (Impara & Plake, 1995; Perry, 2013; Rosas 2014). Additionally, principals’ beliefs and perceptions about their assessment literacy and development thereof have also been analyzed (Hall, 2003; Henry, 2011; Impara et al., 1994;). Studies of assessment literacy have revealed various strengths and weakness with regard to principals’ areas of assessment literacy. Within a national sample, principals outperformed teachers in measures of assessment literacy; however, more recent, contextualized studies comparing both groups reveal mixed results. Research describing the varying levels of assessment literacy is discussed within the synthesis below.

Impara et al. (1993) compared the assessment literacy of administrators and teachers using a national sample. Respondents were analyzed within seven strands of assessment literacy, including: *choosing assessment methods*, *developing assessment methods*, *administering*, *scoring and interpreting assessment results*, *using assessment results for decision making*, *using assessment results in grading*, *communicating assessment results*, and *recognizing unethical practices*. When compared to teachers
within the scope of this study, administrators had significantly higher scores in the areas of choosing assessment methods, using assessment results for decision making, and using assessment in grading and communicating assessment results; however, the authors noted that these results should be interpreted cautiously, as the differences between the groups was small, even though they were statistically significant.

In addition to analyzing differences between administrators and teachers, Impara et al. (1993) further analyzed relative strengths and weaknesses for administrators’ assessment knowledge. Results indicated that administrators scored highest on their ability to administer, score and interpret results, and their relative weakness was their ability to develop assessment methods. Specific areas where administrators scored relatively high included: “selecting assessment strategies; validly interpreting classroom test scores; and understanding that it is inappropriate to base a student’s grade on a single test score” (p. 520). Additionally, specific survey items that administrators had the most difficulty with included the differences between reliability and validity, understanding standardized test scores, and selecting the appropriate measure for specific contexts.

In addition to examining performance across the various standards, Impara et al. (1993) examined the relationship between assessment literacy and self-efficacy. Self-efficacy was related to administrator’s levels of assessment literacy (Impara et al., 1993). Administrators who indicated that they were “most comfortable” with various components of assessment literacy also had the highest scores (p. 516). From the administrators surveyed, “90% agree that classroom tests should be used extensively to enhance instruction”; however, differences existed among administrators regarding their perception of the role standardized assessments play in classroom instruction (p. 516).
Additionally, those administrators who supported the utility of classroom assessment for instructional purposes had significantly higher scores in assessment literacy (p. 516). Small but statistically significant differences in assessment literacy existed among administrators who had received training in assessment and those who had not. The implications of this study further highlight that “in addition to the necessity of serving as a resource to teachers in this area, administrators must have sufficient knowledge to protect themselves and their teachers from the potential unethical or improper use of test scores” (p. 520).

In addition to comparing teachers’ and administrators’ assessment literacy, Impara and Plake (1995) further analyzed differences among administrators, teachers, and counselors within a sample of Virginia educators. When comparing the three groups, administrators scored lower than counselors but higher than teachers. Elementary principals scored higher than secondary principals. Administrators demonstrated that their strongest areas of assessment literacy included their ability to choose appropriate assessments, analyze the validity of an assessment, share the results of assessments with other stakeholders, and identify unethical practices. Administrators’ lowest scores were in the area of understanding standardized test results. Again, the authors acknowledged that administrators would also benefit from additional training in assessment practices if they must continue to support teachers in their roles. Additionally, this study collected descriptive statistics on the number of participants who received formalized training in assessment, and only 2.6% administrators responded that they had not engaged in such training.
Other studies similarly explored the assessment literacy of principals in a contextualized state sample. Rosas (2014) compared the assessment literacy levels of elementary teachers and principals in a sample of Central California participants. Principals had significantly higher levels of assessment literacy than teachers, a finding that is consistent with the Impara et al. (1993) and Impara and Plake (1995) studies. Within the seven assessed standards, there was a significant difference between principals and teachers in their ability to administer, score, and interpret assessment results.

Another contextual study of Montana secondary principals and teachers almost two decades later compared the assessment literacy of principals and teachers, and revealed that teachers scored higher than principals. Perry (2013) found that teachers outperformed principals on six out of seven areas of assessment literacy, as measured by the Classroom Assessment Literacy Inventory (CALI) (p. 74). Principals only outperformed teachers in the area of recognizing unethical or illegal practices. Principals’ relative strengths within this survey involved their capacity to use assessment in decision-making; by comparison, principals’ lowest areas were recognizing unethical or illegal practices. Consistent with Impara et al. (1993) and Impara and Plake (1995), this study further signified the need to specifically analyze and address the assessment literacy of principals who are the instructional leaders within a building.

One major difference among the studies has been the use of the instrument to measure assessment literacy. DeLuca, Lapointe-Mcewan, and Luhanga (2016b) compared the utility of the various instruments, as described in the following section. The original instrument designed to measure assessment literacy was the Teacher Assessment Literacy Questionnaire (TALQ), which was first used in the Impara et al. (1993) study. It
contained 35 content-based items and was based on 1990 Standards (Plake, Impara, Fager, 1993). The Classroom Assessment Literacy Inventory (CALI) (Mertler, 2003) also had 35 content-based items based on 1990 Standards. Mertler and Campbell (2005) revised a previous version of the ALI, known as the Revised Assessment Literacy Inventory (ALI), which includes 35 scenario-based items and is also based on the 1990 Standards. A commonality among the instruments is their alignment with the 1990 Standards for Teacher Competence in the Educational Assessment of Students. Additionally, the psychometric principles for each of the instruments were determined with in-service and/or pre-service teachers. A comparison of instruments revealed Mertler and Campbell’s (2005) ALI had the highest psychometric principals when used with pre-service teachers ($\alpha=0.74$) (DeLuca et al., 2016b, p. 258).

While the aforementioned studies of principals utilized an objective measure of principals’ levels of assessment literacy, the literature also explored principals’ perceptions and beliefs about their assessment literacy. In a sample of superintendents from The School Superintendents Association (AASA) and principals from the National Association of Elementary School Principals (NAESP) and the National Association of Secondary School Principals (NASSP), administrators rated their own perceived levels of assessment literacy (Impara et al., 1994). Although different organizations ranked their skillsets somewhat differently, the overall three highest areas of strength included their knowledge of assessment-related vocabulary from standardized assessment, knowledge of different tests and usages, and finally, comprehending assessment-related principles (such as validity/reliability). Impara et al. (1993) found a relationship between administrator’s self-efficacy in assessment literacy and their actual scores on the
measure; therefore, the results from Impara et al. (1994) highlighted an important area of investigation with regard to principals’ beliefs in their own capabilities.

Also using self-analysis as a means for exploring principals’ assessment beliefs and practices, Henry (2011) examined secondary principals in Orange County, California. In a self-report of their preparedness, principals indicated that they were better equipped to use and analyze information from summative assessments rather than formative assessments. In a quantitative survey, 97% of principals indicated that they felt formative assessment is a valuable indicator of student learning, compared to 68% for summative assessment. With regard to supports and barriers to employing assessment data, principals revealed that their personal ability to use assessment data were an overall strength (94%); whereas time for staff to use data were a barrier (80%) as well as their own time to use data (74%).

In a focus group of principals, Hall (2003) found that principals attributed their development of assessment literacy to coursework, interactions with other educational stakeholders, professional development, and experience in the field. Areas of assessment literacy that principals wished they had greater knowledge of were those most closely related to accountability, especially with regard to curricular alignment with assessments.

**A Model for Assessment Leadership**

As described in the aforementioned research, the role of assessment leader is complex and multi-faced, and assessment literacy plays a significant role in instructional leadership, and more specifically, assessment leadership. Assessment leaders are able to guide the instructional vision of their school through their own understanding of assessment, support of teachers in their knowledge and use of assessment for learning
practices, utilization of assessment information to guide the school improvement process, and ability to communicate assessment results with the broader audiences. Chappuis et al. (2004) articulated 10 competencies for assessment leadership that encompass these skills and understandings:

1. The leader understands the standards of quality for student assessment and how to ensure that these standards are met in all assessments.
2. The leader understands the principles of assessment for learning and works with staff to integrate them into classroom instruction.
3. The leader understands the necessity of clear academic achievement targets, aligned classroom-level achievement targets, and their relationship to the development of accurate assessments.
4. The leader knows and can evaluate the teacher’s classroom assessment competencies and helps teachers learn to assess accurately and use the results productively.
5. The leader can plan, present, and/or secure professional development activities that contribute to the use of sound assessment practices.
6. The leader accurately analyses student assessment information, uses the information to improve curriculum and instruction, and assists teachers in doing the same.
7. The leader develops and implements sound assessment and assessment-related policies.
8. The leader creates the conditions necessary for the appropriate use and reporting of student achievement information, and can communicate effectively
with all members of the school community about student assessment results and their relationship to improving curriculum and instruction.

9. The leader understands the attributes of a sound and balanced assessment system.

10. The leader understands the issues related to the unethical and inappropriate use of student assessment and protects students and staff from such misuse. (p. 125)

Chappuis et al. (2004) describe the function of assessment leadership qualities through the application of these 10 competencies:

When school leaders put into practice the skills underlying the competencies, they promote the intentional use of accurate, student-involved classroom assessment on a daily basis to improve student learning, and in doing so, they also address the need to raise test scores as measured on standardized tests. (p. 124)

These 10 competencies have been categorized into four overarching areas: “knowing why something is important, knowing what we need to do, knowing how to do it, and knowing when we do it” (Chappuis, 2004, p. 20). These four overarching areas are based on Waters and colleagues’ (2003) knowledge taxonomy referred to as the balanced leadership framework. Using these overarching categories, Chappuis (2004) developed four domains of assessment leadership for principals, as referred to in Figure 2. Chappuis (2004) specifically categorized some of these competencies within this framework; however, competencies that were not specifically categorized were further delineated into one of the four domains by the researcher for the purpose of this study.
Figure 2. Chappuis (2004) conceptual framework for assessment leadership is categorized into four domains. The first domain addresses principals’ knowledge of best practices in assessment as well as their ability to evaluate student assessments. The second domain addresses the use of assessment for learning to improve student outcomes. Because the research questions align with these two domains, they are described in greater depth in the sections below.

**Knowing What to Teach and How to Assess**

The first competency states: “The leader understands the standards of quality for student assessments and how to ensure that these standards are met in all assessments” (Chappuis et al., 2004, p. 127). This competency implies that principals are knowledgeable of the criteria that constitute a quality assessment, which can be better described as the assessment literacy of a principal. Secondly, this competency suggests that principals as assessment leaders must provide assistance to teachers in the
appropriate use of assessments and subsequently making applicable inferences from the information provided. It is an administrator's responsibility to support teachers in their use of assessments, including their ability to judge the purpose of an assessment, determine the type of assessment that will be used, define what constitutes an adequate sample, and examine any threats to validity will affect the quality inferences that can be made (Chappuis et al., 2004).

Competency three states, “The leader understands the necessity of clear academic achievement standards, aligned classroom-level achievement targets, and their relationship to the development of accurate assessments” (Chappuis et al., 2004, p. 203). Competency three addresses the need of leaders to ensure that all teachers are able to clearly understand learning intentions and that these learning intentions are transparent for students. Competency three addresses the need for alignment between learning intentions, standards, curriculum, instruction, and assessments (Chappuis et al., 2004).

Competency five states, “The leader can plan, present, or secure professional development activities that contribute to the use of sound assessment practices” (Chappuis et al., 2004, p. 225). Assessment leaders must be able to provide professional development that will facilitate a teacher’s development or improvement of assessment literacy. Assessment leaders should be able to organize professional development for teachers to improve their use and understanding of formative classroom assessments (Chappuis et al., 2004).

Competency nine states, “The leader understands the attributes of a sound and balanced assessment system” (Chappuis et al., 2004, p. 284). This competency reiterates the need for both formative and summative assessments to provide information about
student learning. Leaders should recognize the implications of both standardized assessments as well as classroom assessments. Additionally, assessment leaders recognize and support the different forms of assessments to measure student outcomes.

Competency 10 states, “The leader understands the issues related to ethical and inappropriate use of student assessment and protects students and staff from such misuse” (Chappuis et al., 2004, p. 289). This involves ensuring assessment data are used to make appropriate interpretations based on student performance while following ethical practices, such as confidentiality and ethical test administration.

**How We Use Assessment as Instruction and Involve Students in the Process**

This domain encompasses competency two, “The leader understands the principals of assessment for learning and works with staff to integrate them into classroom instruction” (Chappuis et al., 2004, p. 167). There are nine principles that underpin this competency:

1. Teachers understand and can articulate in *advance of teaching* the achievement targets students are to hit.
2. *Students are informed regularly* about those targets in terms they can understand, in part throughout the study of the criteria by which their work will be evaluated, and samples of high-quality work.
3. *Students can describe what targets they are to hit* and what comes next in their learning.
4. Classroom teachers can transform those targets into *dependable assessments* that yield accurate information.
5. Both the teacher and the student use classroom assessment information to revise and guide teaching and learning.

6. Feedback given to students is descriptive, constructive, frequently, and timely; helping students identify their strengths and know how to plan and improve their work.

7. Students are actively, consistently, and effectively involved in assessment, including learning to manage their own learning through the skills of self-assessment.

8. Students actively, consistently, and effectively communicate with others about their achievement status and improvement.


Rationale for competency two is addressed within the research on formative assessment (e.g., Black & Wiliam, 2010). This domain specifically addresses the skills and strategies needed for classroom teachers to improve student outcomes through the formative assessment process that is intertwined within the teaching and learning process (Chappuis et al., 2004). “Teachers who develop useful assessments, provide corrective instruction, and give students second chances to demonstrate success can improve their instruction and help students learn” (Guskey, 2003, para. 1). Popham (2003) described the types of decisions that can be made using assessments including information about what students already know, the curricular objectives, the time needed to teach content, and the impact of the instruction (pp. 5-6).
The following sections describe the standards that dictate quality assessment. Additionally, the evolution of formative assessment is described, as well as the impact of assessment on student achievement.

**Classroom Assessment Standards**

To guide an understanding of what criteria constitute quality assessments, The Joint Committee on Standards for Educational Evaluation (2015) developed Classroom Assessment Standards for PreK-12 Teachers. These standards have been categorized into three overarching areas, including Foundations, Use, and Quality.

**Foundations Standards**

The Foundations standards address the need for assessment to guide instructional decisions. The Foundations standards also emphasize the function of curricular alignment. The Foundations standards underscore the need for all stakeholders to recognize the instructional purpose behind classroom assessment. The Foundations standards state:

- **F 1 Assessment Purpose**: Classroom assessment practices should have a clear purpose that supports teaching and learning.
- **F 2 Learning Expectations**: Learning expectations should form the foundation for aligning classroom assessment practices with appropriate instruction and learning opportunities for each student.
- **F 3 Assessment Design**: The types and methods of classroom assessment used should clearly allow students to demonstrate their learning.
F 4 Student Engagement In Assessment: Students should be meaningfully engaged in the assessment process and use of the assessment evidence to enhance their learning.

F 5 Assessment Preparation: Adequate teacher and student preparation in terms of resources, time, and learning opportunities should be part of classroom assessment practices.

F 6 Informed Students and Parents/ Guardians: The purposes and uses of classroom assessment should be communicated to students and, when appropriate, parents/ guardians. (Joint Committee for Standards for Educational Evaluation, 2015, Foundations section)

These standards provide the background for classroom assessment, its overarching purpose, and the outcomes of assessment (Joint Committee for Standards for Educational Evaluation, 2015).

**Use Standards**

The standards related to assessment use indicate that assessments should be used for the primary purpose of improved student outcomes. Assessments should be used to inform the teaching and learning process for students and teachers. Additionally, student assessments should be used in a manner that can be easily understood by intended stakeholders. The Use standards specifically state:

U 1 Analysis of Student Performance: The methods for analyzing evidence of student learning should be appropriate for the assessment purpose and practice.

U 2 Effective Feedback: Classroom assessment practices should provide timely and useful feedback to improve student learning.
U 3 Instructional Follow-Up: Analysis of student performance should inform instructional planning and next steps to support ongoing student learning.

U 4 Grades and Summary Comments: Summative grades and comments should reflect student achievement of the learning expectations.

U 5 Reporting: Assessment reports should be based on a sufficient body of evidence and provide a summary of a student’s learning in a clear, timely, accurate, and useful manner. (Joint Committee for Standards for Educational Evaluation, 2015, Use section)

These standards provide direction about the ways in which classroom assessments should be used, with regard to the formative and summative feedback that is provided to students and other stakeholders. (Joint Committee for Standards for Educational Evaluation, 2015)

**Quality Standards**

The standards related to assessment quality govern the reliability and validity of assessments. Quality standards affect the inferences that can be drawn about student learning based on the evidence gathered from an assessment. Quality standards also dictate the need for assessments that minimize threats to bias and distortion. Quality assessments are regularly updated to ensure alignment and appropriateness. These Quality standards specifically state:

Q 1 Cultural and Linguistic Diversity: Classroom assessment practices should be responsive to and respectful of the cultural and linguistic diversity of students and their communities.
Q 2 Exceptionality and Special Education: Classroom assessment practices should be appropriately differentiated to meet the specific educational needs of all students.

Q 3 Unbiased and Fair Assessment: Classroom assessment practices and subsequent decisions should not be influenced by factors unrelated to the intended purposes of the assessment.

Q 4 Reliability and Validity: Classroom assessment practices should provide consistent, dependable, and appropriate information that supports sound interpretations and decisions about each student’s knowledge and skills.

Q 5 Reflection: Classroom assessment practices should be monitored and revised to improve their overall quality. (Joint Committee for Standards for Educational Evaluation, 2015, Quality section)

**History of Formative Assessment**

An understanding of the role and implications of formative assessment has developed over time in the literature. Scriven (1966) distinguished between the role and purposes of formative and summative evaluation. Summative evaluation was likened to a “terminal evaluation” with regard to the effectiveness of an intervention (p. 5). Conversely, formative evaluation was described as a process that practitioners “automatically” engage in an effort to make adjustments and improvements based on “feedback” (p. 6). This conceptualization of the differences between formative and summative assessment suggest that formative evaluation involves an improvement process, whereas summative evaluation results in a judgment regarding effectiveness.
Although previously used interchangeably in the literature, Bloom (1968) distinguished among the roles of measurement, evaluation, and assessment. Evaluation involves “the identification of learning experiences and educative environments which produce significant changes in individuals and for the creation of instruments and methods of testing which will best reveal these changes” (p. 9). The primary objective of measurement entails “a small number of dimensions or measures which will completely account for the variance of a criterion when put together in some additive or summative combination” (p. 4). Assessment involves “the search for evidence on both individual and environment” (p. 12). Assessment can be comprised of qualitative and quantitative information, and a variety of instruments can be employed to collect such evidence. The specific focus of assessment is the interplay between individual people and their respective contexts.

This understanding of evaluation was extended with Natriello’s (1987) literature on evaluation. Natriello’s (1987) review provided a conceptual framework for evaluation processes. It began with “establishing the purposes for evaluating students” followed by “assigning tasks to students,” “setting criteria for student performance,” and “setting standards for student performance” (p. 156). Next steps involved “sampling information on student performance,” “appraising student performance,” “providing feedback to student performers,” and finally “monitoring outcomes of the evaluation of students” (p. 156). The cyclical nature of this process then began again with determining the reason for assessing students. His review further highlighted four competing functions of evaluation, which included “certification, selection, direction, and motivation” (p. 157). This extensive review of the literature was weakened as a result of methodological flaws.
within the studies included within the review. Many of the studies included lacked empirical evidence that supported the impact of classroom evaluation on student achievement.

Crooks (1988) also reviewed the literature on classroom evaluation. This review defined classroom evaluation as:

activities that students undertake as an integral part of the educational programs in which they are enrolled. These activities may involve time spent both inside and outside the classroom. This definition includes tasks such as formal teacher-made tests, curriculum-embedded tests (including adjunct questions and other exercises intended to be an integral part of learning materials), oral questions asked of students, and a wide variety of other performance activities (cognitive and psychomotor). It also includes assessment of motivational and attitudinal variables and of learning skills. (p. 439)

Like Natriello (1987), findings from this review identified that the literature failed to receive the necessary attention in analyzing the impact of classroom evaluation on student performance and outcomes (Crooks, 1988). Additionally, this review highlighted the importance of the role in evaluation to provide feedback to students within the learning process, as opposed to a measure to evaluate student performance. The review emphasized the need to provide more feedback to students regarding progress as opposed to summative appraisals of performance. The authors cited evaluation as “one of the most potent forces influencing education” (p. 467). These reviews served as the springboard for Black and Wiliams’ (1998) literature review on the impact of formative assessment on student learning.
Impact of Formative Assessment on Student Achievement

Black and Wiliam’s (2010) review highlighted the link between classroom learning and assessment. In this review, assessment was defined as “all those activities undertaken by teachers—and by their students in assessing themselves—that provide information to be used as feedback to modify teaching and learning activities” (p. 82). The term formative assessment was used, which describes a process in which teaching strategies are modified as a response to what students have or have not mastered.

The impact of formative assessment yielded effect sizes that ranged between 0.4 and 0.7 (Black & Wiliam, 2010, p. 83). In addition to improving student outcomes, formative assessment also became a powerful tool for providing feedback to students about his or her performance. Formative assessment should provide specific feedback about a student’s work and the ways in which he or she can improve.

