


Spring 2017

Examining the Relationship Between Selected Grade 3-12 Teachers' Perceived Assessment Literacy and Their Classroom Assessment Practices

Mark E. Hudson
Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/efl_etds

 Part of the [Educational Assessment, Evaluation, and Research Commons](#), and the [Educational Leadership Commons](#)

Recommended Citation

Hudson, Mark E.. "Examining the Relationship Between Selected Grade 3-12 Teachers' Perceived Assessment Literacy and Their Classroom Assessment Practices" (2017). Doctor of Philosophy (PhD), dissertation, Educ Foundations & Leadership, Old Dominion University, DOI: 10.25777/gze1-wb31
https://digitalcommons.odu.edu/efl_etds/37

This Dissertation is brought to you for free and open access by the Educational Foundations & Leadership at ODU Digital Commons. It has been accepted for inclusion in Educational Foundations & Leadership Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

EXAMINING THE RELATIONSHIP BETWEEN SELECTED GRADE 3-12 TEACHERS'
PERCEIVED ASSESSMENT LITERACY AND THEIR CLASSROOM ASSESSMENT
PRACTICES

by

Mark E. Hudson
B.B.A. December 1990, Howard University
Teaching Certification December 1994, Norfolk State University
M.S. Ed. December 2001, Old Dominion University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Concentration in

EDUCATIONAL LEADERSHIP

OLD DOMINION UNIVERSITY

May 2017

Approved by:

Dr. Steven Myran (Director)

Dr. Karen Sanzo (Member)

Dr. Linda Bol (Member)

ABSTRACT

EXAMINING THE RELATIONSHIP BETWEEN SELECTED GRADE 3-12 TEACHERS' PERCEIVED ASSESSMENT LITERACY AND THEIR CLASSROOM ASSESSMENT PRACTICE

Mark E. Hudson
Old Dominion University,
Director: Dr. Steven Myran

The study examined the self-reported understandings of selected 3rd to 12th grade teachers had of assessment and the assessment practices they reported implementing in their classrooms along with evidence extracted from written lesson plans. The literature on classroom assessment supports the idea that teachers who create meaningful assessments, offer corrective action, and give students multiple opportunities to demonstrate success can improve their instruction and increase student learning (Guskey, 2003). McMillan, Myran, and Workman (2002) argued that teachers' understanding of assessment matters are inadequate; although there is common use of standardized math and reading tests, there is limited knowledge of how the assessments are scored, what inferences can be drawn, and even less knowledge of issues involving reliability and validity. The mixed-method study examines the relationship between these two phenomena using a model of assessment literacy that holds the student as the most essential variable in a data-driven practice involving interdependent actions that results in increase achievement. To achieve this goal, a teacher questionnaire was distributed to teachers working in 10 different schools serving varied student populations. As a method of triangulation, the data gathered by the questionnaire was reconciled with a systematic analysis of lesson plans to establish consistent themes. Findings indicate an implementation gap as it relates to teachers' surface knowledge of assessment and what they practice daily in the classroom. Additionally, the study found that teachers do not consistently integrate assessment activities into daily instruction

ACKNOWLEDGMENTS

Our Sun is the source of all life and it has been that kind of life energy that my family, friends, classmates, and colleagues have provided me that allowed me to complete my doctoral journey. In that sense, Sunshine has been my inspiration and I hope it continues to shine in my life. My committee members, Dr. Myran (Chair), Dr. Sanzo, and Dr. Bol provided me with guidance, wisdom and knowledge throughout my process, and I cannot thank you all enough for your help and encouragement. I would also like to thank all of the students and teachers I have come in contact with during my career—you have been my motivation to learn and grow in my knowledge of teaching and learning. I would like to further thank the central office personnel, past and present, working in my system of schools, for their support, encouragement and advice. They have served as my models of good character, ethics, professionalism, and continuous learning. Finally, I would like to thank my parents, Edward Hudson and Barbara Hudson (deceased). My father has been my conscious and unconscious role model during my adult years, and has provided me with the advice and support that have proved invaluable. My mother was my first teacher, molder of my character and my protector. Her work ethic and love for teaching blessed generations of students at the Eva C. Mitchell Laboratory School at Hampton University (Institute), and I can only hope that I can affect the lives of young people a fraction of how she was able to do so in her life.

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	v
LIST OF Figures.....	vi
Chapter 1: Introduction.....	1
Statement of Purpose and Research Questions.....	8
Significance of the Study.....	10
Conceptual Framework.....	11
Definition of Terms.....	14
Chapter 2: Review of the Literature.....	16
Evolution of Assessment and Accountability.....	17
Assessment Literacy.....	21
Empirical Evidence Linking Assessment and Student Learning.....	25
Defining the Forms of Assessment.....	29
Assessment Practice.....	32
Summary.....	34
Chapter 3: Research Design and Methodology.....	36
Research Design.....	37
Participants.....	39
Instrumentation.....	43
Data Collection.....	45
Data Analysis.....	47
Limitations and Biases.....	48
Chapter 4: Research Findings.....	50
Description of the Analysis.....	52
Final Themes from Triangulation.....	69
Summary.....	73
Chapter 5: Discussions and Implications.....	76
Study Overview.....	76
Statement of the Problem.....	76
Methodology.....	77
Discussion of Themes and Implications.....	79
Implications for Leadership.....	87
Implications for Teacher Practice.....	88
Implications for Research.....	90
Conclusion.....	92

REFERENCES94

APPENDICES

A. Teacher Questionnaire108
 B. Document Analysis Rubric112
 C. Approval to Conduct Research113
 D. Consent Letter to Participants114
 E. Professional Resume115

LIST OF TABLES

Table	Page
1. Counter Examples of Effective Assessment Practices.....	33
2. Research Design Process	38
3. Selected Schools for the Teacher Questionnaire	40
4. Continuum of Assessment Literacy	53
5.1. Descriptive Statistics: Assessment Design Practice	60
5.2. Descriptive Statistics: Assessment Design Understanding.....	60
6.1. Descriptive Statistics: Alignment Practice.....	61
6.2. Descriptive Statistics: Alignment Understanding.....	62
7.1. Descriptive Statistics: Student Involvement/Formative Use Practice	63
7.2. Descriptive Statistics: Student Involvement/Formative Use Understanding.....	63
8.1. Descriptive Statistics: Communicating a Clear Purpose Practice.....	65
8.2. Descriptive Statistics: Communicating a Clear Purpose Understanding.....	65
9.1. Descriptive Statistics: Communicating Results Practice	66
9.2. Descriptive Statistics: Communicating Results Understanding	67
10. Teacher Questionnaire Summary Statements	68
11. Questionnaire Cluster Grand Means.....	71
12. Research Questions with Matched Themes	78

LIST OF FIGURES

Figure	Page
1. Holistic Model of Assessment Literacy	12

CHAPTER 1

The Need for Change

The current debates over the merits of accountability systems in public K-12 schools may lead one to believe that accountability is a new concept. However, the idea of measuring success based on student assessment results has had a long history in public schools (Bloom, Hastings, & Madaus, 1971). Not very long ago, assessment to many teachers may have meant administering a multiple-choice, criterion-referenced test that measured whether students had mastered designated competencies. However, more recently, even standardized assessments have changed to require greater student involvement in assessment – shifting from static summative testing on discreet content to an increased emphasis on the students’ active role in assessment and learning. The shift to a more prominent role of the student in their own learning can be seen in an emphasis on self-assessment, peer assessment and the use of formative feedback (Brown, 1994; Darling-Hammond, Aneess, & Falk, 1995; Hattie, 2016; National Research Council, 2001; Sadler, 1989). Similarly, a greater emphasis on higher levels of cognition in the design of multiple-choice questions, and increased use of open-ended questions to determine student content mastery can be seen (Herrington, Reeves, & Oliver, 2014; Linn, Baker, & Dunbar, 1991).

Systems of accountability are being redefined, but the high stakes attached to the assessments are still a reality to schools and school divisions across the country. Despite numerous educational reform efforts, students, nationwide are failing to meet the achievement expectations. In the 2016-17 SY, 81%, or 1,482 of the 1,825 of the public schools in one Mid-Atlantic state were accredited (based on 2015-16 data), however, there are still concerns about the level of achievement in schools and questions about what efforts are needed to ensure that all

schools meet at least the minimum requirements for full accreditation. Hattie (2003) suggests that what is powerful in the learning equation is what teachers know, do, and care about. With this in mind, educators put careful thought into the questions being asked about why students are not achieving on state-created assessments. An accurate and purposeful focus on the reasons students are not able to master standards of learning at the minimum level of mastery will help stakeholders craft solutions that provide schools with strategies for addressing their students' needs as well as building capacity in their staffs to facilitate continuous improvement and establish sustainable best practices. One such question should address how teachers are using student data to plan and deliver instruction on a daily basis.

The Impact of Assessment on Student Learning

There is an abundance of research surrounding the use of classroom assessment and how using data to inform instruction can foster improvements student performance (Black & Wiliam, 1998 and 2004; Marzano & Marzano, 2012; Pitler, & Stone, 2012; Hattie, 2013). However, because formative assessment isn't any one strategy, measuring its impact on achievement and learning presents some notable challenges. As such, a common research methodology for estimating the effects of assessment on achievement and learning has been meta-analysis. Meta-Analysis has allowed assessment scholars to conduct systematic syntheses of the empirical research of individual studies that address the various aspects of assessment, i.e. goal setting, short term assessments, forward feedback etc. by using the results from more than one study as the unit of analysis and calculating effect sizes (Card, 2011).

With an understanding of the limitations and methodological concerns pointed out by some scholars (e.g. Dunn & Mulvenon, 2009; Natriello, 1987), there is a substantive body of evidence that demonstrates the links between formative assessment practices and students'

learning (e.g., Andrade & Cizek, 2010; Bangert-Drowns, Kulik, Kulik and Morgan, 1991; Black & Wiliam, 1998; Dempster, 1991; Fuchs & Fuchs, 1986; Hattie, 2009; 2012; Reeves, 2007).

Overall this empirical evidence suggests that the various strategies that fall under the formative assessment umbrella can have a modest to strong effect on student achievement. More recently, Wiliam (2011) emphasized that given the challenges of measuring the complex role that assessment plays in student learning and performance, there is no single ideal model or theory of formative assessment that policy and practice can be built from. Instead he argues that a common set of guiding principles from the research highlights needed changes in classroom practices and teacher outlooks that emphasize the active role of the student in their own learning and that such changes will take significant investments in time and professional development. Taken together the empirical literature on the impacts of formative assessment on student learning reveals mixed results. While research has shown specific learning and achievement benefits from various assessment strategies that fall under this umbrella, both conceptual and methodological challenges limit the strength of assertions about its impact on learning and achievement.

Making Data-Driven Decisions

Similarly, there is a vast amount of research that supports the idea that teachers have the greatest influence on student achievement (Hattie, 2013; Dufour, 2004), suggesting that what teachers know, and how they deliver instruction, assess student mastery and monitor progress are all essential questions to be considered. However, there are limited studies that examine the specific relationship between a teacher's knowledge of assessment and how their level of knowledge influences their classroom practices. With so many different assessments available to teachers, and many different opinions about the benefits and significance of each, it is important

that teachers maximize their instructional time and use the available student data to both inform their lessons and evaluate student progress. The decisions to be made with all forms of assessment involve considering the validity of the inferences that can be drawn from the results as well as what decisions, if any, should be made based on the results. However, before inferences can be made, the purpose of any assessment must be clear and therefore, teachers must have a working knowledge of a variety of assessment methods.

Identifying Assessment Methods

It is generally accepted that there are two main forms of assessment most frequently used by teachers—summative and formative. Both have a role in measuring student learning and assist in guiding instructional planning and delivery. Each form of assessment has its role in education and traditionally, the primary function of summative assessment has been to check for mastery following the instruction while formative assessment focuses on informing teachers on ways to improve student learning during lesson delivery (Gualden, 2010). Stiggins (2005a; 2005b) further identifies and defines three kinds of assessment; 1) assessment for learning where teachers and students are constantly informed about student progress while the learning is happening, 2) formative assessment where teachers are informed about student progress in terms of mastery of standards frequently, and 3) infrequent summative assessment that serve to verify students' mastery of standards, after teaching is complete. Zhang & Burry-Stock (2003) found a strong correlation (.71) between teachers' perceived assessment skills and the nature of their assessment practices as it relates to four of the dimensions used as a framework of the study (paper-pencil tests; standardized testing, test revision, and instructional improvement; performance; and non-achievement based grading). This leads one to believe that the perceived skill level of the teacher influences the amount of time dedicated to assessment, the kinds of

assessments being administered, and each assessment's purpose and use. Without a clear purpose for each daily activity and a plan for using the student progress data from each, affords teachers the opportunity to get the most instructional value out of their allotted time.

The Use of Time in the Classroom

Assessments in today's classrooms consume a great deal of the time designated for instruction as teachers and students devote a considerable amount of time preparing for and completing various assessments. The American Federation of Teachers (AFT) reports that an analysis of the time spent for test preparation and test taking, 3rd to 10th grade students in one school district spent an equivalent of 3 school days taking required state and district assessments (AFT, 1990). Additionally, it was found that students in grades 3-8 spent approximately 16 full days preparing for mandated assessments, and the district's calendar indicated 19 days were devoted to test preparation and administration. Totals were even higher in the second school district that was studied (Testing More, Teaching Less AFT). In light of the emphasis placed on preparing for and completing assessments, there appears to be a need for teachers to be aware of and execute assessment practices daily that will help them integrate instruction and assessment in ways that maximize their time with students.

Popham (2009) argues that the majority of standardized assessments being used to measure student achievement are *instructionally insensitive*; meaning they are not designed in a manner that allows the results to determine whether the students have truly mastered the content. Student learning in K-12 education is a function of countless variables, and constant in any learning formula is the importance of the teacher. According to Hattie (2003), the teacher has the largest effect size (.5) in student achievement variance, and that maximizing teacher effectiveness is the key to optimizing student achievement. Simply put, teachers matter and what

teachers know about the content they teach and pedagogy, including assessment, makes a difference in how effectively students learn and achieve. Teaching is a process of decision-making, and to reach a quality instructional decision, teachers must obtain and interpret robust evidence about student progress gleaned through some form of student assessment. Teacher behaviors vary based on their beliefs, conceptions, and competencies (Brown, 2004); therefore, how assessments are viewed and use in daily decision-making is dependent upon the level of “*assessment literacy*” each teacher attains. Gareis and Grant (2008, 2013) suggests that assessment literacy includes a teacher’s knowledge and skills to create, choose and use assessment tools appropriately and that assessment literacy involves the ability to evaluate the data gained from an assessment and use it to inform instructional decision-making.

Teachers and Assessment Literacy

Students rely on teachers to provide quality instruction that assesses their learning fairly and accurately. This requires not only content knowledge, but also sound decision-making based on student data. In order to make effective data-driven decisions, teachers must first become assessment literate, much like a physician needs to become knowledgeable about the effects of different medicines before writing prescriptions. It is widely accepted that when positive teacher efficacy (feeling confident and effective) is coupled with the successful implementation of effective instructional strategies, there is an increase in student achievement. Riggan and Oláh (2011) state that at the classroom level, teachers are asked to draw from a multitude of data sources to inform instruction. Thus, the importance for using accurate and meaningful assessment data cannot be overstated. Popham (2009) argues:

the more importance that a teacher ascribes to classroom assessment, the more profound will be the impact of such assessment on a classroom’s

day-to-day instructional activities. Effective teachers prioritize what they value and what they feel will positively impact their students. (p. 7)

Learning science theory suggests that the more involvement students have in their learning, ranging from constructing learning intentions to developing the assessment, the more success they will achieve (Pitler, & Stone, 2012; Vosniadou, 2001). Nicol and Macfarlane-Dick (2006, p. 15) concluded that, "...if students are to be prepared for learning throughout life they must be provided with opportunities to develop the capacity to regulate their own learning as they progress through higher education." This self-regulated concept relies heavily on formative assessment, feedback and reflection—all vital to the idea of effective learning. Knowing what we do about how children and adults' learn, the reaction to standardized curriculum design, common pacing guides, summative assessments and high-stakes consequences have raised a multitude of questions about the fairness, accuracy, and reliability of the results. Perhaps even more importantly, is the question of how accountability systems influence school leaders, teachers, and most importantly, students. It is important that teachers align their instruction to the cognitive level of the standards to ensure student learning, and not just to the structure and format of the test.

Response to High Stakes Testing

Assessment practices that focus on test preparation are restrictive by nature because creating a test that can be reasonably administered to students is limiting in terms of the ability to capture and evaluate the totality of what even the standards require. Korte (2002) states that the incomplete measurement of achievement is manifested in the process of sampling used to construct an achievement test. Essentially, this means that tests are frequently created without including all of the material students have learned during the period prior to the test

administration—a sampling of material is chosen to be included. As perhaps a very simplified example, a math teacher may teach students a unit on place value from tenths to one-thousandths, yet not have a problem on the assessment that requires students to specifically identify the hundredths place. Although necessary, this sampling limits the ability to evaluate the breadth of knowledge a student has attained. In addition, changes in test design or content can change the result for tests that are supposed to measure achievement over time. Changing the format or sampling of a test over years makes comparing results difficult at best and unreliable at worse. According to Kortez, (2002, p. 758) in the 1980s, simply changing the relative weight of algebra and geometry on the National Assessment of Educational Progress (NAEP) altered the gap between black and white students. Due to the limitations of high-stakes testing as a means of accountability, it is hard to make inferences about the effectiveness of schools and teachers based on their outcomes alone. In fact, there are some questions about how strong the links are between the accountability assessments and the assessments used more frequently in the classroom. Popham (2009) argues that too often, classroom assessments are not aligned with accountability assessments because they are selected from a textbook, teachers' manual, or created by individuals that are not assessment literate.

The Statement of Purpose and Research Questions

What is now commonly known about teaching and learning is, in many ways, in direct conflict with how many schools and school districts assess and measure student achievement. According to the American Psychological Association (APA) there are psychological principles that maximize the learning process. “The [14] principles are divided into those referring to cognitive and metacognitive, motivational and affective, developmental and social, and individual difference factors influencing learners and learning. Finally, the principles are

intended to apply to all learners—from children, to teachers, to administrators, to parents, and to community members involved in our educational system” (APA, 1995, p. 3).

In addition, the current climate in education that widely promotes the value of formative assessments in the measurement of student learning has led to making important distinctions between the purpose and design of these assessments and the traditional use of assessments that are more summative in nature. Formative assessments are a regular part of the daily instructional program and are considered to be “practice” for students. As teachers and students work towards achieving their learning intentions, formative data is used to monitor learning and adjust instruction as needed. Summative assessments are given at longer intervals to measure what students do and do not know. State standardized tests, district benchmark tests, chapter tests, and end-of-course exams are the forms of assessment usually considered to be summative. Many of the practices and policies that have developed as a result of the current trend towards using summative assessments to measure the value of schools, administrators, teachers and the amount of student learning are in contrast with some concepts we know are vital to student learning.

The primary purpose of the study is to examine the relationship between grades 3-12 teachers’ perceived assessment literacy (understanding of assessment) and their classroom assessment practices. A secondary purpose is to compare the teachers’ perceptions of assessment literacy to the holistic model that includes; establishing a clear purpose, aligning instruction and assessment, assessment design, communicating results, and student involvement (Stiggins, Arter, Chappuis, & Chappuis, 2006; Chappuis, Stiggins, Arter, Chappuis, 2011). The following research questions will be used to frame this study:

- 1) To what extent are selected grade 3-12 teachers’ perceived understandings of assessment aligned with the holistic model?

- 2) What assessment practices do selected grade 3-12 teachers report using in their classrooms?
- a. How do grade 3-12 teachers use assessment to inform instruction?
 - b. How do grade 3-12 teachers select and design assessments?
 - c. What assessment practices do grade 3-8 teachers report regularly using in their classrooms?

Significance of the Study

The significance of the teacher's role in their students' learning was referenced earlier in this chapter and there are very few who would argue that the skill level and sense of self-efficacy a teacher has greatly impacts the learning experiences of their students—teachers matter. Also, it is generally accepted that the increased use of strategies that focus on assessment for learning help to facilitate student learning and increase achievement—assessment matters (William, Lee, Harrison, & Black, 2004). Thus, how teachers conceptualize assessment, master the key assessment competencies, and implement effective assessment practices should be considered essential knowledge needed in determining how schools and school districts can meet the challenges of maximizing the authentic learning experiences of their students.

