



W&M ScholarWorks

Dissertations, Theses, and Masters Projects

Theses, Dissertations, & Master Projects

2019

The Influences on K-2 Teachers' Approaches Towards Assessment and Developmentally Appropriate Practice

Leah K. Shy

William & Mary - School of Education, leahshywilson@gmail.com

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

 Part of the [Elementary Education Commons](#)

Recommended Citation

Shy, Leah K., "The Influences on K-2 Teachers' Approaches Towards Assessment and Developmentally Appropriate Practice" (2019). *Dissertations, Theses, and Masters Projects*. Paper 1563898905.

<http://dx.doi.org/10.25774/w4-kmqm-1571>

This Dissertation is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

**THE INFLUENCES ON K-2 TEACHERS' APPROACHES TOWARDS
ASSESSMENT AND DEVELOPMENTALLY APPROPRIATE PRACTICE**

A Dissertation

Presented to

The Faculty of the School of Education
The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

by

Leah Kathleen Shy

March 2019

**THE INFLUENCES ON K-2 TEACHERS' APPROACHES TOWARDS
ASSESSMENT AND DEVELOPMENTALLY APPROPRIATE PRACTICE**

by

Leah Kathleen Shy

Approved March 20, 2019 by

Christopher R. Gareis, Ed.D.
Chairperson of Doctoral Committee

Leslie W. Grant, Ph.D.
Committee Member

Megan Tschannen-Moran, Ph.D.

Dedication

To my family

Thank you to my sisters for your encouragement and especially the support with the details to get over the finish line. I won't make you call me doctor after this. Mom and Dad, there is no way we can repay you for all your help (so I hope Dad isn't keeping track on a spreadsheet somewhere), other than to share the research that shows how babysitting your grandchildren keeps you young and healthy. Thank you for everything. Joshua, you have inspired and informed so much of this work. May you always be as imaginative and curious as you are now. Ellie, some people say doing a dissertation is like giving birth, but it does not come close to the joy of having birthed you during this time period. May you grow to know you can be anything you want to be. And to Frazer, to live in this country during these years on our budget, you have made incredible sacrifices. You've proven yourself in times *for worse*; I look forward to many more *for better* times in life with you ahead. We did it.

Table of Contents

Acknowledgements.....	xiv
List of Figures.....	xv
Abstract.....	xvi
Chapter 1: Introduction.....	2
Background.....	2
Problem Statement.....	5
Research Questions.....	6
Significance of the Study.....	7
Definition of Terms.....	8
Chapter 2: Review of Literature.....	12
Role of the Literature Review in Grounded Theory.....	12
Start of the Standards Movement.....	13
High-Stakes Accountability.....	15
Impact of Assessment Practices.....	16
Classroom Assessment as a Construct.....	17
Assessment Literacy.....	18
Approaches to Assessment.....	21
Early Childhood.....	23
Developmentally Appropriate Practice.....	23
DAP in the Early Elementary Years.....	25
Developmentally Appropriate Assessment in K-2.....	28
Summary of Literature Review.....	30
Chapter 3: Method.....	31

Research Approach	31
Rationale for Choosing Grounded Theory (Methods)	32
Rationale for Selecting Constructivist Grounded Theory (Methodology)	33
Researcher as Instrument	35
Reflexive Journal	37
Conceptual Framework	38
Participants	39
Target Participants	39
Sample Size Based on Saturation	40
Recruiting Participants	42
Building a Database of Willing Participants	45
Incentives	46
Addressing Challenges	47
Sample Characteristics	49
Data Generation	51
Interviews in Constructivist Grounded Theory	51
Interview Structure	52
Stages of the Interview	54
Logistical Issues	57
Data Analysis	60
Constant Comparative	61
Phases of Coding	61
Initial coding	61
Focused coding	62

Theoretical coding	63
Memoing	64
Theory Generation	66
Theoretical Sampling	66
Theoretical Saturation	67
Model	68
Trustworthiness.....	69
Credibility	69
Transferability	70
Dependability	71
Confirmability	72
Ethical Considerations	73
Delimitations, Limitations, and Assumptions.....	74
Summary of Methods.....	77
Chapter 4: Findings.....	78
Framing Teachers’ Perspectives	80
Limitations	80
Construct of Classroom Assessment.....	81
Developmentally Appropriate Practice.....	82
Different Vocabulary and Terms Used.....	82
Perceptions of the Purposes of Assessment.....	85
Description of Figure	85
Assessment for Compliance.....	86
Assessment for Reporting.....	86

Requirements for reporting.....	87
Concerns with reporting practices	88
Assessment for Instruction.....	89
Assessment for Multiple Purposes.....	91
Seeking clarity of purpose	92
Problematic overlap	93
Summary of Purposes of Assessment.....	94
Theory of Influences.....	94
Overview of Theory	95
The “I” Phase	98
Personal traits.....	99
Prior experience	99
Parenting	100
Teacher preparation	102
The “They” Phase	105
State and district policymakers	108
Central office personnel.....	110
Emphasis on content	111
Experience of administrators	112
Developmentally inappropriate practices	114
Openness to feedback	115
Need for support	117
Principal	118
Positive influence.....	119

Negative influence	120
Impact of turnover.....	120
Trickle-down effect.....	122
Unknown “They”	123
The “We” Phase.....	125
Shared practice.....	127
Collaborative work.....	128
Reluctant teammates	129
Nature of the team.....	131
Grade level teams.....	131
Data teams.....	132
Collegial learning.....	134
Mentors	134
Supporting colleagues.....	135
Remaining “I”	136
Learning orientation.....	137
Closed doors.....	138
Autonomous practice	139
Time	140
Experiential time.....	141
Elapsed time.....	142
Cyclical time	143
Resources	145
Teaching assistants.....	145

Curriculum and pacing guides	146
Supplies.....	146
Technology	147
Requirements to use technology	147
Access to technology	149
Professional development	149
Foundations.....	150
Training.....	150
Kindergarten	151
Summary of influence of resources	152
Examples of Teachers in the Theory	152
Veteran and Early-career Teachers.....	153
Veteran teacher	153
Early-career teacher	154
Two Teachers, One School.....	156
Kindergarten teacher.....	156
Second grade teacher	158
One School District, Two Teachers	159
Lyn’s story	159
Maria’s story	161
Unique Teacher Perspectives.....	164
Influence of parenting.....	165
Dual language immersion teacher.....	166
Teacher with strong philosophical bent	167

Summary of Theory	168
Summary of Findings.....	169
Chapter 5: Discussion	170
Influences on Approaches to Assessment.....	171
The school district.....	172
Separation of assessment and instruction	173
“Formative” assessment.....	175
Definitions of “formative” assessment	176
Effects of misunderstanding formative assessment	178
Grading practices	179
The principal	182
Sensemaking of policy	182
Assessment leadership	183
Leadership of data teams	184
Formal learning.....	185
Teacher preparation	186
Professional development	186
Professional learning communities	187
Data teams.....	188
Community of Practice	190
Summary of influences on approach to assessment.....	191
Influences on Developmentally Appropriate Practice	192
Individual factors	193
Self-confidence	194