Multiple challenges exist in formative assessment practices. One of these challenges is that current assessments do not lead to deeper levels of student learning. (Black & William, 2010). Additionally, when using grades along with assessment, students have a tendency to devalue the feedback provided. And finally, feedback often satisfies “managerial functions” in lieu of learning (p. 84). Another obstacle in effective formative assessment is that teachers do not have a solid understanding of formative assessment and how to effectively use it to promote student learning (Black & Wiliam, 1998). Additionally, teachers often employ a “normative rather than a criterion approach,” which encourages comparisons among students as opposed to focusing on individual progress and growth (Black & Wiliam, 1998, p. 18).
The teacher plays a pivotal role in assessment for learning. In order to create a classroom environment that promotes assessment as an integral part of the learning process, Stiggins (2005) recommends that teachers:

1. Become competent masters themselves of each of the standards their students are to master.
2. Understand how those standards transform into the curriculum that forms the scaffolding students will climb on their journey up to each standard.
3. Transform classroom-level achievement targets into student-friendly versions.
4. Transform the classroom targets into high-quality classrooms assessments capable of accurately reflecting student achievement.
5. Use those assessments over time in collaboration with their students to help motivate them to keep learning. (p. 1)

With a strong literature base supporting the integration of formative assessment to improve student learning, principals must ascertain how to balance formative assessment within their system of assessment. In doing so, there must be a clear reason for each assessment, and the assessment types must correspond to the data that is needed to make appropriate instructional decisions and inferences.

**Balanced Assessment Approach**

Assessment leaders are responsible for ensuring that the reason for assessing is clear. In addition to understanding the purpose for the assessment, it should also be clear how the assessment will be used in decision making. Balanced assessment systems will rely on varied forms of data that support instructional decision making (Stiggins, 2008; Huebner, 2009; Cizek, 1995; Chappuis et al., 2004). In essence, “the goal of a balanced
The assessment system is to ensure that all assessment users have access to the data they want when they need it, which in turn directly serves the effective use of multiple measures” (Chappuis, Chappuis, & Stiggins, 2009, p. 17).

A balanced assessment system should be inclusive of multiple measures. “The challenge for schools is designing a balanced assessment system using the strengths of summative, interim, and formative assessments to address instructional, accountability, and learning needs” (Huebner, 2009, p. 85). Chappuis et al. (2009) described the different use of assessments based on their frequency. For instance, ongoing classroom assessments provide regular information to students and teachers about student progress and may be formative or summative in nature. Benchmark assessments may be used to identify students in need of remediation but may also support the goals of program evaluation. Additionally, standardized assessments also provide accountability information and may serve administrative functions. In some instances, the line between formative and summative assessments may be blurred as some assessments may serve multiple purposes. “Formative and summative assessments support each other and should be viewed as in sync. They can be the exact same thing—only the purpose and time of the assessment determine its label” (Burke, 2010, p. 24).

Summary

Rigorous state standards place accountability demands on principals. As a result of challenging standards and assessments, instructional leadership is warranted, specifically assessment leadership. Assessment literacy is a prerequisite to assessment leadership. Assessment literacy requires an understanding of the nature of and use of assessments and standards for quality assessment creation.
CHAPTER THREE: METHODOLOGY

The literature on formative assessment supports its positive impact on student achievement when teachers effectively employ assessment for learning principles (Black & Wiliam, 1998, 2010). As the instructional leaders within a building, principals are charged with the continuous growth and development of staff with a continued focus on student outcomes (Leithwood et al., 2004; Hallinger, 2005). As assessment leaders, principals must be cognizant of what constitutes quality assessment and provide support to teachers in the implementation of assessment for learning practices (Chappuis et al., 2004).

The purpose of this study was to investigate the level of assessment literacy of building principals in Virginia as well as describe their support of assessment for learning practices within their school. Previous research specifically analyzing assessment literacy and assessment practices of Virginia principals has not been conducted in a post-NCLB context. There have been nationwide samples of principals’ assessment literacy (Impara et al., 1993) and contextual studies of Virginia principals’ assessment literacy (Impara & Plake, 1995); however, the relationship between principals’ assessment literacy and their assessment leadership practices that support assessment for learning practices have not been examined in this context.
The findings of this research will serve to guide professional development for administrators as well as inform principal preparation programs to ensure principals are equipped to serve as assessment leaders within their respective contexts. Additionally, because of the potential for formative assessment to improve student outcomes when implemented with fidelity, it examined the degree to which principals supported assessment for learning practices within their buildings. The study provided a rich description of the current state of assessment leadership practices within Virginia public schools.

**Research Design**

This study employed a mixed methods design. The first research question stated:

*To what degree are Virginia principals knowledgeable of classroom assessment practices as measured by the Assessment Literacy Inventory?* This question was addressed through quantitative measures. Principals were administered a two-part online survey. The first part of the survey collected relevant demographic information, see Appendix A. Participants were asked to provide their years of experience as a classroom teacher, years of experience as an administrator, and primary method for training in assessment. Additionally, school specific information was collected, including school level (elementary versus secondary) and school accreditation status. The second part of the online survey was used to evaluate principals’ knowledge of classroom assessment.

Principals were administered the revised Assessment Literacy Inventory, see Appendix B (Mertler & Campbell, 2005). The survey instrument has previously established reliability measures. The intent of the survey was to analyze and describe principals’ knowledge of classroom assessment practices.
The second research question stated: *What are the differences, if any, among principals’ assessment literacy related to level assignment (elementary versus secondary) and type of assessment training?* The two parts of this question were also answered through analysis of survey information.

A qualitative design was used to investigate the practices principals employ that support assessment for learning within their setting. The third research question stated: *What is the relationship between principals’ knowledge of classroom assessment practices and leadership practices that support assessment for learning?* Table 2 summarizes data collection and data analysis by research question.

**Table 2**

<table>
<thead>
<tr>
<th>Question</th>
<th>Data collection</th>
<th>Data analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1: To what degree are Virginia principals knowledgeable of classroom assessment practices as measured by the Assessment Literacy Inventory?</td>
<td>Revised Assessment Literacy Inventory (ALI) (Mertler &amp; Campbell, 2005)</td>
<td>Descriptive statistics</td>
</tr>
<tr>
<td>Question 2: What are the differences, if any, among principals’ assessment literacy related to level assignment (elementary versus secondary) and type of assessment training?</td>
<td>Revised Assessment Literacy Inventory (ALI) (Mertler &amp; Campbell, 2005) and survey using demographic questions</td>
<td>t-test ANOVA</td>
</tr>
<tr>
<td>Question 3: What is the relationship between principals’ knowledge of classroom assessment practices and leadership practices that support assessment for learning?</td>
<td>Semi-structured interview</td>
<td>Coding</td>
</tr>
</tbody>
</table>

Principals were interviewed and asked questions about their school’s assessment practices using a semi-structured interview. The interview focused specifically on two
domains of the Chappuis (2004) framework, specifically: Knowing What to Teach and How to Assess and How We Use Assessment as Instruction and Involve Students in the Process.

The combination of quantitative information from the survey as well as qualitative information from the interview provided insight into the assessment knowledge and practices of Virginia principals.

**Research Questions**

Question 1: To what degree are Virginia principals knowledgeable of classroom assessment practices as measured by the Assessment Literacy Inventory?

Question 2: What are the differences, if any, among principals’ assessment literacy related to level assignment (elementary versus secondary) and type of assessment training?

Question 3: What is the relationship between principals’ knowledge of classroom assessment practices and leadership practices that support assessment for learning?

**Participants**

The Virginia Department of Education Public School Directory contained a list of all public schools in Virginia. Only traditional public schools were invited to participate. All Virginia public school principals were invited to participate in the study. There were 1152 elementary schools, 299 high schools, and 303 middle schools. Combined schools were excluded from this study. Non-traditional public schools, including Governor’s Schools, alternative schools, vocational schools, and charter schools were excluded from participating in this study. The email addresses of principals were obtained using an online company, *The Email List*. The list included principals’ names, the school name,
address, zip code, type (elementary, middle, and high), phone number, fax number, and the principal’s email addresses. The list was cross-referenced with the Virginia Department of Education Public School Directory for accuracy of information. In instances of discrepancies between the two databases, the division or school webpages were accessed to verify the principal’s information. All K-12 public school principals that met the criteria for this study were sent an informative email describing the study and encouraging participation in the study. In total, 1742 emails were distributed to Virginia principals, two emails failed, there were 22 duplicates, and 32 were returned as undeliverable. Four participants completed the survey using an anonymous link. One hundred thirty-three participants completed the survey, with an overall response rate of 7.6%. Participants who completed 85% of the survey were included in data analysis. This threshold was chosen because it ensured participants had not skipped more than five questions, which would have negatively skewed the results because an omitted response received a score of zero. Eleven participants were excluded from data analysis because they did not complete a minimum of 85% of the Assessment Literacy Inventory. Some of the principals declined participation due to division protocols requiring division permission and protocols for participation in scholarly research; therefore, these participants and divisions were excluded from follow-up emails.

The survey was distributed using email via Qualtrics, an online data collection application supported by The College of William and Mary. There were no anticipated risks associated with participation in this study. Participation in this study entered principals into a drawing for one of five $100 Visa gift cards. Principals were sent a reminder email if the survey had not been completed within one week. The survey also
asked if participants consented to be contacted for a phone interview if selected. All survey responses remained confidential but for those that agreed to be contacted for a phone interview, they provided their name, phone number, and email address within the survey; however, the results of their survey also remained confidential.

Participation in the qualitative interviews was selected based on purposeful criterion sampling. Two criteria were considered when selecting candidates, including principals’ ALI scores as well as their level assignment. Principals scoring highest on the ALI were selected for interviews and principals who scored lowest on the ALI were selected for interviews while also ensuring there was equal representation of elementary and secondary principals within the sample. If principals declined participation in the interview, the next participants within the score range on the list (either scoring highest or lowest) were invited to participate. The intention behind this purposeful sampling method was to analyze the assessment practices of principals who ranked higher in their knowledge of assessment literacy and also compare them to practices of principals who ranked lower in their knowledge of assessment literacy. All identifying names and information were removed for the purpose of reporting. Participants in the interview phase of the study were also be entered into a second drawing for one of two $100 Visa gift cards, in which their chances of winning were one in six.

**Instrumentation**

**Survey.** The survey instrument used within this study was Mertler and Campbell’s (2005) revised Assessment Literacy Inventory (ALI). Demographic questions within the survey can be found in Appendix A and a copy of the survey can be found in Appendix B. Permission to use the instrument was provided by Dr. Craig Mertler. The
ALI provided specific information regarding educators’ strengths or deficits within the seven Standards for Teacher Competence in the Educational Assessment of Students. The purpose of the measure is described by the authors below:

The Assessment Literacy Inventory provides a mechanism for educators to measure assessment literacy (i.e., their knowledge of and abilities to apply assessment concepts and techniques to inform decision-making and guide practice). Considering the current state of high-stakes accountability in education, the ALI could provide school districts an effective, as well as efficient way to allocate resources for developing or otherwise selecting teacher professional development opportunities on the topic of classroom assessment. Because the ALI is based entirely on the Standards for Teacher Competence in the Educational Assessment of Students, its use could provide educational leaders with a diagnostic tool for identifying areas (i.e., as represented by a given standard) where teachers may be deficient and in need of further remediation and training. (Mertler & Campbell, 2005, pp. 15-16)

Information collected from the ALI provided an overall M and SD of assessment literacy. Scores ranged from [0-35] for an overall assessment literacy score. For each standard, participants could have received a maximum score of five and a minimum score of zero. The ALI includes five classroom-based scenarios, providing a more authentic context to the examinee. “The ALI consisted of 35 items, embedded within five classroom-based scenarios, featuring teachers who were facing various assessment-related decisions” (Mertler & Campbell, 2005, p. 9). The instrument was developed because “the original instrument was difficult to read, extremely lengthy, and contained items that were
presented in a decontextualized way. The mean item difficulty for items on the ALI was .681; however, difficulty values ranged from .212 to .992.

The revised ALI was administered to 249 pre-service teachers across two institutions and had an overall reliability of $r = .74$ (Mertler & Campbell, 2005). The decision to use the ALI was made because of its reliability measures, when compared to other instruments. Additionally, its use of scenarios provided a more contextualized measure of assessment literacy. Although designed for pre-service teachers, this inventory was used with principals because principals as instructional leaders are responsible for overseeing and supporting the teaching and learning process within their schools.

The psychometric principles of other instruments of assessment literacy were another consideration in the selection of the instrument for the purpose of this study. DeLuca et al. (2016b) provided a comparison of these instruments, as described below. The original measure of assessment literacy, the TALQ, when administered to 555 in-service teachers and had a reliability score of $r = 0.54$ (Plake, Impara, & Fager, 1993). Mertler (2003) developed the Classroom Assessment Literacy Inventory (CALI), which was inclusive of 35 content-based items. It was administered to 197 in-service teachers, with a reliability score of $r = 0.57$, and it was also administered to 67 pre-service teachers, with a reliability score of $r = 0.74$. Mertler and Campbell (2005) revised a previous version of the ALI, known as the Revised Assessment Literacy Inventory (ALI), which includes 35 scenario-based items and is also based on the 1990 Standards.

The Approaches to Classroom Assessment Inventory (ACAI), currently in its infancy stages of development, presents an opportunity to employ updated instruments to
measure assessment literacy, as it is aligned to the Joint Committee on Standards for Educational Evaluations (2015)’s updated Classroom Assessment Standards. DeLuca, LaPointe-McEwan, and Luhanga (2016a) identified the urgency for revised assessment literacy instruments because “current assessment literacy instruments do not fully reflect current transformations in the assessment landscape and remain predicated on dated standards for teacher classroom assessment practice” (p. 2). While the revised instrument represents great potential as a measure of assessment literacy, it was unable to be used in this study because it was in its early stages of refinement within the time frame of this study. In the absence of a revised instrument, the decision was made to use the ALI based on its known psychometric principles and use of contextual scenarios.

The literature identified eight different instruments to measure assessment literacy between 1993 and 2012 (DeLuca et al., 2016b). Of the eight instruments, six instruments, including the ALI, are aligned to the seven Standards for Teacher Competence in the Educational Assessment of Students (American Federation of Teachers, National Council on Measurement in Education, & National Education Association, 1990). These standards have been used as the guiding framework for other instruments measuring assessment literacy. The earliest measure of assessment literacy, the Teacher Assessment Literacy Questionnaire (TALQ) (Plake et al., 1993) used these standards as a basis for its creation.

The ALI contains five classroom scenarios for a total of 35 items. Seven multiple-choice items follow each scenario, and each question is aligned to one of the seven Standards for Teacher Competence in the Educational Assessment of Students (Mertler & Campbell, 2005). The standards state:
1. Teachers should be skilled in choosing assessment methods appropriate for instructional decisions.

2. Teachers should be skilled in developing assessment methods appropriate for instructional decisions.

3. The teacher should be skilled in administering, scoring and interpreting the results of both externally-produced and teacher-produced assessment methods.

4. Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement.

5. Teachers should be skilled in developing valid pupil grading procedures, which use pupil assessments.

6. Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.

7. Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information. (AFT, NCME, & NEA, 1990, pp. 4-6)

Table 3 summarizes the alignment of standards with items on the ALI (Mertler & Campbell, 2005). The decision to focus on the domains Knowing What to Teach and How to Assess and How We Use Assessment as Instruction and Involve Students in the Process was made because each of the standards assessed on the ALI correlated with one of these two domains. Although the domains How We Monitor Our Practices and How We Communicate about Student Learning were addressed within the ALI, all of the standards addressed one of first two domains. Table 3 represents a crosswalk between the seven

Table 3

Alignment of Standards with Items on the ALI and Alignment of Standards to Chappuis (2004) Framework

<table>
<thead>
<tr>
<th>Standard</th>
<th>Items</th>
<th>Domain(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1, 8, 15, 22, 29</td>
<td>Knowing What to Teach and How to Assess</td>
</tr>
<tr>
<td>2</td>
<td>2, 9, 16, 23, 30</td>
<td>Knowing What to Teach and How to Assess</td>
</tr>
<tr>
<td>3</td>
<td>3, 10, 17, 24, 31</td>
<td>Knowing What to Teach and How to Assess</td>
</tr>
<tr>
<td>4</td>
<td>4, 11, 18, 25, 32</td>
<td>How We Use Assessment as Instruction and Involve Students in the Process</td>
</tr>
<tr>
<td>5</td>
<td>5, 12, 19, 26, 33</td>
<td>Knowing What to Teach and How to Assess &amp; How We Monitor our Practices</td>
</tr>
<tr>
<td>6</td>
<td>6, 13, 20, 27, 34</td>
<td>How We Use Assessment as Instruction and Involve Students in the Process &amp; How We Communicate About Student Learning</td>
</tr>
<tr>
<td>7</td>
<td>7, 14, 21, 28, 35</td>
<td>Knowing What to Teach and How to Assess</td>
</tr>
</tbody>
</table>

**Interview.** The interviews were guided using a semi-structured interview protocol. Participants were asked a series of questions based on their implementation of assessment-related activities. Each of the questions aligned to one of the 10 competencies presented by Chappuis et al. (2004) that reflect assessment leadership. The questions reflect standards within the domain of *Knowing What to Teach and How to Assess* and *How We Use Assessment as Instruction and Involve Students in the Process* (Chappuis, 2004). The questions were reviewed by a panel of assessment experts for evidence of content and construct validity and modifications to the protocol were made based upon
recommendations of the panel. The instrument was piloted with a former Director of Curriculum, Assessment and Technology and further revised based on feedback. A copy of the interview protocol can be found in Appendix C. See Table 4 for questions as well as their alignment to specific competencies.
Table 4

*Interview Questions*

<table>
<thead>
<tr>
<th>Questions</th>
<th>Competency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What strategies, if any, do you employ to help staff use assessment to support student learning?</td>
<td>2</td>
</tr>
<tr>
<td>2. What evidence, if any, do you look for in the classroom to determine if assessment is guiding the learning?</td>
<td>2</td>
</tr>
<tr>
<td>3. What evidence do you look for to determine if an assessment is aligned to achievement targets?</td>
<td>3</td>
</tr>
<tr>
<td>4. What strategies, if any, has your school employed to ensure alignment of standards, learning intentions, and assessments?</td>
<td>3</td>
</tr>
<tr>
<td>5. What professional development opportunities, if any, are provided for teachers to contribute to their use of sound assessment practices?</td>
<td>5</td>
</tr>
<tr>
<td>a. Probing question: What is your role in the professional development?</td>
<td></td>
</tr>
<tr>
<td>b. Probing question: What is your role after the professional development?</td>
<td></td>
</tr>
<tr>
<td>6. How are specific professional development assessment topics chosen?</td>
<td>5</td>
</tr>
<tr>
<td>7. What formative assessment practices are consistently used as part of your school’s overall assessment system, if any, and what is the role of the assessment?</td>
<td>9</td>
</tr>
<tr>
<td>8. What summative assessment practices are consistently used as part of your school’s overall assessment system, if any, and what is the role of the assessment?</td>
<td>9</td>
</tr>
<tr>
<td>9. What strategies, if any, do you use to prevent unethical and inappropriate administration of assessments and unethical and inappropriate use of assessments and assessment results?</td>
<td>10</td>
</tr>
<tr>
<td>10. What strategies, if any, do you use to ensure stakeholders, including students, parents, and school community make appropriate interpretations from various assessments?</td>
<td>10</td>
</tr>
</tbody>
</table>
**Procedure**

**Survey.** A total of 1742 emails were distributed to Virginia principals, two emails failed, there were 22 duplicates, and 32 were returned as undeliverable. K-12 Virginia principals using Qualtrics. Participants were provided an opportunity to provide informed consent before beginning the survey. An email was sent to all participants’ school web address soliciting participation in the study. The introductory email contained the purpose of the study, an opportunity to provide informed consent, confidentiality information, and a link to the actual survey. Within the consent section of the survey, participants also had a space to agree to participate in a follow-up phone interview.

Participants were given approximately three weeks to participate in the survey. The first email was sent to participants soliciting participation in the survey. A second email was sent one week after the initial email in which participation was again requested and participants were also acknowledged and thanked if they have already participated and given an opportunity to request results of the study.

**Interview.** After the conclusion of the survey window, participants who provided consent to participate in the phone interview were selected for participation in the interviews. Six participants scoring in the higher range on the ALI were selected for the interview, and six participants scoring in the lower range on the ALI were selected for the interview. The intention of using purposive sampling for selection of interviewees was to have representation of principals spanning the spectrum of assessment literacy. Additionally, both secondary and elementary principals were selected for the interview phase, with the intention to have equal representation of both groups reflected within the sample.
Analysis

Survey. All survey information was collected using Qualtrics and subsequently analyzed using the Statistical Package for Social Science (SPSS). SPSS was used to present and analyze demographic information about the participants. Additionally, all statistical analyses were made using SPSS.

Part I of the survey was analyzed using the demographic information collected. Participants were asked regarding their years of experience as a classroom teacher, and descriptive information was provided. Clark, Martorell, and Rockoff (2009) did not find a relationship between principal effectiveness and years of teaching experience; however, this information was collected to provide additional information about the background experiences of the sample of principals. Information regarding principals’ years of experience in education was, however, collected in Perry (2013) for the principals surveyed. Additionally, this information may be informative because the ALI was originally designed for use with pre-service and in-service teachers.

Participants were also asked regarding years of experience as an administrator, and descriptive information was provided. Years of experience as an administrator was collected because years of principal experience were related to improved student performance (Clark et al., 2009). Again, this information was collected in Perry (2013) as it provided greater information regarding the sample of principals.