This study takes place in an urban community located in the mid-Atlantic region of the United States. The state department of education reports that based on 2015-16 standardized testing data, 81% of the 1,826 public schools in the state earned full accreditation and the goal of local, state and federal educational policies is to have 100 percent of school accredited. Thus, the gap between the current reality and the goal must be addressed with urgency.

The *Improving the State of America's Schools' Act of 1994*, strongly encouraged states to establish content and performance standards that are demanding (Linn, 2000, p. 8). Tests should

be aligned with these standards and students should be keenly aware of what they will be tested on. This is what Fenwick English refers to as *the doctrine of no surprises*, which states that students should not be surprised by any assessment because in an instructional program, curriculum and assessment alignment are essential (Lunenburg, 2011). Most teachers believe that they employ solid, if not, superior assessment practices (Gareis & Grant, 2013), however, Popham (2009) found that most assessments are not used in a manner that helps to inform and adjust instruction in ways that enhance the learning experience of students and there continues to be a lag in student performance on standardized tests leading many educators ponder over why students are not achieving. Stiggins (2004) argues that if teachers use assessments effectively, students' achievement will improve, and this study is intended to examine what teachers' perceive as their understanding of effective assessment practices and how they employ these practices in a climate of accountability.

Conceptual Framework

The concept of assessment literacy has become more defined in recent years and distinctions have been made between formative and summative assessment. The idea of administering better assessments has emerged as a component of educational reform and the format of standardized state and national assessments have changed in an effort to better measure authentic student learning. The research has used a process-oriented conception of assessment rather than some of the traditional models that involve a linear process of planning, instruction, assessment and analysis, or a concept of assessment that involves a cycle of planning, simultaneous teaching and assessing student progress, evaluating student performance data, reflecting on instruction, making adjustments based on data and reflections, continued instruction and assessment, and back to planning.

Holistic Model of Assessment Literacy

The conceptual framework of this study examines teachers' perceptions of their assessment literacy through the lens of a holistic model (Figure 1) and builds from, one, Stiggins (2002a) the keys to effective formative assessment and two, incorporates the role of the learner which may be underemphasized in Stiggins's keys by focusing on establishing a clear purpose, aligning instruction and assessment with curriculum standards, assessment design, communicating results, student involvement and formative use (Stiggins, Arter, Chappuis, & Chappuis, 2006; Chappuis, Stiggins, Arter, Chappuis 2012). Where other models look at assessment as a checklist or menu of key factors, this study is framed in a model that involves interdependent considerations that all revolve around student learning as the intentional driving purpose of every action.

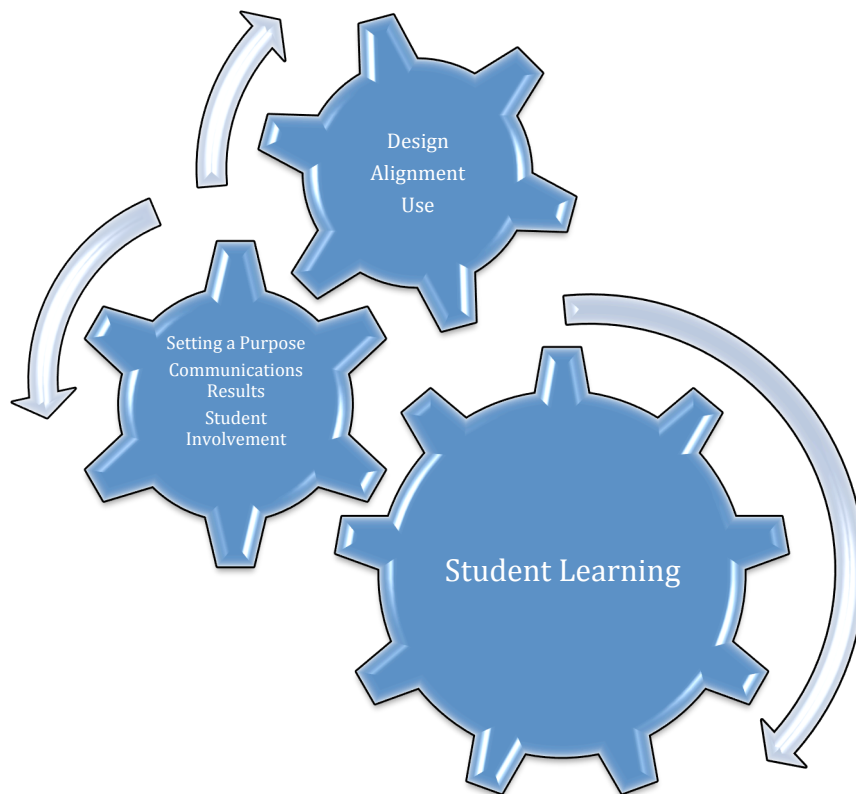


Figure 1: Holistic Model of Assessment

Motivation has been reported in primary, secondary and college education to influence academic performance (Vansteenkiste, Zhou, Lens, Soenens, 2005). Establishing a clear purpose allows everyone involved to understand *why* any form of assessment is being administered and entails understanding that students' motivations are connected to their assessment experiences (both formative and summative), establishing what the assessment data will be used for, and developing a comprehensive plan for; integrating formative assessment, assessment for learning and summative assessment. Aligning instruction and assessment with curriculum standards involves teachers turning the essential knowledge and skills written in curriculum documents into lesson plans, learning intentions and creating assessments that reflect the same level of cognitive demand. In his study, Squires (2012) demonstrated a strong, positive and significant correlation (.49) between the instructional content delivered and student achievement gains and Schmidt, McKnight, Houang, Wang, Wiley, Cogan, & Wolfe (2001) found there was a connection between achievement and the alignment between the taught and the tested curriculum. A sound assessment design requires that the method of assessment is aligned with the essential knowledge and skills, includes an appropriate amount of varied questions, serves its stated purpose, and avoids factors that may lead to mismeasurement and bias. Mehrens and Lehmann (as cited by Alade & Omoruyi, 2014) identified this practice as a means of helping to maximize the teacher's efficiency and student's learning. In order to effectively communicate results, teachers have to decide on the best tool, and a decision must be made on whether the data should be reported through grades, narratives, or student conferences. Effective communication means that results should be reported in a timely manner and should be completely transparent and available to the students, their parents, colleagues, and other stakeholders. CTB/McGraw-

Hill LLC (as cited by Clark & Smitherman, 2013) reasons that communication with all stakeholders becomes easier when schools share information about purposes, meanings and results of assessment programs regularly. Using assessments formatively have proved to have significant effect sizes ranging from .40 to .70. Involving students in the assessment process requires teachers to view students as partners by sharing the learning intentions with the students at the beginning of instruction, providing students with practice and frequent opportunities to reflect on their progress based on given models of success and specific feedback, and adjusting instruction based on any gaps between their current level of success and the desired student outcomes. Hattie (1999) found that providing feedback to cue and reinforce progress towards learning intentions carried an effect size of .94 (Stiggins, 2007, Rutherford, 2013). The distinction of the holistic model is that it interconnects all of the recognized concepts of effective assessment practice and considers them as interdependent parts of a whole system, focusing on the identified learner as the driving purpose behind all actions and reactions, and measuring success based on the desired learning outcomes.

Definition of Terms

Assessment literacy: a dynamic, context dependent, social practice that involves teachers articulating and negotiating classroom and cultural knowledge with one another and with learners, in the initiation, development, and practice of assessment to achieve the learning goals of the students (Willis, Adie, & Klenowski, 2013).

Learning Intentions: The skill or knowledge the teacher wants the students to be able to demonstrate as a result of the lesson and related activities

Success Criteria: The measurable behavior or knowledge required of each student to determine

whether the learning intention has been achieved

Assessment Design (How): Using the best method, avoiding bias, and

Appropriately sampling the content.

Alignment (What): Focusing on learning intentions that are aligned with the content standards.

Student Involvement and Formative Use (How/Who): Students tracking their progress, reflecting on their own knowledge, and using student data to construct the assessment.

Setting a Clear Purpose (Why): Deciding on a clear purpose and determining how the results will be used.

Communicating Results (Who/How): Communicating the results to the appropriate stakeholders and deciding the best form and method of providing the information

Holistic: relating to or concerned with whole or with complete systems rather than with the analysis or treatment of parts.

CHAPTER 2

REVIEW OF LITERATURE

This review of the literature describes how assessment in education has evolved over the past century, exploring the changes in the generally accepted purpose of education and the increasing layers of accountability systems that hold schools responsible for student achievement. My review also explored the similarities and differences between formative and summative assessment, as well as, how assessment practices can motivate students and inform instruction for teachers. Despite the research that supports the use of effective assessment practices, Riggan and Olah (2011), asserts that teachers do not frequently assess their students for conceptual understanding and provide only token feedback. Brown (2004) argues that all pedagogical acts, including teachers' perceptions of evaluations of student behavior and performance, are affected by the conceptions teachers have about their own confidence to teach, the act of teaching, the nature of curriculum or subjects, the process and purpose of assessment, and the nature of learning among many educational beliefs. The review of literature has five main sections as outlined below:

1. Evolution of Assessment and Accountability
 - a. A Brief History of Assessment
 - b. The Changing Perspective on Assessment
2. Assessment Literacy
 - a. An Definition of Assessment Literacy
 - i. *Teacher Competencies.*
 - ii. *Decision Making.*
3. Empirical Evidence about the Links between Assessment and Student Learning
4. Defining the Forms of Assessment
 - a. *Summative Assessment*
 - b. *Formative Assessment*
 - c. *Formative Assessment for Learning*
 - d. *Balancing Forms of Assessment.*
5. Assessment Practices
 - a. Classroom Assessment
 - b. Student Accountability

The first section of recalls the evolution of assessment and accountability. The second section addresses the definition of assessment literacy and the competencies currently held by teachers. The third section reviews the empirical evidence about the links between assessment and student learning and performance. The fourth section defines the forms of assessment that are common in the field and the final section reviews some common assessment and grading practices teachers use in their classrooms on a regular basis.

Evolution of Assessment and Accountability

A Brief History of Assessment

Gathering information about student performance has been a part of education for centuries. Formal and informal assessment have had a long-standing role in education in various different forms, however, during the turn of the 20th Century, the evaluation of students and schools began to take a more prominent place as continued industrialization resulted in a movement towards universal schooling (Earl, 2003). Prior to this time, extended education was available only to a small segment of the population—the wealthy and the talented. Others were offered just enough instruction focused on the basics of reading, writing, and math to provide the skills and knowledge thought necessary to function productively in society. Earl (2003) suggests that the idea that schools should serve the academic, social and emotional needs of all students, not just the wealthy and talented, meant that opportunities for extended education needed to be offered based on merit rather than privilege. This paradigm shift led to the birth of a system that would intend to assess student achievement for the expressed purpose of distinguishing those who performed at higher levels.

During WWI, the army successfully used the Alpha Intelligence Test to identify those individuals that showed the potential to serve as officers. Using this test as a model, educational

standardized achievement tests were developed to establish a means of comparing students with a relative score-based measurement. Popham (2003) argues that this marked the initial stages of “high-stakes testing” because the test results determined what students’ access to higher education would be and ultimately, what position they would hold in society. Since this change, assessment has been an integral part of the education system, yet the role assessment has played, and the understanding teachers have had of its uses and values have not been as consistent. Stiggins (2002b) describes a timeline of assessment that he suggests, demonstrates the direction assessment has taken over the years:

- In the 1940s, nationally standardized college admission practices began.
- In the 1950s and 1960s, standardized testing extended into every grade level with K-12 accountability testing programs.
- In the 1970s, the statewide testing program movement began.
- In the 1970’s and 1980s, national assessment programs added another layer.
- In the 1990s, national and international engrossment in standardized testing became the accountability measure of choice.

Currently, all states have some formal testing program, as testing trends are the topic of journal articles, news stories, and national conferences. Today, testing has become so prominent that the United States Department of Education publishes an annual report detailing state-by-state testing results with breakdowns that include subject area, student demographics and grade level performance. Furthermore, national testing results are commonly compared to the performance of students in other countries as a measure how effective the public education system is in the United States (Stiggins, 1991). Earl (2003) points out that there are mixed messages being given with regard to assessment and accountability. On the one hand, large

incentives and sanctions placed on schools based on standardized assessment results appear to advocate a performance goal orientation that places priority on outcomes. On the other hand, debates loom about the importance of social justice with respect to providing access to opportunities for all and developing experiences for students that are based on the mastery goal orientations that allow for value-added measures for students and teachers.

As there continued to be a growing reliance on assessment to measure student performance, Stiggins (2002b) found that there was little to no attention given to the quality of classroom assessment, leading some to call for reformed thinking as it relates to the evaluation of student learning. Bloom, Hastings and Madaus (1971) urged that both teaching and evaluation must undergo changes if we are to meet the vast needs of our youth. Black and Wiliam (1998) argued that assessment reform was missing something—the connection between the curriculum standards, what is being taught, what is being assessed, and what students learn. Additionally, Linn (2000), who served as a major developer of standardized tests for many years made the following statement about the their impact on the improvement of instruction and learning:

As someone who has spent his entire career doing research, writing and thinking about educational testing and assessment issues, I would like to conclude by summarizing a compelling case showing that the major uses of tests for student and school accountability during the past 50 years have improved education and student learning in dynamic ways. Unfortunately, that is not my conclusion (p. 4).

The Changing Perspective on Assessment

High-stakes testing continues to dominate school cultures in PreK-12 educational settings. It is generally accepted that this climate of accountability has led to some unintended, unfavorable classroom practices. Many teachers began ignoring curriculum content that was not heavily weighted on tests and changed their instruction and assessment practices to mimic the format of standardized tests. Popham (2001) referred to this practice of selective content inclusion as “item-teaching”. As school system administrators began to realize that once-a-year external assessments did not provide enough information about achievement, districts began adopting interim tests to serve as predictors for future performance, and classroom practices began to improve as the stigma of “teaching to the test” grew and it became more common for teachers to engage in “curriculum-teaching”, which involves focusing instruction on content standards or cognitive skills (Popham, 2001). However, the focus remained on the ultimate goal of improving test scores, and not the measurement of student learning.

The current changing perspective does not dismiss the necessity for external standardized testing measures; rather it addresses the significance of classroom assessment as a vital component of a balanced student assessment system (Stiggins, 2001). Standardized assessments are norm-referenced, and provide information (and sometimes misinformation) about how a student compares to age or grade level counterparts. However, what teachers and parents are most concerned about is if students are learning what they are being taught in school—and to what degree are they learning. This requires data yielded from criterion-referenced assessments that are based on setting learning goals and working towards their attainment (Brookhart, 2001). Bloom (et al, 1971) conceded that it is not possible or desirable to completely omit summative assessments (assessment that measures what students have learned after instruction). However,

formative assessment (assessment that informs instructional decisions during the learning process) provides teachers and students information that helps them improve teaching and learning respectively. In this vein, assessment is currently viewed in a broader context—one that involves both summative assessing to measure outcomes and formative assessing to inform and guide instruction. Stiggins (2004) argues that a shift from traditional thinking about assessment to a more productive perspective will involve the following:

- High-stakes testing must be complimented by a supportive assessment climate in the classroom
- Assessment results and decisions must be shared with and involve the students
- The decisions teachers and students make each day about instruction have the greatest impact on learning.
- Teachers must have knowledge of effective assessment practices and use them daily

According to Stiggins (2004), research has shown that the use of effective assessment practices by teachers have yielded some gains higher than a full standard deviation on assessments and he further suggests that gains in achievement realized by effective assessment practices are similar to those that result from one-on-one tutoring (Stiggins, 2004). Thus, it is of the utmost importance that teachers are *assessment literate* in order to create, deliver and evaluate assessments in a manner that maximizes students' learning.

Assessment Literacy

A Definition of Assessment Literacy

Teacher competencies. “Assessment literacy is a dynamic, context dependent, social practice that involves teachers articulating and negotiating classroom and cultural knowledge with one another and with learners, in the initiation, development, and practice of assessment to

achieve the learning goals of the students” (Willis, Adie, & Klenowski, 2013, p. 2). Those who are assessment literate have a deep understanding of the standards that will be measured and can communicate them to the learners. Furthermore, this understanding is evident when seeking and creating assessments, and the use of assessments as a tool for making short and long-term instructional decisions should be integrated into teachers’ daily practices. Assessing student learning is one of the most important facets of teaching, however, most teachers complete their preparation programs without having to demonstrate competencies in educational assessment even though the American Federation of Teachers (AFT), National Council on Measurement in Education, and National Education Association (NEA) developed the Standards for Teacher Competence in Educational Assessment of Students in 1990 (Brookhart, 2001, p. 3). There were seven standards established:

- 1) Teachers should be skilled in choosing assessment methods appropriate for instructional decisions.
- 2) Teachers should be skilled in administering, scoring, and interpreting the results of both externally produced and teacher-produced assessment methods.
- 3) Teachers should be skilled in developing assessment methods appropriate for instructional decisions.
- 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 communicating assessment results to students, parents, other lay audiences, and other educators.
- 6) Teachers should be skilled in developing valid pupil grading procedures that use pupil assessments.

- 7) Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.

These competencies are consistent with Stiggins' (1991) argument regarding assessment literate teachers while adding that assessment literate teachers are capable of understanding the implications of improper assessments and recognizing external factors that confound assessment results. Stiggins (1991) contends that assessment literate educators ask two key questions that consider what the chosen assessments tell students about the achievement outcomes we value, and what effect the assessment is likely to have on students. The answers are vital because those who are assessment literate look to find and use assessments that relate clear, specific and robust delineations of the outcomes they have identified as important. In addition, he suggests that student assessment takes on three primary forms; paper-pencil assessments (quizzes and tests that are teacher made, quizzes and tests that are included in textbooks and workbooks, homework, classwork, etc.), performance assessments (behavior observation or evaluating products), and assessments that involve direct communication with students (questioning during instruction, interviews and conversations, or inferences made about a student). Recognizing the different ways that assessments may be administered is only a surface level of understanding for teachers. By-and-large, teachers can readily appreciate the continuous assessments they make as they carry out their duties, however, the next level in effective assessment practice is the decisions that are made, if any, when choosing an assessment and what actions need to be taken based on the information obtained.

Decision making. Popham (2009) presented two decision clusters related to assessment. He argued that these clusters have a significant impact on what teachers do and how students are taught. The first cluster focuses on the concept of assessment for learning and formative

assessment and involves what assessments should be given and how they should be utilized. An assessment literate teacher is more likely to create and select better classroom assessments and make more accurate inferences from the results. The second cluster is centered on summative assessment. Popham (2009) used the term instructionally insensitive to describe those assessments—standardized or otherwise, that “...are unable to distinguish between students who have been skillfully taught and those students who have been shabbily taught” (p. 7). The decisions to be made with these assessments are about considering the validity of the inferences we can draw from the results and what decisions, if any, should be made based on the results. Hattie (2003) suggested that it is what teachers know, do, and care about, which is very powerful in the learning equation because students rely on teachers to provide quality instruction and assess their learning fairly and accurately. To do this, teachers not only need content knowledge, but also sound decision-making skills based on student data—assessment literacy.

It is widely accepted that when positive teacher efficacy is coupled with the successful implementation of effective instructional strategies, there is an increase in student achievement (Guskey, 1986; 2002; Guskey & Huberman, 1995). Riggan and Oláh (2011) state that at the classroom level, teachers are asked to draw from a multitude of data sources to inform instruction. The importance for using accurate and meaningful assessment data cannot be overstated, thus Popham (2009) suggested:

The more importance that a teacher ascribes to classroom assessment, the more profound will be the impact of such assessment on a classroom’s day-to-day instructional activities. Effective teachers prioritize what they value and what they feel will positively impact their students (p. 6).

Sheppard (2000) contended that assessments used strategically, would transform the way classrooms feel and suggested that the process of assessment should be more of a partnership between teachers and students where the goal is to assess for insight rather than an opportunity to reward and punish. Additionally, she supported the idea of assessment being an on-going part of the learning process instead of something that occurs at the end of instruction. Students and teachers should look at assessment and learning as a chance for providing the feedback needed to make self-evaluations necessary for knowledge transfer. With regard to instructional decision-making, its effect on student learning goes beyond “doing the right things” and requires that teachers are ensuring that they are maximizing their time and optimizing their instruction. To accomplish this, teachers will have to acquire a deep understanding of the content in order to be able to ask the right questions at the right time, plan for the any anticipated stumbling blocks students may encounter, and use assessments effectively (Sheppard, 2000). The keys to being assessment literate have varied as the topic has become increasingly significant as an explicit component of instruction, however, the aspects of effective assessment practices generally have been founded on the following dimensions; clearly establishing a purpose for the assessment, sound assessment design, effectively communicating results, engaging students in the assessment process, and aligning instruction and assessment with curriculum standards (Stiggins, Arter, Chappuis, & Chappuis, 2006; Chappuis, Stiggins, Arter, Chappuis 2011). The dimensions take on varying forms when practiced within the context of the different recognized forms of assessment being administered.