“Old school”	195
Time	196
Administration	197
Curriculum	200
Summary of Influences on Developmentally Appropriate Practice	202
Interaction of Approaches to Assessment and DAP	203
Conceptualizations of “Assessment”	203
Assessment <i>as, for</i> and <i>of</i> learning	204
Fundamental roles	206
Continuum of conceptions	207
Four frames	208
Early childhood perspective.....	208
The ACAI.....	210
Summary of conceptualizations of assessment.....	213
Integrating purposes of assessment.....	213
DAP and Assessment as Pedagogical Approaches.....	217
Summary of Interaction of Approaches to Assessment and DAP	220
Summary of Discussion	221
Chapter 6: Implications for Practice and Recommendations for Research	222
Implications for Practice	223
Implications for “I”	223
Emphasize early childhood in licensure programs	223
Provide preservice assessment instruction.....	224
Increase experiences	225

Implications for “They”	226
Acknowledge that early childhood is different	226
Improve “early childhood literacy”	227
Develop assessment literacy	228
Value teacher expertise	229
Share learning and leadership	231
Implications for “We”	232
Broaden the understanding of assessments	232
Build shared assessment literacy	233
Focus on the child	234
Advocate for early childhood	235
Implications for Time and Resources	237
Summary of Implications for Practice	239
Recommendations for Research	239
Knowledge and Experience in Central Office	239
Historical Look at Teachers’ Development	240
K-2 Teachers’ Approaches to Early Childhood Pedagogy and Assessment	241
Teacher Preparation Programs	241
Reporting Systems in Early Elementary	242
What Works in Early Elementary Assessment Practices	243
Summary of Implications for Research	244
Summary of Study	245
Appendices	247

Appendix A: Researcher as Instrument	247
Appendix B: Email to Principals Seeking Recommendations	249
Appendix C: Letter to Teachers Seeking Participants (Attached to Email to Principals)	251
Appendix D: Summary of Research Plan	253
Appendix E: Courtesy Email to Principals About Study	258
Appendix F: Questions for Qualtrics Survey	259
Appendix G: Response Email to Teachers After Submitting Survey	261
Appendix H: Template for Email to Request an Interview	263
Appendix I: Email to Teachers in Database to Maintain Interest	264
Appendix J: Email to All Teachers When Research is Done	265
Appendix K: Interview Guide	266
Appendix L: Informed Consent	269
References.....	272
Vita.....	288

Acknowledgments

To the teachers who kindly volunteered their time to share their experiences and so richly contribute to this study: I am honored to have been trusted by you and inspired by the perspectives and commitment you have shared.

To the real women whose names I borrowed to provide my participants' pseudonyms: each of you is a teacher who has influenced and shaped me in my professional journey as we worked together, and I am forever grateful for the ways in which you touched my life and encouraged me in California, Japan, and Hong Kong.

To Maya Nelson: through your leadership and friendship I learned more than I could have imagined, and your fingerprints are all over this dissertation. Thank you.

To my professors, especially Dr. Denise Johnson: you revealed to me what it is to be part of the Tribe by showing me flexibility and grace through some challenging days.

To Dr. Leslie Grant: the first time we spoke at an Open House, you struck me as someone I would want to work with, but I could not have envisioned what an incredible teacher, leader, and advisor you would be. Thank you for the wonderful model you are.

To Dr. Megan Tschannen-Moran: when we met a decade ago in Manila I was in awe of your work, but when we reconnected recently, I discovered your compassion at a time I was most in need. Thank you for sharing both these gifts with me.

To Dr. Chris Gareis: I might be tempted to stay in school longer if there were any other classes you taught that I have not yet taken. I hope one day to be half the professor you are. I am grateful that you saw something in me worth cultivating, and my career path has forever been shaped by your wisdom and support. Of all your helpful advice, perhaps the most important to me at the moment is, "good writing never gets finished. It simply gets turned in." Thank you for helping make this happen.

List of Figures

1. Teachers’ Perceptions of the Purposes of Assessment	84
2. The Influences on Teachers’ Approaches Towards Classroom Assessment and Developmentally Appropriate Practice, Three Phases.....	96
3. Influences on the “I” (Individual) of Teachers’ Approaches Towards Classroom Assessment and Developmentally Appropriate Practice	98
4. Members of “They” Influencing Teachers’ Approaches Towards Classroom Assessment and Developmentally Appropriate Practice	106
5. Influences on the “We” of Teachers’ Approaches Towards Classroom Assessment and Developmentally Appropriate Practice	126
6. The Influences on Teachers’ Approaches Towards Classroom Assessment and Developmentally Appropriate Practice, Three Phases in Detail.....	140
7. The Influences on Teachers’ Approaches Towards Classroom Assessment and Developmentally Appropriate Practice, Three Phases in Detail with Added Resources	145

Abstract

Kindergarten-second grade teachers often have to navigate conflicting paradigms as they attempt to honor the developmentally appropriate practices best suited for their young learners while working within the demands of the current educational paradigm of high-stakes testing and standardization. This challenge is acutely experienced in the assessment of young children, yet little research has been done to look at how teachers in the early elementary years approach or use assessment in developmentally appropriate ways. The purpose of this study was to use a constructivist grounded theory approach to address the overarching question: How do K-2 teachers come to their conceptualizations regarding developmentally appropriate practices and strong classroom assessment practices? Thirty-five teachers were interviewed in reaching theoretical saturation. Through the constant comparative process of data generation, data analysis, and extensive memoing, the researcher generated a theory with *I*, *They*, and *We* phases to explain the influences on teachers. The study also revealed how teachers conceptualize assessment and the ways teachers' knowledge, beliefs, and practices interact with each other in regard to classroom assessment and developmentally appropriate practice. The influences of school district administration, teaching colleagues, and experience through time were some of the most considerable influences. Among its implications, the theory suggests a need for more dissemination of knowledge of best practices in early elementary education. The theory also provides a framework for future research to improve assessment decisions and inferences in early elementary classrooms.

LEAH KATHLEEN SHY

PROGRAM IN EDUCATIONAL POLICY, PLANNING, AND LEADERSHIP

WILLIAM AND MARY SCHOOL OF EDUCATION IN VIRGINIA

**THE INFLUENCES ON K-2 TEACHERS' APPROACHES TOWARDS
ASSESSMENT AND DEVELOPMENTALLY APPROPRIATE PRACTICE**

CHAPTER 1

INTRODUCTION

Teachers of children in kindergarten through second grade face a unique problem as they attempt to navigate conflicting educational paradigms. With the developmentally appropriate practices of early childhood pulling in one direction and the standardized practices and measurements of the accountability movement pulling in another, K-2 teachers often have to navigate a tricky balancing act in order to best support their students and honor policy demands of the school and state. One area where this tension is particularly pronounced is in creating, delivering, and interpreting developmentally appropriate classroom assessments.

Background

The early years in a child's life (the National Association for the Education of Young Children [NAEYC] defines early childhood as birth through age 8) are a time of rapid growth and development unique to other periods of life (Copple & Bredekamp, 2009). During this time, the foundation is laid for future growth across the domains of learning—physical, cognitive, social-emotional—and optimal periods for some types of development occur. The early years are crucial for later success in school, and thus it is vital that young children's learning environment is designed for the specific needs of children's development. Educators trained specifically in early childhood development are best suited for supporting children's growth in ways that are most appropriate, using "developmentally appropriate practice" (DAP; Buchanan, Burts, Bidner, & White, 1998; File & Gullo, 2002; NAEYC, 2009).