Participants also indicated their primary means of formal assessment training and descriptive statistics was provided. Impara and Plake (1995) found that 97.4% of administrators surveyed within their sample had taken a course in assessment; however, the format for the assessment training was not specified. Perry (2013) found almost split
results between administrators who had taken a course in assessment and those that had not. This question was designed to find out greater specificity about the type of training principals received in assessment and if certain methods of assessment training yielded differing scores on the ALI.

An analysis of variance was conducted to determine whether groups differed significantly in terms of assessment literacy according to predominant method of assessment training. If $F_{crit}$ had been significant at the $p < .05$, the post hoc comparisons would have been made using the Tukey HSD. The decision to use analysis of variance instead of multiple $t$-tests was made to minimize the threat of a Type I error.

Participants also denoted their school’s accreditation status during the 2016-2017 school year and descriptive information was provided. This information was collected as it provides information about the school’s overall performance, as measured by the Virginia accreditation ratings. Additionally, this information was compared to the state accreditation ratings across schools. Because of the variety of factors that may influence school accreditation, only descriptive information was provided. This information also helps describe the school context and performance that was collected within the sample.

Additionally, school specific information was collected, including school level (elementary vs. secondary). Impara and Plake (1995) found elementary administrators to be more knowledgeable of assessment than secondary administrators. Additionally, this information also provided additional contextual information about the sample. The ALI score for both levels was compared using $t$, $p < 0.05$. A $t$ test was chosen because only two groups were compared in this study, specifically elementary versus secondary.
The data were also used to identify strengths and weakness of specific areas of principals’ assessment literacy. In addition to overall performance on the ALI, this information was designed to help discern how principals perform within each of the seven standards.

**Interview.** The interviews were recorded using a voice recorder and then transcribed for analysis. A coding process was used to categorize information from the interviews. Four a priori codes were established prior to interviews, including: *assessment for learning, balanced assessment, support for teachers, and professional development*. Upon further refinement of interview questions, these codes were expanded to also include *alignment* because questions three and four aligned to competency three which addressed alignment. Additionally, the codes *support for teachers* and *assessment for learning* were consolidated to align with competency two, and the revised code now states: *support of assessment for learning principles*.

The codes continued to expand and develop upon a review of the information and transcripts. A “selective or highlighting approach” was also employed, which involved multiple reading of the transcripts to identify themes and key information (Manen, 1990, p. 94). Multiple readings of the interviews resulted in themes in which ideas or concepts were grouped similarly for shared meaning.

To protect the quality and rigor of the interview, three validation strategies were employed. These include clarifying researcher bias, member checking, and providing a rich, thick description. To clarify researcher bias, the researcher drafted a researcher as instrument statement to identify the personal biases associated with conducting qualitative interviews, see Appendix D. Identifying researcher bias and providing a
statement prior to conducting the interviews enabled the reader to identify subjectivity that may influence the validity of results (Merriam, 1998). Additionally, after the interviews were transcribed, an opportunity for member checking was employed to ensure that participants accurately represented their perspectives (Merriam, 1998). Both member checking and clarifying researcher bias were strategies to strengthen the internal validity of the study. To address external validity, a rich, thick description accompanied each interview with the intention of providing the reader with an understanding of the context of the participant (Merriam, 1998). A rich, thick description is imperative because school contexts as well as leadership styles may differ within varying settings, and it will impact the transferability of results to a different context and leader.

Ethical Considerations

This study was submitted, reviewed, and approved by the W&M Education Institutional Review Board (EDIRC). There were no anticipated risks associated with participation in this study. All survey results remained confidential and only known to the researcher. Upon completion of the survey, all survey participants were assigned a number that was subsequently used for data analysis. The key linking survey results to the participant was only be available to the researcher. Participants who agreed to be contacted for a follow-up interview provided their name, phone number, and email; however, these individuals’ actual scores on the ALI remained confidential and only available to the researcher. Moreover, all data and records were stored on a password-protected computer. Participation was voluntary. Participants were free to withdraw at any time without penalty or loss of benefits. Participants were permitted to skip any
question or opt to not participate in the follow-up interview portion of the investigation. See Appendix E for the Research Participation Informed Consent Form

**Delimitations**

Assessment leadership is a relatively underdeveloped area in the literature. The decision to explore assessment literacy arose from the accountability demands placed on the principalship. Because accountability systems differ from state-to-state, the decision was made to explore assessment leadership within the context of Virginia standards. Additionally, because of the challenges Virginia schools and principals face in light of increased rigor within standards and assessments, the topic warranted further investigation. Although assessment literacy has been explored on a national scale and within Virginia principals in a pre-NCLB context, the implications of principals’ assessment literacy in this context have not been explored, specifically how their assessment literacy relates to their support of assessment for learning practices. The Chappuis et al. (2004) competencies for assessment leadership outline the competencies needed for assessment leadership. Additionally, this study would inform the professional development needs for Virginia principals.

**Limitations**

One specific limitation of this study was that the context was restricted to Virginia principals. Virginia public schools operate using the Standards of Learning curriculum and achievement is measured using Standards of Learning assessment. Because of this, generalizability may be limited in other states operating within a different accountability system.
Another limitation of this study was survey response bias. There may be a certain tendency or bias based on those who agree to complete the survey, as well as those who were willing to participate in the interview portion of the study.

Because the survey was administered in a non-controlled setting, participants may have had access to resources or information that would skew their results on the ALI. This may positively skew the results if participants use outside resources to assist them with item responses on the survey.

Additionally, this study excluded principals in non-traditional public schools, including Governor’s Schools, alternative schools, vocational schools, and charter schools. Additionally, combined schools were excluded; therefore, the results of this study will not reflect the assessment literacy or practices of these principals.

This study was also limited to only principals’ knowledge of assessment and does not take into consideration the ways in which other stakeholders use assessment information. It is possible that the viewpoints and practices of a principal may not be indicative of the knowledge or practices of the rest of the faculty.

**Assumptions**

It is assumed that all principals would complete the ALI without outside assistance or resources. If this assumption was not met, interpretations of the ALI will be invalid. It will provide a measure of a principal’s personal assessment literacy only if the assessment is taken without outside assistance.

It is assumed that participants honestly reported their support and utilization of assessment for learning principles within their respective contexts. It was important that
principals were honest and clear about how assessment is used within their building and they not cloud their responses with what they would instead like to implement.

It was assumed that the principal is the primary instructional leader who determines which assessment practices are employed within the school. It is assumed that the principal has the authority to make instructional decisions and oversee the curriculum, instruction, and assessment within their context.

Summary

This study was designed to measure Virginia principals’ overall knowledge of classroom assessment practices. Additionally, this study was designed to determine specific strengths and weaknesses of principals’ assessment literacy across the seven Teacher Standards in Student Assessment. This study was also designed to use qualitative interviews to develop an understanding of the ways in which principals with higher and lower levels of assessment literacy engage in assessment leadership practices that support assessment for learning within their respective settings.
CHAPTER FOUR: RESULTS

The results section is a summary of major findings from this study. This section is organized by first presenting demographic information, followed by a presentation of analyses for each major research question.

Demographic Information

All Virginia principals in traditional, public school settings were included within this survey. Nontraditional schools (alternative, charter, combined, and vocational) were excluded from participation. An email distribution list was purchased and downloaded from the website: Email List. Subsequently, this list was cross-referenced with the VDOE Education Directory. Discrepancies between contact information contained within the purchased database and the VDOE Education Directory were verified on school and division webpages. Upon verification of accurate email addresses, 1742 emails were distributed to Virginia principals, two emails failed, there were 22 duplicates, and 32 were returned as undeliverable. Four participants completed the survey using an anonymous link. A total of 133 participants completed the survey, with an overall response rate of 7.6%. Participants who completed 85% of the survey were included in data analysis. This threshold was chosen because it ensured participants had not skipped more than five questions, which would have negatively skewed the results because an omitted response received a score of zero. Eleven participants were excluded from data
analysis because they did not complete a minimum of 85% of the Assessment Literacy Inventory (ALI). Some of the principals declined participation due to division protocols requiring division permission and protocols for participation in scholarly research; therefore, these participants and divisions were excluded from follow-up emails.

The revised ALI, with an overall reliability of $r = .74$, was used to assess principals’ assessment literacy (Mertler & Campbell, 2005). The decision to use the ALI was made because of its established reliability measures, when compared to other instruments. Additionally, its use of scenarios provided a more contextualized measure of assessment literacy. Although designed for pre-service teachers, this inventory was used with principals because principals as instructional leaders are responsible for overseeing and supporting the teaching and learning process within their schools. Information collected from the ALI provided an overall $M$ and $SD$ of assessment literacy. Scores ranged from [0-35] for an overall assessment literacy score. Additionally, a mean and standard deviation for each standard was calculated.

Participants provided relevant demographic information before beginning the Assessment Literacy Inventory. As Table 5 shows, classroom experience varied widely with a minimum of no classroom experience to a maximum of 28.5 years of experience. The mean years of classroom experience was 11.51 years ($SD=5.50$), reflecting a wide range of teaching experience across participants. Table 5 describes participants’ years of classroom teaching experience.
Table 5

Participants’ Years of Classroom Teaching Experience

<table>
<thead>
<tr>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>122</td>
<td>0</td>
<td>28.50</td>
<td>11.51</td>
<td>5.50</td>
</tr>
</tbody>
</table>

Participants were additionally asked to provide their years of administrative experience. As Table 6 shows, administrative experience varied widely with a minimum of one year of experience to a maximum of 32 years of experience. The mean years of administrative experience was 11.09 years ($SD=5.16$) reflecting a wide range of years of administrative experience across participants.

Table 6

Participants’ Years of Administrative Experience

<table>
<thead>
<tr>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>122</td>
<td>1.00</td>
<td>32.00</td>
<td>11.09</td>
<td>5.16</td>
</tr>
</tbody>
</table>

Participants provided their school’s accreditation status. A majority of schools (85.2%) represented within the sample were Fully Accredited Schools. This mirrors the Virginia Department of Education 2016-2017 School Accreditation ratings, in which 82% of schools were fully accredited (VDOE, 2017b). Additionally, within the sample, 4.1% of schools were denied accreditation, which again paralleled the state comparison of 5% (VDOE, 2017b). Participants’ school accreditation status and associated mean scores on the Assessment Literacy Inventory are presented in Table 7.
Table 7

*Accreditation Status and Principals’ ALI Scores*

<table>
<thead>
<tr>
<th>Accreditation status</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully Accredited</td>
<td>104</td>
<td>22.84</td>
<td>3.33</td>
<td>85.2</td>
</tr>
<tr>
<td>Partially Accredited: Approaching Benchmark Pass Rate</td>
<td>4</td>
<td>24.00</td>
<td>4.83</td>
<td>3.3</td>
</tr>
<tr>
<td>Partially Accredited: Warned School-Pass Rate</td>
<td>5</td>
<td>24.40</td>
<td>2.88</td>
<td>4.1</td>
</tr>
<tr>
<td>Partially Accredited: Warned School-GCI</td>
<td>1</td>
<td>26.00</td>
<td>0.00</td>
<td>0.8</td>
</tr>
<tr>
<td>Partially Accredited-Reconstituted School</td>
<td>3</td>
<td>16.00</td>
<td>4.58</td>
<td>2.5</td>
</tr>
<tr>
<td>Accreditation Denied</td>
<td>5</td>
<td>23.40</td>
<td>4.97</td>
<td>4.1</td>
</tr>
</tbody>
</table>

**Principals’ Knowledge of Classroom Assessment Practices**

The first research question stated: *To what degree are Virginia principals knowledge of classroom assessment practices as measured by the Assessment Literacy Inventory?* Scores could range from zero to 35 on the ALI. The data reflected a range of composite scores between 12 and 31, with the mean score of 22.82 (SD=3.58). Table 8 provides participants’ overall performance on the Assessment Literacy Inventory.

Table 8

*Participants’ Mean Scores on ALI*

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>122</td>
<td>12.00</td>
<td>31.00</td>
<td>22.82</td>
<td>3.58</td>
</tr>
</tbody>
</table>

The range and frequency of ALI scores is indicative of a normal bell curve. Measures of central tendency reflect normal distribution of scores in which the median score is 23, the mode is 21, and the mean is 22.82. Table 9 provides the range of scores on the Assessment Literacy Inventory, as well as the frequency and percentage of participants who scored within the respective range.
Table 10 provides participants’ mean scores according to the seven Standards for Teacher Competence in the Educational Assessment of Students (Mertler & Campbell, 2005). The scores range from zero to five in each of the standards, as each standard is inclusive of five questions, therefore, one point is assigned for each correct response. A score of five is indicative of a perfect score within each of the standards.
Table 10

Principals’ Mean Scores by Standard

<table>
<thead>
<tr>
<th>Standard</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard 1</strong>: Teachers should be skilled in choosing assessment</td>
<td>3.38</td>
<td>1.04</td>
</tr>
<tr>
<td>methods appropriate for instructional decisions.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standard 2</strong>: Teachers should be skilled in developing assessment</td>
<td>2.52</td>
<td>1.05</td>
</tr>
<tr>
<td>methods appropriate for instructional decisions.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standard 3</strong>: The teacher should be skilled in administering, scoring</td>
<td>3.17</td>
<td>0.98</td>
</tr>
<tr>
<td>and interpreting the results of both externally-produced and teacher-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>produced assessment methods.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standard 4</strong>: Teachers should be skilled in using assessment results</td>
<td>3.70</td>
<td>0.85</td>
</tr>
<tr>
<td>when making decisions about individual students, planning teaching,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>developing curriculum, and school improvement.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standard 5</strong>: Teachers should be skilled in developing valid pupil</td>
<td>3.05</td>
<td>1.03</td>
</tr>
<tr>
<td>grading procedures, which use pupil assessments.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standard 6</strong>: Teachers should be skilled in communicating assessment</td>
<td>3.12</td>
<td>0.81</td>
</tr>
<tr>
<td>results to students, parents, other lay audiences, and other educators.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standard 7</strong>: Teachers should be skilled in recognizing unethical,</td>
<td>3.89</td>
<td>0.98</td>
</tr>
<tr>
<td>illegal, and otherwise inappropriate assessment methods and uses of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>assessment information</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11 summarizes the percentage correct by question for Standard One: *Teachers should be skilled in choosing assessment methods appropriate for instructional decisions.*

Question 1 assessed principals’ knowledge of *authentic assessment*. Question 8 required participants to choose among assessment strategies to assess students’ problem solving abilities. Question 15 assessed knowledge of multiple-choice and true-false assessments and included terminology such as validity and reliability. Question 15 was a relative weakness for participants within this standard. Question 22 involved assessing students’ writing skills, and required principals to choose from the following assessment types: selected response methods, true/false statements, completion items, and essay prompts.
Question 22 was a relative strength for participants within this standard. Question 29 included terminology such as: diagnostic assessment, informal assessment, standardized assessment, and summative assessment.

Table 11

*Percentage Correct by Question for Standard One: Choosing Assessment Methods*

<table>
<thead>
<tr>
<th>Question</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>52.5%</td>
</tr>
<tr>
<td>Question 8</td>
<td>81.1%</td>
</tr>
<tr>
<td>Question 15</td>
<td>40.2%</td>
</tr>
<tr>
<td>Question 22</td>
<td>89.3%</td>
</tr>
<tr>
<td>Question 29</td>
<td>74.6%</td>
</tr>
</tbody>
</table>

Table 12 summarizes the percentage correct by question for Standard Two: *Teachers should be skilled in developing assessment methods appropriate for instructional decisions*. Question two referenced accurate and consistent grading practices. Question two was a relative strength within this standard. Question nine assessed principals’ knowledge of using multiple choice assessments to predict performance on state assessments. Question 16 assessed knowledge of: item analysis, item difficulty values, item discrimination values, and reliability coefficients. Question 16 was a relative weakness within this standard. Question 23 referred to developing story-based math assessment questions. Finally, Question 30 assessed knowledge of item analysis and discrimination values.
Table 12

**Percentage Correct by Question for Standard Two: Developing Assessment Methods**

<table>
<thead>
<tr>
<th>Question</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 2</td>
<td>89.3%</td>
</tr>
<tr>
<td>Question 9</td>
<td>49.2%</td>
</tr>
<tr>
<td>Question 16</td>
<td>15.6%</td>
</tr>
<tr>
<td>Question 23</td>
<td>59.8%</td>
</tr>
<tr>
<td>Question 30</td>
<td>37.7%</td>
</tr>
</tbody>
</table>

Table 13 summarizes the percentage correct by question for Standard Three: *The teacher should be skilled in administering, scoring and interpreting the results of both externally-produced and teacher-produced assessment methods.* Question three related to comparison groups. Question 10 assessed knowledge of percentile rank and was a relative weakness. Question 17 related to score interpretation and was a relative strength within this standard. Question 24 addressed using standardized assessments, means, and standard deviations to interpret scores. Question 31 involved scoring responses using an analytic rubric.

Table 13

**Percentage Correct by Question for Standard Three: Administering, Scoring and Interpreting Results**

<table>
<thead>
<tr>
<th>Question</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 3</td>
<td>81.1%</td>
</tr>
<tr>
<td>Question 10</td>
<td>30.3%</td>
</tr>
<tr>
<td>Question 17</td>
<td>96.7%</td>
</tr>
<tr>
<td>Question 24</td>
<td>50.0%</td>
</tr>
<tr>
<td>Question 31</td>
<td>59.0%</td>
</tr>
</tbody>
</table>

Table 14 summarizes the percentage correct by question for Standard Four: *Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement.* Question
four was scored using a reverse scoring procedure. The original *Assessment Literacy Inventory* was designed to ask: *Which of the following is an inappropriate use of the results from this standardized math test?* However, the question was incorrectly entered, and instead asked: *Which of the following is an appropriate use of the results from this standardized math test?* As a result, all responses that included an appropriate use of the standardized math test were marked as correct. Question 4 related to use of standardized testing information. Question 11, a relative weakness of this standard, related to standard error of measurement. Question 18 assessed knowledge of criterion-referenced information. Question 25 related to the role of formative assessments in classroom instruction. Question 32 referred to alignment of instruction and assessment.

Table 14

*Percentage Correct by Question for Standard Four: Using Assessment Results When Making Decisions*

<table>
<thead>
<tr>
<th>Question</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 4</td>
<td>100.0%</td>
</tr>
<tr>
<td>Question 11</td>
<td>25.4%</td>
</tr>
<tr>
<td>Question 18</td>
<td>56.6%</td>
</tr>
<tr>
<td>Question 25</td>
<td>91.0%</td>
</tr>
<tr>
<td>Question 32</td>
<td>96.7%</td>
</tr>
</tbody>
</table>

Table 15 summarizes the percentage correct by question for Standard Five: *Teachers should be skilled in developing valid pupil grading procedures, which use pupil assessments*. Question 5 related to the concept of weighting assessments. Question 12 assessed principals’ knowledge of appropriate grading practices. Question 19 related to using multiple pieces of information to determine grades. This was a relative strength within this standard. Question 26 addressed consistent scoring practices. This was a
relative weakness within this standard. Question 33 referenced a criterion-referenced grading system.

Table 15

*Percentage Correct by Question for Standard Five: Developing Valid Pupil Grading Procedures*

<table>
<thead>
<tr>
<th>Question</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 5</td>
<td>67.2%</td>
</tr>
<tr>
<td>Question 12</td>
<td>68.9%</td>
</tr>
<tr>
<td>Question 19</td>
<td>84.4%</td>
</tr>
<tr>
<td>Question 26</td>
<td>16.4%</td>
</tr>
<tr>
<td>Question 33</td>
<td>68.0%</td>
</tr>
</tbody>
</table>

Table 16 summarizes the percentage correct by question for Standard Six: *Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.* Question 6 related to explaining the meaning of percentile. Question 13 related to explaining and comparing student performance across assessments. Question 20 addressed formal and informal assessments. Question 27 related to grade equivalency and was the overall lowest scoring item on the ALI as well as the relative weakness within this standard. Question 34 related to the concepts of raw scores and percentile ranks.

Table 16

*Percentage Correct by Question for Standard Six: Communicating Assessment Results*

<table>
<thead>
<tr>
<th>Question</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 6</td>
<td>99.2%</td>
</tr>
<tr>
<td>Question 13</td>
<td>72.1%</td>
</tr>
<tr>
<td>Question 20</td>
<td>69.7%</td>
</tr>
<tr>
<td>Question 27</td>
<td>0.80%</td>
</tr>
<tr>
<td>Question 34</td>
<td>70.5%</td>
</tr>
</tbody>
</table>
Table 17 summarizes the percentage correct by question for Standard Seven: *Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.* Question 7 related to using assessment information when making decisions about student learning. This was a relative strength within this standard. Question 14 related to unethical practices to increase student performance. This was a relative weakness within this standard. Question 21 related to unethical grading practices. Question 28 referred to standardization of practices during assessment administration. Question 35 referred to The Family and Education Rights and Privacy Act.

Table 17

<table>
<thead>
<tr>
<th>Question</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 7</td>
<td>98.4%</td>
</tr>
<tr>
<td>Question 14</td>
<td>54.1%</td>
</tr>
<tr>
<td>Question 21</td>
<td>67.2%</td>
</tr>
<tr>
<td>Question 28</td>
<td>77.9%</td>
</tr>
<tr>
<td>Question 35</td>
<td>91.0%</td>
</tr>
</tbody>
</table>

**Differences in Assessment Literacy**

The second question stated: *What are the differences, if any, among principals’ assessment literacy related to level assignment (elementary versus secondary) and type of assessment training?* Participants selected elementary, secondary, or other within the survey to describe their present school level. Five participants selected *Other*, and were subsequently categorized as elementary or secondary, as described by Table 18.
Table 18

School Level Designations for Participants Who Denoted Other for Level Assignment

<table>
<thead>
<tr>
<th>School level</th>
<th>Level designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-6</td>
<td>Elementary</td>
</tr>
<tr>
<td>8-12</td>
<td>Secondary</td>
</tr>
<tr>
<td>PK-7</td>
<td>Elementary</td>
</tr>
<tr>
<td>6-8</td>
<td>Secondary</td>
</tr>
<tr>
<td>PK-6</td>
<td>Elementary</td>
</tr>
</tbody>
</table>

Descriptive statistics for participants’ ALI Score across levels is summarized in Table 19.