Empirical Evidence Linking Assessment and Student Learning

Because formative assessment isn't any one strategy, measuring its impact on achievement and learning presents some notable challenges. As Robinson, Myran, Straus, &

Reed (2014) emphasize, the construct of formative assessment is grounded in multiple theories of learning which incorporates behavioral, cognitive, and constructivist approaches and draws from self-regulated learning (Zimmerman, 1990), motivational theories (Bandura, 1982) and learning orientations (Stipek, 2002). As such, a common research methodology for estimating its effect on achievement and learning has been meta-analysis. This approach allows researchers to look at individual studies that address the various aspects of assessment, i.e. goal setting, short term assessments, forward feedback etc. that can generate combined estimates. Meta-Analysis has allowed assessment scholars to conduct systematic syntheses of the empirical research by using the results from more than one study as the unit of analysis and calculating effect sizes (Card, 2011).

While it has been widely accepted within both the professional and academic literature that there exists a positive impact of formative assessment practices on students' learning and achievement, a number of scholars have pointed out methodological concerns for this work (e.g. Dunn & Mulvenon, 2009; Natriello, 1987). For example, Natriello (1987) pointed out the limited number of high quality studies, that most of the studies only address one or two aspects of the assessment process and lastly, the individual strategies that fall under the larger formative assessment umbrella are used in different contexts and for different purposes and as such comparisons may be misleading. Similarly, Crooks (1988) meta-analysis concluded that teachers place too much emphasis on the grading of assessment. Further challenges have been noted which include some studies used in meta-analyses are merely theoretical inventories of classroom practices and do not address the domain specific nature of learning across disciplines (Brookhart, 2004), that most of the available studies are relatively small (Kluger & DeNisi, 1996), that many of the studies use measures of content knowledge and lower-level skills to

assess achievement (Dempster, 1991) which limit what can be said about assessment impact on meaningful or substantive learning, and lastly that few studies are found in common across the various meta-analyses (Black and Wiliam, 1998).

Similarly others have pointed out that experimental and quasi-experimental research on the effectiveness of formative feedback on student learning cannot fully account for results being delayed or masked by other factors (Sadler, 1998). Specifically the long-term exposure and socialization to more summative patterns of assessment practices and expectations, the students focus on meeting individual teachers' assessment expectations may promote "accommodating survival habits" among students rather than the metacognitive and self-regulated learning behaviors formative assessments are designed to promote (Sadler, 1998). Likewise, how educators conceptualized learning has also changed, emphasizing that learning is a process that involves the student taking an active role in constructing their own meaning based on their prior knowledge and shaped by their metacognitive and self-regulated behaviors (Dunn & Mulvenon, 2009). Also worth mentioning is Shute's (2008) distinction between achievement, performance and learning indicating that while there is wide acceptance that formative assessment does impact achievement and performance, it's impact on learning itself is not as clear.

With an understanding of these limitations and challenges in methodology, there is evidence that demonstrates the links between formative assessment practices and students' learning (e.g., Andrade & Cizek, 2010; Black & Wiliam, 1998; Hattie, 2009; 2012; Reeves, 2007). Early work in this area includes Fuchs & Fuchs (1986) meta-analysis that investigated the effect of systematic formative evaluation with high functioning special education populations and showed the positive impacts on student learning. Their study reviewed 21 empirical studies on the use of feedback with students with and without mild to moderate learning disabilities.

Their study yielded a mean effect size of .70 for students with disabilities and .63 for students without disabilities. In a meta-analysis of over 250 relative studies that examined the impact of formative assessment practices on student learning, Black and Wiliam (1998) showed an average effect size ranging from .4 to .7. Hattie's (2007) meta-analysis of over 500 meta-analytic studies (including 180,000 individual studies) examined the many factors influencing student achievement. This research revealed an average effect size of .79 for the effective use of feedback. Of the over 100 factors examined, the effective use of feedback, which Hattie compares with formative assessment, was in the highest 5 of the 10 effect sizes reported.

Bangert-Drowns, Kulik, Kulik and Morgan (1991) and Dempster (1991) in separate meta-analyses on the effects of feedback found a number of important mediating variables which included building feedback mechanisms into instruction, reviewing test results with students and the quality and nature of the feedback provided. A key finding was that feedback strategies that prompted students seeking to correct misconceptions in subsequent learning had a greater impact on learning than strategies that focused on correct answers. Elshout-Mohr's (1994), study confirmed facets of these findings that highlight a focus on correct answers is only effective for learning simple content and suggest that feedback needs involve more dialog, inviting students to be active agents in their own learning.

More recently, Wiliam (2011) emphasized that given the challenges of measuring the complex role that assessment plays in students learning, there is no single ideal model or theory of formative assessment that policy and practice can be built from. Instead he argues that a common set of guiding principles highlight that what is needed in formative assessment is substantive changes in classroom practices that emphasize the students active role in learning, that assessment can foster this self-regulated behavior and these outlooks and skills sets need to

be are incorporated into teachers' daily practice and that these changes will take significant investments in time and professional development.

Taken together the empirical literature on the impacts of formative assessment on student learning reveals mixed results. While research has shown specific learning and achievement benefits from various assessment strategies that fall under this umbrella, both conceptual and methodological challenges limit the strength of assertions about its impact on learning and achievement. Because formative assessment is not any one strategy, but a collection of related constructs and strategies, no grand theory supported by empirical evidence from well-controlled studies exists. Given the limitations of the studies, it appears that formative assessment practice can foster moderate to large improvements to learning and achievement. Perhaps most importantly, the research suggests that educators who employ these practices grounded in an understanding of the importance of the students' own active role in learning, prompting them to reflect and correct misconceptions over giving correct answers and engaging them in dialog are more likely to nurture greater student learning and achievement.

Defining the Forms of Assessment

Summative Assessment

Garrison and Ehringhaus (2007) suggest that summative assessments are more familiar to the majority of teachers and are given in intervals as a measure of what students know—or don't know. Although these kinds of assessments have been associated more frequently with state and national standardized assessments, they are an important component of classroom assessment and grading. They usually take the form of a quiz, unit test, semester or quarterly test, or a culminating project to measure a students' knowledge of a given content at a specific time. Because of their periodic nature, summative assessments do not provide information that

teachers can use to adjust instruction or provide interventions in a timely manner. Similarly, students are not able to use feedback or results from these assessments to monitor their own learning in a way that they can during instruction. Boud (2000) theorized that, as they are currently utilized, summative assessment acts as a means to gain or prohibit access to privilege or honors in society (i.e. accreditation, college acceptance, grade promotion, scholarship awards). He adds that this gives evaluation powers to others while weakening the ability of students to recognize and monitor their own progress towards standards.

Formative Assessment

William, Lee, Harrison, & Black (2004) found that, as it relates to external assessments, there are some real benefits to improving formative assessment practices (approximately one-half of the measured unit of the assessment). According to Stiggins and Chappuis (2006), formative assessment is often used as a method of benchmark testing to determine where students are in relation to achieving specific standards, differentiating themselves from summative assessments only by the frequency in which they are administered and not so much by the way they are created and used. They suggest that while this helps to identify students in need of extra attention prior to the end of a specified time period (unit, quarter, semester, and year) and, it does not provide the essential components that will help students experience more success on a daily basis. This success requires teachers to think about assessment as a tool *for* learning. As the idea of formative assessment has become more imbedded into the daily instruction researchers have seen more definitive positive effects. Fuch and Fuch's (1986) meta-analysis of formative assessment found that it yielded an average weighted effect size of .70.

Formative Assessment for Learning

Formative assessment for learning is a less familiar concept and these assessments are differentiated from other forms of assessment based on when they are given and how they are used. Riggan and Oláh (2011) stated that at the classroom level, teachers are asked to draw from a multitude of data sources to inform instruction. Stiggins (2002b) suggested that true formative assessment for learning involves the following: 1) understanding and articulating, in advance of teaching, the achievement targets that students are to hit, 2) informing students about those learning goals in terms that students understand, from the very beginning of the teaching and learning process, 3) becoming assessment literate and thus able to transform their expectations into assessment exercises and scoring procedures that accurately reflect student achievement, 4) using classroom assessments to build students' confidence in themselves as learners and help them take responsibility for their own learning, so as to lay a foundation for lifelong learning, 5) translating classroom assessment results into frequent descriptive feedback (versus judgmental feedback) for students, providing them with specific insights as to how to improve, 6) continuously adjusting instruction based on the results of classroom assessments, 7) engaging students in regular self-assessment, with standards held constant so that students can watch themselves grow over time and thus feel in charge of their own success, and 8) actively involving students in communicating with their teacher and their families about their achievement status and improvement. The research supports the idea that assessment *for* learning is not the prevailing practice in classrooms in schools today.

Balancing Forms of Assessment

Summative assessment has taken a more prominent role in the debates in public education and formative assessment has received far less attention (Boud, 2000). Some have

called for a balance of summative and formative assessment (NRC, 2001), acknowledging that serves a different purpose. Thus, Boud (2000) argues, the purposes of each form of assessment should be evaluated with respect for its impact of student learning and further states that, as contrary to how it may seem, different forms of assessment are intertwined and are difficult to separate. Furthermore, because teachers have to manage the various forms of assessment tensions between them can be significant (Bol & Strange, 1996).

Assessment Practices

Classroom Assessment

Studies reveal that behaviors that positively impact learning are not consistent with what we know about the most prevalent classroom practices. Teachers assess for conceptual understanding infrequently, provide only perfunctory feedback, and do not respond to feedback instructionally with consistency (Herman, Osmundson, Ayala, Schneider, & Thomas, 2006). Hattie (2003) posits that typically, classroom assessments only require students to demonstrate surface knowledge on the content. Similarly, Cizek and Others (2005), argued that based on the highly individualistic nature of assessment practices, many teachers seem to have assessment policies based on their idiosyncratic values and conceptions of teaching. Even as a greater awareness of the importance of aligning assessments to the cognitive demand of the standards influences teachers' decision making, practices that are counter-productive to achieving this goal are still frequently taking place in classrooms. Strage, Tyler, Thomas, and Rohwer (1987) called practices that are in conflict with the principles of effective assessment, "compensations". An example of a compensation practice is when the students are given the same item on an assessment that they have used to practice the skill. Strage (1987) and colleagues argue that the students' familiarity with the test item limits the amount of critical thinking required to correctly

answer the question and Black and Wiliam (2001) cited a United Kingdom study (Table 1) that categorized daily practices that serve as counter examples of what they considered to be effective use of assessment.

Table 1: *Counter Examples of Effective Assessment Practices*

Difficulty	Indicators
Effective Learning	<ul style="list-style-type: none"> •Teachers' tests encourage rote and superficial learning understanding. •The questions and other methods used are not discussed with other teachers and are not critically reviewed. •There is a tendency to emphasize quantity and presentation of work and to neglect its quality in relation to learning.
Negative Impact	<ul style="list-style-type: none"> •Grading is over-emphasized, while providing feedback and the functions of learning are under-emphasized. •Use of approaches promote performance rather than mastery, so poor performance de-motivates students.
Managerial Role	<ul style="list-style-type: none"> •Teachers' feedback to pupils often seems to serve social and managerial functions, often at the expense of the learning functions. •Teachers' tests mimic external tests, yet the know little about their pupils' learning needs. •Assessments are used more to fill grading reporting criteria than to discern learning needs.

Bol (2004) suggests that teachers may feel that there is a dichotomy between assessment practices that promote higher order thinking skills and deeper learning and those that prepare students for state and district standardized tests. In her study that examined teachers' assessment practices, teachers reported using multiple-choice and true/false tests more frequently. High-stakes testing has certainly continued to impact the decisions teachers make as it relates to creating and assigning assessments for their students. Bol (2004) found that teachers feel as if their assessments, as well as their instructional strategies, are strongly influenced by standardized tests.

Student Accountability

“The major principle of student accountability is that assessment holds students accountable for their learning through grade or score assignment, checking performance against established criteria, and reporting grades to parents, and other stakeholders” (McMillan, Myran and Workman, 2002, p. 211). Despite the changes in education and the evolution of assessment, current grading still resembles traditional practices and not what research has shown to be most effective. Additionally, educators are now realizing that the grades students are receiving in their classes are, many times, inconsistent with their standardized assessment scores (Jung & Guskey, 2011). Thus, the task of how to best measure, assign, and report grades has become a source of trepidation. Stiggins, Frisbie, & Griswold (1989) pointed out that teachers employ a variety of grading methods and want their grades to motivate students along with accurately reflecting effort and achievement. Curwin (2014) asserts that students’ motivation to learn is actually diminished by an assessment process that is attached to rewards, incentives, threats, or punishment.

Summary

Mandates and reforms have led to a heightened sense of accountability in education. There is no doubt that student learning is the overall goal of teachers and administrators, however, there is some contention about whether how we currently measure students’ learning is more aligned with authentic learning and critical thinking or rote memorization and surface level knowledge. Marzano (2003) argues that placing a priority on student learning involves using data to make decisions relating directly to student achievement and even when curriculum and assessments are well designed, no factor is as important than the data teachers gather while interacting with students daily. Additionally, (Popham (2001) argues, to fulfill the purpose of

assessment, teachers should be able to make accurate inferences based on the data that is gleaned. The literature supports the idea that assessments designed to align with content standards, and used specifically to inform instruction and provide feedback to students will facilitate this purpose (Black and Wiliam, 2004). When teachers are assessment literate and employ the principles of effective assessment, research has shown that achievement can be increased as much as seven-tenths of a standard deviation (Black and Wiliam, 1998).

There does seem to be some disconnection between what the literature says about effective assessment practices and what teachers do in class. External pressures have been credited with being the major reason why there is a difference between what teachers know about quality assessment and the practices they employ in class. Another reason may be a miscalibration by teachers of their own knowledge of assessment (Kohn, 2000). This study intends to extend the data gathered by previous research on the topic of assessment literacy by looking through a holistic lens—moving beyond the process itself, and considering the learner as the focus of all decision-making.

CHAPTER 3

RESEARCH DESIGN AND METHODOLOGY

In order to examine the relationship between the perceived assessment literacy of grade 3-12 teachers and their assessment practices, an exploratory, mixed method design was utilized. This chapter details the research questions, the design rationale, the population and sample, the instruments that will be used for data collection, the data collection procedures, the methods that will be used to analyze the data, and the limitations of the design.

The review of the literature provided insights into assessment literacy and the competencies teachers reportedly demonstrate in the area of assessment and instruction. The literature also cited the benefits of effective assessment practices and the practices commonly found in classrooms throughout the United States and abroad. However, the link between teachers' perceived understandings about assessment practices and how those specific understandings are aligned to their assessment practices has not been examined as intensively. The literature addresses what an assessment literate teacher should know in theory, but has not set specific criteria for what its application looks like in classrooms that include many challenges and limitations. As discussed in Chapter 2, truly effective assessment practices are more holistic in nature and involve, among other things, student involvement as well as the considerations teachers take about the effect the assessment will have on the students. Based on these findings, this study was designed to examine the relationship between grade 3-12 teachers' perceived assessment literacy and their classroom assessment practices using this holistic model. The following research questions guided the examination of this purpose:

- 1) To what extent are selected grade 3-12 teachers' perceived understandings of assessment aligned with the holistic model?

- 2) What assessment practices do selected grade 3-12 teachers report using in their classrooms?
- a. How do grade 3-12 teachers use assessment to inform instruction?
 - b. How do grade 3-12 teachers select and design assessments?
 - c. What assessment practices do grade 3-12 teachers report regularly using in their classrooms?

Research Design

Using an exploratory, mixed-method design, both quantitative and qualitative approaches were employed to achieve the research goals. Obtaining statistical information and then following up with participants to further investigate the results in more depth, supports the rationale for using this design (Creswell, 2009).

The study used data gathered to describe and measure the lived experiences of grade 3-12 teachers and assess their perceived knowledge as it relates to their experiences and reasoning about student assessment. The contextual factors and how the participants realize the concepts presented will be relevant the study was constructed and the data was interpreted. Johnson and Onwuegbuzie (2004) posit that mixed method research offers a consolidated approach to data collection and analysis, rather than restricting researchers to a single approach. Qualitative studies consider the lived experiences of the participants, allowing for the context to be factored into the analysis of the data in order to develop an in depth understanding of the phenomenon being studied. On the other hand, quantitative research provides statistical data that can improve the validity of the research tools and makes the data easier to analyze (Sun, Pan, & Wang 2011). Using both qualitative and quantitative methods in a single study is justified because the goal of each research question should be the driving force behind the method of data collection. Hence

using different philosophies may be appropriate and even necessary (Johnson & Onwuegbuzie, 2004).

The purpose of my exploratory design was to examine the depth of teachers' understanding of assessment and their classroom assessment practices through a mixed-method design using a teacher questionnaire and an analysis of lesson plans and authentic assessment tools (Table 2). Creswell and Plano (2007) suggest that this design is a good fit for exploring areas where there exists very little published research. This approach facilitated the development of themes and descriptions of the relationship between teachers' perceived assessment literacy and their assessment practices not currently readily available because of the underrepresentation of research that directly examines this phenomenon through a holistic lens.

Table 2: *Research Design Process*

Method	Research Question	Rationale
Quantitative Teacher Questionnaire	To what extent are selected 3-8 teachers' perceived understandings of assessment aligned with the operational definition of the holistic model?	The responses will provide statistical data about what teachers understand/believe about assessment. The responses will provide statistical data about teachers' assessment behaviors
Qualitative Document Analysis	What assessment practices do selected teachers report using in their classrooms?	To gain insights into and expand on the responses from the questionnaire and the themes formed from the document analysis.

The study was designed to collect and analyze data that represents the authentic experiences of teachers; planning documents, assessment tools, and questionnaire responses, which served the purpose of examining the relationship between teachers' perceived assessment literacy and their classroom assessment practices. The perceptions in this study were measured

through the unique holistic model of assessment literacy. This model requires that each essential component is active simultaneously and constantly considered during every stage of planning, lesson delivery and assessment. With the students' learning serving as the primary gear in the process, teachers adhering to the holistic concepts are able to communicate the daily learning intentions in a manner that is easily understood and relevant to the students. Teachers also are able to plan purposeful and meaningful assessments that are integrated into daily instruction and intentionally communicate the results to all the identified stakeholders including students, parents and colleagues. Additionally, teachers who subscribe to the holistic model of assessment ensure that students are involved in each stage of assessments; providing choice based on the needs of the students, providing specific feedback to students, and facilitating opportunities for self-reflection. Finally, teachers thinking holistically will design assessment activities that are aligned with curriculum standards, and they will consider the impact each will have on the student's academic and emotional well-being. In the holistic model, no component is optional. Each component is depended on the other and omitting, misusing, or underusing one, creates a void, negatively impacting the entire system.

Participants

Teacher Questionnaire Participants

A teacher questionnaire was administered as a means of collecting quantitative data. Purposeful sampling was used to achieve the goal of maximizing the information gained by selecting participants that have experience in, and knowledge of the central issues involved in the study (Patton, 1990). The population for gathering the quantitative data was the 1,530 teachers working a small, Mid-Atlantic urban school district that serves approximately 20,700 Prek-12 students. To gather data on a large percentage of diverse students in the district, a purposeful

sample consisting of teachers in grades 3-12 that teach at selected schools was the goal. Based on the staffing at the selected high school, middle schools, and elementary schools, the sample on teachers invited to participate equaled 186 ($n=186$) and the grade levels were chosen because these students are involved in state standard and End-of-Course (EOC) testing that are administered using standardized tools. Additionally, these grade levels were chosen with the assumption that teachers that work in these grades where testing is prominent will have more exposure to a variety of assessments in general, although benchmark and standardized tests are administered in Pre-K through second grade. I did not include these primary teachers because the curriculum development and assessment of primary students has evolved differently due to the limited influence of public accountability systems (Geva, Blenkin, & Kelly, 1992, p. 3), and based on this reasoning, it was decided that the data from primary school teachers would obscure the findings. The selected schools (Table 3) involved in the study were a combination of zoned schools, choice schools, Title I schools, and non-Title I schools chosen to achieve a diverse representation of teachers' contextual experiences with regard to student population, school structures, and external influences. Therefore, consideration was given to the schools' demographics and their academic status as reported by the Virginia Department of Education (VDOE).