However, most K-2 teachers are not trained in early childhood practices, despite the definition of early childhood including children in these classes. The majority of early childhood degree and licensure programs focus on the period before formal school, from birth through preschool only, while elementary degree and licensure programs underemphasize or neglect issues of developmentally appropriate practice (File & Gullo, 2002). This is in large part due to the structure of elementary schools typically being K-fifth grade, and teaching licenses similarly being an elementary credential covering approximately K-fifth grade. Though licenses and teacher preparation programs vary across and within states, the majority of teachers in K, first, and second grade have been trained in programs oriented towards elementary school rather than early childhood (File & Gullo, 2002). Licensure programs leave these teachers without the depth of knowledge required to adequately deliver a developmentally appropriate educational program in their classrooms (Buchanan et al., 1998; File & Gullo, 2002; Vartuli, 1999).

Furthermore, regardless of the age or subject matter they teach, teachers tend to have relatively weak assessment literacy (Brookhart, 2001; Campbell & Evans, 2000; Campbell, Murphy, & Holt, 2002; DeLuca, Chavez, Bellara, & Cao, 2013; DeLuca & Klinger, 2010; Guskey & Jung, 2013; MacLellan, 2004; McMillan, 2013; Mertler, 2004; Mertler & Campbell, 2005; Popham, 2011; Stiggins, 1999). The standards movement ushered in full force by No Child Left Behind shifted emphasis onto the importance of classroom assessment, but teacher preparation programs still under-emphasize assessment (DeLuca et al., 2013; Popham, 2011). The focus on assessment has been primarily on standardized assessments, technical measures administered by, but not created by, classroom teachers. Due to the nature of these tests, standardized assessment results are rarely utilized in the planning for learning activities. Teachers must create their

own assessments (collaboratively or individually) for ongoing assessing that can inform planning for instruction. Targeted instruction based on results from ongoing assessments is a significant lever in improving educational outcomes, yet most teachers lack the assessment literacy skills to adequately pull that lever (DeLuca et al., 2013; DeLuca & Hughes, 2014; Gareis & Grant, 2015; Mertler, 2003; Mertler & Campbell, 2005; Popham, 2011; Stiggins, 1999).

This situation is compounded when considering the challenges facing teachers of young children. While early childhood teachers working in preschool settings may have had training in an early childhood degree or licensure program, teachers in K-2 will likely have had the same training all elementary education degree or license holders received, which does not focus on developmentally appropriate practices for early childhood, despite the fact that teachers in K-2 are teaching children within the early childhood age band (Buchanan et al., 1998; File & Gullo, 2002; Vartuli, 1999). These teachers may not be aware of how children develop within and across domains, typical features of child development, appropriate benchmarks at different ages, and best practices for creating and maintaining rich learning environments. The added pressures of high-stakes accountability tests approaching in third grade add to the push down of upper years' demands and expectations into lower elementary years (Cooper, 2009; Jeynes, 2006; Michnick-Golinkoff, Hirsh-Pasek, & Singer, 2006; Popham, 2011). So rather than having developmentally appropriate classrooms through age 8, early elementary years are often just a "younger" version of the overarching elementary framework, driven by the end goal expectations rather than building from the bottom up. This situation has created two "potentially conflicting paradigms" (File & Gullo, 2002, p. 126), though there is evidence

that teachers find ways to negotiate the needs of the two paradigms (DeLuca & Hughes, 2014).

Even if K-2 teachers have received training in the foundations of classroom assessment, and even if they are familiar with guidelines for developmentally appropriate practice, teachers may still have a difficult time weaving these ideas together, as preservice teachers tend to have a fragmented and compartmentalized knowledge base (Fives, Barnes, Dacey, & Gillis, 2016). NAEYC has released a position statement on Curriculum, Assessment, and Program Evaluation, published a number of books on assessment, and created a teacher training program on classroom-based assessment of preschoolers. However, none of these resources exist specifically for K-2 teachers, addressing their needs of bridging between purer early childhood programs centered on child development and more standardized late-elementary school programs centered on grade level benchmarks. The Classroom Assessment Standards published by the Joint Committee on Standards for Educational Evaluation do not address developmentally appropriate practices in the early years, and there is no framework that integrates the key concepts of developmentally appropriate practice and best practices in classroom assessment (Klinger et al., 2015). K-2 teachers are not getting sufficient pre-service training in developmentally appropriate practice or classroom assessment practices, and there are few resources that help bridge the gap for teachers to utilize for professional development once in the field, as the field is driven by standardized test-based accountability.

Problem Statement

At present, there is little research to support the development of K-2 teachers' assessment literacy in developmentally appropriate ways. Not enough is known about

how teachers develop their approaches to assessment while trying to piece together elements from differing paradigms. The pre-service period provides the most structured, formal, and focused period of learning for a teacher, but in-service professional development, school- and district-level factors, individual experiences and orientations, and other unidentified factors may also greatly influence a teacher's approach towards developmentally appropriate assessments. In addition to illuminating how teachers conceptualize developmentally appropriate classroom assessment, the purpose of this study of K-2 teachers' approaches towards DAP and classroom assessment was to develop a theory to explain the process by which teachers come to their conceptualizations.

Research Questions

As there is little research in the area of K-2 teachers' use of developmentally appropriate classroom assessment practices, there are a range of research questions that are important to examine the problem. The overarching question that drove this dissertation was *How do K-2 teachers come to their conceptualizations regarding developmentally appropriate practices and strong classroom assessment practices?* The research questions that guided this inquiry are as follows:

- (1) Through what processes and influences do experienced K-2 teachers report that they have developed their *approaches to assessment*?
- (2) Through what processes and influences do experienced K-2 teachers report that they have developed their knowledge and implementation of *developmentally appropriate practice*?
- (3) How do experienced K-2 teachers' approaches towards classroom assessment and philosophical orientations towards DAP interact with each other?

Significance of the Study

Findings have implications for teacher preparation programs as well as in-service professional development of K-2 teachers and administrators. It should be of interest to anyone who works in developing and delivering teacher education programs and materials. The study also illuminates the need for in-service professional development for both teachers and administrators. The findings highlight a number of contextual factors at the district and building level that heavily influence K-2 teachers' approaches towards and practices in classroom assessment use, including implications for professional teams. Furthermore, the results show areas in need of improved clarity and communication in the area of assessment literacy.

Having a theory to explain how various influences help shape teachers' conceptualizations and practices, rather than simply a description of findings, provides richer opportunities for future research to build upon—or challenge—findings in a structured way, and the theory provides a frame for discussions in the field around an emerging area of study. In addition to offering a framework for future qualitative discovery, it highlights key factors deserving quantitative analysis. The focus of *this* project, however, was not to validate specific assessment methods or techniques nor measure the impact of specific factors on teachers' development. It was to examine the interplay of influences on teachers' development within the construct of the overlapping paradigms of early childhood and the standards movement in assessment. The findings have the potential to impact both the field of practice and the research on the topic, but perhaps the most important contributions of this study are the ways it draws attention to the neglected needs of early elementary teachers' assessment literacy.