Average scores across elementary ($M=22.91, SD=3.52$) and secondary ($M=22.64, SD=3.74$) were comparable with participants’ overall scores ($M=22.82, SD=3.58$).

An independent samples t-test was used to determine if there were significant differences in assessment literacy by school level, as summarized in Table 20. Significance was set at the $p < .05$. There was no significant difference between secondary and elementary principals’ knowledge of assessment literacy, $t (120) = .393, p = .695$.

Table 19

ALI Scores by School Level

<table>
<thead>
<tr>
<th>Level</th>
<th>n</th>
<th>% of sample</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary</td>
<td>80</td>
<td>65.6</td>
<td>22.91</td>
<td>3.52</td>
</tr>
<tr>
<td>Secondary</td>
<td>42</td>
<td>34.4</td>
<td>22.64</td>
<td>3.74</td>
</tr>
<tr>
<td>All</td>
<td>122</td>
<td>100</td>
<td>22.82</td>
<td>3.58</td>
</tr>
</tbody>
</table>

Table 20

Independent Samples t-test for School Level and ALI Score

<table>
<thead>
<tr>
<th>$t$</th>
<th>df</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.393</td>
<td>120</td>
<td>.695</td>
</tr>
</tbody>
</table>
Descriptive statistics for participants’ ALI Score across types of assessment training is summarized in Table 21. A majority (53.3%) of principals received assessment training through professional development as an administrator. The least common form of assessment training reported was administration preparation coursework (9.8%).

Table 21

<table>
<thead>
<tr>
<th>Primary Method of Assessment Training and ALI Score</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial teacher preparation coursework</td>
<td>14</td>
<td>23.07</td>
<td>2.92</td>
<td>11.5</td>
</tr>
<tr>
<td>Administration preparation coursework</td>
<td>12</td>
<td>23.83</td>
<td>2.25</td>
<td>9.8</td>
</tr>
<tr>
<td>Professional development as a teacher</td>
<td>31</td>
<td>22.06</td>
<td>3.84</td>
<td>25.4</td>
</tr>
<tr>
<td>Professional development as an administrator</td>
<td>65</td>
<td>22.94</td>
<td>3.78</td>
<td>53.3</td>
</tr>
</tbody>
</table>

An analysis of variance was conducted to determine whether groups differed significantly in terms of assessment literacy, as described in Table 22. There were no significant differences among participants’ level of assessment literacy as a result of predominant method of assessment training. Since the F-ratio was not significant at the $p < .05$, the post hoc comparisons were not conducted, $F(3, 118) = .822, p = .484$.

Table 22

<table>
<thead>
<tr>
<th>Analysis of Variance</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>31.81</td>
<td>3</td>
<td>10.604</td>
<td>.822</td>
<td>.484</td>
</tr>
<tr>
<td>Within groups</td>
<td>1522.22</td>
<td>118</td>
<td>12.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Interview Participants

In total, 12 principals were selected using purposive sampling to be interviewed. Principals scoring highest and lowest on the ALI who also agreed to participate in an interview were contacted for participation. When participants did not accept an invitation...
to participate, a new invitation was sent to additional participants within the next scoring range. The highest scoring principals that participated in interviews scored between 31 and 26 on the ALI. The lowest scoring principals scored between 15 and 19 on the ALI.

The original sample was designed to be inclusive of 10 principals (five high scoring and five low scoring) with equal representation of elementary and secondary; however, two additional principals were added to ensure equal representation of secondary principals within the sample and to balance the sample with equal numbers of high and low scoring principals.

In an effort to provide a rich, thick description of the context of each of the principals, information about the participants’ years of experience in their current context, their highest degree earned, and their teaching experienced was included. Additionally, following the interview, school accreditation status was reviewed using the VDOE School Quality Profiles for each of the principals interviewed. As a part of this review, it was revealed that one of the 12 participants served as an assistant principal at the secondary level; however, this individual’s data were included within the analysis to provide greater representation of administration at the secondary level. Table 23 provides contextual information about participants and their corresponding ALI scores. Participants are listed by ALI score in descending order with highest scoring participants listed first.
<table>
<thead>
<tr>
<th>ALI score</th>
<th>School level</th>
<th>Years in present position</th>
<th>Highest degree earned</th>
<th>Teaching experience</th>
<th>School accreditation status</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>Elementary</td>
<td>6</td>
<td>Masters in School Administration</td>
<td>7th Grade: Language Arts, Social Studies and Math</td>
<td>Fully Accredited</td>
</tr>
<tr>
<td>30</td>
<td>Elementary</td>
<td>2</td>
<td>Ed.S. in Administration and Supervision</td>
<td>Pre-K, 4th through 6th</td>
<td>Accreditation Denied</td>
</tr>
<tr>
<td>28</td>
<td>Elementary</td>
<td>1</td>
<td>Master’s in School Administration</td>
<td>3rd Grade; Elementary Math Specialist</td>
<td>Fully Accredited</td>
</tr>
<tr>
<td>27</td>
<td>Elementary</td>
<td>3</td>
<td>Doctorate in Administration and Supervision</td>
<td>10th and 11th Grade English</td>
<td>Fully Accredited</td>
</tr>
<tr>
<td>26</td>
<td>Secondary</td>
<td>4</td>
<td>Doctorate in Educational Leadership</td>
<td>6th, 7th, and 8th Grade English, Reading, and Science</td>
<td>Fully Accredited</td>
</tr>
<tr>
<td>19</td>
<td>Elementary</td>
<td>4</td>
<td>Post Graduate Certificate in Administration and Supervision</td>
<td>Pre-K, 3rd Grade and 4th Grade</td>
<td>Fully Accredited</td>
</tr>
<tr>
<td>19</td>
<td>Secondary</td>
<td>2</td>
<td>Doctorate in Educational Leadership</td>
<td>6th-12th History and Social Sciences</td>
<td>Fully Accredited</td>
</tr>
</tbody>
</table>

(Currently an Assistant Principal)
<table>
<thead>
<tr>
<th></th>
<th>Level</th>
<th>Program</th>
<th>Specialization</th>
<th>Accreditation</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>Elementary</td>
<td>Educational Specialist in 1st-3rd Traditional and Montessori</td>
<td>Supervision</td>
<td>Fully Accredited</td>
</tr>
<tr>
<td>19</td>
<td>Secondary</td>
<td>Post Graduate Certificate in Educational Leadership</td>
<td>Special Education</td>
<td>Fully Accredited</td>
</tr>
<tr>
<td>17</td>
<td>Elementary</td>
<td>Masters in Educational Leadership</td>
<td>Music: K-12</td>
<td>Fully Accredited</td>
</tr>
<tr>
<td>15</td>
<td>Secondary</td>
<td>Masters of Education</td>
<td>Elementary Physical Education</td>
<td>Fully Accredited</td>
</tr>
</tbody>
</table>

**Assessment Leadership Themes**

The sections that follow describe major themes that emerged across the 12 interviews. Four a priori codes were established prior to interviews, including:

*assessment for learning, balanced assessment, support for teachers, and professional development*. Upon further refinement of interview questions, these codes were expanded to also include *alignment* and *ethical considerations*. Additionally, the codes *support for teachers* and *assessment for learning* were consolidated to align with competency two, and the revised code stated: *support of assessment for learning principles*. Table 24 describes the assessment leadership competencies and their corresponding codes.
Table 24

*Assessment Leadership Competencies and Corresponding Codes*

<table>
<thead>
<tr>
<th>Competency</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. The leader understands the principles of assessment <em>for</em> learning and works with staff to integrate them into classroom instruction.</td>
<td>Support of Assessment <em>for</em> Learning Principles</td>
</tr>
<tr>
<td>3. The leader understands the necessity of clear academic achievement standards, aligned classroom-level achievement targets, and their relationship to the development of accurate assessments.</td>
<td>Alignment</td>
</tr>
<tr>
<td>5. The leader can plan, present, or secure professional development activities that contribute to the use of sound assessment practices.</td>
<td>Professional Development</td>
</tr>
<tr>
<td>9. The leader understands the attributes of a sound and balanced assessment system.</td>
<td>Balanced Assessment</td>
</tr>
<tr>
<td>10. The leader understands the issues related to the unethical and inappropriate use of student assessment and from such misuse.</td>
<td>Ethical Considerations</td>
</tr>
</tbody>
</table>

Within these five overarching codes, themes also emerged upon data analysis. A description of these themes is further described in Table 25.
Table 25

Assessment Leadership Codes and Themes

<table>
<thead>
<tr>
<th>Codes</th>
<th>Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support of Assessment for learning principles</td>
<td>Use of Professional Learning Communities to Support Instruction, Grouping Students Based on Formative Assessment Data, Support from Specialists</td>
</tr>
<tr>
<td>Alignment</td>
<td>Use of Pacing Guides, Alignment to Learning Targets, Digital Item Banks</td>
</tr>
<tr>
<td>Professional Development</td>
<td>Focus on Learning Intentions, Training in Assessment Administration</td>
</tr>
<tr>
<td>Balanced Assessment</td>
<td>Use of Common Assessments, Use of Benchmarks, Student Growth Assessments, Classroom Examples of Formative Assessment, Use of Literacy Screenings</td>
</tr>
<tr>
<td>Ethical Considerations</td>
<td>Unethical Practices a Nonissue, Training on Appropriate Administration of SOLs, SOL Practices to Minimize Testing Irregularities, Common Practices to Ensure Valid Results, Interpretation of Assessment Results</td>
</tr>
</tbody>
</table>

The sections that follow are organized by the five overarching codes related to assessment leadership and the themes that emerged within each of the codes. This section is designed to describe the relationship between principals’ knowledge of classroom assessment practices and leadership practices that support assessment for learning.

Support of Assessment for Learning Principles

The second competency states: *The leader understands the principles of assessment for learning and works with staff to integrate them into classroom instruction.*

Professional Learning Communities were described as a mechanism for supporting formative assessment practices, during which specialists supported teachers as they analyzed student assessment data to subsequently group students for instruction.
Across principals with both higher and lower levels of assessment literacy was the use of professional learning communities to analyze student assessment data and differentiated instruction. Additionally, principals described the use of grouping practices to provide targeted instruction. A supporting factor across principals included the use of specialists, specifically within professional learning communities and in their roles of supporting individualized instruction for students. Table 26 provides a description of each of these subthemes related to use of professional learning communities to support instruction, grouping students based on formative assessment data, and support from specialists as well as their relative strength across principals scoring higher and lower on the ALI along with illustrative examples of each theme.
### Table 26

**Support of Assessment for Learning Principles Themes**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Frequency of higher scoring principals</th>
<th>Frequency of lower scoring principals</th>
<th>Illustrative examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of Professional Learning Communities to Support Instruction</td>
<td>5</td>
<td>3</td>
<td>“We have a pretty tight PLC framework here at our school. I wouldn't say we are at, like, level six with implementation. We're probably at about a level three on a six-point scale, but we require teachers to use common formative assessments as well as more formal division level assessments to analyze student learning and make decisions about next steps for instruction.” (High)</td>
</tr>
<tr>
<td>Grouping Students Based on Formative Assessment Data</td>
<td>6</td>
<td>3</td>
<td>“We frequently look at formative assessment and then group and regroup students for intervention work. We have regular common assessments. We work through a PLC, Professional Learning Community model and we have regular common assessments that teachers give and then we meet about the data.” (High)</td>
</tr>
<tr>
<td>Support from Specialists</td>
<td>5</td>
<td>4</td>
<td>“In our PLCs, we have an instructional coach, we have a math specialist, a reading specialist, a gifted education teacher, a special education teacher, and administrators, so that all the various hats and disciplines can be looking together and trying to piece together what, what student work is telling us, what numbers on assessments are telling us and where we need to go from there.” (High)</td>
</tr>
</tbody>
</table>

**Professional learning communities.** One of the primary mechanisms of support for teachers in employing assessment for learning practices was through the use of defined professional learning communities, or PLCs. PLCs provided a structure for teachers to create assessments, analyze data, identify students’ strengths and weaknesses,
adjust instruction, discuss needs with specialists and/or administration, and/or formulate
groups of students based on similar instructional needs.

Professional learning communities in some instances served as an opportunity to
support the need for alignment between assessment and instruction. One elementary
principal with higher levels of assessment literacy described using the PLC framework to
guide classroom instructional practices and decision making:

The framework of our PLCs is that, you know, teachers work together and try to
use assessment to guide their practice, so they're constantly talking about, ‘What
is it we want to teach? How might we go about teaching it? And then, how did
students do and what do we need to do next?’ You know, the guiding questions
of the PLC. (High)

The work conducted in professional learning communities was primarily dictated by
student performance on assessments; therefore, the PLCs were orchestrated to support
students’ instructional needs. An elementary principal with higher levels of assessment
literacy described the focus of her PLCs:

For us at the school, it's really about what our own data is telling us. So, for
example, at the beginning of the year…We had decided for a number of reasons
that we were shifting our PLC focus from math to literacy this year and when we
gave our beginning of the year assessments, we were confirmed that really our
focus needed to be writing because…our student pre-test scores in writing were
abysmal. So then, really, what we started with was the assessment of the writing.
We gathered and tweaked a new writing rubric that we wanted teachers to use.
We taught them how, what it means, how to use it. We practiced analyzing
student work with this rubric. We calibrated with one another so that we were confident that when teachers were going back and doing their whole class sets that they were aligned with their peers and with our specialists on how everyone would be assessing the student work. And that was directly related to the fact that our kids were not writing well. We had evidence of that. And we needed to help teachers understand exactly where their kids were, and what would be, and using a developmental rubric to help them figure out how to really stretch them much farther. (High)

In addition to analyzing student work, PLCs also served to analyze data from common assessments and structure blocks of intervention and remediation. One principal with lower levels of assessment literacy described the questions that guided her PLC meetings:

We talk about who those kids are. Where they are academically? What assessments are you using to determine? What, strategies are you putting in place as far as intervention is? What are you doing to address that particular area that it seems like the child is not performing? So, for example, if it's kindergarten and it's concept of word, what are you doing specifically to target that particular area? Not generalities, but if that is the area that the child is struggling in, what are you doing to target that area and what support do you need from us as administration or the reading or the math specialist? (Low)

PLCs were also described as an avenue for developing assessments. One participant with lower levels of assessment literacy commented on using professional learning communities as an opportunity to create common assessments and also use them as “horizontal and vertical planning teams.”
As illustrated, the theme that emerged is that professional learning communities were described as an avenue for teachers to analyze student assessment and student performance and consequently provide supports or remediation for students based on their instructional needs. Additionally, they are designed to ensure there is a formative cycle of instruction and assessment in which students are grouped according to differentiated needs.

**Grouping students based on formative assessment data.** The need for small group instruction was evident in nine out of the 12 principals. Principals described the need for students to be grouped for instructional purposes to meet students differentiated needs within the classroom and as a response to student performance on assessments. All six of the higher scoring principals discussed the need for teachers to group and regroup students for instruction. Within the sample of lower scoring principals, three of the six principals addressed the need for small group instruction, and out of this three, they were all elementary principals.

The principals described small group instruction as a look-for when observing teachers’ classrooms and when reviewing lesson plans. For instance, an elementary principal with higher levels of assessment literacy shared, “We look at lesson plans, and they also turn in their RTI lesson plans and what they're working on. We look at the grouping of students, how often they change their grouping, what strategies they’re targeting in their grouping.”

In multiple instances, grouping practices existed in tandem with professional learning communities as teachers reviewed assessments and consequently grouped
students based on instructional needs. One elementary principal described the role of assessments in designing small group instruction:

> Basically the assessment is so very important. You know, it's that whole backwards design. You have to know where the kid’s going…You have to be able to check where they are. So it's basically having a pulse on if your instruction is effective. If, you're meeting the need for the kids. How to differentiate that instruction? How do I group these children? Who should be working with who? What is my reading group? What are the different levels in my class? Who's on the low or above? Like so, it plays such an important role in what we do every day. (Low)

The use of small group instruction was not evident across the three secondary principals with lower levels of assessment literacy. One of these principals referenced differentiation within the interview but did not describe what that differentiated instruction looked like in the secondary classroom. The principals who did, however, describe grouping practices also shared the role that specialists have in ensuring teachers are using assessment data to appropriately group students for instruction.

**Support from specialists.** Across principals with higher and lower levels of assessment literacy was the need for instructional specialists and coaches to support assessment for learning practices. In a majority of instances, the coaches played an integral role in professional learning communities; however, the principals described an array of supports that specialists provide to teachers. Some of these related supports included: unpacking curriculum, analyzing assessment data, creating assessments, supporting teachers through professional development, and ensuring teachers know how
to administer assessments. Additionally, the principals also described instances in which specialists provide targeted intervention based on students’ needs. They provided both direct support through intervention for students or indirect support by providing professional development to teachers.

For instance, one principal with higher levels of assessment literacy described the role of specialists in analyzing the curriculum and its relationship to assessments. “Our reading specialist and our math specialist are working really hard with the classroom teachers to make sure that the assessments are getting at the unpacked learning targets, at the level of rigor that we want to see” (High).

Specialists also played a role in supporting the professional growth of teachers. For instance, a high school principal with higher levels of assessment literacy shared that specialists play a role in determining what topics will be used in professional development. She indicated that topics are selected based on “what our specialists and our department leads identify as areas of weakness with the teachers that they work with in particular.” In other instances, specifically with one participant with lower levels of assessment literacy, he relied more extensively on specialists to identify areas for professional development because he shared that he is unable to attend all department meetings. He shared, “I would say the organization of instructional specialists, they actually are able to attend more of the departmental meetings than the building level administrators.” In this instance, the participants described that specialists also helped to ensure pacing is on target and that teachers are utilizing similar instructional practices.

In a more direct capacity, some of the specialists worked in conjunction with classroom teachers to provide remediation within a small group setting:
Every classroom teacher has a group. In addition, our math specialist has a group, our gifted education specialist has a group, and our special education resource teacher has a group so that our tier three and tier two groups are really small, and then our tier one groups are much larger. Then our enrichment group is our biggest group often. (High)

Despite their multifaceted roles, principals with both higher and lower levels of assessment literacy described the role of specialists working alongside teachers and the administration through professional learning communities to support the needs of students and student learning. Additionally, through direct or indirect needs, specialists and coaches ensured teachers had the necessary resources to support students’ needs in small group instruction.

Alignment

The fourth competency states: The leader understands the necessity of clear academic achievement standards, aligned classroom-level achievement targets, and their relationship to the development of accurate assessments. Alignment was addressed across principals with higher and lower levels of assessment literacy through use of pacing guides and digital resources that support development of standards-based assessments through the use of test-item banks; however, a focus on alignment of curriculum and instructional learning targets was only addressed among principals with higher levels of assessment literacy.

There was a disparity of practice among principals with higher and lower levels of assessment literacy related to this competency. For principals with the higher levels of assessment literacy, there was a concerted effort to address alignment of learning targets
and instruction to match the rigor of standardized assessments with a specific focus on learning intentions. However, principals with both higher and lower levels of assessment literacy referenced the use of test item banks to develop aligned assessments. Principals with both higher and lower levels of assessment literacy addressed the utility of pacing guides. Table 27 provides a description of each of these themes related to alignment of learning targets and use of digital item banks and pacing guides as a resource, as well as their relative strength across principals scoring higher and lower on the ALI along with illustrative examples of each theme.
Alignment Themes

<table>
<thead>
<tr>
<th>Theme</th>
<th>Frequency of higher scoring principals</th>
<th>Frequency of lower scoring principals</th>
<th>Illustrative examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alignment to Learning Targets</td>
<td>4</td>
<td>0</td>
<td>“I'm always looking for some sort of alignment between what I know they're shooting for and what they're actually assessing the kids on.” (High)</td>
</tr>
<tr>
<td>Use of Digital Item Banks</td>
<td>4</td>
<td>3</td>
<td>“It's pretty easy to just find the standards you want to test, and find the question that fits your kiddos best for that.” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“Us and two other schools were the schools who bought the program for our schools to help us with questioning, because we found it to be so successful, last year the county bought it for the entire county. So, now we use that to pull questions.” (Low)</td>
</tr>
<tr>
<td>Use of Pacing Guides</td>
<td>3</td>
<td>5</td>
<td>“We have a division pacing guide. And so what that helps us do is look at what standards need to be covered in a given quarter that will then, we know in theory, be tested on the benchmark assessment at the end of the quarter.” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“For us, the SOL or the pacing guide is what defines what we should be covering. So based on the assessment or based on the content being covered, the assessment should reflect that, and have an accurate tool to ascertain if the students are getting that content from that lesson or that unit that they're reviewing.” (Low)</td>
</tr>
</tbody>
</table>

Alignment to learning targets. Through the discussion of aligned assessment, four of the principals with higher levels of assessment literacy referenced the need to ensure there was alignment between assessment and intended learning targets or objectives. In one instance, one of the principals had provided trainings for teachers on how to write learning intentions and subsequently aligned assessments. Two of these four
principals additionally shared that their school participates in a university partnership that addresses alignment of content to ultimately support classroom instruction. Within this group, some of the principals addressed the need for rigorous instruction to match the level of rigor addressed within the assessment. For instance, one principal with higher levels of assessment literacy shared:

We used to assess pretty basic math skills as opposed to flexible problem solving and the ability to communicate in math, and so we really have done a nice job of redesigning those assessments so that the level of teaching is leading to success on those assessments. (High)

Additionally, others have used the format of professional learning communities to unpack standards to ensure they have appropriately addressed the standard within the instruction and assessment. For instance, an elementary principal with higher levels of assessment literacy shared:

We create the assessments together often in the PLCs and so … our reading specialist and our math specialist are working really hard with the classroom teachers to make sure that the assessments are getting at the unpacked learning targets at the level of rigor that we want to see. (High)

Principals with lower levels of assessment literacy did not address the need for alignment among curriculum, instruction, and assessments. The issue of alignment of intended learning intentions was not raised across principals with lower levels of assessment literacy. Alignment was, however, addressed across principals through the use of digital item banks to support teachers in developing assessments.
**Digital item banks.** Seven of the 12 principals with varied levels of assessment literacy described technological resources that were available to teachers to support their development and integration of assessments aligned to standards. These resources were varied across principals; however, their use was primarily to develop assessments that aligned with the SOLs.