Table 3 *Selected Schools for the Teacher Questionnaire*

Secondary Schools	Elementary Schools
1. One choice middle school	1. Two choice elementary school
2. One middle school with an attendance zone	2. Four Title I elementary schools
3. One high school	3. One non-Title I elementary school

In the chosen school district, choice schools are those schools that have an application process. Students are accepted on a first-come, first-served space available basis and no transportation is provided for students who attend these schools. Additionally, parents sign a school contract upon registration and students' continued enrolment is contingent on their continued adherence to the tenants of the contract. Schools with attendance zones are those who register all students that reside within a particular area and there is no application necessary. Title I provides financial assistance to support instructional programs in schools with high numbers or percentages of low-income students to ensure that all children meet challenging content and achievement standards (US Department of Education, n.d.). Elementary schools in this district enrol students from kindergarten to grade 5, middle schools enrol students from grade 6 to grade 8, and high schools enrol students from grade 9 to grade 12.

District Focus on Assessment

The school district involved in the study has made multiple efforts to improve assessment practices in recent years. District-wide adoption of professional learning communities have been the standard for over a decade, school teams and curriculum leaders have been trained in examining standards and creating valid assessments, school teams have been trained on creating learning intentions and success criteria based on curriculum standards, quarterly benchmark testing protocol has been revised so that the data is looked at in a more formative nature, and lesson plan templates have been redesigned to promoted using instructional time for assessment and intervention. Despite these efforts, schools have continued to struggle with meeting the minimum state standards for acceptable achievement. Over the past five reporting cycles, the percentage of accredited schools in the district have has ranged from 40% to 55%.

Document Analysis Sample

Concurrent to the questionnaire completion window, an examination of lesson plans and assessment data from teachers at the selected schools was conducted as a method of document analysis. Document analysis involves reviewing material (written or electronic) and, similar to other forms of analysis in qualitative research, analysing documents involves making interpretation to extrapolate meaning and gain understanding (Corbin & Strauss, 2008). Bowen (2009) posits that documents for analysis may take many different forms. They can include, meeting agendas, books, letters, journals, reports, photos, and memorandums.

This study included an analysis of teacher lesson plans along with assessment tools that were used or that teachers planned to use in the classrooms. These data were used to provide examples and evidence to compare and contrast with the quantitative data gathered by the questionnaire as a means of triangulation—combining methods when examining the same phenomenon and is supported by the idea that qualitative researchers are expected to draw upon a variety of (at least two) sources of evidence to seek synthesis and verification by using different methods and using different sources of data (Bowen, 2010).

To collect these documents a convenience sample of 8 teachers was used—3 elementary school teachers, 2 middle school teachers and 3 high school teachers. The schools will be chosen based on my relationships with the principals and their willingness to release the lesson plans their teachers submit for review. Marshall (1996) argued that convenience sampling is the “least rigorous technique, involving the selection of the most accessible subjects” (p. 523), however, I chose this method of sampling gaining access and participation thru personal networks.

Instrumentation

Two separate instruments were used to collect information for this study. Data collection involved administering a questionnaire to quantify teachers' perceived understandings of assessment as well as their reported practices, while qualitative data collection involved a document analysis of lesson plans and assessment tools. A theoretical blueprint was used to address content validity of the questionnaire and to ensure the instrument was designed in a manner that thoroughly addressed the research questions. In addition, doctoral candidates and university professors served as reviewers and revisions were made based on their feedback. This section describes all of the instruments that were used in the study. Additionally, Results from the Cronbach's Alpha (.738) indicated an acceptable level of internal consistency from the data collected by the questionnaire.

Teacher Questionnaire: Quantitative

Gavin T.L. Brown (2004) designed a Conception of Assessment Inventory that categorizes conceptions using four assumptions; improvement of teaching and learning, accountability of teachers, accountability of students, and irrelevance. Brown (2004) used a tool consisting of 65 statements by which teachers indicated their level of agreement or disagreement with each statement using a 6-point rating scale. Others have since modeled their surveys to collect similar data for examination, and several surveys can be found that seek to measure or identify teachers' classroom assessment practices and assessment in general. However, after an extensive search, an established survey tool that was constructed around the more holistic assessment criteria I used in this study as the basis for measuring assessment literacy could not be found. As discussed in Chapter 2, this study involved a view of assessment that differs from the majority of the literature reviewed in that it delves deeper into assessment knowledge and

practice. Building on the commonly questionnaires that address the structure of the questions used and how assessments are used, this questionnaire sought to examine how clearly the assessment goals are communicated, how students are engaged in the assessment process and how the students are affected by the assessment, in its design.

The voluntary questionnaire (Appendix A) concerning teachers' perceived understanding of assessment and their reported assessment practices was administered in 2 middle schools, 1 high school, and 7 elementary schools, to core content (math, reading, social studies, and science) teachers of grades 3-12 in an effort to achieve a robust sample size. Hence, the questionnaire was created using a theoretical blueprint framed around the aforementioned holistic model. The questionnaire was organized in accordance to the criteria being used to define assessment literacy and followed the framework of the holistic model that includes; establishing a clear purpose, aligning instruction and assessment with curriculum standards, assessment design, communicating results, and student involvement, (Stiggins, Arter, Chappuis, & Chappuis, 2006; Chappuis, Stiggins, Arter, & Chappuis, 2012) and consisted of closed-end statements. Additionally, the questionnaire was divided into two strands; one that examined teachers' understanding of assessment literacy using a 4-point positively weighted scale: Strongly Disagree, Disagree, Agree, and Strongly Agree, and teachers', and the other section that reported teachers' assessment practices using a 4-point frequency scale; Never, Seldom, Some of the time, and Most of the time. To secure some demographic information about the teachers, additional questions were asked about the core subject they taught, their grade level, along with their schools current accreditation status. The questionnaire was reviewed by a team of experts and revised based on feedback and discussion.

Document Analysis: Qualitative

To analyze the lesson plans and assessments used by the teachers, a rubric was used to determine the alignment of the documents with the five components identified as defining assessment literacy; Alignment, Formative Use, Assessment Design, Communicating Results, Communicating Purpose, and Student Involvement. Rubrics are charts, lists or tables that describe the criteria being used to examine or measure performance or compliance. They are helpful in assessing behaviors or performances that can be difficult to capture through surveys (Salkind, 2006). The rubric was designed to measure assessment literacy as either high, medium, or low as related to each of the 5 components of the holistic model. The rubric also included a column to record examples and counter examples of each component as well as a means to provide specific evidence in support of the derived measurement of assessment literacy. To address trustworthiness, doctoral students from other institutions, as well as Old Dominion University subject area experts, reviewed the questionnaire to evaluate coherence-does it align with the purpose and design of the study.

Data Collection**Teacher Questionnaires**

Questionnaires for the selected schools were distributed, completed, and collected electronically using Qualtrics, an Old Dominion University endorsed survey tool, and after approval by the district's Research Authorization Committee, invitations to participate were sent to teachers by email with a link to the questionnaire and a letter of informed consent. Participants were able to complete the questionnaire at their convenience over a 4-week period from March 30, 2016 to April 29, 2016. Multiple reminders were sent to participants in an effort to obtain as many responses as possible. From the total number of questionnaires distributed, 73

participants completed the entire survey while at least 80 completed the questionnaire through one strand. Questionnaires that did not include at least one completed strand were not considered for analysis.

Document Analysis

Administrators in the selected school district are required to provide lesson plan feedback to teachers systematically, and as part of the information provided to participants about the purpose of the study, teachers were informed that their lesson plans and assessment materials may possibly be reviewed to provide additional insights into the other data being collected. The documents were accessed from 1 high school, 1 middle school, and 1 elementary school, based on the established relationships with the administrators, and electronic or printed copies of the documents were obtained either by email, document sharing, or the school interoffice mail system.

Document Analysis Trustworthiness

As the primary researcher, I recorded my thoughts and feelings in an effort to accurately document the results and to limit biasing effects on the processes of data collection and analysis. A reflective journal was maintained throughout the analysis of the documents to chronicle my thoughts about the process, the emerging themes, notes about the documents, and my reflections as I examined the data being collected.

Non-Specific Strategies of Trustworthiness

Verification and trustworthiness strategies were employed wherever possible to meet the criteria for credibility, validity and believability (Harrison, MacGibbon, & Morton, 2001).

Triangulating sources of data, maintaining an audit trail during the collection process, and a

simultaneous data collection and analysis, along with the strategies previously mentioned, addressed the threats to the trustworthiness of the study.

Protection of Human Subjects

An application to the Institutional Review Board to conduct research on human subjects was submitted and subsequently approved. After this approval, I received permission to conduct research by the school district's Research Authorization Review Committee. Collaborative Institutional Training Initiative (CITI) mandates that all of the participants be given notice of the risks and benefits before interviews, questionnaires, and document analysis are conducted. Therefore, consideration was given to the potential risks participants may be taking when disclosing information and documents that reveal their planning, instructional, assessment and grading practices. All of the data collected was stored, either electronically on a password-protected secure drive, or in a locked file that was only be accessible to my dissertation chair or me. Additionally, all identifiers were removed after document analysis coding.

Data Analysis

Teacher Questionnaire

The primary purpose of the questionnaire was to address the following research question: To what extent are selected grade 3-12 teachers' definitions and understandings of formative and summative assessment aligned with the operational definitions? Therefore the analysis of the quantitative questionnaire data involved looking at descriptive statistics to include the mean scores of each statement along with the grand means of the different clusters, and the standard deviation of each of the criteria being used to define assessment literacy. Using SPSS, mean scores were calculated for each statement and for the each question cluster in both strands (Practice and Understanding). The means were then evaluated based on the high, medium and

low assessment literacy scale previously established.

Document Analysis

The analytic process involves finding, selecting, interpreting (making sense of), and synthesizing the information gathered. Documents, whether authentic material or interview transcripts provide data—excerpts, quotations, or entire passages—that are then organized into major themes, categories, and case examples specifically through content analysis (Labuschagne, 2003). Lesson plans and assessments were coded using the rubric based on a concept map to indicate whether evidence of the criteria for assessment literacy was present. Specific examples and counter examples were recorded to support the various components and levels of assessment literacy as defined by the holistic model used for the study. After the initial review of the lesson plans, a face-to-face meeting was held with a colleague and doctoral student, who served as an independent evaluator. This meeting provided an opportunity to introduce the holistic model, explain the rubric being used for analyzing the lesson plans, and provide the initial lesson plan excerpts, absent the assigned ratings. This allowed for a comparison of the same examples for inter-rater reliability. Following the independent analysis, the notations were compared to the initial codes, characterizations based on levels of assessment literacy. The rate of agreement the original ratings and those of the independent evaluator was 86 percent. New excerpts and notations made by the independent evaluator were then assessed and the rate of agreement was 60 percent. Those items that were not in agreement were given the higher of the 2 ratings.

Limitations and Biases

Study Design Limitations

Based on the design of the study, the findings will not be externally generalizable. Even though a mixed-method design will be used to triangulate the data, the nature of self reporting

questionnaires lead to the potential for social desirability (participants providing answers that they feel are desirable rather than being completely honest) to influence the answers given. In addition, as noted in the teacher questionnaire sample section of this chapter, there is a chance that the data gathered from the questionnaire did not glean the diversity desired, based on the number of responses obtained. Furthermore, the demographic data gathered was considered beyond the sampling stage of the research and no analysis was conducted using that information concerning subject taught, grade level, or school accreditation status and the review of the literature did suggest that grade level and content area did influence assessment practice. It is suggested that future research consider these and other factors such as gender, ethnicity, and years of experience to examine whether they have any significance in the assessment practices of teachers based on the holistic model being used in this study.

Researcher Bias

Going into my research, my belief was that there exists, a strong positive correlation between teachers' assessment literacy and their instructional practices, and my interests and motivation were managed throughout the research by consistent collaboration with my research chairperson. Furthermore, my feeling was that teachers have a limited technical knowledge of the holistic model of assessment and that limited knowledge is reflected in their instructional practices. However, I do feel teachers aspire to be effective and want to use the best strategies possible to help their students learn. Additionally, the participants in the study were familiar with me as an administrator in the district, so that may have had some influence on the responses to on the teacher questionnaire.

Chapter 4

RESEARCH FINDINGS

Overview of Research Process

The purpose of this study was to examine the relationship between teachers' perceived assessment literacy and their classroom practices, guided by the following research questions:

- 1) To what extent are selected grade 3-12 teachers' perceived understandings of assessment aligned with the holistic model?
- 2) What assessment practices do selected grade 3-12 teachers report using in their classrooms?
 - a. How do grade 3-12 teachers use assessment to inform instruction?
 - b. How do grade 3-12 teachers select and design assessments?
 - c. What assessment practices do grade 3-8 teachers report regularly using in their classrooms?

This chapter will present an analysis of the data collected through a mixed method design by a review of teacher lesson plans and a self-report teacher questionnaire. Additionally, this chapter will describe the process involved in constructing the three major themes from the analysis and explain how they relate to the components of the Holistic Assessment Model and the stated research questions. The themes emerged after a systematic review of written lessons, including an analysis from a colleague and doctoral student at another institution for inter-rater reliability, coupled with a triangulation of the findings using the data from the teacher questionnaire.

Review of Lesson Plans

The lesson plans were examined to find evidence of each component of the rubric and excerpts from the plans were colored coded to indicate high, medium or low assessment literacy. In addition, while examining the written plans, notations were made to capture ideas, conclusions, and trends that may have emerged as themes, and to document assessment literacy counterexamples, missing evidence, and points in the described lesson where opportunities for documenting evidence of high assessment literacy were missed by the teacher. Specific examples of assessment literacy were clustered, with the goal of identifying trends as they related to lesson sequence, selected activities, assessment tools, verbiage, and learning intentions. Additionally, plans were also compared to curriculum documents on the state and district level to support analyzing how closely the learning intentions were aligned with the required standards.

An independent analysis of the plans by a colleague and doctoral student at Virginia Polytechnic and State Institute provided additional insights and added to the validity of the findings. In a face-to-face meeting, The Holistic Model of Assessment Literacy and the rubric being used to examine the lesson plans were reviewed and discussed with none of the prior observations, notations, ideas, or classifications being shared with the independent reviewer. After eleven days, the plans were returned and the observations and notations made by the independent reviewer were examined. Observations that were similar were given the appropriate previously established color-code and new observations and notations provided by the independent analysis were added to the original list of items—preliminary themes were constructed at this stage.

Review of the Questionnaire

Finally, the mean data gathered by the teacher questionnaire was analyzed by separating each cluster of statements and considering both individual and grand means. Summary statements were developed based on the trends that emerged from the responses of the selected grade 3 through 8 teachers and were compared to the preliminary themes developed from the analysis of the lesson plans. From this comparison, final themes were determined by tabulating the frequency of examples or counterexamples found in the lesson plans, along with the means calculated from the questionnaire and comparing the aligned summary statements. The following themes emerged from the analysis of the data.

1. There is a surface level understanding and use of assessment practices.
2. Assessment is not viewed or practiced as an integral component of instruction.
3. There is a gap with regard to what is understood about assessment and what is practiced.

Description of the Analysis

Lesson Plan Analysis

The analysis of the lesson plans were guided by a rubric (Appendix A) that rated the quality of the written learning intentions and activities as demonstrating high, medium, or low assessment literacy, and categorized based on the holistic model of assessment literacy that considers; alignment, formative use, student involvement, communicating results, communicating purpose, and assessment design. Table 4 illustrates the continuum of assessment literacy established by the rubric.

Table 4: *Continuum of Assessment Literacy*

Alignment	Classroom activities and assessments are completely aligned to the curriculum standard(s)	Classroom activities and assessments are somewhat aligned to the curriculum standard(s)	Classroom activities and assessments are not aligned to the curriculum standard(s)
Formative Use	There is clear evidence of using formative assessment to make guide instruction	There is some evidence of using formative assessment to make guide instruction	There is little or no evidence of using formative assessment to make guide instruction
Student Involvement	There is clear evidence of teacher feedback, student reflection, and student input	There is some evidence of teacher feedback, student reflection, and student input	There is little or no evidence of teacher feedback, student reflection, and student input
Communicating Results	Plans indicate a method for the timely communication of the results of an assessment	Plans indicate a method for the timely communication of the results of an assessment	Plans do not indicate a method for the timely communication of the results of an assessment
Communicating Purpose	Plans indicate clear communication of the assessment activity's purpose and how the results will be used.	Plans indicate some level of communication of the assessment activity's purpose and how the results will be used.	Plans do not indicate clear communication of the assessment activity's purpose and how the results will be used.
Assessment Design	Assessment activities are chosen or created that meet the needs of the students and are appropriate in length	Assessment activities are chosen or created that somewhat meet the needs of the students and are appropriate in length	Assessment activities are chosen or created that do not meet the needs of the students and are appropriate in length

The Initial Review of the Lesson Plans

The plans of eight different teachers were reviewed for the document analysis. At the elementary level, science, reading, and math plans from three different teachers were examined. At the middle school level, science, and language arts plans prepared by two different teachers were analyzed, along with the US Government, biology, and algebra II plans written by three high school teachers. Each teacher's plans spanned a 2 to 3 week period of instruction to provide an opportunity to track their instructional practices over time and examine assessment trends. Lesson plans were analyzed to determine how consistently the written plans were aligned with the design of the rubric (Appendix A) and preliminary themes were developed based on the evidence extracted and the inferences drawn from the overall set of plans. The following sections will provide insight into the process of analyzing the written lesson plans as they relate to the Holistic Model of Assessment Literacy.

Alignment. Assessment literacy in the area of alignment was evaluated by how well the learning intentions, classroom activities, and assessments are completely aligned to the

curriculum standards. An example of high assessment literacy came from the following excerpt documenting the learning intention for the day. *I can understand the difference between what the Earth's rotation and Earth's Revolution means when I explain it on a graphic organizer with support from my teacher.* This showed the teacher's understanding of aligning the curriculum standards, learning intentions and classroom activities. An example of medium assessment literacy came from the following excerpt documenting the daily learning intention. *Model the formation of the eight moon phases, sequence the phases in order.* This learning intention shows only medium assessment literacy as it relates to alignment because it falls short of meeting the standard that also requires that students describe how the phases of the moon occur. An example of low assessment literacy was extracted from a reading lesson plan that listed a learning intention focused on identifying main idea and summarizing using supporting details. The success criteria was listed as follows: *The student will be able to complete a web on characteristics of tall tales with support from the teacher.* This misalignment demonstrates a lack of understanding of the essential relationship between the goals of the lesson and the measure of its success because completing a web focused on the characteristics of tall tales does not match the stated goal of identifying main idea and summarizing using supporting details. In general, the alignment of lesson plans only went as far as the curriculum standard and did not extend to the learning activities students were engaged in or the assessments used to measure their success.

Formative use. Assessment literacy in the area of formative use was evaluated by the presence of evidence of using formative assessment to guide instruction. The following extract shows a high level of formative use assessment literacy. *Following a Common Assessment, we will break up into 1 Enrichment group and 2 groups for remediation and retest the skill.*

Garrison & Ehringhaus (2007) emphasizes that formative assessment allows teachers to plan their next steps during the learning process leading up to the final summative assessment. This excerpt from the lesson plans demonstrates an understanding of the use of assessment to inform future instructional plans. An example of medium formative assessment literacy came from the following extract. *Teacher and student will engage in discussion until the majority of students understand.* This indicates an understanding of using student data to make decisions, however, there were no measurable criteria for determining what constituted understanding and no plan for addressing the needs of the students who did not learn at the desired rate. The extract, *Bell Ringer will be graded for completion and discussed for understanding,* indicates a low level of formative assessment literacy. It provides no evidence any action will be taken if students do not understand, and further suggests that no specific feedback will be provided to students about their mastery, while providing them with a favorable grade for completing the assignment whether their work is correct or incorrect. The excerpts taken together suggest there are varying levels of understanding when considering how teachers document formative assessment. However, it is not clear if the implementation of the written plans meet the criteria required to reflect effective formative assessment practices.