Definition of Terms

Approach to assessment: the way in which a teacher conceives assessment as a construct and a practice, guiding his or her own practice. The approach may be framed by beliefs, knowledge, attitudes, and/or conceptions.

Assessment literacy (AL): “an individual’s understandings of the fundamental assessment concepts and procedures deemed likely to influence educational decisions” (Popham, 2011, p. 267).

Classroom assessment (CA): “a process that teachers and students use in collecting, evaluating, and using evidence of student learning for a variety of purposes, including diagnosing student strengths and weaknesses, monitoring student progress toward meeting desired levels of proficiency, assigning grades, and providing feedback to parents” (McMillan, 2013, p. 4).

Classroom assessment standards: 16 standards on assessment organized into three areas (foundations, use, and quality) designed to guide educators and serve as a framework for teacher decision making regarding classroom assessment practices. These standards were published by the Joint Committee on Standards for Educational Evaluation in 2015 (Klinger et al., 2015).

Concept of Word (COW): an understanding of language being constituted by units called words that are represented by clusters of letters with space on either side in printed language.

Developmentally appropriate practice (DAP): a framework outlining research-based best practices on child development and learning for early childhood. It was first written by the NAEYC in 1986 and revised most recently in 2009.

Early-career teacher: a teacher having completed between 4-9 years of full-time teaching; designation created for the purposes of this study.

Early childhood: the period in a child's life from birth through age 8.

Early elementary: kindergarten through Grade 3; synonymous with primary grades.

Experienced teacher: a teacher in at least his or her fifth year of full-time classroom teaching.

Formative assessment: "a set of skills and activities that are undertaken by teachers to provide feedback to students to enhance their motivation and learning by designing instruction to meet student needs" (McMillan, 2007, p. 1). Often referred to as "assessment *for* learning."

In-service: a description of the period of a teacher's career when she is working in the field, currently employed in a school and serving students.

International Baccalaureate Primary Years Programme (IB PYP): a Swiss-based international educational foundation with a unique mission and curricular orientation, including the use of a learner profile that describes the intended outcomes for its education. The PYP is its program for children ages 3-12.

Large school district: a district with 20 or more schools; designation created for the purposes of this study.

Mid-career teacher: a teacher having completed 10-17 years of full-time teaching; designation created for the purposes of this study.

Mid-sized school district: a district with more than one and fewer than 20 schools; designation created for the purposes of this study.

National Association for the Education of Young Children (NAEYC): a non-profit professional association that focuses on promoting the well-being and education of young children, defined as birth-age 8.

Phonological Awareness Literacy Screening (PALS): a tool used for screening, diagnostics, and progress monitoring of the components of foundational literacy. The PALS was developed at the University of Virginia and is used throughout the state (and elsewhere) with kindergarten-third grade students.

Pre-service: description of the period of time prior to starting a teaching job; refers to a student teacher or future teacher still in the preparation stage. He or she may be anywhere in the journey from the first college class in education to completing a student teaching internship but is yet to be employed in the field.

Primary grades: kindergarten through Grade 3; synonymous with early elementary.

Professional Learning Community (PLC): a group of teachers and an administrator(s) who meet regularly to address ways to improve student learning. The group's work includes analyzing student work and student data as everyone shares responsibility for ensuring all students learn.

Response to Intervention (RtI): a framework for supporting students based on three tiers of instruction. Rather than waiting for students to fail, the model is designed to proactively identify students' needs and flexibly move students between tiers in order to best target specific instructional needs.

School division: the name for a Local Education Agency (LEA) in Virginia; in most states, the term district is used rather than division

Small school district: a district with just one elementary or primary school; designation created for the purposes of this study.

Summative assessment: an assessment “conducted mainly to monitor and record student achievement, and is used for school accountability” (McMillan, 2007, p. 1).

Technology Enhanced Items (TEI): items on computer-based tests that use a variety of formats to assess different types of learning objectives or assess in different ways. They include specialized interactions between learner and assessment in ways that traditional paper and pencil or standard multiple choice style computer tests cannot support.

Veteran teacher: a teacher having completed 18 or more years of full-time teaching; designation created for the purposes of this study.

CHAPTER 2

REVIEW OF RELATED LITERATURE

The purpose of this study was to help shed light on the overriding question *How do K-2 teachers come to their conceptualizations regarding developmentally appropriate practices and strong classroom assessment practices?* Because there is a dearth of research identifying relevant variables that impact K-2 teachers' developmentally appropriate classroom assessment practices, the methodology of grounded theory was a good fit to address the questions of this study. However, this also meant the initial literature review presented here has unique characteristics. Conducted prior to generating data in order to situate the questions, it could not be, and necessarily *should not* have been, thorough at that stage.

Role of the Literature Review in Grounded Theory

When initially developing the methods of grounded theory, Glaser and Strauss (1967) suggested that the researcher must approach the inquiry without knowledge of the existing literature in order to remain open to the data and not apply preexisting conditions or frameworks to it. Charmaz (2014) supported a delayed literature review, stating that “the intended purpose of delaying the literature review is to avoid importing preconceived ideas and imposing them on your work” (p. 306), while also acknowledging like many later grounded theorists that “most researchers today cannot begin their research without prior knowledge of the scholarship about their field” (p. 59), and they cannot truly be a *tabula rasa*. In order to undertake inquiry on a meaningful question, the

researcher must have at least some awareness of how the question is situated in what is already known. The key, however, is to avoid over-exposure to extant theories that may shape one's interpretation of the data generated in the study. Rather than attempting to strive for theoretical innocence, theoretical agnosticism is a more realistic goal (Charmaz, 2014).

This initial literature review was therefore intentionally narrow and served the purpose of justifying the research questions and highlighting concepts to consider (Gibbs, 2015). It could not fully support a theory that had not yet been written, but Chapter 4 of this dissertation includes an additional review of literature relevant to the findings and categories that make up the theoretical explanations in this study.

Start of the Standards Movement

A historical analysis of the standards movement shows how the nation's political machinery used arguments of equality, student achievement, an improved economy, national security, and global competitiveness as levers to drive educational policy reform. The Elementary and Secondary Education Act of 1965 (ESEA) represented the first substantial federal legislation designed to improve educational outcomes. In the name of equality and closing the opportunity gap as part of Lyndon Johnson's War on Poverty, ESEA provided federal funds to states to distribute to local educational agencies. By 1981, the dominant argument for educational reform had shifted its focus from equal opportunity amongst children within the country, to a competitive drive among countries. The War on Poverty had given way to the Cold War.

In 1983, the National Commission on Excellence in Education, formed by Terrell Bell, the United States Secretary of Education at the time, under President Ronald Reagan, released the landmark report *A Nation at Risk: The Imperative for Educational*

Reform. Using strong rhetoric, the report reframed problems with education as ones that threaten American prosperity, security, and dominance. Furthermore, the primary driver for change was that:

Our once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world. . . . the educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people. What was unimaginable a generation ago has begun to occur—others are matching and surpassing our educational attainments. (para. 1)

Now rather than using education to help ensure equity among its citizens, the nation was working to secure its place in the global standings. This required an emphasis on high standards of achievement. In 1994, President Bill Clinton signed into law Goals 2000, which emphasized outcome-based education. The focus shifted from ESEA's focus on funding to a focus on academic standards and measuring student progress towards these standards (Manna, 2011). Although some national professional organizations wrote academic standards during this time, according to Ravitch (2010), these efforts gave way to conflicts of political ideology, and the fledgling standards movement made little headway until the next administration replaced it with the accountability movement.