One elementary principal with higher levels of assessment literacy described the relative ease associated with this type of resource when developing assessments. “It's pretty easy to just find the standards you want to test, and find the question that fits your kiddos best for that.” Another elementary principal with lower levels of assessment literacy described the successful integration of this type of resource within her setting and how its implementation was later expanded to schools throughout the division. She described how the division “bought the program for our schools to help us with questioning. Because we found it to be so successful, last year the county bought it for the entire county. So now we use that to pull questions.”

Another secondary principal with higher levels of assessment literacy described the use of a consortium where lesson plans and instructional activities are shared and developed by teachers across divisions. Through this consortium, teachers examine lesson plans and assessments for alignment. In addition to using digital item banks, principals with both higher and lower levels of assessment literacy described the integration of pacing guides to support alignment to standards and continuity of instruction across teachers’ classrooms.

**Use of pacing guides.** The utilization of pacing guides was described across both principals with higher and lower levels of assessment literacy. Pacing guides were
described as a means to ensure there was alignment across classrooms with content being taught and when it was being taught. Pacing guides were frequently used in conjunction with benchmarks and common, formative assessments to ensure appropriate content had been covered and mastered. One of the elementary principals with higher levels of assessment literacy described the utility of assessments and pacing guides:

Formative assessments help us A) they help us stay aligned with our pacing as a grade level because we all know we're giving them on such and such a date and we want our kids to all have the same access to the curriculum and the instruction so that when we do regroup them, we're confident that they've all at least had initial exposure to this curriculum. Our more formal quarterly assessments are formative in that they help us monitor our pacing for the whole year. You know that that stuff was coming, you better have covered that information and the kids better know it. (High)

An elementary principal with higher levels of assessment literacy described the role of pacing guides in ensuring teachers were on pace with content that would be covered in benchmark assessments. She shared, “Our more formal quarterly assessments are formative in that they help us monitor our pacing for the whole year.” Another elementary principal with higher levels of assessment literacy described the utility of pacing guides:

We have a division pacing guide. And so what that helps us do is look at what standards need to be covered in a given quarter that will then, we know in theory, be tested on the benchmark assessment at the end of the quarter. So we have- In
second, third, and fourth grade, we have reading benchmark assessment and math benchmark assessments at the end of every quarter. (High)

Similarly, a principal with lower levels of assessment literacy described how pacing guides are used to guide what is taught each quarter:

> For us, the SOL or the pacing guide is what defines what we should be covering. So based on the assessment or based on the content being covered, the assessment should reflect that, and have an accurate tool to ascertain if the students are getting that content from that lesson or, that unit that they're reviewing.

Principals monitored and supported alignment for teachers through the use and integration of pacing guides and digital item banks for test development; however, principals with higher levels of assessment literacy additionally focused on alignment of instructional practices with intended learning outcomes.

### Professional Development

The fifth competency states: *The leader can plan, present, or secure professional development activities that contribute to the use of sound assessment practices.*” Two areas of professional development emerged. For principals with higher levels of assessment literacy, they provided staff development that focused primarily on instruction within the classroom. Across principals, however, was professional development specifically related to training in assessment administration. Table 27 provides a description of each of these themes related to: focus on learning targets and training in assessment administration, as well as their relative strength across principals scoring higher and lower on the ALI along with illustrative examples of each theme.
Focus on learning targets. Principals with the highest levels of assessment literacy prioritized the need for staff to align learning targets to state standards and focused professional development on the instruction that addressed learning targets. In an effort to address the increased rigor in state assessments, the principals recognized the necessity of focused learning targets within classroom instruction. An elementary principal with higher levels of assessment literacy led trainings on how to break down curriculum standards to fully understand learning intentions. To do so, she described how she led staff in “unpacking those assessments, on looking at writing learning intentions and then corresponding assessments. I think we used the coaching model a lot, in terms

<table>
<thead>
<tr>
<th>Theme</th>
<th>Frequency of higher scoring principals</th>
<th>Frequency of lower scoring principals</th>
<th>Illustrative examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus on Learning Targets</td>
<td>3</td>
<td>0</td>
<td>“What we focused on in our specific professional development sessions, though, was on the, the instruction that they needed to be providing in order for the kids to be successful on these more rigorous assessments.” (High)</td>
</tr>
<tr>
<td>Training in Assessment Administration</td>
<td>2</td>
<td>3</td>
<td>“Prior to them doing any of those assessments, we do a retrain every year on, you know, ‘This is how to administer a running record. Here are some of the resources that you might use.’” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“Sometimes the division will offer training, or, as we have new staff members, you know, I might have my coach train them on how to administer, you know, COW, if they're from out of state, or our reading specialists can help with that as well.” (Low)</td>
</tr>
</tbody>
</table>
of coaches working individually with either grade levels or individual teachers on assessments.”

Rather than begin with professional development that specifically addressed assessment related topics, another principal with higher levels of assessment literacy addressed the need to ensure there was a parallel between classroom instruction and learning intentions. She described this prioritization of need with her faculty. “More of what we focused on in our specific professional development sessions, though, was on the instruction that they needed to be providing in order for the kids to be successful on these more rigorous assessments.”

This need to refocus professional development on learning targets was again addressed at the elementary level with a principal with higher levels of assessment literacy. She shared that her faculty needed to focus on reviewing learning targets before beginning to focus on assessment-related topics in professional development:

We're not even talking about the criteria that kids would need to be able to do to show the teacher that they've met the learning target. I don't even think we're doing the general incorporation of the learning target throughout the teaching of the lesson well enough, let alone looking at the assessment piece. (High)

The aforementioned principals indirectly focused on the alignment of instruction and assessment by honing in on learning intentions. Although assessment was not a unitary focus, the need to focus on learning intentions and targeted instruction was a necessary aspect of addressing the need for alignment and ultimately assessment for learning.

There was a lack of consistency in assessment-related professional development topics across principals with lower levels of assessment literacy, and professional
development topics were not driven by assessment-related priorities. For instance, an elementary principal with lower levels of assessment literacy shared her plan for delivering professional development topics within her setting:

So, every first Tuesday of the month we have what's called a learning meeting and it's usually some kind of PD based on what the teachers need. So, for example, this year we had a lot of newer teachers so we started the year with K-two had concept of word training and then you know, three, five, had a different training.

So, each Tuesday we do that. (Low)

In comparison, the principals with lower levels of assessment literacy did provide opportunities for teachers to engage in professional development; however, the role of assessment in the learning process appeared more disjointed. Professional development topics were based on surveys of teachers’ needs, observations, and data; however, the scope of topics was broader and less targeted.

**Training in assessment administration.** Principals with varied levels of assessment literacy supported the work of teachers through trainings on how to administer various assessments. For instance, principals provided opportunities for teachers to engage in content-specific training, such as how to assess reading levels or assess concept of word in reading.

One principal ensured teachers were trained in the proper administration of running records. Additionally, the principal trained staff to ensure that they understood the instructional implications of running records. In this instance, the focus of this professional development and other initiatives emerged through the school improvement plan but ultimately was based on observation and data that had been collected.
An elementary principal with lower levels of assessment literacy highlighted the role of her specialists in ensuring teachers were trained on appropriate administration of a reading assessment:

Basically, based on the need for the teachers. I have conversations with my reading and math specialists often and they'll come and say, you know, we really need to train these teachers on this because I saw this happening with concept of word and I know that this particular teacher's really struggling with teaching it.

So, it's based on observation pretty much. (Low)

These trainings were necessary to ensure teachers implemented assessments as part of the school’s overall assessment system with fidelity.

**Balanced Assessment**

The ninth competency states: *The leader understands the attributes of a sound and balanced assessment system.* Principals demonstrated evidence of a systematic method of collecting data on student performance through a balanced assessment system that encompassed using benchmark assessments, common assessments, classroom assessments, student growth assessments, and reading inventories. Across principals with higher and lower levels of assessment literacy was the use of common assessments and benchmark assessments to monitor and measure student performance. Additionally, across elementary principals was the need to implement reading inventories to measure students’ reading levels. Table 28 provides a description of each of these themes related to: common assessments, benchmark assessments, classroom assessments, and reading assessments, as well as their relative strength across principals scoring higher and lower on the ALI along with illustrative examples of each theme.
<table>
<thead>
<tr>
<th>Theme</th>
<th>Frequency of higher scoring principals</th>
<th>Frequency of lower scoring principals</th>
<th>Illustrative examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of Common Assessments</td>
<td>5</td>
<td>4</td>
<td>“We work through a PLC, Professional Learning Community, model and we have regular common assessments that teachers give and then we meet about the data.” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“Most of our teachers kind of, they make common assessments according to their PLCs. So for example, you know, our 10 or 12 geography teachers will get together and they'll combine their heads to make assessments. Same thing with government, US History.” (Low)</td>
</tr>
<tr>
<td>Use of Benchmarks</td>
<td>5</td>
<td>2</td>
<td>“I would even consider our benchmark assessments that are done by- more by the division, to be very formative in that we really do disaggregate the data, look at it by strand, and then work to use it to inform what we need to do next.” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“We, our school district does quarterly assessments and so, what we do is we have weekly CLT meetings, Collaborative Learning Team meetings” (Low)</td>
</tr>
<tr>
<td>Student Growth Assessments</td>
<td>4</td>
<td>3</td>
<td>“We use Interactive Achievement as a student growth. We use the pre- and the post-student growth assessment out of there” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“We have a lot of meetings with individuals, and we meet pre-assessment, mid-assessment, we look at the student growth, make adjustments” (Low)</td>
</tr>
<tr>
<td>Use of Literacy Screenings</td>
<td>5</td>
<td>3</td>
<td>You know, we recommend doing an FMP at the beginning of the year, midyear, end-of-year” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“And then with the PALs throughout the year we have what's called a quick check and we use that periodically to see how the students are doing to keep them on path or target.” (Low)</td>
</tr>
<tr>
<td>Classroom Examples of Formative Assessment</td>
<td>3</td>
<td>4</td>
<td>“I like to see exit tickets employed. I like to see really quick measures of understanding, whether it's a thumbs up, thumbs middle, thumbs down.” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“I've just seen so many different ways in which teachers just kind of do those daily checks to make sure that the kids are getting what they're teaching.” (Low)</td>
</tr>
</tbody>
</table>
**Use of common assessments.** In addition to using benchmark assessments, principals across both higher and lower levels of assessment literacy described the utility of common, formative assessments to assess student progress. For instance, an elementary principal with higher levels of assessment literacy described using PLCs to create common assessments that addressed appropriate levels of rigor. Another principal with lower levels of assessment literacy at the secondary level described his school and division’s progression towards using common, formative assessments to ensure there is continuity and alignment of practice across classrooms and schools. The value of common assessments was described by a principal who shared that teachers from the same content areas used them to meet to review data from the assessments. For instance, a secondary principal with lower levels of assessment literacy shared how he has prioritized time to review common assessments in his school’s schedule:

> So we have the common assessments, but we have broken into our schedule or we have established in our schedule, set times where grade levels can meet, but also the time where teachers who teach the same content can meet. (Low)

Principals described this use of common assessments to ensure there was appropriate coverage of content and the ability to target instruction for students in need. An elementary principal with higher levels of assessment literacy shared:

> We have common formative assessments in math and we give about two to three of those a quarter. They help us figure out: Are we teaching what we need to be teaching, are kids getting it, who needs more support, who needs to be enriched? (High)
On a less centralized level, the need for greater integration of common, formative assessments was addressed across principals. They addressed the necessity for common assessments to ensure there was alignment with pacing guides, standards, and classroom expectations. One principal described the need for common assessments within her division:

And then this past year, as a division, I helped them implement a county wide where all…the eighth grade math teachers in the entire county met together and looked at curriculum and alignment of assessments and that sort of thing. And then this next year, they are working on creating common assessments at the county level for the teachers to use for every assessment, whether it's a test for a chapter or a concept, or whether it's the county benchmarks that we already have created. (High)

While common assessments typically addressed achievement or mastery of specific concepts or units, principals also described their school’s integration of benchmark assessments to assess student performance using quarterly intervals.

**Use of benchmark assessments.** The use of benchmark or quarterly assessments was pervasive across principals with higher levels of assessment literacy but it was also used with principals with lower levels of assessment literacy. Benchmarks were aligned with pacing and used to guide instruction and remediation practices. One principal jokingly referred to benchmarks as a monitoring tool by central office, “and what is the role of the assessment? I mean, well, the role is that it tells central office that they need to come breathe down my neck.” In this instance, benchmarks were used to be predictive of student performance on Standards of Learning assessments.
Benchmarks were administered at quarterly intervals. For instance, a principal with lower levels of assessment literacy described the frequency of benchmarks. “Quarterly assessments, benchmarking, we do fall, mid-year and spring testing.” Some principals described the benchmarks as cumulative.

Another secondary principal with higher levels of assessment literacy shared how central office reviews benchmark assessments. “They all look at the county benchmarks that our kids take, and then use that to create intervention and enrichment programs to make sure that we are supporting the learning intentions that we claim.”

In other instances, benchmark assessments were described to be used to drive instruction and support the needs of students. A secondary principal with higher levels of assessment literacy shared:

We have county benchmarks that the division has created, and we implement them at the end of first, second, and third quarters. And then those results are shared with the teachers, for question by question analysis, as well as they look at school by school analysis to compare if another school is doing much better in a certain strand than in other schools. We'll try to share what kind of lessons that they're doing that seem to be having a better outcome than others. (High)

Similarly, an elementary principal described how benchmarks were disaggregated and subsequently used to guide instructional decision making. Benchmark assessments were described as both formative and summative, depending on the principal. While benchmarks covered mastery of content within a specified period of time, principals also described their school’s integration of growth measures to assess student performance across time.
**Student growth assessments.** Seven of the principals across levels of assessment literacy shared assessments that are used to measure student growth. Some of the assessments incorporated a pre-post model during which students were administered the same assessment at designated times throughout the year to see how much content has been mastered over time. One elementary principal with higher levels of assessment literacy described a math assessment used to analyze student growth over the course of the year. “We do a pre and a post. So that is also more of a summative assessment to look at how much students have grown over the course of the year in their understanding of numbers and operations.”

Other assessments such as running records were used as a tool to examine and reflect on student growth. An elementary principal with higher levels of assessment literacy shared the impact of running records, “I think it's purely for the purpose of sort of tracking and looking and tracking that growth over time.”

Student growth measures were often used in conjunction with teacher evaluation and student goal setting. For instance, one secondary principal with lower levels of assessment literacy discussed meeting with teachers about their student’s progress on the same assessment over time. “We have a lot of meetings with individuals, and we meet pre-assessment, mid-assessment, we look at the student growth, make adjustments. I won't let them change their smart goal mid-year.” Just as student growth measures were common, using literacy screenings to measure students’ instructional reading levels was isolated primarily to principals with higher levels of assessment literacy and elementary principals with lower levels of assessment literacy.
Use of literacy screenings. In addition to the aforementioned assessments, eight of the 12 principals described using literacy inventories or assessments to pinpoint students’ instructional reading levels. Samples of these types of assessments included: Virginia Phonological Awareness and Literacy Screening (PALS), running records, Qualitative Reading Inventories (QRIs), spelling inventories, and Rigby assessments. Out of these eight principals, seven were elementary principals. One principal described the triangulation of reading assessments to develop a “literacy profile.” He shared:

Our county has some guidelines as far as a literacy profile, where we look at, any, you know, depending on what the level the child is, we'll look at PALS, we'll look at Rigby, and we'll look at QRI. Those are kind of our three main data points. We collect data three times a year on that, and teachers then examine that at those intervals and make adjustments as needed. (High)

Additionally, all three of the elementary principals with lower levels of assessment literacy described using a literacy screening to measure students’ progress and instructional levels in reading. For instance, one principals with lower levels of assessment literacy shared their school wide expectation for administering running records, “They’re supposed to do running records on students that are not meeting grade level benchmarks for reading. They have to do this every week.” In addition to assessments previously mentioned, principals also discussed less formalized measures to gauge student understanding, as described in the following section.

Classroom examples of formative assessment. Seven of the principals described less prescribed uses of classroom assessment to formatively assess student understanding. They described a variety of formats for classroom assessment, such as exit tickets,
thumbs-up/down, and observation. The types of classroom assessment described varied by principal. One principal with lower levels of assessment literacy shared the flexibility of strategies to assess student understanding and mastery of content, “I’ve just seen so many different ways in which teachers just kind of do those daily checks to make sure that the kids are getting what they're teaching.”

Another principal with higher levels of assessment literacy described the range of formative assessment practices she looks for when observing classrooms:

I like to see exit tickets employed. I like to see really quick measures of understanding, whether it's a thumbs up, thumbs middle, thumbs down…teachers walking around, observing, checklists. I look for all of those things in addition to the big and heavy assessments. (High)

To illustrate these quick checks for understanding, one elementary principal with lower levels of assessment literacy described an example from a literacy lesson:

You know, for example I was in the teacher's classroom last week and when she's doing her guided reading groups, kids are given a sticky note. So, as they are reading to themselves they are highlighting words of difficulty. They are writing their own questions based on what they're reading. So, she's assessing their comprehension based on what they find to be difficult. So, I look for those types of things. What, how are you determining if your kids got what you just taught or what you just know, covered? (Low)

Across principals, it was evident that structures for assessment have been established although the types of assessments varied more between elementary versus secondary contexts.
Ethical Considerations for Student Assessment.

Competency 10 states: The leader understands the issues related to ethical and inappropriate use of student assessment and protects students and staff from such misuse. This standard relates to the “interpretation, use and communication of results that leads to appropriate inferences about student learning and proper action on behalf of student success” (Chappuis et al., 2004, p. 289). Principals described ethical considerations as a nonissue with their staff but also demonstrated leadership practices to prevent unethical conduct through the use of trainings on appropriate administration of SOL assessments, changes in SOL practices to minimize opportunities for testing irregularities, standard administration of assessments to ensure valid data, and opportunities to assist parents in appropriately interpreting assessment results for their children. Table 29 provides a description of each of these themes, including: unethical practices a nonissue, training on appropriate administration of SOLs, SOL practices to minimize testing irregularities, common practices to ensure valid results, and stakeholder interpretation as well as their relative strength across principals scoring higher and lower on the ALI along with illustrative examples of each theme.
<table>
<thead>
<tr>
<th>Theme</th>
<th>Frequency of higher scoring principals</th>
<th>Frequency of lower scoring principals</th>
<th>Illustrative examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unethical Practices a Nonissue</td>
<td>4</td>
<td>1</td>
<td>“I've never even really thought about that because if- I feel like most of our ethics are kind of like above the board.” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“We sort of trust our teachers they will be ethical and appropriate when it comes to administrating their assessments.” (Low)</td>
</tr>
<tr>
<td>Training on Appropriate Administration of SOLs</td>
<td>3</td>
<td>1</td>
<td>“We have strategies for specific training of how to give the specific tests and how to utilize accommodations for special education students and, you know, what teachers should and shouldn't do.” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“We're lucky to have a great testing coordinator who you know, weighs that out a couple of times of the year, what the expectations are, the legal ramifications, and the expectations are.” (Low)</td>
</tr>
<tr>
<td>SOL Practices to Minimize Testing Irregularities</td>
<td>2</td>
<td>3</td>
<td>“No teacher is allowed to be in the room where there are kids testing the subject that they taught. For example, the English teachers is probably giving the math test and the math teachers are giving the English test. And you know that way that just eliminates the possibility of wanting to provide help.” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“So for example, for the SOL, which is our big assessment, the teachers don’t administer the test to their own students. So we rotate the person that administers the test.” (Low)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“There are some assessments we will do, and we'll have a reading specialist administer it, and it'll be the same reading specialist, so that there's no inconsistencies in administration.” (High)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“So, that's why we kind of made the reading department give it because then it's even across the plane so it's not based on the experience of the teacher. It's based on the trained individual who's trained to get this assessment.” (Low)</td>
</tr>
<tr>
<td>Common Practices to Ensure Valid Results</td>
<td>3</td>
<td>2</td>
<td>“Whenever we have our, special education meetings or 504 meetings or, you know, IT meetings, the team will go over the results with the parents, and ensure that they have a deeper understanding for what that growth means for their children, or if there's a lack of growth, why that might be occurring.” (High)</td>
</tr>
<tr>
<td>Interpretation of Assessment Results</td>
<td>5</td>
<td>5</td>
<td>“Well, we have parent teacher meetings. We have open house, so the parents come and meet with the teachers. We, also have a website. We have email that they can talk with the teachers back and forth to discuss child's progress. If a child is not</td>
</tr>
</tbody>
</table>
Unethical practices a nonissue. Five of the 12 principals interviewed shared that they do not consider unethical practices to be a major concern regarding teachers’ use of assessments and assessment results. To further illustrate this, an elementary principal with higher levels of assessment literacy shared, “Hmm, that's a good question. I don't know. I don't have any specific strategies that I use particularly for that. I… You know, certainly we try to make sure that's not happening.”