Student involvement. Assessment literacy in the area of student involvement was evaluated by evidence of student feedback, student reflection, and student input using a scale of Clear Evidence, Some Evidence, or Little or No Evidence. As they are written, the following excerpts from the lesson plans show a high level of assessment literacy; *Students will discuss the results of their modules and will correct any problems with context* and *The teacher will have the students that understand reteach the lesson by using their own creative examples in the front of the classroom.* These excerpts demonstrate high assessment literacy because they show a

deliberate effort to provide students with an opportunity share their answers and thought processes with the teacher and each other. They also indicate an understanding that students benefit from having the chance to revise their work prior to submission. A medium level of assessment literacy was evident from the following excerpts; *The students will complete as many words as possible, using their knowledge of level 3 and 4 roots* and *The students will exchange papers with peers and will fill in any words that their classmates have left out.* These excerpts indicate the understanding that students should have a level of involvement in the assessment process and be afforded opportunities to reflect on their work and engage in discussions about their progress. However, the lack of a set criteria for the minimum number of words to complete fails to set a standard, and in the second excerpt falls short of describing what formative value will be gleaned from the activity and how the students will be allowed the close the loop and have discussions about the words they added or omitted, whichever the case may be. As collective evidence, these excerpts, along with other evidence from the lesson plans suggest that there is knowledge of the importance of student involvement, however, the parameters of that involvement appear to be limited to the traditional strategies used to involve students in classroom lessons, failing to reach the level of ownership in assessment described by the Holistic Model.

Communicating results. Assessment literacy in the area of communicating results was evaluated based on evidence that indicated a method for the timely communication of the results of an assessment. The entire analysis of the lesson plans produced no explicit reference to communicating results and, in fact, missed multiple opportunities to use timely and specific feedback to promote student learning. For example, where references were made to assigning homework, classroom assignments to grade, and students completing exit tickets, there was no

subsequent reference to indicate that the results were communicated to any stakeholder. The best examples of providing feedback evidenced in the lesson plans were instances where the class reviewed the answers to assignments and when assignments were graded in class. When these actions occurred however, there was no indication that the activity was purposed as an assessment. The lack of evidence of communicating assessment results to students in the lesson plans demonstrates a tendency to devalue its significance and, thus suggests that it is not viewed as an integral component in instructional delivery.

Communicating purpose. Assessment literacy in the area of communicating purpose was evaluated based on evidence that indicated clear communication of the assessment activity's purpose and how the results were used. The lesson plans lacked evidence of this component of the Holistic Model of Assessment Literacy as well. The one instance that related to an attempt to communicate purpose was documented when a language arts teacher noted that the students would be told that they would be taking a quarterly assessment during the class period. The lesson plan did not indicate that the students were told the purpose of the assessment—how it related to them and how the information gained would be used in the future. As with communicating results, the lack of evidence taken from the lesson plans that indicate the students were informed of the purpose of each assessment activity suggests that the teachers whose plans were analyzed did not consider this action to be one they thought was essential.

Assessment design. Assessment literacy in the area of assessment design was evaluated based on evidence that assessment activities were chosen or created that met the needs of the students and were appropriate in length. The examples of assessment design indicated an overall low level of assessment literacy in this area. One set of lesson plans listed the following learning intentions:

- The students will put texts in order, according to sequence, when given a piece of nonfiction text.
- The students will identify the pattern of organization used by an author in a piece of nonfiction text
- The students will identify common signal words used inside nonfiction text and will analyze how those signal words contribute to the text's organization

The lesson plan that included the above learning intentions did not provide an assessment measure to indicate the level of success for each, and the listed student activities only indicated that students worked with identifying signal words and nothing else. A different lesson plan indicated that the chosen assessment tool was a chapter test provided by the textbook company without any revisions to address the specific needs of the students. A review of the chapter test showed sections one and two of the assessment was not aligned with the learning intention that stated, *Identify the two houses of Congress and identify the role of Congress in checks and balances*. Hence, the use of a pre-made assessment that did not match the content students needed to learn demonstrated a low level of assessment literacy. Taken together, the observations and notations from the lesson plans suggest that pre-made or standardized tools were used in class more than those created by the teachers and were not constructed to meet the specific needs of the students.

Teacher Questionnaire Analysis

To analyze the teacher questionnaire, SPSS was used to calculate the mean scores for each question as well as the grand mean for each question cluster based on the five components of the Holistic Assessment Model. The questionnaire was divided into High, Medium and Low assessment literacy determined based on the following scale; High Assessment Literacy = \geq

3.50, Medium Assessment Literacy = < 3.50 and ≥ 3.00 , and Low Assessment Literacy = < 3.00 . The following sections will report the findings by question clusters aligned with the Holistic Model of Assessment Literacy.

Assessment design. With regard to assessment design, the grand means fell in the low assessment literacy range with the mean for practice being 2.56 and the mean for understanding being 2.88. The data presented in Table 5.1 and Table 5.2 indicate that statements that were consistent with current standardized, summative assessment practices and inconsistent with the assessment best practices based on the Holistic Model of Assessment yielded the lowest mean scores in this cluster. For example, the statements, *My assessments mirror the design of standardized test* (1.52) and *My assessments cover multiple skills and concepts* (1.43) were both reverse coded and the responses indicate that the teachers employ assessment practices that are aligned with the format of the Standards of Learning examination. Similarly, the statements, *Multiple choice tests can address all levels of cognitive demand* and *It is difficult to create/decide on an assessment until you know what students have learned*, had a means of 1.93 and 2.77, respectively. The questionnaire results also indicated that teachers understand that assessments should be tightly aligned to curriculum standards (3.22) and they felt collaboration with other teachers could help to improve the quality of the assessment (3.42). Interestingly, when reporting their classroom practices, teachers indicated that they create their own assessments (3.42), design assessments to meet the individual needs of students (3.02), and design and select assessments differently based on the specific purpose (3.47). Taken together these data suggest that there is a lack of understanding about the interrelatedness of design, format and purpose and that how assessments are used are equally important as how they are created.

Table 5.1: *Descriptive Statistics: Assessment Design Practice*

Questionnaire Statement	<i>n</i>	<i>M</i>	<i>SD</i>
I create the assessments I use in class	80-84	3.42	.587
My assessments mirror the design of standardized test (R)	80-84	1.52	.633
My assessments cover multiple skills and concepts (R)	80-84	1.43	.567
My assessments are designed to match the individual needs of students	80-84	3.02	.841
I design/select assessments differently based on its specific purpose	80-84	3.47	.801
Grand Mean	80-84	2.56	.340

Table 5.2: *Descriptive Statistics: Assessment Design Understanding*

Questionnaire Statement	<i>n</i>	<i>M</i>	<i>SD</i>
Assessments should be lengthy to accurately gauge students' knowledge (R)	71-73	3.11	.591
It is difficult to create/decide on an assessment until you know what students have learned. (R)	71-73	2.77	.773
Multiple choice tests can address all levels of cognitive demand	71-73	1.93	.751
Assessments are should be tightly aligned with the essential skills students will be held accountable for based on the state standards	71-73	3.22	.610
Creating assessments collaboratively can help improve the quality of the design	71-73	3.42	.644
Grand Mean	71-73	2.88	.318

Alignment. The alignment cluster of the questionnaire examined how curriculum standards were considered as a component of assessment. Table 6.1 and Table 6.2 provide supporting data that indicate statements involving curriculum standards yielded the highest mean scores. As an example, the statement, *I assess students strictly based on the content standards* (3.61) and the statement, *Assessments should be aligned with the written and taught curriculum* (3.44) indicated some of the highest levels of assessment literacy. Conversely, within the same cluster, statements yielded some of the lowest means, indicating the lack of depth in assessment literacy that evolved into a major theme. For example, reverse coded statements, *I create*

assessment items that are more challenging than the content standards so that students will be ready for easier questions on future assessments and previously taught content should be part of each assessment to make sure students remember yielded means of 1.98 and 1.95 respectively, indicating a low level of assessment literacy. Graue (1993) posits that just matching the curriculum content to the assessment as it relates to standards is not enough, supporting the Holistic Model of Assessment that requires that a deeper understanding of the interconnections between each component. These data suggest that the teachers associate alignment mainly with curriculum documents and do not reconcile alignment with the purpose of the assessment as prescribed by the Holistic Model, while simultaneously reporting their belief that grades should be consistent with the students' achievement on summative and formative assessments.

Table 6.1: *Descriptive Statistics: Alignment Practice*

<i>Questionnaire Statement</i>	<i>n</i>	<i>M</i>	<i>SD</i>
I assess students strictly based on the content standards	80-84	3.61	.561
I include content in my assessments that think students need to know, even though it may not be included in the curriculum standards (R)	80-84	2.07	.833
I create assessment ideas that are more challenging than the content standards so that students will be ready for easier questions on future assessments (R)	80-84	1.98	.737
I decide how I will assess a skill/concept prior to beginning my instruction	80-84	3.40	.700
I use the same questions/problems/prompts on my assessments that I have used during the course of my instruction	80-84	3.06	.775
Grand Mean	80-84	2.82	.360

Table 6.2: *Descriptive Statistics: Alignment Understanding*

<i>Questionnaire Statement</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Assessments should promote student success over rigor in order to foster motivation (R)	71-73	2.49	.715
Previously taught content should be part of each assessment to make sure students remember (R)	71-73	1.95	.575
Assessments should be aligned with the written and taught curriculum	71-73	3.44	.500
Using a blueprint to design an assessment helps to improve its alignment with the curriculum standards	71-73	3.32	.524
Classroom grades should be consistent with both formative and summative assessment scores	71-73	3.10	.670
Grand Mean	71-73	2.85	.306

Student involvement/formative use. Four of the five statements that measured the teachers' assessment practices as they relate to student involvement and formative use of assessment yielded means in the range of low assessment literacy (Table 7.1). For example, the reverse coded statements addressing formative assessment, *I wait to assess my student until after I am finished all of my instruction* (1.89) and *I make sure I follow the curriculum pacing regardless of whether all students are mastering the content* (2.09) suggest assessment practices that are far less than student-centered. In addition, statements that related to student involvement also had means that indicated low assessment literacy. Students setting learning goals (2.33) and student documentation of reflection of their own progress (2.45) were not reported as being frequently practiced in classrooms, and the lone statement that reflected practices that reached the level beyond low assessment literacy, *I incorporate interventions into my instructional block based on the results of my assessments* (3.31), measured in the medium assessment literacy range. Looking at student involvement and formative use of assessment from the perspective of teachers' understanding resulted in different findings. Table 7.2 shows three of the five statements yielded means that fell in the medium assessment literacy range, and the grand mean of 2.96 was higher than the 2.41 grand mean of the statements addressing classroom practices.

The highest mean was associated with the statement referencing the idea that when students talk to one another, it encourages new ideas and different perspectives (3.47). The concepts of using assessments students fail as a data source (3.31) and the use of rubrics to help students understand the requirements of an assignment (3.24) were the two other statements that reflected medium assessment literacy. However, when considering students being involved in the creation of the assessment (2.65) and the importance of assigning a value to students' conceptual knowledge (2.18), teachers were in less agreement with these statements focused on their understanding of effective assessment practices. Collectively, these data suggest that there is a gap between what is understood and what is practiced as it relates to formative assessment. Moreover, students do not appear to be viewed as an active participant in classroom assessment.

Table 7.1: *Descriptive Statistics: Student Involvement/Formative Use Practice*

<i>Questionnaire Statement</i>	<i>n</i>	<i>M</i>	<i>SD</i>
I wait to assess my students until after I am finished all of my instruction (R)	80-84	1.89	.955
I require students to document reflection of their progress	80-84	2.45	.884
I make sure I follow the curriculum pacing regardless of whether all students are mastering the content (R)	80-84	2.09	.845
Student set learning goals for skill/concepts being taught	80-84	2.33	.938
I incorporate interventions into my instructional block based on results from my assessments	80-84	3.31	.789
Grand Mean	80-84	2.41	.510

Table 7.2: *Means and Standard Deviations: Student Involvement/Formative Use Understanding*

<i>Questionnaire Statement</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Understanding Student Involvement/Formative Use Grand Mean	71-73	2.96	.323
Assessments are most useful for assigning value to what skills students have mastered (R)	71-73	2.18	.657
An assessment that the majority of students fail can still provide usable data	71-73	3.31	.521
Students talking to one another encourages new ideas and different perspectives	71-73	3.47	.530
Rubrics help students understand what is required of them when completing assignments	71-73	3.24	.617
Student involvement creating assessments can bias the results (R)	71-73	2.65	.609
Grand Mean	71-73	2.96	.323

Setting a Clear Purpose. The grand mean for the portion of the teacher questionnaire concentrated on measuring the teachers' practices as they relate to setting a clear purpose for assessments was 2.88 (Table 8.1). Teachers also reported means in the low assessment literacy range for statements addressing grading for completion rather than accuracy (2.51) and administering pop quizzes as a student accountability measure (2.83). They reported a lack of use of pre-assessments to guide their planning to present new information to students (2.72), falling in the low assessment literacy range as well. Responses in the medium assessment literacy range indicated that teachers show students examples of work that meet the success criteria (3.40) and that teachers use rubrics to communicate the criteria for success for daily assignments. Four of the five statements measuring the teachers' understanding of setting a clear purpose for assessment (Table 8.2) yielded means in the medium assessment literacy range. Teachers agreed with the following concepts that are aligned with assessment best practices; *Students should be informed about what will be included on their assessments* (3.01), *Assessments should be used to measure student progress and guide instruction* (3.22), *There should be a comprehensive plan for integrating assessments into the classroom over time* (3.04), and *Assessments can be used for the sole purpose of evaluating and adjusting instruction* (3.03).

Contrary to effective assessment practice, teachers felt that assigning a grade to an assessment is the best method for monitoring student progress. The mean for this response was 2.65, falling in the low assessment literacy range, and overall, the responses for this cluster of statements further supported the existence of gaps between what is known about assessment and what is practiced, and further suggests that grading policies may significantly influence assessment practices.

Table 8.1: *Descriptive Statistics: Communicating a Clear Purpose Practice*

<i>Questionnaire Statement</i>	<i>n</i>	<i>M</i>	<i>SD</i>
I show students examples of work that meets the criteria for success	80-84	3.40	.683
I use pre-assessments to determine where I need to begin my instruction on a new topic	80-84	2.72	.855
Rubrics are used to communicate the success criteria of daily assignments and/or projects	80-84	3.00	.791
I give pop quizzes in order to make sure my students have completed their assignments (R)	80-84	2.83	.863
I grade assignments for completion (effort) rather than for accuracy (R)	80-84	2.51	.744
Grand Mean	80-84	2.88	.352

Table 8.2: *Descriptive Statistics: Communicating a Clear Purpose Understanding*

<i>Questionnaire Statement</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Informing students what will be included on an assessment makes it difficult to accurately analyze the results (R)	71-73	3.01	.661
Most assessments should be used to measure student progress and guide future instruction	71-73	3.22	.610
There should be a comprehensive plan for integrating assessments into the classroom over time	71-73	3.04	.491
To best monitor students' progress, a grade is should be assigned for each assessment (R)	71-73	2.63	.680
Assessments can be used for the sole purpose of evaluating and adjusting instruction	71-73	3.03	.676
Grand Mean	71-73	2.98	.363

Communicating results. Table 9.1 shows that what teachers reported as their practices for communicating results of assessments had a grand mean of 2.69, which fell in the range of low assessment literacy. The following statements yielded response means that were in the low assessment literacy range; *I only discuss the assessment results with other teachers and/or administrators* (2.28), *I use rubrics to give specific feedback to students on their progress* (2.64), and *I use reporting options other than grades to communicate the results of assessments* (2.40). The idea of students revising their work based on feedback and allowing students to correct each others' papers to save time (a reverse coded item) yielded means of 3.11 and 3.08, respectively

and were in the range indicating a medium level of assessment literacy. Table 9.2 shows understanding the significance of communicating the results of assessments had the highest grand mean on the questionnaire and indicated a medium level of assessment literacy at 3.04. Teachers reported believing that students require feedback to promote success (3.14), non-numerical or graded feedback should be provided (3.08), and that more than grades are needed for students to accurately reflect on their progress (3.03). However, teachers did not agree that students should be told how the assessment results will be used (2.89) and further disagreed with the idea of predetermining who receives results based on the stated purpose of the assessment (2.99). These data suggest that teachers may have a limited understanding of the scope of communication and that there again exist a gap between knowledge and practice. The glaring contradiction between the responses to several of the statements in this cluster and the lesson plan analysis will be discussed in more detail in the explanation of Theme 2.

Table 9.1: *Descriptive Statistics Communicating Results Practice Means*

<i>Questionnaire Statement</i>	<i>n</i>	<i>M</i>	<i>SD</i>
My students revise their work based on teacher or peer feedback	80-84	3.11	.612
I use reporting options other than grades to communicate the results of assessments	80-84	2.40	.890
I use rubrics to give specific feedback to students on their progress	80-84	2.64	.747
To save time, I allow my students to exchange and correct each others' papers (R)	80-84	3.08	.823
I only discuss the assessment results with other teachers and/or administrators (R)	80-84	2.28	.941
Grand Mean	80-84	2.69	.398

Table 9.2: *Descriptive Statistics Communicating Results Understanding Means*

<i>Questionnaire Statement</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Students should frequently receive feedback on completed assignments that DOES NOT include a numerical or alphabetical grade	71-73	3.08	.599
Students do not require frequent feedback o their performance to them succeed (R)	71-73	3.14	.737
The numerical or alphabetical grade provides enough information for students to accurately reflect on their progress towards learning goals (R)	71-73	3.03	.609
Students should be told how the results of each assessment will be used.	71-73	2.89	.571
How and to whom the results of an assessment is reported depends on its pre-determined purpose	71-73	2.99	.643
Grand Mean	71-73	3.04	.370

The lesson plans and the questionnaire analysis provided a variety of observations and responses that could have been included in multiple themes. To further assist in triangulating the data, summary statements were derived from the mean scores of each cluster (Table 10) with their assigned level of assessment literacy based on the established scale. Additionally, due to the interdependent nature of the Holistic Assessment Model, there is an implicit overlap between themes and each theme was derived after categorizing the evidence, considering the questionnaire summary statements, and making distinctions based on common verbiage, as well as perceived intentions and understandings. For example, assessment alignment and assessment design are interrelated and an assessment created or chosen that is not aligned with standards would demonstrate low assessment literacy in both areas of the model. Additionally, the purpose of an assessment should consider who the results will be shared with, what form the feedback should take, when it will be provided and how the data will be used.

Table 10: *Teacher Questionnaire Summary Statements*

Assessment Practices	Assessment Literacy	Understanding of Assessment	Assessment Practices
<p>Assessment Design Summary: Teachers report creating assessments that cover multiple skills and that consider the specific need of students but design them to match the format of standardized tests</p>	Low	<p>Assessment Design Summary: Teachers agreed that multiple-choice tests can address levels of high cognition assessments are best created after knowing what students have learned. Teachers agreed that they benefit from collaboration and that assessments should be relatively brief and in tight alignment to with the state standards.</p>	Low
<p>Alignment Summary: Teachers report assessing strictly based on standards, beginning with the end in mind, but claim to increase the rigor on purpose and also add content they believe is important even if it is not in the standards.</p>	Low	<p>Alignment Summary: Teachers agreed that assessments should be aligned to the written and taught and that using state documents helps to improve alignment when creating assessments. Teachers agreed that grades should be aligned with formative and summative assessment results, but also agreed that success should be promoted over rigor and previously assessed material should be included in each assessment.</p>	Low
<p>Student Involvement and Formative Use Summary: Teachers report assessing more summatively and stressing pacing over teaching for mastery. They report providing interventions, but do not involve students in setting learning goals or require students to reflect on their own progress.</p>	Low	<p>Student Involvement and Formative Use Summary: Teachers agreed that student conversation and rubrics are beneficial to learning and that failed assessments can provide useful data. However, teachers also agreed that the best use of assessments is to assign value to knowledge and that involving students in the creation of an assessment can bias the results.</p>	Low
<p>Setting a Clear Purpose Summary: Teachers report using pop quizzes for student accountability, grading for completion rather than accuracy and not using pre-assessments. However, teachers also report showing students exemplars and using rubrics to assist students understanding success criteria.</p>	Low	<p>Setting a Clear Purpose Summary: Teachers agreed that there should be a plan for incorporating assessments into instruction, measuring learning, and adjusting instruction. However, teachers also agreed that assigning grades to assessments are the best way to monitor progress.</p>	Low
<p>Communicating Results Summary: Teachers report using grades as the primary means of communicating grade (sharing results with only teachers and administrators), and seldom using rubrics to provide students with feedback. In addition, teachers report students exchanging papers to grade as a means to save time and allowing students to revise their work based on feedback</p>	Low	<p>Communicating Results Summary: Teachers agreed that feedback is important students should receive non-numerical feedback. However teachers disagreed with the idea that students should be informed on how results will be used and who is informed should be based on the purpose of the assessment.</p>	Medium

Final Themes from Triangulation

Theme 1: Data Results Indicate a Surface Level Understanding and Use of Assessment

Practices

This theme was developed from the inconsistent relationship between the questionnaire responses and the lesson plans. When examining the grand means of the all of the statement clusters on the questionnaire listed in Table 11, ranging from 2.41 to 3.04, it was determined that the overall level of assessment literacy fell in the low, to low medium range. Excerpts and evidence from lesson plans provided additional information to support the theme. Different plans examined included what seemed to be generic statements that consistently appeared in their plans. The statements did not include the degree of detail that would indicate that the activity would be implemented with fidelity. As an example, plans from a math class included the following statement on each day. *Select groups will share their answers with the class. Any mathematical concerns will be addressed.* A separate math lesson plan included the statement, *Following a Common Assessment, we will break up into 1 Enrichment group and 2 groups for remediation and retest the skill. This will occur until success is achieved,* on each day of plans with no alterations or specificity about which students would be involved, what strategies would be used, or what triggers would lead to determining when this actions would occur. Another set of lesson plans included the following written description in the section of the plan labeled reteach. *Students will receive additional instruction from the teacher and/or peers to focus on areas that need improvement.* Similar to the previous lesson plan notations, there was no specificity to indicate that criteria were established and used to trigger the additional instruction of areas needing improvement. More importantly, plans for the following day showed no evidence that there was any area that needed improvement based on the previous day's

instruction and assessment. More evidence that supported this theme appeared in a language arts lesson plan that listed formative assessment every day as, *teacher observation and a prodigy report*. Consistent with the evidence previously reported, there was no criteria established or reaction noted for the class or for individual students based on the data gathered from the observation or the prodigy report.