In 2001, ESEA was again reauthorized and the result, the No Child Left Behind Act (NCLB), tried to correct areas in which legislators felt the previous version of ESEA had faltered: measurement and accountability (Manna, 2011). Schools became accountable not just for inputs or dollar figures, but for the outputs of students' achievement.

High-Stakes Accountability

In order to hold states accountable for student achievement, there had to be a way to measure achievement, and it came to be narrowly defined by standardized test scores. NCLB was more than guidelines, it was a system that allowed more control over compliance through the issuing of sanctions. It demanded that students show adequate yearly progress through state testing, and the results carried with them high stakes. Popham (2004) defined a high-stakes educational test as “one that has serious consequences—either for the students who take the test or for those who prepared students for that test” (p. 63). These consequences could include student promotion, tracking, and graduation; teacher salaries, retention, and firing; and school grading, closure, and takeover by the state (Popham, 2004). Much like the business world was being transformed by the use of data to measure performance and improve efficiency, so was educational reform shifting towards an efficiency model (Groen, 2012).

This shift was amplified by Race to the Top, a competitive grant program out of the Obama administration that not only held schools accountable for low test scores, it held teachers accountable for test outcomes (Ravitch, 2013). Then Secretary of Education Arne Duncan even utilized some of the same rhetoric first heard in *A Nation at Risk*, calling the U.S. performance on the 2012 Programme for International Student Assessment “a picture of educational stagnation” and continuing to say: “That brutal truth, that urgent reality, must serve as a wake-up call against educational complacency and low expectations” (Duncan, 2013, para 10). Although the latest reauthorization of ESEA, the Every Student Succeeds Act (ESSA) of 2015, narrows the federal government’s role in education by shifting more control back to states, it still requires

annual standardized testing and, thus, is unlikely to dramatically shift the dominant paradigm away from the high-stakes tests of the accountability movement.

Impact of Assessment Practices

With the sanctions and high stakes for states, schools, and even individual teachers on the line, the accountability era has turned assessment from a pedagogical tool into a political one. Zhao (2009) critiqued NCLB in how it narrowly defined student success: “While the intention is to ensure every child receives a good education, the problem is that NCLB practically defines good education as being able to show good scores in a limited number of subjects” (p. x). Not only has curriculum been narrowed, but teachers have been forced to focus on improving outcomes on standardized tests rather than emphasize classroom assessment practices that can more directly impact student learning.

The potential harm of fixating on standardized tests has been a concern for some long before NCLB. The National Center for Fair and Open Testing has advocated for alternative assessments since the mid-1980s. Its 1989 report “Standardized testing: Harmful to educational health” asserts that the nation’s misuse of standardized tests narrows what is taught, limits student opportunity and deprives the public of its power to determine what matters (Neill & Medina, 1989). The authors further contended that standardized tests often do greatest damage to minority and low socio-economic students. More than a decade after that report, NCLB was ushered in on the premise that it would improve schooling for these subgroups. However, it has had greater impact on reducing that which is taught to what is tested. As Neill pointed out in 2006, “the higher the stakes, the more schools focus instruction on the tests” (p. 30).

A perhaps unintended side-effect of the standards and accountability movement was that its birth wrested assessment from the hands of professional educators and placed it in the hands of policymakers and test publishers. Rather than assessment being viewed purely as a pedagogical tool, it also became an accountability tool and, to some, a weapon.

Classroom Assessment as a Construct

Due to these negative effects and sometimes punitive impacts of high-stakes testing, the term “assessment” has come to be a dirty word to many classroom teachers. There has been “an extraordinary emphasis on assessment as something *external* to the teaching and learning that takes place in the classroom” (Gareis & Grant, 2015, p. 10). In addition to narrowing the taught curriculum, the over-reliance on standardized tests has also narrowed how teachers think about assessment. Gareis and Grant stated:

Although our purer understanding of teaching and learning has us value assessment as *integral* to the educational process, the current era of accountability has shifted the thinking of many teachers to view assessment as something that is *done to* their students, and, by extension, is *done to* teachers, too. (p. 10)

Therefore, it is crucial that we define classroom assessment as a construct different from standardized assessment, with all its negative connotations.

Classroom assessment (CA) as a construct is separate from the psychometric measurements of traditional testing and instead focuses on what teachers do in the classroom with students. The construct has not always been clearly defined and emphasized in the field or in research, so McMillan (2013) edited the *SAGE Handbook of Research on Classroom Assessment* for the purpose of signaling “the beginning of CA as

a separate and distinct field of study with an identifiable research base” (p. 3). He provided the following description of classroom assessment:

CA is a broad and evolving conceptualization of a process that teachers and students use in collecting, evaluating, and using evidence of student learning for a variety of purposes, including diagnosing student strengths and weaknesses, monitoring student progress toward meeting desired levels of proficiency, assigning grades, and providing feedback to parents. ...It is distinguished from large-scale or standardized, whether standards-based, personality, aptitude, or benchmark- or interim-type tests. It is locally controlled and consists of a broad range of measures. ...CA connects learning targets to effective assessment practices teacher use in their classrooms to monitor and improve student learning. (McMillan, 2013, p. 4)

Using this understanding of CA, it is clear that assessment should be an integral part of a teacher’s classroom practices, intertwined with his or her instructional practices. The reason this is not always the case, however, lies at least in part on teachers’ weakness in assessment literacy.

Assessment Literacy

In 1991, Rick Stiggins introduced the concept “assessment literacy” as a way to support the idea that “teachers and school leaders absolutely must understand the basic principles of sound assessment practice” (Stiggins, 2014, p. 67). Assessment literacy (AL) is used to describe the professional competency made of the knowledge and skills teachers need in order to create, analyze, and use classroom assessments to improve student learning (Gareis & Grant, 2015). Popham (2011) contended, “assessment literacy consists of an individual’s *understandings* of the *fundamental assessment concepts and*

procedures deemed likely to influence educational decisions” (p. 267). Popham further explicated this definition by saying that *understandings* refers to a basic knowledge and awareness rather than technical mastery; *fundamental assessment concepts and procedures* refers to central, not exotic, content; and *likely to influence educational decisions* refers to the concepts and procedures most likely to impact teachers’ decisions (and thus, learning) as opposed to every small detail about assessment.

Significant research supports the importance of teachers having strong AL as a key responsibility of their professional competency (DeLuca et al., 2013; DeLuca & Hughes, 2014; Fives et al., 2016; Gareis & Grant, 2015; Mertler, 2003; Mertler & Campbell, 2005; Popham, 2011; Stiggins, 1999). The reason for this is, naturally, because the power of CA to improve student learning is well-documented (Black & Wiliam, 1998; S. Chappuis, Chappuis, & Stiggins, 2009; Earl, 2003; Fives et al., 2016; Gareis & Grant, 2015; McMillan, 2013; Popham, 2011; Stiggins, 2004). However, teachers’ general weakness in AL is also well-documented (Black, 2013; Brookhart, 2001; Campbell & Evans, 2000; Campbell et al., 2002; DeLuca & Bellara, 2013; DeLuca et al., 2013; DeLuca & Klinger, 2010; Fives et al., 2016; MacLellan, 2004; McMillan, 2013; Mertler, 2004; Mertler & Campbell, 2005; Popham, 2011; Schmoker, 2006; Stiggins, 1999).