Another principal with lower levels of assessment literacy shared a similar sentiment with regard to the trust in teachers’ ethical practices. He shared, “We look at teachers giving their own test in the classroom on a day-to-day basis. We sort of trust our teachers they will be ethical and appropriate when it comes to administrating their assessments.”

Although these principals did not report that they had ethical concerns about teachers’ uses of assessments within their respective contexts, they did, however, share strategies they employed to prevent ethical misconduct related to SOLS and common practice.

Training on appropriate administration of SOLs. Four out of the 12 principals referenced SOL training for staff to ensure appropriate administration of standardized assessments. In both principals with higher and lower levels of assessment literacy, a concern regarding the ethical implementation of assessments related to training staff on appropriate administration of the state standardized assessments, or SOLs. Principals with
both higher and lower levels of assessment literacy addressed the need for training to ensure these assessments were implemented appropriately and according to state guidelines and protocol. Additional precautions included: the use of scenarios to support teacher trainings, use of school testing coordinators to provide training, and training regarding appropriate accommodations for students. One elementary principal with higher levels of assessment literacy addressed the training provided to teachers during SOL assessments:

We do a lot of training around SOLs and what you can and can't do. We are actually parked outside those classrooms and peeking in during the administration of the test to make sure there aren't any unethical things. We're very clear with teachers about required accommodations for students and making sure that students have access to those accommodations all year long as well as during the assessment.

For instance, another elementary principal with higher levels of assessment literacy shared “We do obviously, all the trainings and go through case studies and scenarios of things that…have happened across the state to make sure everybody knows, you know, what is and is not okay.” Additional consideration was given to ensure teachers were aware of appropriate accommodations for students with disabilities for standardized assessments. For instance, an elementary principal with higher levels of assessment literacy shared:

We have strategies for specific training of how to give the specific tests and how to utilize accommodations for special education students and, you know, what
teachers should and shouldn't do and all of those kinds of things are—They're all in place, so we would certainly use strategies there. (High)

A participant with lower levels of assessment literacy described the role of a testing coordinator in ensuring staff were appropriately trained in the administration of the SOLs. He shared that the testing coordinator trains staff about “what the expectations are, the legal ramifications” to ensure appropriate administration of SOLs. Training for SOLs was also accompanied with heightened attention to ethical practices during the actual administration of the SOLs.

**SOL practices to minimize testing irregularities.** In addition to providing specific training about the appropriate administration of assessments, 5 of the 12 principals referenced additional safeguards to ensure ethical administration of assessments during SOL testing. Some of these precautions included: additional proctors within the testing environment and SOL proctors outside of the content area. In further instances, measures were taken to ensure teachers did not assess their own students.

This heightened concern for fidelity in implementing test regulations was consistent across levels. Although none of the principals suggested that teachers would consciously attempt to provide inappropriate assistance to students, they attempted to minimize opportunities for testing irregularities. For instance, one secondary principal with lower levels of assessment literacy reported that he did not believe teachers would behave unethically during SOLs. Instead, he was attempting to minimize opportunities for “potential accusations” as well as “non-intended consequences.”

Another elementary principal with higher levels of assessment literacy addressed the need to be present and visible during SOL testing to ensure ethical administration of
assessments. She reported, “We are actually parked outside those classrooms, and peeking in during the administration of the test to make sure there aren't any unethical things.”

Although there was heightened vigilance during SOL testing, another consideration during the year was ensuring that assessments across classrooms were administered using common practices to ensure valid results.

**Common practices to ensure valid results.** Another ethical consideration that emerged across groups was the need to have consistency in test administration for assessments outside of the SOLs, which in some instances involved ensuring there was consistency in test administration practices. Five principals across both higher and lower levels of assessment literacy addressed the need for common practice. An elementary principal with lower levels of assessment literacy described this practice, “Like in this case the reading department is going to give the assessment to all the kindergarten…children. Just because they are trained to do it. They are trained to do it correctly.” In this particular circumstance, the principal addressed the need for consistency in test administration because the results were linked to the teacher’s SMART goal as part of the teacher evaluation process.

Another principal with lower levels of assessment literacy described the importance of consistency in administration of assessments in order to ensure appropriate growth was captured by the assessment. She described the use of a trained reading specialist to administer the assessment at the end of year in order to ensure fidelity of implementation:
Each classroom teacher does their own PALs assessment at the beginning of the year. So then they can see where the students are and how they’re doing and also to interact with the student on an individual basis. The midyear and end of the year test is administered by the PALs teacher. (Low)

Another principal shared the implications of inappropriate assistance to students when a student’s data reflects an overestimate of his or her ability due to inappropriate assistance provided to a student. The principal described the example of using running records at the end of the year to measure students’ instructional reading levels; however, when common practice was not used for implementation of the running records, there were implications for students the subsequent year. This elementary principal with higher levels of assessment literacy shared:

We had a couple students whose reading level was called to be higher than what the next year's teacher anticipated. You can imagine that. ‘I can't believe that this kid was reading at that level this year, 'cause they're only doing this right now’. And it seemed like that the previous year teacher did give them a passage that gave them an advantage, more than maybe just a totally cold passage. I don't think it was done out of trying to gain the system. I think it was done more out of just not even thinking about the ramifications. (High)

The principals consistently reported the need for accurate data, had confidence in their school’s practices and individual teacher’s use of ethical judgment, and the overarching conclusion that inappropriate assistance to students would not be beneficial to them in the future. A secondary principal with higher levels of assessment literacy shared:
And then as far as inappropriate use of assessment, I don't really, there would be no gain for them to use the kind that we use, inappropriately…It doesn't ... help us to try to make our kids look better because in the long run, when the state assessment happens, somebody would figure out that you must've cheated all along on the county assessments if all the sudden you know, you have this issue.

(High)

With a wealth of assessments available to teachers and students, principals additionally addressed the need for educators to assist stakeholders in the appropriate interpretation of assessment results.

**Interpretation of assessment results.** The tenth competency also relates to the ways in which stakeholders make appropriate interpretations from assessments:

This standard of ethical practice underpins all of the previous nine competencies, and is accomplished when leaders promote interpretation, use, and communication of results that leads to appropriate inferences about student learning and proper action on behalf of student success (Chappuis et al., 2004, p. 289).

Practices across principals suggested established initiatives to ensure parents were informed of student performance and progress through various means and degrees. Additionally, principals with higher levels of assessment literacy and one principal with lower levels of assessment literacy focused on communicating measures of student growth and also involving students in goal setting related to their progress.

A concerted effort was made by principals of varying levels of assessment literacy to provide opportunities for parents to discuss progress. A principal with lower levels of
assessment literacy used standards based report cards for parents and encouraged staff to help parents interpret student performance, as parents may have a more rudimentary understanding of this type of grading practice, as compared to the more traditional letter-grading system. She shared the value of “conversations to make sure that everyone involved, those stakeholders, know exactly what this means. This is where we are and this is what we're going to do to get them to where they need to be.”

Principals required teachers to send home assessment results and encouraged teachers to give parents a context for understanding the scores. Again, this occurred across levels. A high school principal with higher levels of assessment literacy shared:

We send the results home, but if a parent doesn't understand they are allowed to call the school or come up here and we'll talk about the assessment. And talk about what the score means and help parents understand where their children are and hopefully help them help their kids become better students, better test-takers and care more about their education. (High)

Across interviews, it was evident that principals utilized multiple assessments to measure student progress; however, many parents may not have been adept at understanding how to interpret assessment results. A secondary principal with lower levels of assessment literacy discussed helping parents interpret assessment results, specifically for parents of students with disabilities. He shared, “So a lot of times, we will talk to the parents and really spend that time to discuss what that data means and how we can use that database to better support the child in the future.”

Some of the principals with higher levels of assessment literacy one with lower levels of assessment literacy, however, involved students in the process of interpreting
their assessment results and goal setting. For instance, an elementary principal with higher levels of assessment literacy shared how specialists teach students to chart and track their own progress and performance and focus on their individual growth:

I think our teachers are really thoughtful about helping kids understand that perfection is not required and that you are gonna have ups and downs because everybody has a good and a bad day and everybody learns, you know, their rates of learning changed over the course of the year. (High)

As students tracked their progress, this school also adopted a strategy in which students facilitated their own parent-teacher conferences. Students discussed their learning targets and shared their individual growth. She shared the outcome of student-led parent-teacher conferences:

You know, actually, we weren't talking about their test scores. We were talking about their own assessment of their learning based on the evidence that they had collected over the course of the year from their writing, from their reading. (High)

For the principal with lower levels of assessment literacy, she described the use of a conferencing model to provide individualized feedback to students about their writing performance.

From the interviews, it was evident that schools have a wealth of assessment data that is subject to interpretation from various stakeholders. Principals with both higher and lower levels of assessment literacy seek strategies to engage and inform parents; however, only a selected number of principals have extended this practice to engage students in examining and discussing their own learning and growth.
CHAPTER FIVE: DISCUSSION

This study was designed to examine Virginia principals’ knowledge of classroom assessment and their support of assessment for learning practices across the commonwealth. Using the Assessment Literacy Inventory (ALI), participants completed a 35-item inventory to gauge their level of assessment literacy. Assessment literacy was compared across levels of participants (secondary vs. elementary), and there were no significant differences in terms of assessment literacy based on level assignment. Additionally, participants reported their primary method of assessment training. The predominant method of training for principals was professional development as an administrator; however, there were no significant differences among principals’ levels of assessment literacy as a result of training. The study further analyzed principals’ performance when disaggregated by the Seven Standards for Teacher Competence in the Educational Assessment of Students. Additionally, participants with higher and lower scores on the ALI were selected using purposive sampling and subsequently interviewed regarding their assessment leadership practices in the areas of: support for teachers, alignment, professional development, balanced assessment, and ethical practices. The following section describes these results in greater depth, their relationship to previous research and findings, limitations of this study, implications for practice, and recommendations for future research.
Summary and Discussion of Findings

There was a wide range in participants’ classroom experience across the sample of principals. The mean years of teaching experience was 11.51 years ($SD=5.50$). Similarly, there was a wide range in participants’ years of administrative experience across the sample of principals. The mean years of administrative experience was 11.09 years ($SD=5.16$). Clark et al. (2009) found that principals’ years of experience was related to improved student performance and was “especially steep over the first few years of principal experience” (p. 26). Given the range of administrative experience within this sample, central office leadership should consider the impact of administrative experience and how it could impact student performance if there are not appropriate supports for instructional leaders. In addition to demographic information, the following section summarizes principals’ performance on the ALI by standard.

Scores across the Seven Standards for Teacher Competence in the Educational Assessment of Students were fairly uniform. The lowest standard was Standard Two: *Teachers should be skilled in developing assessment methods appropriate for instructional decisions*, and the highest standard was Standard Seven: *Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information*. An item analysis was used to determine relative strengths and weaknesses according to each question; however, inferences related to these item analyses should be interpreted with caution. Inferences related to topics within specific standards are limited to one question only and should be interpreted cautiously in the absence of multiple, consistent measures. The sections that follow discuss each of the seven standards and interpretations of each of the scores.
Standard one states: *Teachers should be skilled in choosing assessment methods appropriate for instructional decisions*. Administrators with assessment literacy understand how to select an assessment method when presented with various formats based on the intended learning outcomes (Stiggins, 1991, 1995). Within this standard, there was a higher frequency of incorrect responses that involved application and knowledge of assessment-related terminology. The 52.5% score on Question 1 of the ALI suggested that principals might not have a solidified understanding of the differences and utility of varied formats of assessments (performance assessment, authentic assessment, extended response assessment, and standardized test). Similarly, the 40.2 percentage correct score on Question 15 related to other assessment-related terminology, including validity and reliability. Performance across the five questions in standard one suggests that principals understood when to select appropriate assessments based on contextual information; however, when faced with assessment-related terminology in isolation, their understanding decreased. This suggests that principals may not have had concrete knowledge and familiarity with assessment-related terminology, but they were, however, more adept at applying their knowledge when using contextual information. Within the scope of the interviews, only one principal referenced performance assessment and none of the principals referenced authentic assessments, suggesting that alternative forms of assessment may not be widely utilized as a part of each school’s balanced assessment system.

Standard two, which states: *Teachers should be skilled in developing assessment methods appropriate for instructional decisions*, was the lowest scoring standard ($M=2.52$, $SD=1.05$). Stiggins and Duke (2008) highlighted the impact of principals in the
formative assessment process by describing their role in “helping teachers develop and use sound classroom assessment that strengthens instruction and student learning” (p. 286). This suggests that principals’ relative weakness in assessment literacy related to their knowledge of assessment creation and the instructional implications. Specifically, items that addressed item analysis and discrimination values reflected an area of weakness for administrators. A possible explanation for this low scoring standard may be principal’s inexperience with the technical aspect of assessment development.

Additionally, some of the principals mentioned having specialists develop tests or that assessments were developed in central office; therefore, it might not be an area where principals have had extensive experience and training. Principals discussed the use of test item banks to support teachers in assessment development; however, this standard suggests that principals may benefit from additional guidance with how to use this resource with fidelity.

Standard three states: The teacher should be skilled in administering, scoring and interpreting the results of both externally-produced and teacher-produced assessment methods. Black and Wiliam (1998) described challenges that teachers often face when interpreting scores, suggesting that they often focus on comparisons among students when interpreting scores instead of focusing on individual growth (p. 18). Within this standard, however, a strength of principals was score interpretation. Question 10, relating to percentile rank was a relative weakness for principals within this standard. Possible explanation for this is the impact of the state standardized assessment system in Virginia. SOLs are criterion-referenced assessments, not norm-referenced; therefore, principals may have minimal experiences with percentile rank within this overarching framework.
for assessment. Within the interviews, a majority of principals did, however, share that
they involve stakeholders, specifically parents, in the process of understanding and
interpreting student performance.

Standard four states: *Teachers should be skilled in using assessment results when
making decisions about individual students, planning teacher, developing curriculum,
and school improvement.* Popham (2003) articulated the decisions that can be made as a
result of assessments, including a better understanding of what students know, the
curricular objectives, the time necessary for teaching content, and the impact of teaching
(pp. 5-6). Areas of relative strength within this standard included: standardized
assessment, role of formative assessments, and alignment of instruction and assessment.
An area of relative weakness within standard four involved standard error of
measurement. Again, a possible explanation is the technical nature of this concept.
Because Question 4 was incorrectly entered within the survey, a reverse scoring
procedure was used to score the results; therefore, the results for Question 4 and standard
four overall should be interpreted with caution. In order to support school performance on
standardized assessment, the VDOE has provided professional development and support
in the areas of both formative assessment and alignment of instruction and assessment, a
possible explanation for their relative strength within standard four.

Standard five states: *Teachers should be skilled in developing valid pupil grading
procedures, which use pupil assessments.* Relative to the impact of grading practices,
Crooks (1988) described the need for more feedback for students about relative progress
as opposed to summative appraisals of performance. A relative strength within this
standard involved using multiple pieces of information to determine grades. A relative
weakness involved using consistent scoring practices. This suggests that principals understood the need to use multiple pieces of evidence when grading students; however, principals should ensure consistency in grading practices. Two of the principals with higher levels of assessment literacy described the need to analyze student writing through professional learning communities to increase consistency in grading practices.

**Standard six states:** *Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.* Stiggins (2001) highlighted the concern that external stakeholders may not understand the implications of assessment data outside the realm of report cards and standardized assessments. A relative strength within this standard involved an explanation of the concept of percentile. This was inconsistent with participants’ knowledge of percentile addressed in standard three. A relative weakness and overall weakness of the entire assessment was the concept of grade equivalency. This suggests that principals should strengthen their knowledge of and ability to interpret the concept of grade equivalency when communicating scores with stakeholders. Grade equivalencies were not referenced within any of the interviews; however, some of the principals referenced explaining percentiles to stakeholders when interpreting individual students’ scores.

**Standard seven states:** *Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.* Standard seven was a relative strength across standards \((M = 3.89, SD = 0.98)\). Impara et al. (1993) noted the importance of ensuring principals “have sufficient knowledge to protect themselves and their teachers from the potential unethical or improper use of test scores” (p. 520). A strength within this standard relates to using assessment information when
making decisions about student learning. A relative weakness relates to identifying unethical practices to increase student performance. Within the interviews, a theme that emerged was that unethical practices were not a major consideration for principals with higher levels of assessment literacy. With knowledge of ethical practices as a relative strength for the sample of principals, one may have greater confidence in principal’s claims that unethical practices were not a major cause of concern among principals interviewed.

Across all standards, the relative strength was administrators’ knowledge of ethical practices (Standard 7), and the relative weakness was their ability to develop assessment methods (Standard 2). In a national sample of principals, Impara et al. (1993) indicated that administrators scored highest on their ability to administer, score and interpret results, and their relative weakness was their ability to develop assessment methods. With more than two decades separating the research, administrators continue to demonstrate a relative weakness in their ability to develop assessment methods. Impara and Plake (1995) found that Virginia administrators’ strongest areas of assessment literacy included their ability to choose appropriate assessments, analyze the validity of an assessment, share the results of assessments with other stakeholders, and identify unethical practices, and administrators’ lowest scores were in the area of understanding standardized test results. These scores were not consistent with the current sample.

**Assessment Literacy Across Levels**

In addition to an analysis of each of the seven standards, results on the ALI were compared across levels. Current results suggested there were no significant differences among elementary and secondary principals’ levels of assessment literacy, as measured
by the ALI. In a Virginia sample, Impara and Plake (1995) found elementary administrators to be more knowledgeable of assessment than secondary administrators. Current results suggest greater uniformity in assessment literacy in principals across levels throughout The Commonwealth. The surface explanation for this change and greater uniformity in assessment literacy of principals across levels may be attributable to the standardization movement in which principals across levels must be familiar with assessment-related policies and practices. Another explanation is that administrators participate in the same administrator preparatory coursework, regardless of elementary or secondary experiences and therefore receive similar preparation and exposure related to assessment.

Types of Training

In addition to comparing principals’ assessment literacy by level, this study also examined the differences in principals’ assessment literacy as a result of training. Results indicated there were no significant differences among participants’ level of assessment literacy as a result of predominant method of assessment training. One limitation of this interpretation, however, is the impact of small sample sizes across the four areas of assessment training. For instance, teacher preparation coursework and administration preparation coursework had sample sizes smaller than thirty; therefore, these small sample size may limit the ability to make valid comparisons of assessment literacy across types of assessment training.

The primary method of assessment training was professional development as an administrator, suggesting that the strongest mechanism for growing the assessment literacy of administrators in the future is through professional development. Professional
development should involve utilizing assessments in a practical and meaningful way so administrators have an opportunity to support teachers in a similar capacity. Examples of such professional development may involve developing assessments, analyzing assessment-related information to make instructional decisions, observing teachers who effectively employ assessment for learning, and learning strategies for supporting teachers who may not understand how to effectively employ formative assessment practices for student consumption. Impara and Plake (1995) found that 97.4% of administrators surveyed within their sample had taken a course in assessment; however, the format for the assessment training (e.g., undergraduate coursework, graduate coursework, or professional development) was not specified. Perry (2013) found almost split results between administrators who had taken a course in assessment and those that had not. Although Virginia licensure requires administrators to pass the School Leaders Licensure Assessment (SLLA) before obtaining licensure, individual administrative preparatory programs may vary in their level of graduate coursework in assessment. Despite differences in types of assessment training, there were many commonalities in principals’ practices, as evidenced by the analysis of their interviews, as described in the following sections.

Support of Assessment for Learning

Principals referenced multiple support structures that facilitated teachers’ use of assessment for learning practices. Pervasive across principals with higher and lower levels of assessment literacy included the use of professional learning communities, grouping practices to support differentiated instruction, and specialists to provide support to teachers. A fundamental role of the principal is to determine a school’s needs and
utilize resources that will support professional teams (Portin et al., 2003). This sample of principals demonstrated that professional learning communities were a mechanism for support of teachers, regardless of principals’ knowledge of classroom assessment practices. Additionally, Hollingworth (2012) cited professional learning communities as a vehicle for supporting “learning and collaboration” to grow the skillset of teachers (p. 377). The function of these professional learning communities, as described by principals, was to develop assessments, review assessment data, and/or make decisions about necessary supports and groupings for students as a result of student performance on assessments. Through the support of specialists and coaches, principals have an opportunity to further expand the professional knowledge of teachers and develop assessment leadership within their staff.

**Alignment**

In the area of alignment, principals with higher levels of assessment literacy addressed the need for teachers to prioritize the alignment of learning targets; however, this practice was not evident across principals with lower levels of assessment literacy. Absent among principals with lower levels of assessment literacy was the recognition of the integral nature between curriculum, instruction, and assessment. The focus on learning targets by some principals with higher levels of assessment literacy reflected at least a surface level understanding that each of these components exists in tandem and are all necessary to see improved student outcomes. This was addressed through either professional development or university partnerships. Moss et al. (2013) found that “Administrators need to both be part of and provide leadership for the intentional lesson-by-lesson focus on what students are actually doing to develop and produce evidence of
their understanding of essential learning targets” (p. 217). This is an area for growth with principals with lower levels of assessment literacy.