To accompany the lack of detail included in the written plans relative to the use of effective assessment practice, there was a general lack of authentic or performance based assessments noted in lesson plans. Furthermore, although differentiation was listed in some cases as a re-teaching strategy, there was no indication that differentiation occurred in content, process, or product in any of the lesson plans examined. The following excerpts from various plans were meant to serve as examples of differentiation:

- The teacher will pass out vocabulary packets to each student
- The students will glue the vocab and affix cards they have cut out to the slot provided on the sheet
- Students will take notes on a PowerPoint about permutations and combinations

The examples listed do not consider student differences and requires all students to complete the same task in the same manner, providing a common product. Furthermore, lesson plans do not mention any form of differentiation as part of the initial instructional plan, which would demonstrate that teachers use their knowledge of student differences to design instruction that meets the individual needs of their students.

Table 11: *Questionnaire Cluster Grand Means*

Assessment Practice Cluster	<i>M</i>	Assessment Understanding Cluster	<i>M</i>
Assessment Design	2.54	Assessment Design	2.88
Alignment	2.82	Alignment	2.85
Student Involvement/Formative Use	2.41	Student Involvement/Formative Use	2.96
Setting a Clear Purpose	2.88	Setting a Clear Purpose	2.98
Communicating Results	2.69	Communicating Results	3.04

Theme 2: Assessment is Not Viewed or Practiced as an Integral Component of Instruction.

This theme was developed as a direct result of the findings in the lesson plan review and is supported by the self-report teacher questionnaire. The daily lesson plans did not show any evidence that the purpose or the result of assessments were communicated to students, or that there was any reaction to the assessments listed (which served as evidence to support Theme 1). Lesson plans from day to day did not reference previous assignments or activities as a means to inform the current plan of instruction, and most of the activities were tasks that students needed to complete or teacher direct instruction/modeling that did not involve formative assessment intentions. In addition, as shown in Table 6, the grand means for setting a clear purpose for assessment and communicating results on the Assessment Practices section of the survey, and several of the individual statements about setting clear purpose were in the low assessment literacy range.

Theme 3: There is a Gap With regard to What is Understood About Assessment and What is Practiced.

This theme was derived as a natural extension of themes one and two, using evidence from the lesson plans and the teacher questionnaire. The questionnaire was designed using two strands (Understanding Assessment and Assessment Practices) with clusters to match the Holistic Assessment Literacy Model. The grand means for each cluster of the Understanding Assessment strand were higher than the corresponding cluster in the Assessment Practices strand. This is suggestive of a knowing-doing gap as it relates to what is understood about assessment and what practices are implemented in the classroom. More telling is that many statements that yielded medium to high assessment literacy results on the questionnaire were not validated by the examination of the lesson plans, which serve as the blueprint for daily instruction. For example, two statements; *I design/select assessments differently based on its specific purpose* and *Creating assessments collaboratively can help improve the quality of the design*, both yielded means that suggest medium assessment literacy (3.47 and 3.42). However, lesson plans showed no evidence of varied purposes for assessments or specific non-traditional assessment design, but instead heavily referenced worksheets, workbook pages, and standardized assessments. As another example of the gap between what is understood about assessment and what is practiced, the following statements from the questionnaire yielded means of 3.31 and 3.24 respectively, indicating a medium level of assessment literacy; *I incorporate interventions into my instructional block based on results from my assessments*, and *Rubrics help students understand what is required of them when completing assignments*. There was no evidence of cohesion between these understandings and the practices recorded in the lesson plans. Not only was there no written indication that rubrics were used to assist and inform students and teachers,

there was only one reference that included enough specificity to indicate any deliberate attempt to respond to student progress. This reference came from a math teacher, who wrote, *The teacher will re-teach solving rational equations based on the ability of student to complete the sample problems.*

Summary

The purpose of this study was to examine the extent to which, if any, there exists an implementation gap for the participants of the research. This chapter discusses the findings of the teacher questionnaire and the analysis of the teachers' lesson plans. The findings were presented in a manner that emphasized the essential components of the Holistic Model of Assessment that guided this study to provide insight into how each component is essential for maximizing assessment literacy.

As it relates to grade 3 through 12 teachers' understanding of assessment, the data gleaned from the research suggests that they have a conceptual awareness and are familiar with the fundamental ideas and terminology that drive the theories associated with how assessment should be used to maximize learning and student achievement. The higher mean scores from the statements on the questionnaire were noted on the sections that measured teachers' understanding and those statements that used explicit language that represented widely accepted principles of assessment best practices. Lesson plans, while showing less evidence of understanding about favorable assessment practices, did show that the teachers were aware of the importance of curriculum alignment when creating daily learning intentions and that overall, they were cognizant of the need to show cohesion between the learning intentions and learning activities assigned to the students. In addition, taken collectively, the lesson plans showed that

the teachers understood the need for re-teaching when students failed to achieve and that some form of daily assessment are expected to be used as a measurement of the students' progress.

The data collected and analyzed suggests that the classroom practices of the selected grade 3 through 12 teachers lagged behind their concept knowledge. Generally speaking, responses to statements designed to measure their use of effective assessment practice yielded lower mean scores, and in many cases, were in direct contradiction to more favorable responses given to responses that measured understanding. Additionally, lesson plans showed very little evidence of the teachers putting into practice, the knowledge they reported relating to assessment. Plans were generic in nature and did not appear to consider student differences or indicate reactions to students based on data collected from assignments, observations, or discussions. There were no pre-assessments recorded, although the questionnaire results indicated teachers understood the value of this practice, and neither assignments nor exit tickets students completed were mentioned on the proceeding day's plans to indicate that the data gathered was used to inform instruction.

It is significant that the analysis of both the lesson plans and the questionnaire showed, communicating purpose and communicating results were the two areas of the Holistic Model that were least understood and even less practiced by the participants. Student involvement was another area where there was very little evidence found in the lesson plans and where low mean scores were calculated from the teacher questionnaire. Taken together, this shows a tendency to consider less vital, those assessment components that involve students most. Whereas the holistic model holds the student as the most important variable, and driving force behind every instructional decision made. Moreover, based on the data gathered, assessment, in general, did

not appear to be associated with classroom assignments or activities due to the lack of purposeful use of, and reaction to the data teachers collected in class from the work students completed.

Chapter 5

DISCUSSION, IMPLICATIONS, RECOMMENDATIONS

Study Overview

This study examined the self-reported understandings selected grade 3-12 teachers had of assessment and the assessment practices they reported implementing in their classrooms along with evidence extracted from written lesson plans. The literature on classroom assessment suggests that teachers who create meaningful assessments, offer corrective action, and give students multiple opportunities to demonstrate success can improve their instruction and increase student learning (Guskey, 2001). McMillan, Myran, & Workman (2002) argued that teachers' understanding of assessment matters are inadequate; although there is common use of standardized math and reading tests, there is limited knowledge of how the assessments are scored, what inferences can be drawn, and even less knowledge of issues involving reliability and validity. The relationship between these two phenomena was the primary focus of this mixed-method study.

Using a holistic model of assessment literacy specifically designed for this study, a mixed-method research designed was employed, using a teacher questionnaire and lesson plan analysis data in order to collect complimentary information to “offset the weaknesses inherent with one method” (Tashakari & Teddlie, 2003, p. 229). Chapter 5 focuses on the conclusions and inferences drawn from the findings presented in the previous chapter, and the teaching, leadership and research implications of the data analyzed.

Statement of the Problem

Debate continues about the high stakes of state and standardized testing and the essential need to provide instruction to students and respond appropriately to their individual and

collective needs after assessing the progress of learners. Assessment defines what students regard as important, how they spend their time and how they come to see themselves as students and then as graduates (Brown, 1997). If you want to change student learning, then change the methods of assessment. Specifically how teachers interact with students and how they provide students the opportunity to interact with the content, engage in critical thinking, and reflect on their learning is paramount to the learning process. Similarly, how students are assessed and provided feedback are the most important aspects of teaching and will influence them for the rest of their lives (Race, Brown, & Smith, 2005). However, the relationship between what is known to be grounded in research and what is practiced by teachers, is not as simple as one would expect, and teachers have a difficult time converting what they know into consistent and effective classroom practice (Dixon & Haigh, 2009). Even faced with research findings that demonstrate the benefits of effective assessment strategies, they are not widely adopted in classrooms (Black & Wiliam, 1998, Hattie, 2003; Duncan & Noonan, 2007; DeLuca & Bellara, 2013). This study adds to the body of research on assessment literacy and expands the definition by viewing teachers' assessment literacy through lens of a holistic conceptual model, which emphasizes the learner as an active member of the assessment process and the central factor in all decisions made as they relate to planning and delivering instruction.

Methodology

This exploratory, mixed-method design, using quantitative and qualitative approaches involved a two-month period of gathering and analyzing data guided by the following research questions:

- 1) To what extent are selected grade 3-12 teachers' perceived understandings of assessment aligned with the holistic model?

- 2) What assessment practices do selected grade 3-12 teachers report using in their classrooms?
- a. How do grade 3-12 teachers use assessment to inform instruction?
 - b. How do grade 3-12 teachers select and design assessments?
 - c. What assessment practices do grade 3-12 teachers report regularly using in their classrooms?

The research questions were answered by the themes derived from the teacher questionnaire and lesson plan analysis data delineated in Chapter 4 and arranged in Table 12 to illustrate their alignment.

Table 12: *Research Questions Matched with Themes*

Research Question	Theme
To what extent are selected grade 3-12 teachers' perceived understandings of assessment aligned with the holistic model?	<p>Assessment is not viewed as an integral part of instruction</p> <p>There is a gap between what is understood and what is practiced</p>
What assessment practices do selected grade 3-12 teachers report using in their classrooms?	<p>There is a surface level understanding and use of assessment practices</p> <p>Assessment is not viewed as an integral part of instruction</p> <p>There is a gap between what is understood and what is practiced</p>

Discussion of the Themes

Theme 1: There is a Surface Level Understanding and Use of Assessment Practices

Black and Wiliam (1998) described assessment as activities teachers and students use to assess themselves and that supply information to be used as feedback to promote increased levels of teaching and learning. The Holistic Model used in this study considers student involvement, communicating purpose and results, and using the results formatively to be essential to the deep level of understanding that will result in maximizing the effect of assessment for learning. Hattie (2015) states that educators should concentrate on utilizing assessment as a method of feedback to help teachers maximize their effectiveness and that teachers should guide students in interpreting their own success from the results and feedback.

Data from the teachers' questionnaire indicated that they had an understanding of student involvement that reached a medium level of assessment literacy. The results of this study show that teachers consistently lesson planned using the traditional models of instruction and assessment. While they used key words and phrases to indicate an awareness of assessment being an important component of a lesson, there was little to no written evidence that students were explicitly involved in any facet of the assessment process. Moreover, there were rarely any noted actionable reactions to daily independent learning activities that appeared to serve a formative purpose.

This indicates that while the teachers are familiar with the language of assessment and the have knowledge of some key principles, they lack the depth of comprehension that will allow them to deliberately plan routines, procedures, and practices that include activities to address the important role of the student as an active participant in their learning. The difference between surface level and deeper understanding is often presented as the difference between assessment

of learning and assessment for learning. Stiggins and colleagues (Stiggins, Arter, Chappuis, & Chappuis, 2006; Chappuis, Stiggins, Arter, Chappuis, 2011) outlined the attributes of a student-centered assessment for learning model that guided the holistic model framing this study and provided a perspective that considers the needs and knowledge of the learner above the sterile act of assessment. Among several other attributes, these scholars assert that assessment should support ongoing student growth, include students in being informed and reflective of their performance, and imbedded in learning. Furthermore, they suggest that teachers should use the assessment results to adjust instruction, offer descriptive and specific feedback directly to students, and ensure that the students have a clear understanding of the purpose for the assessment. Evidence of these critical attributes were lacking in the data collected in this study. While teachers were able to identify and document some of the language of the holistic model, there was a consistent lack of evidence showing the deeper knowledge of assessment required to effectively and consistently use assessment strategies in the classroom.

Implications of Theme 1: Surface Level Understanding and Use of Assessment Practices

Teacher competency. Assessment literacy is related directly to teacher competencies and how they are put into practice. The degree to which teachers are assessment literate determine their understanding of the standards and their ability to communicate learning intentions to students. The time spent in school and in a particular classroom is finite, and to maximize that time, teachers need to stay focused on the things that matter most. Stiggins (2007) argues that teachers and students should be partners in the assessment process and the students' role is to conceptualize what models of success look like in practice and use specific feedback from the teacher to move closer to the success models from their current level of proficiency. If teachers do not have the depth and the breadth of knowledge necessary to use the time they have

with students in a manner that facilitates this partnership, teaching for mastery will continue to be a challenge and the opportunities by both teachers and students to use failures as sources of information to foster improvement will continue to be sporadic.

Complimenting summative assessment. Stiggins (2004) advocated for a classroom assessment culture that complements high stakes testing. For this to occur, teachers must have a deep knowledge of the curriculum standards and possess the competencies necessary to identify and design quality assessments and be able to link them to more summative assessments in ways that build confidence and promotes responsibility in students. The purpose of assessment is diminished when the process is not maximized and when teachers do not have the requisite knowledge for planning, designing and employing the strategies we know are most effective. Thus, they will not be able to foster a classroom culture that supports students to achieve at the highest levels possible.

Theme 2: Assessment is Not Viewed or Practiced as an Integral Component of Instruction

Pointing out that when classroom assessment becomes an integral part of instruction and drives what teachers do to improve achievement, Guskey (2003) emphasizes that students and teachers will benefit tremendously. He also argues that when assessment is used at its best, teachers are informed on which parts of a lesson went well and what needs improving. The findings presented in Chapter 4 indicate that assessment is, by and large, used as a measure of what students have learned rather than a tool to inform the process of teaching and learning. Despite the lesson plan templates having sections that specifically addressed re-teaching lessons, differentiated instruction, and evaluating the lesson, the written plans supported the same low levels of assessment literacy indicated by the responses on the teacher questionnaire that included statements such as; *I wait to assess my student until after I am finished all of my*

instruction, I grade assignments for completion rather than accuracy, and I make sure I follow the curriculum pacing regardless of whether all students are mastering the content. The most detailed portion of the lesson plans described direct instruction, guided-instruction, and modeling, but did not include any notations about what adjustments were or would be made based on how students progressed. Furthermore, there were no pre-assessments, rubrics, or project-based assessments mentioned in any of the plans, and when the sequence of daily lessons delivered were examined for an entire week for each teacher, there was no instance when the previous day's assignments or activities were referenced as a source of information for the current day's lesson.

Dean, Hubbell, Pitler, and Stone (2012) suggest that teachers and students should clearly be aware of the objectives for each lesson based on information about what students know, what they should know, and what steps are needed to close the gap between the two. The results of the study indicate that teachers are not consistently and systematically using daily activities as a means for assessing student progress, to inform the next stage in the instruction, or to provide students with the feedback they need to effectively measure their own levels of success. This supports the low assessment literacy level indicated by the response to the statements, *I use pre-assessments to determine where I need to begin my instruction on a new topic* and *I make sure I follow the curriculum pacing regardless of whether all students are mastering the content.*

Assessment cannot be considered an integral component of instruction unless it is part of the planning, weaves its way through the lesson, and results in student reflection and adjustments in the course of the instruction.

Implications of Theme 2: Assessment is Not Viewed or Practiced as an Integral Component of Instruction

Decision making. Without the required level of assessment literacy, the decisions about what assessments to give and when to give them will not be grounded in student-centered motives. The evidence gathered in this study indicates the use of counter examples of effective assessment practices, identified by Black and Wiliam (2001), having a negative impact on learning that included; under-emphasizing feedback, promoting performance over mastery, and over-emphasizing grading. Assessment needs to take place day-by-day and minute-by-minute, and is most effective when the results drive the next phase of the instruction. As a result of separating instruction and assessment, a valuable tool for learning is lost and the role of the student in their own learning is marginalized. Moreover, the ability of teachers to engage in critical assessment decisions theorized by Popham (2009), emphasizing the important role teachers play in engaging in critical assessment decisions involving knowing what kinds of assessment tools are best for different purposes and determining the validity of the results is compromised.

Transforming the classroom. The strategic use of assessments has the power to transform classrooms to help create a collaborative relationship between students and teachers (Sheppard, 2000). Transformation will not be realized until holistic assessment becomes the way instruction is planned and delivered. However, the data suggested that assessment was primarily viewed as a means to assign grades or determine what students had gained after instruction. This practice creates a classroom culture that undermines the idea that mistakes are a part of the learning process and has the potential to negatively impact students' motivation (Curwin, 2014).

Moreover, grades do not provide accurate data for teachers to make informed decisions about how to direct their instruction. This is supported by Jung and Guskey (2011), who argue that grades do not serve as dependable indicators of knowledge students have gained or as reliable predictors of how well they will do on standardized assessments. Assignments and activities the students are involved in during the day should serve as opportunities for teachers to communicate purpose to the students, and use the results to provide feedback that promotes reflection and self-assessment. Providing student feedback is essential to the learning loop, and if teachers are not deliberately using classroom activities as an opportunity to provide feedback and to monitor progress, they are not maximizing instructional time. This study indicated that feedback was not frequently and purposely provided to students, aligning with the findings of Riggan and Olah (2011).

Theme 3: There is a Gap With Regard to What is Understood About Assessment and What is Practiced.

Pfeffer & Sutton (2000) suggest the gap between what people know and what they do is more significant than the gap between ignorance and knowledge. This seems to ring true based on the evidence gathered in this study. Assessment is critical to promoting the skills that develop life-long learning (Stiggins, 2001), and with the right assessment practices, student learning will ultimately improve. In addition, Popham (2003) theorizes that when teachers have a strong grasp of the connection between teaching and assessment, the result is an increase in the effectiveness of instruction. Reeves (2009a; 2009b), suggests the difference between what people know and what they actually do can be referred to as “the implementation gap.” Theme 1 suggests that teachers have only a surface knowledge of the holistic concepts of assessments and lack the deeper understanding it takes to maximize the use of the most effective assessment practices.

Theme 3 advances that because of this surface knowledge, the more significant gap that exists with the teachers in this study involves the gap between what they know and what they do, rather than what they know and do not know. Even with the possible limitation of answers on the questionnaire being influenced by social desirability, it was clear that the respondents had enough knowledge to recognize what they thought would be the best answer. For example, the statement, *My assessments are designed to match the individual needs of students* garnered responses that indicated a medium level of assessment literacy and recognition that students are an important part of the assessment equation. In addition, even though the notations appeared to be generic in nature, the lesson plan analysis also indicated that teachers had enough knowledge to document times when they would reteach based on student mastery of the learning activity. The gap between knowing and doing was apparent given the lack of evidence that these practices were actually data-driven or that they were actually employed during the course of instruction.