The combination of the importance of having strong AL and the lack of strong AL in the teaching force is problematic for a number of reasons. Perhaps the most obvious, and certainly most important, reason is that teachers are not able to leverage CA for their students’ learning benefit. It also means there is not a strong counterpoint in the field to push against assessment in its standardized form. Without voices of teachers supporting strong CA practices, the strength of policymakers (and beneficiaries of policies)

supporting the dominant paradigm of assessment as an accountability measure drown out the calls for assessment to be a pedagogical tool in teachers' hands.

One voice working to make an impact on teachers' AL is the Joint Committee on Standards for Educational Evaluation (JCSEE). The Joint Committee published a set of heavily-researched and vetted classroom assessment standards (Klinger et al., 2015) for free access as an e-book in order to readily disseminate the standards into classroom usage. These 16 standards are organized into the three groups of foundations, use, and quality. While the standards serve as a framework through which to consider strong CA practices and analyze practices, they have yet to gain widespread usage across a range of contexts.

Other scholars and researchers have made attempts at developing instruments that would measure teachers' AL. One instrument was Mertler and Campbell's (2005) Assessment Literacy Inventory. The instrument was based on the *Standards for Teacher Competence in Educational Assessment of Students* by the American Federation of Teachers, National Council on Measurement in Education, and the National Education Association (1990), and although it showed promise in its design of embedding items into scenarios requiring application of assessment in real world contexts, the Assessment Literacy Inventory never took hold as an instrument in a way that made a significant impact on the field, and the standards upon which it was based are now largely considered outdated, as they are from the pre-accountability era. The JCSEE's standards have attempted to update the standards themselves, but as yet there is not a comparable instrument through which teachers' AL can be measured. An alternate method to measuring teachers comes from a different angle: considering not teachers' actual

knowledge and skills in assessment, but teachers' *attitudes and approaches* towards assessment.

Approaches to Assessment

In his influential work on teachers' beliefs, Pajares (1992) examined the construct of teacher beliefs and provided a simple definition of a complex concept: "Belief is based on evaluation and judgment; knowledge is based on objective fact" (p. 313). These subjective beliefs teachers carry towards something have a substantial impact on their practices. According to Pajares (1992),

Few would argue that the beliefs teachers hold influence their perceptions and judgments, which in turn, affect their behavior in the classroom, or that understanding the belief structures of teachers and teacher candidates is essential to improving their professional preparation and teaching practices. (p. 307)

Furthermore, Pajares even claimed that "beliefs have stronger affective and evaluative components than knowledge and that affect typically operates independently of the cognition associated with that knowledge. Knowledge of a domain differs from feelings about a domain" (p. 309).

When specifically applying this to assessment, Barnes, Fives, and Dacey (2017) stated, "how teachers conceive of assessment shapes if (and how) they use assessment results and frames their selection and development of assessment tools in their classrooms" (p. 107). In short, teachers' beliefs, attitudes, conceptions, and approaches towards assessment may be just as important as their technical AL.

Barnes et al. (2017) used the Concepts of Assessment Instrument to investigate teachers' concepts of assessment from a person-centered approach. Their 2017 study on the beliefs surrounding assessment of 179 practicing teachers in the United States used a

confirmatory factor analysis to suggest three factors that rest on a continuum from an extreme pedagogical to an extreme accountability conception, along with a factor that sits outside the continuum: Assessment as irrelevant. Much like the ALI that predated it, the COA-III was based on *Standards for Teacher Competence in Educational Assessment of Students*, so while these findings may be pertinent, an instrument designed for use in the accountability era could provide more useful information.

As the JCSEE standards were a redevelopment of previous AL standards, there was a need for a new AL measurement to reflect the new standards. DeLuca, LaPointe-McEwan, and Luhanga (2016) created the Approaches to Classroom Assessment Inventory (ACAI), setting out to “construct a reliable instrument reflective of current assessment practices and contexts. Specifically, we draw heavily on the recently developed Classroom Assessment Standards” (p. 249).

Coombs, DeLuca, LaPointe-McEwan, and Chalas (2018) utilized the ACAI to examine teachers’ beliefs across career stages and found wide variability in teachers’ approaches to assessment across the stages of their careers. The authors called for more research into the various influences on teachers’ assessment identity, stating:

There is a growing body of scholarship, including this study, that suggests assessment literacy is far more than the learning of knowledge and skills in assessment. Rather it is a complex socio-cultural professional disposition tied to teacher identity and other aspects of classroom practice. (p. 142)

There is a clear need for more research into understanding how teachers are influenced in complex socio-cultural ways to develop their identity and approaches towards assessment, which heavily impact their practices and, thus, student learning. But how does this need change when considering the narrowed population of teachers who

work in early elementary classrooms? In order to better understand the contextual factors of teaching K-2 students, we must start by examining early childhood.

Early Childhood

For over 200 years, theorists including Johann Pestalozzi, Friedrich Fröbel, and Maria Montessori have drawn attention to the crucial importance of children's development and education in the early years. Their exact age range of "early childhood" varies slightly, but the current definition widely used is birth-age 8, based on the range given by the National Association for the Education of Young Children (NAEYC). The non-profit NAEYC is the modern leader in the sphere of early childhood education in the United States. The association was founded in 1926 under the name National Association for Nursery Education, and in 1964 it changed its name to NAEYC, but it was its 1986 publication of a framework for outlining best practices for promoting young children's optimal learning that has had a widespread and lasting impact on the field. This framework has been known as *developmentally appropriate practice* (DAP), and has been periodically updated, including its most recently adopted position statement published in 2009.

Developmentally Appropriate Practice

The framework of developmentally appropriate practice was first published to clarify best practices in early childhood and as a response to the growing push down of teaching methods and curricular expectations from the elementary years into early childhood (Copple & Bredekamp, 2009). The framework is, "grounded both in the research on child development and learning and in the knowledge base regarding educational effectiveness" (NAEYC, 2009, p. 1), and is structured as three core considerations, 12 principles of child development and learning, and five guidelines for

DAP. The core considerations address the areas which teachers are to consider with intentionality: what is known about child development and learning, what is known about each child as an individual, and what is known about the social and cultural contexts in which children live. The 12 principles of child development and learning unpack these considerations, and the authors of the statement are careful to note that DAP is “not based on what we think might be true or what we want to believe about young children. [DAP] is informed by what we know from theory and literature about how children develop and learn” (NAEYC, 2009, p. 10). NAEYC acknowledges that “whether or not what actually happens in the classroom is, in practice, developmentally appropriate is the result of myriad decisions at all levels—by policy makers, administrators, teachers, and families about the care and education of young children” (p. 16). The five guidelines address areas within which early childhood professionals must make decisions: creating a caring community of learners, teaching to enhance development and learning, planning curriculum to achieve important goals, assessing children’s development and learning, and establishing reciprocal relationships with families.