With regard to alignment, there was uniform distribution of the need for and utility of digital-item banks to support assessment development; however, given that assessment development was principals’ lowest scoring area on the ALI, measures at the central office level should ensure administrators know how to effectively employ these resources. Additionally, pacing guides were a widely-used practice to support alignment.

**Professional Development**

Although principals described various topics for professional development within their contexts, the use of professional development on assessment-related issues and topics was not common practice across principal. Principals with higher levels of assessment literacy addressed the need to align learning targets with instruction through professional development; however, this practice was not addressed by principals with lower levels of assessment literacy. Both principals with higher and lower levels of assessment literacy focused on ensuring teachers were appropriately trained to administer various assessments. Previous research emphasized the need for support and training for teachers in order for them to use formative assessment to improve student achievement (Heritage, 2007; Moss et al., 2013; Renihan & Noonan, 2012; Stiggins, 2001; Webber et al., 2013). This area would be an opportunity for continual growth for principals with varied degrees of assessment literacy. As a recommendation from this study, professional development related to assessment topics should be provided frequently to teachers to reflect changes in assessment and to ensure that there is shared understanding of practices. Principals with lower levels of assessment literacy may consider the utilization
of instructional specialists or coaches to facilitate the professional development.

**Balanced Assessment**

A strength across administrators, however, was the use of a balanced approach to assessment. Principals addressed the utility of common assessments, benchmark or quarterly assessments, student growth measures, and varied forms of classroom assessments. These balanced assessment systems “incorporated the strengths of summative, interim, and formative assessments to address instructional, accountability, and learning needs” (Huebner, 2009, p. 85). Principals articulated multiple measures designed to serve unique roles within their contexts. Literacy screenings were more common among elementary principals and principals with higher levels of assessment literacy, which may be attributable to the developmental needs of early literacy instruction in which students are still learning to read. Secondary principals with lower levels of assessment literacy did not describe the use of a literacy screening tool to assess students’ progress and performance in reading. A possible explanation is that secondary principals are less concerned with students’ reading levels and instead prioritize criterion-referenced information from assessments.

Additionally, only one principal referenced the utilization of performance-based assessments. One explanation of this limited view of assessment formats is that principals narrowly define assessment to include measures that align more closely to standardized assessment measures. The definition of balanced assessment reflects multiple uses and formats of assessment. While it was evident that multiple assessments were used to collect data on students within this sample of principals, the use of multiple formats of assessment, such as performance assessment and authentic assessment, were not well-
reflected.

In addition to a limited view of assessment format, some of the principals had difficulty distinguishing between formative and summative assessment within the interview. During the interviews, two of the principals with higher levels of assessment literacy reversed the assessments they listed when discussing formative and summative assessments, but they self-corrected themselves and made the adjustment. One principal with lower levels of assessment literacy was not able to accurately distinguish between the use of formative and summative assessments when providing the various types and uses of assessments. This suggests that for principals across levels of assessment literacy, fundamental concepts such as the function and utility of formative and summative assessment should be considered as an area for additional professional development for principals to ensure principals understand the role and purpose of various assessments.

**Ethical Considerations**

With regard to ethical considerations, a majority of principals with higher levels of assessment literacy did not consider unethical practices to be a pervasive issue. Because this was the highest scoring standard within the ALI, it provides greater assurance that administrators understand ethical issues and practices. In an effort to ensure ethical practices, principals reported using trainings and heightened SOL procedures and practices, most likely attributable to the standardization of practices when implementing SOLs and minimizing opportunities for testing irregularities.

A strength, however, was principals use of opportunities to assist stakeholders in interpreting results of varied assessments, especially to parents through various forms of communication. Only principals with the highest levels of assessment literacy and one
outlier principal with lower levels of assessment literacy described how they or teachers involved students in assessment. There were minimal examples in which principals described specific student involvement in the assessment process. Black and Wiliam (2010) cited the range of effect size for formative assessment, between 0.4 and 0.7 but also cited that formative assessment should provide specific feedback about a student’s work and the ways in which a student can improve. If common assessments, benchmarks, and other assessments are administered but student feedback is absent, teachers will miss an opportunity for students to develop ownership in their own learning. This is recommended as an area for further professional development for administrators as instructional leaders in order to translate this practice to staff.

**Limitations**

One limitation of this study is the sample of principals. With a return rate of 7.6%, this study is subject to response bias. The time commitment to complete the survey was estimated at 20-30 minutes, which may have negatively impacted the overall return rate of completed surveys. Additionally, the interview sample was inclusive of one assistant principal. This assistant principal received the link for participation through an anonymous link; therefore, he was not on the original email distribution list. The results of this interview were included within the final results to ensure adequate representation of secondary administrators, but also because many assistant principals aspire to be principals.

In addition to response bias for survey completion, interview completion was also subject to response bias due to the time commitment. Two additional interviews were conducted to ensure greater representation of secondary principals and principals with
lower levels of assessment literacy, for a total of 12 interviews.

Another limitation of this study is that Question 4 was incorrectly worded within the survey instrument and therefore had to be scored with a reverse scoring procedure. Using this reverse scoring procedure, 100% of answers were correct. This may positively skew the overall performance on the ALI as well as the mean score of Standard 4.

Another limitation of this study is that it does not examine the knowledge, beliefs or practices of teachers or the impact of assessment leadership on classrooms or student performance. Although the sample provided information about school accreditation, the indirect influence of leadership does not provide a causal link between leadership and school performance. The sample did, however, closely mirror the state breakdown of accreditation status through its sampling of principals.

Finally, a limitation of this study is the instrument itself. The ALI is designed to measure the assessment literacy of pre-service teachers. Additionally, the ALI is based on antiquated standards and classroom assessment practices (DeLuca et al., 2016b). In the absence of a more appropriate measure, this instrument was used to quantify principals’ assessment literacy. A more comprehensive and valid measure of a principals’ assessment literacy would reflect an understanding of the integral nature of curriculum, instruction, and assessment. Additionally, a revised instrument could also include the use and application of multiple assessments to gauge student understanding. The revised instrument may also have greater content-validity if it reflected standardized assessment practices. Another area neglected by the instrument is the need to interpret various assessment results and communicate to various audiences. For instance, teachers are now involved in the eligibility process for special populations of students; however, this is not
reflected by a teacher’s knowledge of assessment in this instrument.

Additionally, this study determined that administrators currently see evidence of various assessments, such as benchmarks, common assessments, literacy screenings, student growth assessments, and classroom examples of formative assessments. These varied forms and uses of assessment were not reflected within this instrument; suggesting, furthermore, that the current instrument’s construct of classroom assessment is limited. This instrument neglects many of the changes that have occurred in educational assessment, which may ultimately affect the validity and relevance of data used from this instrument.

**Recommendations for Practice**

To further grow the assessment literacy of administrators, universities, the Virginia Department of Education, and division central office staff should seek opportunities to provide professional development to principals, as this is the predominant method of assessment training reported by principals. An area of focus should be on the development of assessments, and since the use of electronic test item banks is pervasive within the sample of participants, professional development should address how to analyze and use these types of resources for assessment development.

Additionally, although there were established balanced assessment systems through the use of benchmarks, common assessments, and growth measures, there were minimal references to varied forms of assessment, such as authentic assessments or performance assessments. The VDOE has provided guidelines for local alternative assessments in areas where SOL tests have been reduced (VDOE, 2017a). Despite these updated guidelines, principals continued to focus heavily on preexisting assessment
measures, such as common assessments and benchmark assessments. Continual state support and division central office support is needed to ensure principals prioritize this change in assessment reform. Additionally, further professional development should address the role of students in the assessment process and how to further involve students in this process.

Another focus should be on program evaluation of the effectiveness of professional learning communities at the school and/or division level. This widespread practice is designed to analyze assessments and subsequently adjust instruction. These should be monitored regularly for effectiveness to ensure fidelity of implementation and gains in student achievement. If deemed effective, principals should further use these learning communities as venues for professional development and should capitalize on these existing school structures to further support teachers’ development of assessment literacy.

Additionally, instructional specialists were a highly regarded support for teachers and administrators, and they played a significant role in the use of assessment and analysis of assessment. In addition to teachers and administrators, instructional specialists should continue to receive support in their roles and should have open lines of communication between the teachers and the administrative team.

And finally, many of the principals discussed measures of student growth and some referenced the role of growth measures within the scope of the teacher evaluation system. Principals described the need to meet with teachers to review results of assessments; however, it is important that they weigh the roles of “ensuring accountability and quality control, on one hand, and nurturing professional empowerment
among teachers, on the other” (Renihan & Noonan, 2012, p. 4). This role may be difficult to balance, given the demands of school’s teacher evaluation systems and expectations. Some principals described the student growth measures as a pre-post assessment whereas others looked at growth using longitudinal assessments over time. When discussing growth measures, principals frequently referred to the teacher evaluation system within their school. There were varying denotations of student growth measures, as some principals defined this using criterion-referenced information and others looked at statistical growth measures. Greater clarity is needed by principals regarding what constitutes a valid growth measure for student performance.

As a recommendation for teacher and administrator preparation faculty, there are areas of assessment literacy that could be further developed in university preparatory work. For instance, only principals with higher levels of assessment literacy referenced the need to align learning targets to improve student performance on assessments. The interrelatedness and alignment of curriculum, instruction, and assessment should receive heightened priority at the university level. Additionally, the administrators referenced using multiple assessments as part of a balanced assessment system; however, attention should be given to ensure that teachers and administrators know how to use data from a wealth of data points to make informed decisions about student progress and performance. Additionally, attention should be given to ensure teachers and administrators know how to develop assessments that reflect reliable and valid measures of student performance.

Furthermore, the state department should continue to support initiatives that reflect supporting teachers and administrators in the utilization of alternative forms of
assessment to measure student performance. Performance assessment was only addressed by one principal, and further attention should be provided to directed to ensuring that these varied assessment formats are integrated across the curriculum and are considered valuable indicators of student performance. Because of the focus on standardized, multiple-choice assessments, the assessments administrators and teachers value as valid measures of student performance appear more limited in scope.

Recent changes in state licensure regulation now require pre-service teachers to complete a stand-alone course in assessment. This initiative may lead to increased assessment literacy for educators across Virginia; however, a limitation to this approach may deny the integral nature of curriculum, instruction, and assessment. Universities have an opportunity to embed assessment across methods courses for pre-service teachers, but conversely, may not adequately address the technical nature of assessment through this approach. Further research should explore the impact of stand-alone courses in assessment on teachers’ assessment literacy and application of appropriate assessment-related practices.

**Directions for Further Research**

Further research should consider exploration of the relationship between participants’ years of classroom experience and knowledge and application of assessment leadership practices. Additionally, do years of classroom experience impact a principals’ assessment literacy? What school or contextual factors in the classroom may impact a principals’ assessment leadership practices? There have been a wealth of assessment-related forms experienced by teachers in light of No Child Left Behind legislation; therefore, future research should also examine administrator’s recent experiences in the
classroom and how their experiences impact their development and application of assessment leadership practices that support assessment for learning.

Within the interviews, some of the participants shared previous education-related experience in addition to classroom teaching, including central office experience and experience as an instructional specialist/coach. Further research should consider the impact, if any, of a leader’s varied educational roles on assessment leadership practices. Principals with these varied roles had a highly-defined knowledge of classroom assessment and how to use assessment results to support the goals of instruction.

Further research is needed in the area of assessment literacy, specifically in designing an instrument to match the current demands of the classroom and accountability system. The Approaches to Classroom Assessment Inventory is currently an instrument still in developmental stages to address the shortcomings of current instruments, which “do not fully reflect current transformations in the assessment landscape and remain predicated on dated standards for teacher classroom assessment practice” (DeLuca et al., 2016a, p. 2). This new instrument was developed to serve as a “reliable instrument reflective of contemporary assessment practices and contexts” (p. 2). Within the current study, shortcomings of the current inventory within the sample were its intended use with pre-service classroom teachers; however, because of principals’ instructional leadership role and the inadequacy of existing measures, it proved to be the most appropriate existing measure of assessment literacy.

Further research should also explore principals’ knowledge and integration of varied assessment forms, such as performance assessment and authentic assessment. As state guidelines reflected a decrease in the number of state assessments, it would be
imperative for Virginia principals to lead classroom reform and increase opportunities for students to have learning opportunities in which assessment exists in conjunction with learning and/or which assessment is a meaningful learning experience.

Additionally, next steps for further research would involve examining the relationship between principals’ assessment leadership practices and the impact on teachers’ professional knowledge and classroom application, and ultimately student performance. Although the area of instructional leadership has been heavily researched, the impact of assessment leadership proves to be an area in need of further research.
Appendix A

Survey Part I

1. Please provide your years of experience as a classroom teacher: ______

2. Please provide your years of experience as an administrator, including this current year ______

3. Please describe your primary method of assessment training
   a. Initial teacher preparation coursework
   b. Administration preparation coursework
   c. Professional development as a teacher
   d. Professional development as an administrator

4. Please indicate your school level
   a. Elementary (grades pk-5)
   b. Secondary (6-12)
   c. Other (Please describe): _____________

5. Please indicate your current accreditation status for the 2016-2017 school year:
   a. Fully Accredited
   b. Partially Accredited: Approaching Benchmark-Pass Rate
   c. Partially Accredited: Approaching Benchmark-Graduation and Completion Index
   d. Partially Accredited: Improving School-Pass Rate
   e. Partially Accredited: Improving School-GCI
   f. Partially Accredited: Warned School-Pass Rate
   g. Partially Accredited: Warned School-GCI
   h. Partially Accredited-Reconstituted School
i. Accreditation Denied

j. To be Determined

6. Please indicate if you would like to be contacted to participate in the phone interview as a follow-up to this survey. Individuals who participate in the phone interview will be entered for a chance to win a $100 Visa gift card. Odds of winning are one in 10. Please provide your name, email, and phone number if you wish to participate. Your score on this assessment will only be identifiable to the researcher and the results from the interview will not reveal personal or school identifiable information within the published study.

a. Yes, please contact me for a follow-up interview

   Name:

   Phone number:

   Email address:

b. I do not wish to participate in a follow-up interview
Appendix B

Assessment Literacy Inventory

Cynthia Campbell, Ph.D.
Northern Illinois University

Craig A. Mertler, Ph.D.
Bowling Green State University

Description of the ALI

The Assessment Literacy Inventory (ALI) consists of five scenarios, each followed by seven questions. The items are related to the seven "Standards for Teacher Competence in the Educational Assessment of Students." Some of the items are intended to measure general concepts related to testing and assessment, including the use of assessment activities for assigning student grades and communicating the results of assessments to students and parents; other items are related to knowledge of standardized testing, and the remaining items are related to classroom assessment.

Directions:

Read each scenario followed by each item carefully; select the response you think is the best one and mark your response on the answer sheet. Even if you are not sure of your choice, mark the response you believe to be the best.

Scenario #1

Ms. O'Connor, a math teacher, questions how well her 10th grade students are able to apply what they have learned in class to situations encountered in their everyday lives. Although the teacher's manual contains numerous items to test understanding of
mathematical concepts, she is not convinced that giving a paper-and-pencil test is the best method for determining what she wants to know.

1. Based on the above scenario, the type of assessment that would *best* answer Ms. O'Connor's question is called a/an
   A. performance assessment.
   B. authentic assessment.
   C. extended response assessment.
   D. standardized test.

2. In order to grade her students' knowledge accurately and consistently, Ms. O'Connor would be well advised to
   A. identify criteria from the unit objectives and create a scoring rubric.
   B. develop a scoring rubric after getting a feel for what students can do.
   C. consider student performance on similar types of assignments.
   D. consult with experienced colleagues about criteria that has been used in the past.

3. To get a general impression of how well her students perform in mathematics in comparison to other 10th graders, Ms. O'Connor administers a standardized math test. This practice is acceptable *only* if
   A. the reliability of the standardized test does not exceed .60.
   B. the standardized test is administered individually to students.
   C. the content of the standardized test is well known to students.
   D. the comparison group is comprised of grade level peers.
4. Which of the following is an inappropriate use of the results from this standardized math test?
A. planning instruction  
B. assigning student grades  
C. determining students' strengths and weaknesses  
D. developing curriculum  

5. Throughout instruction, Ms. O'Connor assesses how well her students are grasping the material. These assessments range from giving short quizzes following introduction to a new topic, to administering an end-of-the-unit final exam. In order to improve the validity of this grading procedure, Ms. O'Connor should
A. make the grading scale the same for all assessments.  
B. consider students' prior performance before assigning a final grade.  
C. weight assessments according to their relative importance.  
D. take into consideration each student's effort when calculating grades.  

6. During a parent teacher conference, one of the parents of a student in Ms. O'Connor's class wants to know what it means that his daughter scored in the 80th percentile in mathematics. Which of the following provides the best explanation of this student's score?
A. She got 80% of the items on the math test correct.  
B. She is likely to earn a grade of 'B' in her math class.  
C. She is demonstrating above grade level performance in math.  
D. She scored the same or better than 80% of the norm group.
7. Which of the following is an appropriate use of assessment information?

A. Utilize information from a variety of assessments when making decisions about student learning.
B. Use scores from standardized tests to determine teacher instructional effectiveness.
C. Use scores from a standardized test as the primary indicator of student retention.
D. Post final grades in order to provide normative information to students in the class.

Scenario #2

Mr. Okawa, a fifth-grade teacher, is planning his instruction for the next grading period, aware of the fact that his students will be taking the statewide achievement test near the end of the grading period.

8. Mr. Okawa's mathematics unit for this grading period will focus on multi-step problem-solving. He wants to assess his students' problem-solving abilities at the end of the unit to determine if any reinstruction will be necessary prior to the statewide test. Which of the following assessment strategies would be the most appropriate choice?

A. He should choose the assessment included in the teacher's manual from the textbook he uses.
B. He should choose an assessment which is consistent with the content and skills he taught.
C. He should choose a different standardized assessment that provides a score on similar skills.

D. He should choose an assessment which covers single-step problem-solving skills.

9. Mr. Okawa decides to develop his own assessment in order to determine if any reinstruction will be necessary. He also wants to use his assessment as a means of anticipating how his students will perform on the statewide assessment. In order for him to accurately approximate his students' performance, which of the following would be the most appropriate type of assessment for him to develop?
   A. a performance assessment
   B. a multiple-choice test
   C. a portfolio assessment
   D. an essay test

10. Julie, one of Mr. Okawa's students, receives a percentile rank of 60 on the problem-solving skills subtest of the statewide assessment. This score is most appropriately interpreted as which of the following?
   A. Julie scored above average.
   B. Julie scored below average.
   C. Julie scored at the national average.
   D. Not enough information to determine.

11. Juan, another student in Mr. Okawa's class, receives a scaled score of 196 on the reading comprehension portion of the statewide assessment. The cut score is 200; therefore, Juan does not pass this subtest. However, the subtest has a standard
error of measurement equal to 6. Which of the following is the best decision for Mr. Okawa to make regarding instruction appropriate to meet Juan's needs?

A. Juan has clearly not achieved the minimum level of reading comprehension and should receive remedial reading instruction.

B. Mr. Okawa knows that Juan could have scored higher, so the results of the test should be ignored.

C. Juan may likely have achieved the minimum level of reading comprehension and nothing different or additional should be done.

D. Mr. Okawa knows that Juan should have scored much lower, so the results of the test should be ignored.

12. Which grading practice being considered by Mr. Okawa would result in grades that would least reflect achievement?

A. grades based on daily homework and chapter tests

B. grades based on daily homework and chapter tests, with points deducted for poor effort

C. grades based on daily homework and chapter tests, where students are permitted to redo assignments in order to meet higher standards

D. grades based on chapter tests, where daily homework is not formally graded

13. Barbara scores at the 60th percentile on mathematics problem-solving and at the 56” percentile on reading comprehension. The percentile bands for each test are five percentile ranks wide. What advice should Mr. Okawa give to Barbara's parents?

A. They should ignore the difference; her performance was essentially the same
on the two tests.

B. They should seek additional tutoring help for Barbara in reading.
C. They should force Barbara to read more at home.
D. They should provide enrichment experiences for Barbara in math, which is her better performance area.

14. Mr. Okawa was worried that his students would not perform well on the statewide assessment. He did all of the following to help increase students' scores. Which was unethical?
A. He instructed students in strategies for taking multiple-choice tests, such as how to use answer sheets.
B. He planned his instruction so that it focused on concepts and skills to be covered on the test.
C. He encouraged the students to do their best, and provided them with a reward after testing was complete.
D. He allowed students to practice with items from an alternate form of the test.

Scenario #3

Ms. Green is an eighth-grade American History teacher. She has just finished teaching a unit on the Industrial Revolution and wishes to make decisions about her students regarding their higher-order thinking skills. Ms. Green has decided to give her students a single assessment in the form of an end-of-unit multiple-choice test. She anticipates that most of her students will perform well on the test.

15. Based on her goal, what can you conclude about her decision to administer a multiple-choice test?
A. This is an appropriate choice for a unit assessment.
B. The test scores may not be valid for this purpose.
C. The test scores may not be reliable for this purpose.
D. A true-false test would be more appropriate.

16. To determine the quality of her multiple-choice test, Ms. Green should conduct an item analysis and examine all of the following except
A. item difficulty values.
B. item discrimination values.
C. reliability coefficients.
D. validity coefficients.

17. Ms. Green decides to score the tests using a 100-percent correct scale. Generally speaking, what is the proper interpretation of a student score of 85 on this scale?
A. The student answered 85% of the items on the test correctly.
B. The student knows 85% of the content covered by this instructional unit.
C. The student scored higher than 85% of other students who took this test.
D. The student scored lower than 85% of other students who took this test.