Implications of Theme 3: There is a Gap With Regard to What is Understood About Assessment and What is Practiced

Teacher behaviors and beliefs. Teachers' choice of pedagogy are guided by their self confidence and their educational beliefs (Brown, 2004) and Rutherford (2013) argues there needs to be an alignment of the curriculum, instruction and assessment if student achievement is to be maximized. In order for this alignment to be present in the classroom, teachers need to implement the knowledge of explicit instruction, the use of high yield strategies, the practice of assessment mapping, and tiered support systems they have gained through professional development and published literature. While school leaders emphasize that teachers use effective planning through collaboration and develop well thought out lessons, there continue to be little to no attention to the quality of classroom assessment (Stiggins, 2002). More

professional development resources should be provided to monitor and nurture the desired assessment practices beyond the training stage, so that behaviors grow to mirror the desired assessment strategy and the teachers' pedagogical beliefs change. Reeves (2007) puts forward that reducing the gap between knowing and doing requires "quick wins" that provide teachers with feedback that reinforces the practices they are expected to implement (much like the formative assessment practices discussed in this study).

Creating a climate for change. Thompson & Wiliam (2007) reasoned, teachers and instructional leaders need to be able to know and understand what strategies have the greatest impact on learning as well as those that have the smallest, in order to adjust school wide plans for continuous improvement. When the practices we know to be effective are not being implemented, it is difficult to make informed decisions about what works and what needs to be improved. Pfeffer & Sutton (2000) argue the knowing-doing gap is lessened when there is little fear that using the knowledge will result in punishment. Teachers must feel confident in their abilities and believe that they have the freedom to operate outside of their comfort zone, make mistakes, and make adjustments as they become proficient with new practices. School leaders must pay close attention to ensure that mixed messages are not being delivered regarding assessment and accountability (Earl 2003), and the school assessment climate must promote and celebrate learning experiences for students are based on the mastery goal orientations recognize growth in both teachers and students. By removing the anxieties teachers may experience when considering putting new knowledge into practice, the implementation gap will be lessened and schools will be able to do a better job discerning those practices that result in sustainable positive impact on student achievement.

Implications for Leadership

There are potential implications that highlight professional development and the implementation of school improvement efforts, including the utilization of formative assessment practices, that require more than simply training for a new set of skills. Professional development resources are used to provide training to teachers and administrators to ensure that instructional time is spent delivering the most effective, research-based strategies possible. When the knowledge gained from professional development is not put into practice with fidelity, not only are the resources being used inefficiently, the impact on student achievement is not fully realized. The holistic model used in this study represents a mastery goal orientation, urging its use is not enough to utilize specific strategies, but that an overarching model that guides the effective use of interdependent strategies within the model must be put in place. School leaders monitoring professional development activities should look closely at how training is translating into consistent practice.

Plan Beyond the Training

When teachers feel a high level of efficacy and experience success when implementing effective strategies, achievement increases (Riggan and Olah, 2001). Before leaders decide on a course for professional development on assessment strategies, it will be important to establish the reason such training is desired or necessary (Pintrich, Marx, Boyle, 1993). After the initial training, school leaders should plan a systematic process of monitoring and support to ensure that assessment strategies are being implemented. This support should include collaborative teaming opportunities to plan and reflect on how assessment strategies are used during instruction and non-directive feedback that provides teachers to ask questions and get clarification on the expectations they are required to meet.

Evaluating Lessons

School leaders emphasize that teachers use effective planning through collaboration and develop well thought out lesson plans. Furthermore, they review plans and observe classrooms as a measure of how instruction is being delivered to students. The lesson plans evaluated in this study, did not appear to serve as an accurate predictor of what actually took place in class. The generic nature of many entries indicated that the teachers used their knowledge of assessment vocabulary and concepts to record activities they knew would be viewed as favorable when the plans were reviewed. School leaders should spend the majority of their day, observing instruction and providing constructive feedback to teachers and lesson plans should be compared to classroom activities to monitor alignment and to see how, if at all, the planning process can be enhanced to improve instruction. In addition, when reviewing lesson plans in isolation, school leaders should ask teachers to provide further explanation for statements or activities in the plans that involve assessment and lack specificity or depth as it relates to their instructional value.

Implications for Teacher Practice

The findings from this study suggests that teachers have knowledge of many of the research-based strategies that positively impact student achievement, yet they lack the deeper understanding of the role students should play in the instructional decisions they make. Schools will continue to be evaluated and measured based on the student achievement data produced from year to year based on summative assessments, therefore it will be incumbent on schools and school leaders to identify and implement those instructional strategies that yield the highest gains. Understanding that the learners must be the basis for all decisions will lead to data-driven decisions and necessitate assessment becoming an integral component of instruction. The teachers that participated in the study demonstrated that they have accrued a significant amount of knowledge about assessment, but have not consistently put their knowledge into practice.

Knowledge is a noun that can be defined as having information in the mind and the word.

Understand is a verb that can be defined as knowing the meaning of something.

Transforming knowledge into the deeper level of understanding will increase the likelihood of closing the implementation gap and result in the Holistic Model of Assessment organically evolving as a common practice in classrooms. Robinson, Myran, Strauss and Reed (2014) suggest there is an on-going need to understand the gap between theory and practice and factors that serve as obstacles to participation at the school-level. The correlations indicating a negative impact from the influence of other professional development pursuits with teachers' use and understanding suggest suboptimal situations when there are numerous competing demands on teachers' time. Fostering more opportunities for educators to have continued discourse in collaborative teams across content areas and grade levels, about new instructional and assessment practices, and compelling them to see the impact of such practices on student learning should broaden the impact of reform efforts.

Focusing on a more deliberate and structured learning team approach linked to school wide staff development, as well as, data on student learning would likely lead to even more effective classroom and school-level changes (Stiggins et al., 2006; Thomson and Wiliam, 2007). Because formative assessment ideas and practices are slow to be fully integrated into teachers' day-to-day classroom practices, a scaffolded training process that is driven by student performance data is recommended. This will provide a sound model for helping teachers make the major changes in their roles necessary and support a fundamental reorientation to the teacher–student learning relationship that will help facilitate embedding the desired practices into the daily work of teachers.

Limitations and Implications for Research

Limitations

There were a number of limitations for this study. First is the limitation of self-report or response bias. One of the challenges here is that research participants may respond with socially desirable answers rather than their authentic beliefs or responses (Fisher, 2000). Similarly, the act of asking a question can prompt reflection and subsequent response to an issue or concept that the research participant might not have ordinarily thought about (Leedy & Ormrod, 2016). In addition is an instrumentation bias where the particular ways questions are phrased and the inclusion and exclusion of certain questions may shape or influence the outcome in one direction or another (Leedy & Ormrod, 2016). Lastly volunteer bias, the idea that those who volunteer to participate in research are different in some ways from those who choose to not participate, may shape the outcomes of the research (Salkind, 2006). If there is a volunteer bias at work then, in effect, the researcher has sampled only a subset of the population and generalizations can't be made about that population.

Teacher Efficacy and Autonomy

Previous research suggests that a limited amount of value is placed on the learner's role in assessment, virtually ignoring the impact of the assessment on students and the importance of their voice in evaluating the results (Stiggins, 2007). The results of this study support these findings and based on the data collected, suggest that there is a pervasive ideological perspective that isolates assessment from instruction, and a gap between what teachers know about assessment and what practices are implemented in classrooms. Additional research should be done to examine teachers' feelings of efficacy as it relates to the autonomy they have in employing effective assessment practices during their daily interactions with students. Darling-

Hammond, (1993) found that teachers were aware of the state and local pressures inherent in the accountability systems and that these pressures, along with time constraints, overwhelmed the educational goals that emphasized student learning beyond achievement measured strictly by standardized testing. Kohn (2000) argues that teachers feel compelled and even obligated to choose methods to improve test scores over those that enhance and promote authentic learning, and this affects the content and format of instruction. The noted pressures of external accountability have the possibility of conflicting with the holistic model of assessment. Therefore, it will be fruitful to examine to what degree external pressures influence teachers' instructional practices.

Conceptual Change and Assessment Practices

School leaders not only look to identify areas of strengths and weaknesses within instructional practices, but also ways to enhance the strengths and improve the weaknesses by determining why both exist. The apparent weaknesses delineated in this study can possibly be related to Conceptual Change Theory that brings to light four conditions that must be present in order for the change to occur which are; a dissatisfaction with current conceptions, the intelligibility of the new concept, the plausibility of implementing the new concept, and impact or success of the conceptual change (Pintrich, Marx, Boyle, 1993). Conceptual change and conceptual ecology theories charge that it is not enough to simply present new information to teachers. Careful consideration must be given to all of the environmental conditions that may shape or constrain teachers' willingness and ability to accept the new concept and commit to embracing new practices and ideologies. While there is research that addresses various aspects of teachers implementation of assessment practices (e.g. Bol, Strange, 1996; Bol, Nunnery, Stephenson, & Mogge, 2000; Gimbert, Bol, & Wallace, 2007;), a literature search revealed a

scarcity of published research that specifically explores the issue of conceptual change as it relates specifically to teachers' assessment practices. It could be a rewarding area for future research to explore what factors impede change in instructional practices. Additionally, as it relates to the Conceptual Change Theory, a robust study in the area of classroom practices using the Holistic Model of Assessment should also be conducted to examine to what degree, if any, grade level and the subject taught impacts how knowledge translates into instructional practice.

Conclusion

To answer the first research question regarding teachers' perceived understandings of assessment aligned with the holistic model, the collective findings suggest that teachers have a process-centered conception of assessment that is misaligned with the holistic model, which considers the learner as the center of all assessment efforts. Most notably lacking is the understanding about the importance of communication and student involvement.

The second research question of, assessment practices teachers report using in their classrooms addresses how teachers design and use assessments and what practices they report using in their classrooms. The findings of the research suggest that the grade 3 through 12 teachers that participated in the study generally used assignments and assessments that were standardized or created at the district level without adaptation to address the needs of their students. In addition, there was little to no evidence in the data collected that indicated activities or assessments were used to inform instructional decisions, leading to the conclusion that assessing students was not viewed as an integral part of instruction. The data taken together is suggestive of an implementation gap as it relates to knowledge and practice. The teachers reported employing assessment practices that were either unfounded by the lesson plan review or unsubstantiated based on reverse coding and follow up statements on the questionnaire.

Beyond the research questions and findings of this study resides the question schools have nationwide—How do educators help students realize their full potential and stay committed to their core purpose of teaching and learning? Teachers are commonly believed to be the most important factor in student learning and are central to any improvement effort a school makes. Teacher driven factors such as providing student feedback, ensuring high instructional quality, and teaching for mastery are among the influences on learning that have the highest effect sizes (Hattie, 2012). Finding a way to grow the capacity in teachers to implement these practices will go a long way in improving student achievement across the board and help schools establish and maintain a data-driven approach to prioritizing instructional needs. In a practical sense, educational institutions must find a way to identify and lessen the gap between knowledge and practice suggested by the findings of this study. The school day and the school year are both constrained by time, creating a sense of urgency to provide students with the material they need to meet curriculum standards. Hence, it is of the utmost importance that teachers and students are spending their time on those practices known to have the greatest positive impact on achievement.

REFERENCES

- Alade, O. M., & Omoruyi, I. V. (2014). Table of specification and its relevance in educational development assessment. *European Journal of Educational and Development Psychology, 2*(1), 1-17.
- American Federation of Teachers, National Council on Measurement in Education, and National Education Association. Standards of teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice*, vol. 9, no. 4, 1990, pp. 30-32.
- American Psychological Association. (1995). *Learner-centered psychological principles: A framework for school redesign and reform*. ERIC Clearinghouse.
- Andrade, H., & Cizek, G. J. (Eds.). (2010). *Handbook of formative assessment*. Routledge.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American psychologist, 37*(2), 122.
- Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of educational research, 61*(2), 213-238.
- Black, P., & Wiliam, D. (1998). Inside the black box-raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139.
- Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 92*(1), 81-90.
- Black, P. & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. In M. Wilson (Ed.), *Towards Coherence Between Classroom Assessment and Accountability: The 103rd Yearbook of the National Society for the Study of Education, Part II*, (20-50). Chicago, IL: The University of Chicago Press.

- Bloom, B. S., Hastings, J. T., & Madaus, G. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill
- Bol, L., (2004). Teachers' assessment practices in a high-stakes testing environment. *Teacher Education and Practice* 17(2), 162-181.
- Bol, L., & Strage, A. (1996). The contradiction between teachers' instructional goals and their assessment practices in high school biology courses. *Science Education*. 80(2), 145-163.
- Bol, L., Nunnery, J., Stephenson, P., & Mogge, K. (2000). Changes in teachers' assessment practices in the new American schools restructuring models. *Teaching and Change*, 7(2), 127-46.
- Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies In Continuing Education*, 22(2), 151-167.
- Bowen, G. A. (2010). From qualitative dissertation to quality articles: Seven lessons learned. *The Qualitative Report*, 15(4), 864-879. Retrieved from <http://www.nova.edu/ssw/QR/QR15-4/bowen.pdf>
- Brookhart, S. M. (2001, March 02). The "Standards" and Classroom Assessment Research. Retrieved from <https://eric.ed.gov/?id=ED451189>
- Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record*, 106(3), 429-458.
- Brown, A. L. (1994). The advancement of learning. *Educational Researcher*, 23(8), 4-12.
- Brown, G. (1997). *Assessing student learning in higher education*. London: Routledge.
- Brown, G. (2004). Teachers' conceptions of assessment: implications for policy and professional development. *Assessment in Education*, 11(3), 301-314.

- Brown, G., & Harris, L. (2009). Unintended consequences of using tests to improve learning: How improvement-oriented resources heighten conceptions of assessment as school accountability. *Journal of Multi-Disciplinary Evaluation*, 6(12), 68-91.
- Burke, G, Wang, Y. (2010). Methods and uses of classroom assessments employed in teaching grades three through five in five school districts in the Mississippi delta. *Education*. 130(4). 657-65.
- Card, N. A. (2015). *Applied meta-analysis for social science research*. Guilford Publications.
- Chappuis, J., Stiggins, R., Chappuis, S., & Arter, J. (2011) *Classroom assessment for student learning: Doing it right - using it well (2nd ed.)*. Pearson
- Chappuis, S. & Stiggins, R. (2008). Finding balance: Assessment in the middle school classroom. *Middle Ground*, 12(2), 12-15.
- Chappuis, S., Chappuis, J. & Stiggins, R. (2009). Supporting teacher learning teams. *Educational Leadership*, 66(5), 56-60.
- Cizek, G., and O. (2005). *Further investigation of teacher's assessment practices*. Annual Meeting of the American Educational Research Association, (pp. 1-43). San Francisco.
- Clark, M., & Smitherman, D. (2013) Communicating achievement test results with parents, *CSE* 16(3), 38-40.
- Corbin, J.M., & Strauss, A.L. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications, Inc.
- Creswell, J. (2009). *Research design: Qualitative, quantitative, and mixed-methods approaches* (3rd ed.) Los Angeles, CA: Sage.
- Creswell, J.W., Plano Clark V.L. (2007) *Designing and conducting mixed methods research*. (1sted.). Thousand Oaks, CA: Sage.

- Curwin, Richard, L. (2014) Can assessments motivate? *Educational Leadership*, 72(1), 38-40.
- Darling-Hammond, L., & Wise, A.E. (1985). Beyond standardization: State standards and school improvement. *The Elementary School Journal*, 8(5), 315-336.
- Darling-Hammond, L. (1993). Setting standards for students: the case for authentic assessment. *NASSP Bulletin*, 77(556), 18-26.
- Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work*. New York: Teachers College Press.
- Dean, C., Hubbell, E., Pitler, H., & Stone, B. (2012). *Classroom instruction that works* (2nd ed.). Alexandria, VA: ASCD.
- DeLuca, C., & Bellara, A. (2013). The current state of assessment education: Aligning policy, standards, and teacher education curriculum. *Journal of Teacher Education*, 64(4), 356-372.
- Dempster, F. N. (1991). Synthesis of research on reviews and tests. *Educational Leadership* 48(7), 71-76.
- Dixon, H., & Haigh, M. (2009). Changing mathematics teachers' conceptions of assessment and feedback. *Teacher Development*, 13(2), 173-186.
- DuFour, R. (2004) What is a Professional Learning Community? *Educational Leadership*, 61(8), 6-11.
- Duncan, R. C. & Noonan, B. (2007) Factors affecting teachers' grading and assessment practices. *The Alberta Journal of Educational Research*, 53(1), 1-21.
- Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation*, 14(7), 1-11.

- Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin.
- Elshout-Mohr, M. (1994). Feedback in self-instruction. *European Education*, **26**(2), 58-73.
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional children*, **53**(3), 199-208.
- Fisher, R. J. (2000). The future of social-desirability bias research in marketing. *Psychology and Marketing*, **17**(2), 73-77.
- Gareis, C., & Grant, L. (2008). *Teacher-made assessments: How to connect curriculum, instruction, and student learning*. Larchmont, NY: Eye On Education.
- Gareis, C., & Grant, L. (2013). *Teacher-made assessments: how to connect curriculum, instruction, and student learning*. Routledge.
- Garrison, C., & Ehringhaus, M. (2007). Formative and summative assessments in the classroom. Retrieved from http://ccti.colfinder.org/sites/default/files/formative_and_summative_assessment_in_the_classroom.pdf
- Gaulden, S. (2010). Classroom assessment techniques. Essex County College. Retrieved from http://sloat.essex.edu/sloat/delete/contentforthewebsite/classroomassessment_techniques.pdf
- Geva, Blenkin, & Kelly. (1992). *Assessment in early childhood education*. Sage.
- Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and teaching in higher education*, (1), 3-31.

- Gimbert, B., Bol, L., & Wallace, D. (2007). The influence of teacher preparation on student achievement and the application of national standards by teachers of mathematics in urban secondary schools. *Education and Urban Society*, 40(1), 91-117.
- Graue, M. E. (1993). Integrating theory and practice through instructional assessment. *Educational Assessment*, 1(4), 283-309.
- Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher*, 15(5), 5-12.
- Guskey, T. R. (2001). Helping standards make the grade. *Educational Leadership*, 59(1), 20-27.
- Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching*, 8(3), 381-391.
- Guskey, T. R. (2003) How Classroom Assessments Improve Learning. *Educational Leadership*. 60(5), 6-11.
- Guskey, T. R., & Huberman, M. (1995). *Professional development in education: New paradigms and practices*. Teachers College Press, 1234 Amsterdam Avenue, New York, NY 10027 (paperback: ISBN-0-8077-3425-X; clothbound: ISBN-0-8077-3426-8).
- Harrison, MacGibbon, and Morton (2001). Regimes of trustworthiness in qualitative Research: The rigors of reciprocity. *Qualitative Inquiry*, 7(3), 323. doi: 10.1177/107780040100700305
- Hattie, J. (1999). *Influences on student learning* (Inaugural professorial address, University of Auckland, New Zealand). Retrieved from <http://www.arts.auckland.ac.nz/staff/index.cfm?P=8650>
- Hattie, J. (2003) *Teachers make a difference: what is the research evidence?* (p.4). Camberwell: Australian Council for Educational Research.

- Hattie, J. (2009). The black box of tertiary assessment: An impending revolution. *Tertiary assessment & higher education student outcomes: Policy, practice & research*, 259-275.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. London: Routledge.
- Hattie, J. (2013). Calibration and confidence: where to next? *Learning And Instruction*, (24) 62-66.
- Hattie, J. (2015). We aren't using assessments correctly: There's a distinction between formative and summative assessments. *Education Week* 35(10) 23, 28.
- Hattie, J. (2016). Know thy impact. *On Formative Assessment: Readings from Educational Leadership (EL Essentials)*, 36.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Herman, J., Osmundson, E., Dai, Y., Ringstaff, C., & Timms, M. (2015). Investigating the dynamics of formative assessment: Relationships between teacher knowledge, assessment practice and learning. *Assessment in Education: Principles, Policy & Practice*, 22(3), 344-367.
- Herrington, J., Reeves, T. C., & Oliver, R. (2014). Authentic learning environments. In *Handbook Of Research On Educational Communications And Technology* (pp. 401-412). Springer New York.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53-64. doi:10.1080/03075070120099368

- Hodgen, J., & Marshall, B. (2005). Assessment for learning in English and mathematics: a comparison. *Curriculum Journal, 16*(2), 153-176. doi:10.1080/09585170500135954
- Hunt, E., & Pellegrino, J. W. (2002). Issues, examples, and challenges in formative assessment. *New Directions For Teaching & Learning, 89*, 73.
- Johnson, B., Onwuegbuzie, A. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7): 14-26.
- Jung, L. A. & Guskey, T. R. (2011). Fair and accurate grading for exceptional learners. *Principal Leadership*. November. 35-37.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 779*(2), 254
- Kohn, A. (2000). Burnt at the high stakes. *Journal of Teacher Education, 51*(4). 315-327.
- Koretz, D. M. (2002). Limitations in the use of achievement tests as measures of educators' productivity. *Journal of Human Resources, 752-777*.
- La Marca, P. M., & ERIC Clearinghouse on Assessment and Evaluation, C. D. (2001). *Alignment of Standards and Assessments as an Accountability Criterion*. ERIC Digest.
- Labuschagne, A. (2003). Qualitative research – Airy fairy or fundamental? *The Qualitative Report, 8*(1).
- Leedy, P. D., & Ormrod, J. E. (2016). *Practical Research*. Pearson.
- Linn, R. (1990). Essential of student assessment: From accountability to instructional aid. *The Teachers College Record, 91*(3), 422-436.
- Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.