While most educators of children birth-age 8 would acknowledge the importance of the specifics outlined in the DAP position statement, the construct of DAP has failed to take a strong hold in elementary schools. Perhaps it is because the public school setting is more heavily impacted by—or controlled by—additional policy actors than independent preschools are (File & Gullo, 2002). But even public preschool contexts embrace DAP more than teachers at the upper end of the birth-age 8 continuum. A closer look at some possible reasons why this is can illuminate the complexity of embedding the framework of developmentally appropriate practice in K-Grade 2 classrooms.

DAP in the Early Elementary Years

Perhaps the biggest current influence on the nature of early elementary education is the same influence that has dramatically shifted views on assessment: the accountability era. In their influential paper *Is kindergarten the new first grade*, researchers Bassok, Latham, and Rorem (2016) used two large nationally-represented data sets (the Early Childhood Longitudinal Study, ECLS-B and ECLS-K) to measure changes in kindergarten between 1998 and 2006, a time period that spanned the passage of NCLB. Their findings empirically supported the crisis narrative found in Miller and Almon's 2009 report that called for a "reversal of the pushing down of the curriculum that has transformed kindergarten into de facto first grade" (p. 63). Although NCLB does not require any testing prior to third grade, the high-stake pressures associated with those results have led to an "accountability shovedown" (Hatch, 2002).

But academic standards and DAP are not inherent opposites. Bassok et al. (2016) stated, "teaching academic content need not be at odds with 'play' and other types of pedagogical approaches considered developmentally appropriate in early childhood" (p. 22). However, in practice, kindergarten teachers reported a heightened emphasis on academic content and didactic teaching at the cost of experiences that support children's growth across developmental domains, which "may be stressful for children and negatively impact their motivation, self-confidence, and attitudes towards school" (p. 3). This study, and many others (Gallant, 2009; Goldstein, 2007; Graue, 2009; Gullo & Hughes, 2011a, 2011b; Minicozzi, 2016), focused on the tensions between developmentally appropriate practices and accountability in policies in kindergarten.

There is a smaller body of research that looks at DAP in the primary grades (Maxwell, McWilliam, Hemmeter, Ault, & Schuster, 2001). Buchanan et al. (1998)

worked to fill this gap in the research by examining teachers' beliefs and practices regarding DAP in first through third grades. Their study surveyed 277 teachers using The Primary Teachers' Beliefs and Practices Survey, which was formed using NAEYC's 1997 revision of DAP. The research showed that practices and beliefs are situated on a continuum from developmentally appropriate to developmentally inappropriate, and that classroom characteristics had more impact on teachers' practices and beliefs than teacher characteristics. These findings are of interest, but the educational landscape has changed so significantly in the accountability era that new research is needed to support their results. The authors also stated the need for more studies into influences on primary grade teachers:

Future research should focus on how primary grade teachers develop their beliefs and what things, other than their beliefs, influence their practice. If we gain additional insight into teachers' worlds' we may be in a better position to understand how to support and assist them. (Buchanan et al., 1998, p. 480)

Another factor that has been examined is the influence of licensure programs on teachers' beliefs about primary classroom practices. File and Gullo (2002) used the Beliefs about Primary Grades Curriculum and Teaching Survey with 119 students in early childhood and elementary education programs. Students earning early childhood licenses reported beliefs more consistent with constructivism and DAP than those earning elementary licenses, who expressed an orientation more aligned with direct instruction, whole-group teaching, and traditional lessons. According to the authors, the results "reflect traditional differences between the two fields, with expectations for homogeneity within the classroom, greater reliance on teacher-directed and whole-group teaching strategies, and more use of testing and workbooks for assessment purposes characterizing

[elementary education] more so than [early childhood education]” (p. 134). These findings support the claim made by Bornfreund (2012) that:

Disparate preparation of early-grade teachers has set up an either/or scenario: Young children either have a teacher who understands how they learn but lacks subject-area expertise, or they have a teacher who understands what knowledge and skills they need but lacks insight on how they soak up new knowledge and skills. (p. 37)

Interestingly, participants in the File and Gullo (2002) study were more likely to report beliefs aligned with DAP prior to student teaching, regardless of the licensure program. This raises questions about the impact of the student teaching placement and experience and what influences could be at play in shaping teachers during this crucial period of the pre-service preparation.

Rather than looking at the influence of licensure programs, Maxwell et al. (2001) examined predictors of DAP in K-third grade classrooms. The purpose of their study of 69 classroom teachers was to test a new observation measure of DAP designed to be applicable across all grade levels, the Assessment of Practices in Early Elementary Childhood measure. The measure found that approximately one quarter of variance in observed practices could be attributed to just grade, and half of variance could be explained by classroom characteristics, teacher characteristics, and teacher beliefs, with the characteristic of teacher education being the factor most amenable to molding or changing. Both developmentally appropriate and developmentally inappropriate beliefs predicted observed classroom practices.

The notion of developmentally appropriate and inappropriate practices operating on a continuum is one supported by Van Horn and Ramey (2004). They used an extant

database, the National Head Start Public School Early Childhood Transition Demonstration Project, to evaluate of A Developmentally Appropriate Practice Template, an observational measurement designed to address DAP in Grades 1-3. They concluded that A Developmentally Appropriate Practice Template has strong psychometric properties and therefore is a useful tool for its purpose; however, they also raised questions about how DAP is enacted in schools. Van Horn and Ramey (2004) declared, “practically, the nesting of classrooms within schools indicates that educational policy and the preparation of both teachers and administrators need to consider the role of the school in enhancing or constraining DAP” (p. 584).

There could be myriad factors at practical play in how DAP is—or is not—utilized in early childhood classrooms. A 2004 study by Stipek found that classrooms with a high percentage of children of color and teachers’ perceptions of the impact of poverty-related problems on students were correlated with didactic teaching, while constructivist teaching more in line with DAP was higher in classrooms that had a higher percentage of White children. While the specific implications of these findings are disturbing, they also raise the concern for more research in the field to further address how a variety of factors may be at play or interplay with each other to impact teachers’ conceptions and practices of developmentally appropriate practice in the primary grades.

Developmentally Appropriate Assessment in K-2

Given the dearth of research into developmentally appropriate practices in the early elementary stage, it comes as no surprise that the bulk of research on developmentally appropriate assessment practices is focused on preschool or kindergarten and does not address first or second grade. According to DeLuca and Hughes (2014), “research on primary assessments is alarmingly sparse, with little

guidance provided to teachers on appropriating traditional developmental assessments to function within accountability contexts” (p. 442). Even NAEYC has very little emphasis on assessment in the primary years. Their 2003 position statement on Early Childhood Curriculum, Assessment, and Program Evaluation delineates best practices in these areas, but its explication of its research-based stance on assessment in high-quality programs leans heavily towards preschool practices and does not take into account the practicalities of public school structures and policies. The statement includes 11 indicators of effectiveness for assessment that reflect developmentally appropriate guidelines but are not all easy to translate into a public second-grade classroom. Furthermore, while NAEYC has published books on assessment in preschool and has a multi-session training program specifically for this age, it does not have a comparable program for primary year teachers.