18. Some of Ms. Green's students do not score well on the multiple-choice test. She decides that the next time she teaches this unit, she will begin by administering a pretest to check for students' prerequisite knowledge. She will then adjust her instruction based on the pretest results. What type of information is Ms. Green using?
A. norm-referenced information
B. criterion-referenced information
C. both norm- and criterion-referenced information
D. neither norm- nor criterion-referenced information

19. The Industrial Revolution test is the only student work that Ms. Green grades for the current grading period. Therefore, grades are assigned only on the basis of the test. What is the major criticism of this practice?
A. The test, and therefore the grades, reflect too narrow a curricular focus.
B. These grades, since based on tests alone, is probably biased against some minority students.
C. She should add extra points to the scores of students who scored low on the test.
D. Decisions like grades should be based on more than one piece of information.

20. Mr. Simpson, another American History teacher, bases his grades primarily on his observations of students during class. The primary distinction between his system of assigning grades and that used by Ms. Green is best characterized as which of the following?
A. Ms. Green uses formal assessment; Mr. Simpson uses informal assessment.
B. Ms. Green uses formative assessment; Mr. Simpson uses summative assessment.
C. Ms. Green uses standardized assessment; Mr. Simpson uses nonstandardized assessment.
D. Ms. Green uses traditional assessment; Mr. Simpson uses alternative assessment.
21. Based on their grades from last year, Ms. Green believes that some of her low-scoring students are brighter than their test scores indicate. Based on this knowledge, she decides to add some points to their test scores, thus raising their grades. Which of Ms. Green's actions was unethical?
   A. examining her student's previous academic performance
   B. adjusting grades in her course
   C. using previous grades to adjust current grades
   D. adjusting some students' grades and not others'

Scenario #4

Mr. Valdez is an English teacher in the newly built middle school. Experienced in issues of classroom assessment, Mr. Valdez is often asked to respond to the district's questions concerning best practices for evaluating student learning.

22. Ms. Franklin, also an English teacher, asks what type of assessment is best for evaluating her 6th graders' writing skills. Which of the following methods is likely to provide the best response to her question?
   A. selected response methods
   B. true/false statements
   C. completion items
   D. essay prompts

23. One of the middle school math teachers is redesigning her tests to make greater use of "story problems" as a way to check students' math understanding. She consults with Mr. Valdez to see what, if any, concerns she should be aware of
when constructing assessments of this type. Which statement is not an appropriate recommendation when designing story-based math tests?

A. make sure that the reading level is grade appropriate
B. avoid scenarios more familiar to certain groups over others
C. check for clarity of sentence construction
D. incorporate scenarios used during instruction

24. Isabel, a student in Mr. Valdez's class, scored 78 points on a standardized English test which had a mean of 80 and a standard deviation of 4. She scored 60 points on the science portion of this test which had a mean of 50 and a standard deviation of 3. Based on the above information, in comparison to her peers, which statement provides the most accurate interpretation?

A. Isabel is better in English than in science.
B. Isabel is better in science than in English.
C. Isabel is below average in both subjects.
D. Isabel is close to average in both subjects.

25. At the end of each class period, Mr. Valdez does a quick "check in" with his students to get an impression of their understanding. In this example, the primary purpose for conducting formative assessment is to

A. identify cumulative knowledge.
B. determine content for the final exam.
C. plan classroom instruction.
D. evaluate curriculum appropriateness.
26. To prepare students for state testing and identify areas of school improvement, all 6th grade English teachers give a common final exam which contains a series of essay items. Recently, however, several teachers have expressed concern that the time and effort necessary to complete grading on a timely basis may result in inconsistent scoring. They consult with Mr. Valdez. Which of the following provides the best response to the teachers' concern for consistency?
   A. grade all responses to essay #1 before grading responses to essay #2
   B. during grading, adjust rubric criteria to reflect exemplary student work
   C. utilize a holistic scoring method to minimize teacher subjectivity in scoring
   D. all things being equal, it is best to limit the use of multiple essay exams

27. Jeremy, a 6th grade student in Mr. Valdez's class, received a grade equivalent score of 7.2 on a standardized reading test. Jeremy's parents wonder what this means. Based on the above information, which of the following statements provides the most appropriate interpretation of this student's score?
   A. Jeremy is reading at the 7th grade level.
   B. Jeremy is reading better than the majority of students in his class.
   C. Jeremy is reading 6th grade material as expected.
   D. Jeremy should be placed in a 7th grade reading class.

28. "To ensure that standardized test results provide an accurate picture of what students really know, it is recommended that teachers clarify items that are confusing to students."

   Based on best practices of assessment, which of the following is an appropriate response to the above statement?
A. The above statement is an acceptable way to reduce error in testing.
B. The above statement is an acceptable way to increase test validity.
C. The above statement is unacceptable because it labels students as poor readers.
D. The above statement is unacceptable because it breaks standardization.

Scenario #5

Ms. Hawkins is responsible for teaching science at the 4th grade level. Over the past couple of years, her students have really seemed to struggle with investigations of how water changes from one state to another (i.e., freezing, melting, condensing, and evaporating), but she is unsure of where the specific difficulties lie. She is aware that her students need to improve their conceptual understanding of this content standard.

29. Ms. Hawkins wishes to conduct some sort of assessment in order to identify the specific difficulties her students are experiencing. Which of the following would best meet her needs?
   A. a diagnostic assessment
   B. an informal assessment
   C. a standardized assessment
   D. a summative assessment

30. In an effort to refine both her instruction and assessment of this content, Ms. Hawkins conducts an item analysis of student scores from last year's final unit test over this material. She should definitely discard or substantially revise a test item that
   A. has a difficulty value between .50 and .75.
B. has a discrimination value equal to +.30.
C. has a discrimination value equal to -.50.
D. has a difficulty value equal to .90.

31. Ms. Hawkins' unit test also includes a restricted-response essay item. She is concerned with the demonstrated level of understanding of several specific criteria in her students' responses. Which of the following would best facilitate her scoring of these responses?
   A. an objective answer key
   B. a holistic rubric
   C. a checklist
   D. an analytic rubric

32. Following the completion of the unit, Ms. Hawkins determines that her students have satisfactorily mastered these concepts. However, when her students take the statewide standardized assessment in the spring, she notices that her students perform very poorly on items addressing these same concepts. Considering the discrepancy between students' classroom performance and their standardized test results, what action is most appropriate when making decisions concerning school improvement?
   A. recommend that classroom instruction be consistent among 4th grade science teachers
   B. ensure alignment between instruction and what is measured on the standardized test
   C. select a standardized test that is more likely to yield higher scores in science
D. identify the percentage of students predicted to perform well in advanced science classes

33. Ms. Hawkins wants to be sure that the term grades she assigns to her students' performance in science reflect each student's respective level of content mastery for that unit. Which of the following grading systems would best accomplish this goal?

A. a criterion-referenced grading system
B. a norm-referenced grading system
C. a pass-fail grading system
D. a portfolio grading system

34. Nolan is a student in Ms. Hawkins' class. He receives a raw score of 12 items answered correctly out of a possible 15 on the physical science portion of a standardized test. This raw score equates to a percentile rank of 45. His parents are confused about how he could answer so many items correctly, but receive such a low percentile rank. They approach Ms. Hawkins for a possible explanation. Which of the following is the appropriate explanation to offer to his parents?

A. "I don't know...there must be something wrong with the way the test company figured the scores."
B. "Although Nolan answered 12 correctly, numerous students answered more than 12 correctly."
C. "Raw scores are purely criterion-referenced and percentile ranks are merely one form of norm-referenced scoring."
D. "Raw scores are purely norm-referenced and percentile ranks are merely one form of criterion-referenced scoring."

35. In an attempt to try to encourage and motivate her students who are struggling academically, Ms. Hawkins decides to share her gradebook, especially test scores, with them in order to demonstrate how well others are performing. Another teacher advises her not to do this, as it is a clear violation of


C. The Standards for Teacher Competence in the Educational Assessment of Students.

D. The No Child Left Behind Act.
Appendix C

Interview Protocol

Participant Name:

School:

School Level:

ALI Score:

Before we begin, I would like your permission to audio record our conversation today. For your information, only researchers on this project will have access to the audio recordings that are stored on a password-protected computer after they are transcribed using a web-based transcription service. Your signature on the Research Participation Informed Consent Form at the beginning of the study stated that (1) all information will be held confidential, (2) your participation is voluntary and you may stop at any time if you feel uncomfortable, and (3) we do not intend to inflict any harm. Thank you for agreeing to participate in the interview portion of this study. This interview includes ten questions, and is designed to last approximately 30 minutes. I may ask follow-up or clarifying questions. By participating in this interview, you will be entered into a second random drawing to win a $100 Visa gift card. Odds of winning are one in ten. The winner of the gift card will be randomly selected among the ten interview participants, and will be selected and notified upon completion of all ten interviews.

Introduction

You have been asked to speak with me today to describe your support of classroom assessment practices in your school. This dissertation is designed to investigate the level of assessment literacy of building principals in Virginia as well as describe their support
of assessment for learning practices within their school. Assessment literacy is a term used to describe principals’ knowledge of assessment practices. This study seeks to describe how principals’ knowledge of assessment practices relates to their use of assessment practices within their respective contexts. Our study does not intend to evaluate your individual practices or knowledge. Rather, we are trying to learn more about the overall knowledge of Virginia principals and their support of assessment practices across The Commonwealth.

A. Interview Background

The next few questions are designed to describe your educational background and experience.

1. How long have you been in your present position including this year?
2. What is your highest degree earned and what was it earned in?
3. What grade levels or content areas have you taught?

B. Interview Questions

The next section seeks to find out more about your assessment leadership practices. For the purposes of this interview, assessment is defined as: “Any systematic method of obtaining information, used to draw inferences about characteristics of people, objects, or programs; a systematic process to measure or evaluate the characteristics of performance of individuals, programs, or other entities, for purposes of drawing inferences; sometimes used synonymously with test” (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014, p. 216).
1. What strategies, if any, do you employ to help staff use assessment to support student learning?

2. What evidence, if any, do you look for in the classroom to determine if assessment is guiding the learning?

3. What evidence do you look for to determine if an assessment is aligned to achievement targets?

4. What strategies, if any, has your school or division employed to ensure alignment of standards, learning intentions, and assessments?

5. What professional development opportunities, if any, are provided for teachers to contribute to their use of sound assessment practices?
   a. Probing question: What is your role in the professional development?
   b. Probing question: What is your role after the professional development?

6. How are specific professional development assessment topics chosen?

7. What formative assessment practices are consistently used as part of your school’s overall assessment system, if any, and what is the role of the assessment?

8. What summative assessment practices are consistently used as part of your school’s overall assessment system, if any, and what is the role of the assessment?

9. What strategies, if any, do you use to prevent unethical and inappropriate administration of assessments and unethical and inappropriate use of assessments and assessment results?

10. What strategies, if any, do you use to ensure stakeholders, including students, parents, and school community make appropriate interpretations from various assessments?
Conclusion

This concludes our interview. Following the interview, a transcript of the interview will be emailed to you for your review. Please verify that the responses accurately reflect your practices and beliefs. Thank you again for your time and participation in this study.
Appendix D

Researcher as Instrument Statement

As a current principal in a primary school context, my experiences and practice are constantly shaped by my former experiences as an administrator, teacher, and student. My role as an instructional leader has been influenced by my passion for quality pedagogy, curriculum, and assessments. Furthermore, my passion for assessment leadership has additionally been influenced by the significant role assessment has played in my K-12 experiences as a former teacher.

As a former teacher at elementary levels, I experienced assessment in various capacities. With regard to summative assessments, I saw the pressure, anxiety and impact of high-stakes testing, including standards of learning assessments and benchmark assessments. I attempted to use the data from each of these assessments to make adjustments to my instruction; however, I more closely associated these forms of assessment as a judgment of my success as a teacher.

I quickly found greater power in the use of formative assessments within my classroom to make “in the moment” adjustments to my instructional practices. A quick scan of my students enabled me to see which students mastered content and which required additional remediation. It was this type of fluid instruction that enabled me to build a climate in which students felt comfortable taking risks, and I was able to match their needs to my instruction.

When I moved into the role of administrator, I felt compelled to strengthen teacher’s use of assessment through alignment of curriculum, instruction, and assessments. I worked with teachers to develop Tables of Specifications to strengthen the
validity and reliability of their unit assessments. Additionally, I was able to develop a new assessment system that relied on strand mastery as opposed to the traditional benchmark system that I found ineffective with my own population of students as a teacher.

As an administrator, I continue to monitor student data through the use of unit assessments, PALS assessment, running records, cold read assessments, and through Measures of Academic Progress. I encourage teachers to analyze student growth as well as achievement when discussing academic progress and performance. These experiences have influenced my current beliefs about the role of principals as assessment leaders within their respective contexts.

Beliefs

I wholeheartedly believe that formative assessment plays a valuable role in the learning process. When used to modify instruction, formative assessment wields tremendous power and potential. I believe it is one of the most powerful instructional interventions that requires teacher with-it-ness and teacher experience. I do, however, believe that this type of assessment is often overshadowed by high-stakes summative tests.

I believe that the pendulum has swung too far in the direction of summative assessment. Teacher and student performance and achievement is narrowly defined by one test on an isolated day covering a sliver of the curriculum. This has resulted in many teachers “teaching to the test” or teaching test taking in isolation. I must clarify; however, that I am not opposed to accountability. Instead, I believe that an assessment system that
uses a balance of formative and summative assessments would paint a clearer picture of what students know and can do.

I also believe that the principal as the instructional leader of a building must set the tone for assessment use within his or her respective context. He or she must communicate the varied roles and purposes of assessment. I believe a principal must highlight the importance of formative assessment in the learning process while also ensuring there is alignment between what is taught and what is assessment.

Values

Because of this belief system, I value the role of instructional leadership above all other roles as a principal. He or she is responsible for setting the instructional vision for the school. Although instructional and assessment are not always described in tandem, I value the role of assessment leadership because of the integral roles of instruction and assessment.

As an assessment leader, I value the role of feedback within the school setting. First, teachers must receive feedback from administrators about the nature of assessment within the classroom. Additionally, this can be reciprocated with students as teachers adapt their instruction to include targeted feedback to students as part of the formative assessment process.

Finally, I value a balanced assessment system. The assessment system should be aligned to curriculum and instructional practices. The assessment system should clearly articulate the reason and purpose for each type of assessment so that assessments are used in an appropriate, ethical manner. Additionally, I value the role of teacher leadership in crafting the assessment system of a school to increase teacher buy-in and understanding.
Expectations

I anticipate that administrators with lower levels of assessment literacy will have underdeveloped assessment systems within their schools. They may overemphasize the role of summative assessments and underemphasize formative assessments. Conversely, I anticipate that principals with higher levels of assessment literacy will place an increasingly greater emphasis on formative assessment and its influence on student learning.

Additionally, I expect that principals at an elementary level will score higher on the Assessment Literacy and have a more thorough understanding of the role of assessment for learning. I have based this expectation based on a review of the literature as well as my own observation of secondary versus elementary teachers.

Willing and Unwilling to Discover

I am willing to discover the multi-faceted role of assessment leadership. Additionally, I am encouraged to determine which assessment leadership competencies principals have mastered and which require additional professional development. I am also willing to see how principals with high levels of assessment literacy influence the practices of their schools and teachers.

I am, however, unwilling to discover that assessment leadership does not have a significant role for K-12 principals. Because of the impact of high-stakes testing on students, I am unwilling to acknowledge that principals ultimately are not responsible for creating and implementing balanced assessment systems within their schools.
Outcomes

I believe this research has tremendous influence to guide the professional learning needs of principals across Virginia. It will describe opportunities for principals to learn more about assessment and what they need to know.

Additionally, I believe this research also serves to inform readers about the impact of assessment literacy on principals’ overall leadership practices. Themes generated from this study could help determine if a principal’s assessment literacy ultimately impacts his or her practice.

Conclusion

Assessment leadership is an underdeveloped area of the literature; however, assessment in K-12 schools is pervasive, expensive, and time consuming. In order to best manage assessments and appropriate use them to guide instruction, we must first acknowledge our own understandings and how our understandings influence our practices.
Appendix E

Research Participation Informed Consent Form

Education Department
College of William and Mary
Protocol # EDIRC-2017-01-10-11712-lwgran

Title: Virginia Principals' Knowledge of Classroom Assessment and Support of Assessment for Learning Practices

By proceeding with this study, this is to certify that I, have been given the following information with respect to my participation in this study:

1. Purpose of the research: To investigate the level of assessment literacy of building principals in Virginia as well as describe their support of assessment for learning practices within their school.

2. Procedure to be followed: As a participant in this study, you will be asked to provide basic demographic information related to yourself as an administrator and related to your specific school context. Following the demographic portion of the survey, you will be asked to complete the Mertler and Campbell (2005) Assessment Literacy Inventory. If you are willing to participate in a phone interview as a follow-up to the survey, you will also provide your email and contact information.

3. Discomforts and risks: There are no known risks associated with participation in the study.

4. Duration of participation: Participation in this study will take approximately 30 minutes for completion of the survey. If you elect to participate in a follow-up phone interview, it will last approximately thirty minutes.
5. Statement of confidentiality: Your participation is confidential. The data you contribute to this research will be identifiable only to the experimenter and will not be linked to you or your school within the published results. Participants who agree to be contacted for a follow-up interview will provide their name, phone number, and email; however, these individuals’ actual scores on the Assessment Literacy Inventory will remain confidential. Moreover, all data and records will be stored on password-protected computers.

6. Voluntary participation: Participation is voluntary. You are free to withdraw at any time without penalty or loss of benefits. You may choose to skip any question or opt not to participate in the follow-up interview portion of the investigation.

7. Incentive for participation: Participants will be entered to win one of five $100 Visa gift cards for participation in the survey. Additionally, participants who agree to participate in the phone interview will be entered to win a second $100 Visa gift card, in which odds of winning are one in ten.

8. Potential benefits: There are no known benefits to your individual participation in the study. However, your participation in this research will contribute to the development of our understanding about the nature of principals’ assessment literacy and support of assessment for learning practices.

9. Termination of participation: Participation may be terminated by the experimenter if it is deemed that the participant is unable to perform the tasks presented.

10. Questions or concerns regarding participation in this research should be directed to: Rachel Ball, (804) 339-6730 or rfperv@email.wm.edu. I am aware that I must be at least 18 years of age to participate in this project. I am aware that I may report dissatisfactions
with any aspect of this study to Dr. Jennifer Stevens, Ph.D., the Chair of the Protection of Human Subjects Committee by telephone (757-221-3862) or email (jastev@wm.edu). I agree to participate in this study and have read all the information provided on this form. By entering my name and clicking this box, I agree to the terms and conditions, as stated in this letter.

PROJECT WAS APPROVED BY THE COLLEGE OF WILLIAM AND MARY PROTECTION OF HUMAN SUBJECTS COMMITTEE (Phone: 757-221-3966) ON [2017-01-17] AND EXPIRES ON [2018-01-17]
References

American Educational Research Association, American Psychological Association, &
educational and psychological testing*. Washington, DC: American Educational
Research Association.

American Federation of Teachers, National Council on Measurement in Education, &
National Education Association. (1990). *The standards for competence in the
educational assessment of students*. Retrieved from ERIC database. (ED323186)

Education: Principles, Policy & Practice, 5*(1), 7-74.
doi:10.1080/0969595980050102

doi:10.1177/003172171009200119

study of teaching and learning*. Retrieved from ERIC database. (ED439123)

Bloom, B. S. (1968). *Toward a theory of testing which includes measurement-evaluation-
assessment* (Vol. 9). Retrieved from Center for the Study of Evaluation of

Brown, L. I. (2001). *A meta-analysis of research on the influence of leadership on
student outcomes* (Doctoral dissertation). Retrieved from ProQuest Dissertations
& Theses Global. (Order No. 3110265).


doi:10.1080/13632430701800060


196


http://www.peecworks.org/peec/peec_research/I01795EFA.0/Marzano%20BalancedLeadership.pdf


VITA

Rachel Previs Ball was born in Richmond, Virginia on December 6, 1985. She was raised in King William, Virginia by her loving parents, Steve and Ellen Previs along with her older sister, Dr. Rebecca Ann Previs.

Rachel completed her K-12 educational journey in King William County Public Schools and Chesapeake Bay Governor’s School in 2004. She attended The College of William and Mary and double majored in Psychology and Elementary Education, and completed her B.S. degree in three years, graduating summa cum laude in May 2007. She extended her education at William and Mary by completing a Master of Education degree in August, 2011 in the Educational Policy, Planning, and Leadership Program. She immediately began her doctoral coursework in the Educational Policy, Planning, and Leadership Program at William and Mary and plans to graduate in August 2017.

Rachel is beginning her eleventh year in public education and tenth year in King William County Public Schools where she grew up. She spent five years in the classroom, teaching second, third, and fifth grades. She became an assistant principal in July 2011 at Acquinton Elementary School, where she had been teaching. After two years as an administrator, she became the principal of Cool Spring Primary School in King William County Public Schools, where she is beginning her fourth year as principal. She completed the SURN Principal Academy in 2016 and was the first recipient of the SURN Virginia L. McLaughlin Collaborative Leadership Award.

Rachel currently resides in Dunnsville, Virginia with her husband of nine years, Carter Ball. They have two dogs and enjoy their quiet life on the Rappahannock River where she raises oysters, goes kayaking, and watches morning sunrises on the water.