- Linn, R., (2000) Assessments and accountability. *Educational Researcher*, February 2000, p. 14.
- Marshall, M. N., (1996) Sampling for qualitative research. *Family Practice* 13(6), 522-525.
- Marzano, R. J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: ASCD.
- Marzano, R. J. (2004). Using data: Two wrongs and a right. *Educational Leadership*, 60(5), 56-60.
- Marzano, R. J., Heflebower, T., & Marzano Research, L. (2012). *Teaching & Assessing 21st Century Skills. The Classroom Strategies Series*. Marzano Research Laboratory.
- McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education: Principles, Policy & Practice*, 10(2), 209-220.
- McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research*, 95(4), 203–213.
- McTighe, J., & O'Connor, K. (2005). Seven practices for effective learning. (Coverstory). *Educational Leadership*, 63(3), 10-17.
- Merriam, S. B. (2009). *Qualitative case study research qualitative research: A guide to design and implementation* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Morgan, D. L. (1988) *Focus groups as qualitative research*. Newbery Park, CA: Sage.
- Morgan, D. L. (2002). *Focus group interviewing*. In J.F. Gubrium & J.A. Holstein, *Handbook of interview research: Context and method* (141-160). Thousand Oaks, CA: Sage.
- National Research Council. (2001). *Classroom assessment and the national science education standards*. National Academies Press.

- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22(2), 155-175.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies In Higher Education*, 31(2), 199-218.
- Nicolaidis, C. S., & Michalopoulos, G. (2004, April). Education, industry and the knowing-doing gap: A knowledge management perspective of business education. *Industry & Higher Education*, 18(2), 101-110.
- Norcini, J., & Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Medical Teacher*, 29(9/10), 855-871. doi:10.1080/01421590701775453
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Thousand Oaks, CA: Sage.
- Pfeffer, J., & Sutton, R. I. (2000). *The knowing-doing gap: How smart companies turn knowledge into action*. Boston, Mass: Harvard Business School Press.
- Pintrich, P. R., Marx, R. W., & Boyle, R. Beyond cold conceptual change: the role of motivation beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, 63(2) (Summer, 1993), 167-199.
- Pitler, H., & Stone, B. J. (2012). *A handbook for classroom instruction that works*. Association for Supervision and Curriculum.
- Popham, W. J. (2001). Teaching to the test?. *Educational Leadership*, 58(6), 16-21.
- Popham, W. J. (2003). *Test better teach better: The instructional role of assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental?. *Theory Into Practice*, 48(1), 4-11.

- Race, P. Brown, S. and Smith, B. (2005) *500 Tips on assessment*: 2nd edition, London: Routledge.
- Reeves, D. (2007). Leading to change/closing the implementation gap. *Educational Leadership*, 64(6), 85-86.
- Reeves, D. B. (2009a). *Leading change in your school: How to conquer myths, build commitment, my and get results*. ASCD.
- Reeves, D. B. (Ed.). (2009b). *Ahead of the curve: The power of assessment to transform teaching and learning*. Solution Tree Press.
- Riggan, M., & Oláh, L. N. (2011). Locating interim assessments within teachers' assessment practice. *Educational Assessment*, 16(1), 1-14.
- Robinson, J., Myran, S., Strauss, R., & Reed, W. (2014). The impact of an alternative professional development model on teacher practices in formative assessment and student learning. *Teacher Development*, 18(2), 141-162.
- Rutherford, M. (2013) *The artisan teacher: a field guide to skillful teaching*. Weddington, NC: Rutherford Publishing Group, Inc.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment In Education: Principles, Policy & Practice*, 5(1), 77-84.
- Schmidt, W. H., C. C. McKnight, R. T. Houang, H. C. Wang, D. E. Wiley, L. S. Cogan, and R.G. Wolfe. 2001. *Why schools matter: A cross-national comparison of curriculum and learning*. San Francisco: Jossey-Bass.
- Salkind, N. J. (2006). *Encyclopedia of measurement and statistics*. Sage Publications.

- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Shute, V. J. (2008). Focus on formative feedback. *Review Of Educational Research*, 78(1), 153-189.
- Squires, D. (2012). Curriculum alignment research suggests that alignment can improve student achievement. *Clearing House: A Journal Of Educational Strategies, Issues And Ideas*, 85(4), 129-135.
- Stiggins, R. (1991). Assessment Literacy. *Phi Delta Kappan*, 72(7), 534-39.
- Stiggins, R. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77, 238-246.
- Stiggins, R. (2001). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Stiggins, R. (2002a). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-765.
- Stiggins, R. (2002b). Where is our assessment future and how can we get there from here. *Assessment in educational reform: Both means and ends*. Boston, MA: College Division, Allyn & Bacon, Inc, 18.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86(1), 22-27.
- Stiggins, R. & Chappuis, J, (2005a). Using student-involved classroom assessment to close achievement gaps. *Theory Into Practice*, 44(1), 11-18.
- Stiggins, R. (2005b). Assessment for learning defined. ETS Assessment Training Institute's International Conference: Promoting Sound Assessment in Every Classroom.

- Stiggins, R. (2006). Assessment for learning: A key to motivation and achievement. *Edge: The Latest Information For The Education Practitioner*, 2(2), 1-19.
- Stiggins, R. (2007). Assessment through student eyes. *Educational Leadership*, 64(8), 22-26.
- Stiggins, R. & Chappuis, J. (2008). Enhancing student learning. *District Administrator*, January.
- Stiggins, R. J., & Popham, W. J. (2008). Assessing students' affect related to assessment for learning. *Washington, DC: Council of Chief State School Officers*.
- Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice*, 8(2), 5-14.
- Stiggins, R., & Chappuis, S. (2005). Putting testing in perspective: It's for learning. *Principal Leadership*, 6(2), 16-20.
- Stiggins, R., and Chappuis, J. (2006). What a difference a word makes. *Journal of Staff Development*, 27(1), 10-14.
- Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S (2006). *Classroom assessment for student learning: Doing it right using it well*. Princeton, NJ: Educational Testing Service.
- Stipek, D. J. 1996. "Motivation and instruction". In *Handbook of educational psychology*, Edited by: Calfee, R. C. and Berliner, D. C. 85–113. New York: Macmillan.
- Strage, A., Tyler, A. B., Rohwer, W. D., & Thomas, J. W. (1987). An analytic framework for assessing distinctive course features within and across grade levels. *Contemporary Educational Psychology*, 12(4), 280-302.
- Sun, S., Pan, W., & Wang, L. L. (2011). Rethinking observed power. *Methodology*.
- Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage Publications Inc.

- Thompson, M, and Wiliam, D. 2007. Tight but loose: a conceptual framework for scaling up Reform. *Paper presented at the annual meeting of the American Educational Research Association (AERA)*. Chicago, IL. April 9-13.
- US Department of Education (n.d.) <https://www2.ed.gov/print/programs/titleiparta/index.html>. Retrieved October 15, 2015.
- Vansteenkiste M, Zhou M, Lens W, Soenens B. (2005) Experiences of autonomy and control among Chinese learners: Vitalizing or immobilizing? *Journal of Educational Psychology*. 97(3):468–483. doi: 10.1037/0022-0663.97.3.468.
- Vosniadou, S., International Bureau of Education, G. (Switzerland)., & International Academy of Education, B. (Belgium). (2001). *How Children Learn. Educational Practices Series-7*.
- Wiliam, D. (2011). What is assessment for learning?. *Studies in Educational Evaluation*, 37(1), 3-14.
- Wiliam, D., Lee, C., Harrison, C., & Black. P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education: Principles, Policy & Practice*, 11(1). 49-65.
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualizing teachers' assessment literacies in an era of curriculum and assessment reform. *Australian Educational Researcher*, 40(2), 241-256.
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self perceived assessment skills. *Applied Measurement In Education*, 16(4), 323-342.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1), 3-17.

APPENDIX A

Please rate how frequently you practice each of the following assessment strategies by placing a check mark in the appropriate box

Assessment Design	Never	Seldom	Some of the time	Most of the time
1) I create the assessments I use in class				
2) My assessments mirror the design of standardized test (R)				
3) My assessments cover multiple skills and concepts (R)				
4) My assessments are designed to match the individual needs of students				
5) I design/select assessments differently based on its specific purpose				

Alignment	Never	Seldom	Some of the time	Most of the time
6) I assess students strictly based on the content standards				
7) I include content in my assessments that think students need to know, even though it may not be included in the curriculum standards (R)				
8) I create assessment ideas that are more challenging than the content standards so that students will be ready for easier questions on future assessments (R)				
9) I decide how I will assess a skill/concept prior to beginning my instruction				
10) I use the same questions/problems/prompts on my assessments that I have used during the course of my instruction				

Student Involvement/Formative Use	Never	Seldom	Some of the time	Most of the time
11) I wait to assess my students until after I am finished all of my instruction (R)				
12) I require students to document reflection of their progress				
13) I make sure I follow the curriculum pacing regardless of whether all students are mastering the content (R)				
14) Student set learning goals for skill/concepts being taught				
15) I incorporate interventions into my instructional				

ASSESSMENT LITERACY AND ASSESSMENT PRACTICES109

block based on results from my assessments				
--	--	--	--	--

Setting a Clear Purpose	Never	Seldom	Some of the time	Most of the time
16) I show students examples of work that meets the criteria for success				
17) I use pre-assessments to determine where I need to begin my instruction on a new topic				
18) Rubrics are used to communicate the success criteria of daily assignments and/or projects				
19) I give pop quizzes in order to make sure my students have completed their assignments (R)				
20) I grade assignments for completion (effort) rather than for accuracy (R)				

Communicating Results	Never	Seldom	Some of the time	Most of the time
21) My students revise their work based on teacher or peer feedback				
22) I use reporting options other than grades to communicate the results of assessments				
23) I use rubrics to give specific feedback to students on their progress				
24) To save time, I allow my students to exchange and correct each others' papers (R)				
25) I only discuss the assessment results with other teachers and/or administrators (R)				

Please rate how strongly you agree or disagree with each of the following statements by placing a check mark in the appropriate box

Assessment Design	Strongly Disagree	Disagree	Agree	Strongly Agree
26) Assessments should be lengthy to accurately gauge students' knowledge (R)				
27) It is difficult to create/decide on an assessment until you know what students have learned. (R)				
28) Multiple choice tests can address all levels of cognitive demand				
29) Assessments are should be tightly aligned with the essential skills students will be held accountable for based on the state standards				
30) Creating assessments collaboratively can help improve the quality of the design				

ASSESSMENT LITERACY AND ASSESSMENT PRACTICES110

Alignment	Strongly Disagree	Disagree	Agree	Strongly Agree
31) Assessments should promote student success over rigor in order to foster motivation (R)				
32) Previously taught content should be part of each assessment to make sure students remember (R)				
33) Assessments should be aligned with the written and taught curriculum				
34) Using a blueprint to design an assessment helps to improve its alignment with the curriculum standards				
35) Classroom grades should be consistent with both formative and summative assessment scores				

Student Involvement/Formative Use	Strongly Disagree	Disagree	Agree	Strongly Agree
36) Assessments are most useful for assigning value to what skills students have mastered (R)				
37) An assessment that the majority of students fail can still provide usable data				
38) Students talking to one another encourages new ideas and different perspectives				
39) Rubrics help students understand what is required of them when completing assignments				
40) Student involvement creating assessments can bias the results (R)				

Setting A Clear Purpose	Strongly Disagree	Disagree	Agree	Strongly Agree
41) Informing students what will be included on an assessment makes it difficult to accurately analyze the results (R)				
42) Most assessments should be used to measure student progress and guide future instruction				
43) There should be a comprehensive plan for integrating assessments into the classroom over time				
44) To best monitor students' progress, a grade is should be assigned for each assessment (R)				
45) Assessments can be used for the sole purpose of evaluating and adjusting instruction				

Communicating Results	Strongly Disagree	Disagree	Agree	Strongly Agree
46) Students should frequently receive feedback on completed assignments that DOES NOT include a				

ASSESSMENT LITERACY AND ASSESSMENT PRACTICES111

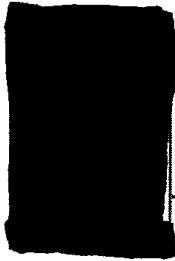
numerical or alphabetical grade				
47) Students do not require frequent feedback o their performance to them succeed (R)				
48) The numerical or alphabetical grade provides enough information for students to accurately reflect on their progress towards learning goals (R)				
49) Students should be told how the results of each assessment will be used.				
50) How and to whom the results of an assessment is reported depends on its pre-determined purpose				

APPENDIX B

Document Analysis Rubric

Assessment Literacy Area	High Literacy	Medium Literacy	Low Literacy	Examples	Counter Examples
Alignment	Classroom activities and assessments are completely aligned to the curriculum standard(s)	Classroom activities and assessments are somewhat aligned to the curriculum standard(s)	Classroom activities and assessments are not aligned to the curriculum standard(s)		
Formative Use	There is clear evidence of using formative assessment to make guide instruction	There is some evidence of using formative assessment to make guide instruction	There is little or no evidence of using formative assessment to make guide instruction		
Student Involvement	There is clear evidence of teacher feedback, student reflection, and student input	There is some evidence of teacher feedback, student reflection, and student input	There is little or no evidence of teacher feedback, student reflection, and student input		
Communicating Results	Plans indicate a method for the timely communication of the results of an assessment	Plans indicate a method for the timely communication of the results of an assessment	Plans do not indicate a method for the timely communication of the results of an assessment		
Communicating Purpose	Plans indicate clear communication of the assessment activity's purpose and how the results will be used.	Plans indicate some level of communication of the assessment activity's purpose and how the results will be used.	Plans do not indicate clear communication of the assessment activity's purpose and how the results will be used.		
Assessment Design	Assessment activities are chosen or created that meet the needs of the students and are appropriate in length	Assessment activities are chosen or created that somewhat meet the needs of the students and are appropriate in length	Assessment activities are chosen or created that do not meet the needs of the students and are appropriate in length		

APPENDIX C



March 25, 2016



Dear Mr. Hudson,

Thank you for your request to conduct research in Hampton City Schools. The Research Committee has granted approval of your proposal on "Examining the Relationship Between Grade 3-12 Teachers' Perceived Assessment Literacy and their Classroom Assessment Practices". We are in receipt of all required documentation, including your IRB approval documentation. You may proceed with the data collection plan as outlined in your application/proposal.

Please be advised that your approval will expire on March 24, 2017. If you are unable to complete your research within this time frame, please contact me. An extension will need to be requested and granted by the Research Committee in order to proceed with the research.

It is our expectation to receive a copy of your findings once the research has been completed. We wish you success in your research.

Sincerely,



Executive Director of Research, Planning and Evaluation



APPENDIX D

Request for Research Study Participation and Consent

March 22, 2016

Greetings:

I am currently working towards my doctor of philosophy in Educational Leadership at Old Dominion. My research topic is a direct result of my interactions teachers, administrators and students in public schools. The title of my dissertation is: **Examining the Relationship Between Grade 3-12 Teachers' Perceived Assessment Literacy and their Classroom Assessment Practices**. This study is designed to examine what relationship, if any, exist between perceived knowledge and practice

The primary purpose of the study is to learn whether there is a relationship between what grade 3-12 teachers' perceive as their level of assessment literacy and their classroom assessment practices. A secondary purpose is to compare the teachers' perceptions of assessment literacy to the holistic model that includes; establishing a clear purpose, aligning instruction and assessment, assessment design, communicating results, and student involvement. The final goal will be to determine what perceptions teachers have about their role in implementing effective assessment practices in the climate of accountability.

It is my hope that you will agree to be a participant in this study. The survey will only take approximately 20 minutes to complete. The first section of the survey will ask for demographic information followed by the survey questions.

In order to complete my study, I am surveying 3-12 math, language arts, science and social studies teachers you will have access to the survey that I would like for you to complete. Lesson plans and assessment materials (*anonymously provided by your administrator*) used by participating schools and teachers **may** also be examined. Completing and submitting the survey will imply your willingness to participate in this study.

This study is highly confidential and information obtained will be kept strictly private. No identifying information linking you or your school name to this study will be included in the data reporting. You may withdraw from the study at any time. If you have questions regarding this study you may contact me at 757-706-4456 and my committee chairperson is Dr. Steven Myran and can be reached at 757-683-6694. Additionally, I would like you to know that I have obtained official approval from the Institutional Review Board (IRB) at Old Dominion to conduct this study.

The deadline for submission of the survey is April 29, 2016. Gentle email reminders will be sent after April 20th if the survey has not been completed in order to remind you of the April 29th deadline.

I want to thank you in advance for taking the opportunity to respond to this survey and if at any time during this process you have questions or concerns, I stand ready to answer your questions.

Sincerely,

Mark Hudson

Mark Hudson
Old Dominion University Doctoral Candidate

MARK EDWARD HUDSON

Hampton City Schools-Phoebus High School Principal
markhudson@hampton.k12.va.us
office: 757-727-1004
mobile: 757-706-4456

Education:

Old Dominion University
Ph. D., Education Leadership (2017)

Old Dominion University
Master of Science in Education Leadership (2001)

Norfolk State University
Teacher Certification Program (1994)

Howard University
Bachelor of Business Administration-Marketing (1990)

Professional Experience:

2015-Present Principal-Aberdeen Elementary School

2010-2015 Principal- Eaton Fundamental Middle School

2008-2010 Principal-Spratley Middle School

2004-2008 Principal-Aberdeen Elementary School

- Organize and evaluate the implementation of the instructional program.
- Responsible for analyzing school data to serve as a guide for planning improvement initiatives.
- Responsible for hiring, observing, evaluating, and coaching staff members
- Providing support for personal and professional development.
- Responsible for creating the school's master schedule to insure that content areas are allotted the required amount of time, staff members have time for planning lessons, and classroom hours are maximized.
- Oversee the school's budget and monitor all aspects of its financial allocations.
- Responsible for interpreting and enforcing the student discipline program based on the division policies to provide and maintain a safe and orderly school environment.
- Communicate with students, parents, staff members, and community agencies to secure resources that supplement our instructional program.
- Responsible planning staff professional development based on school needs

2002-2004 Assistant Principal-Wythe Elementary School

2001-02 Assistant Principal-Armstrong/Cary Elementary School

- Assist with supervising and evaluating classroom teachers, instructional assistants, and clerical personnel.
- Responsible for monitoring and administering the intersession budget and coordinating all intersession programs.
- Communicating with parents and faculty regarding the goals and objectives of the school.
- Assist the principal and staff in improving the instructional program through school improvement plan development and data analysis.
- Administer rules and regulations regarding student conduct based on teacher referral and confer with parents concerning discipline cases.
- Compile, maintain and interpret records of students.
- Monitor attendance and referred students with truancy matters.
- Act as facilitator in meetings concerning individual students such as Child Study, IEP, 504, and Student Intervention Team meetings.

1994-2001 Teacher-Wythe Elementary School

- Present instructional program and utilized teaching methods, which considered the individual needs, interests, abilities, and maturity levels of the students.
- Provide individual student progress reports and evaluations on a regular basis
- Keep requisite records to justify promotions, retentions and special program recommendations
- Meet with parents to advise them concerning student needs and progress.
- Serve on committees and attended staff meetings

Community Involvement:

Youth Football Coach-Hampton Cavaliers Athletic Association (2002-2005)
Youth Basketball Coach-Hampton Cavaliers Athletic Association (2001-2005)
Youth Football Coach-Kappa Cardinals Athletic Association (2006-present)
Youth Soccer Coach-Hampton Soccer Club (2002)
AAU Basketball Coach (1997-2015)

Professional Organization/Committee Membership

ASCD Member

Teacher Evaluation Workteam (2017)

HCS Division Professional Learning Communities Leadership Committee

Division Response to Intervention Workteam

Discipline Work Team (2004)

HAESP-Secretary (2007-08)

Principal's Advisory Committee (2007-08)