A few studies have looked at assessment in K-2 classrooms, even if not explicitly through a lens of DAP. One researcher who has been prolific in the last few years addressing this area is Chris DeLuca (DeLuca & Hughes, 2014; DeLuca, Ogden, & Pero, 2015; Pyle & DeLuca, 2013, 2017). His work aims to address the need to understand this area better, as he has stated, “across assessment and early years literature, there remains little empirical research on how teachers navigate these assessment approaches in relation to developmental and academically oriented stances” (Pyle & DeLuca, 2013, p. 375). While I initially delayed a thorough review of all his studies in order to remain open to emerging theory in my study, DeLuca’s work substantiates the importance of asking the questions of my research study, as he asserted “the integration of assessments in response to developmental and standards-based orientations remains a challenge for many teachers” (DeLuca & Hughes, 2014, p. 442).

Summary of Literature Review

The initial literature review in a grounded theory study is necessarily incomplete, as too much research can narrow the analysis of the researcher and force data into preconceived patterns and categories (Charmaz, 2014). Nonetheless, it was important to recognize how existing literature frames, guides, and supports the question that steered my inquiry: *How do K-2 teachers come to their conceptualizations regarding developmentally appropriate practices and strong classroom assessment practices?*

Certainly, the current literature base shows a need for more research into the influences on early elementary teachers in assessment practices. McMillan (2013) advised, “the research that is needed is extensive, especially given the differences that exist between subject areas, grade level, and students with different abilities, special needs, and dispositions” (p. 12). Further, research into the importance of enhancing developmentally appropriate assessment practices within classrooms is critical. At the dawn of the accountability era, Maxwell et al. (2001) advised:

By marketing the concept of developmentally appropriate practice as one that applies *only* to young children, supporters of developmentally appropriate practice may inadvertently be doing more harm than good. People may dismiss the ideas as relevant only for very young children when, in fact, many of the principles apply to children and adults of all ages. (p. 446)

It is my hope that this grounded theory study will help expand others’ awareness surrounding the unique situations early elementary teachers face when trying to traverse conflicting paradigms in their assessment practices. The next chapter details the methods I utilized to complete this study.

CHAPTER 3

METHODS

The purpose of this study was to develop a theory to explain the development of K-2 teachers' conceptualization of the overlap of developmentally appropriate practice (DAP) and classroom assessment literacy. In addition to improving understanding regarding the pathways to development, it also aimed to serve the purpose of illuminating how teachers approach assessment in the context of the early elementary years. The study used three guiding research questions to address this:

- (1) Through what processes and influences do experienced K-2 teachers report that they have developed their *approaches to assessment*?
- (2) Through what processes and influences do experienced K-2 teachers report that they have developed their knowledge and implementation of *developmentally appropriate practice*?
- (3) How do experienced K-2 teachers' approaches towards classroom assessment and philosophical orientations towards DAP interact with each other?

Research Approach

To answer these research questions, I used a constructivist grounded theory approach as described by Charmaz (2014). Glaser and Strauss (1967) defined the components of grounded theory (GT) and launched the methods into broader acceptance during a period in which positivistic research was the expected norm. Their book *The Discovery of Grounded Theory* is considered a seminal text, though critics note that

Glaser and Strauss “did not write about grounded theory as a methodological/methods package; rather, they wrote only about the various strategies and techniques (methods) that could be used” (Birks & Mills, 2011, p. 5). Birks and Mills (2011) explained the difference between methodology and method in the following manner: “stemming from a congruent philosophy, a methodology is a set of principles and ideas that inform the design of a research study. Methods, on the other hand, are practical procedures used to generate and analyze data” (p. 4). Next, I will briefly discuss both the overarching methodology and methods used in this study.

Rationale for Choosing Grounded Theory (Methods)

Corbin and Strauss (2015) summarized the difference between qualitative and quantitative research as the prior being about discovering rather than testing variables. In distinguishing grounded theory from other qualitative means of discovering variables, they expressed the difference in this way: “description tells about an event or happening while theory offers explanations for why events or happenings occur. ...What makes theory different from descriptive qualitative research is the overarching structure—the skeleton or framework that explains why things happen” (p. 12).

Due to the lack of research identifying relevant variables impacting K-2 teachers’ developmentally appropriate classroom assessment practices, I did not think a quantitative study evaluating and testing variables would be particularly useful at this time. Furthermore, I believe an explanation for how teachers navigate the two conflicting paradigms is valuable towards the goal of opening discussion and focusing on the unique needs of the teachers on the lower age range of the K-12 spectrum. A structure or framework that explains why teachers approach assessment in the way they do seems more meaningful at this time than just a description of their practices.

Birks and Mills (2011) gave reasons for a researcher to utilize grounded theory, stating it is appropriate to use when, “little is known about the area of study, the generation of theory with explanatory power is a desired outcome, [and] an inherent process is imbedded in the research situation that is likely to be explicated by grounded theory methods” (p. 16). Corbin and Strauss (2015) provided a similar set of reasons for the application of GT, declaring it is:

used to study new and emerging areas in need of investigation. The procedures can be used to uncover the beliefs and meanings that underlie action, to examine rational as well as non-rational aspects of behavior, and to demonstrate how logic and emotion combine to influence how persons respond to events or handle problems through action and interaction. ...[It] provides a strong foundation for further studies using quantitative measures. (p. 11)

This study fits well with grounded theory given that there is little research about K-2 teachers’ approaches toward classroom assessment, the study aimed to uncover the meanings behind the ways teachers integrate developmentally appropriate practice with their assessment literacy, and I wanted to construct a theory that will explain how teachers navigate conflicting paradigms in this area. While constructing an explanatory theory was a significant undertaking, I believe this outcome has greater potential to encourage further dialogue and future research. Indeed, even challenges to the theory would be welcome, as ensuing discussions and findings would advance our knowledge about the ways in which K-2 teachers approach classroom assessment in light of DAP.

Rationale for Selecting Constructivist Grounded Theory (Methodology)

Following their 1967 work, Glaser and Strauss each went different directions in their own descriptions and use of GT. While Glaser wrote about grounded theory as a set

of methods, Strauss later developed a methodological approach towards GT based on his own philosophical background in pragmatism and interactionism (Corbin & Strauss, 2015). Second-generation grounded theorists, such as Charmaz, have described GT within different philosophical and methodological positions (Birks & Mills, 2011), and Charmaz's (2014) constructivist approach best aligns with my own philosophical viewpoints.

The constructivist paradigm emerged after Glaser and Strauss (1967) published their seminal text on GT as an alternative to the dominant positivist paradigm. Originally referring to the new paradigm as the naturalistic paradigm or naturalistic inquiry, Yvonna Lincoln and Egon Guba published extensively throughout the 1980s and beyond on what they later came to call the constructivist paradigm (the term I use in this study except when directly citing an author's use of the term naturalistic). They stated that constructivism "begins with the assumption that realities are not objectively 'out there' but are constructed by people, often under the influence of a variety of social and cultural factors that lead to shared constructions" (Guba & Lincoln, 1989, p. 12).

Lincoln and Guba (1989) further described social realities as "social constructions, selected, built, and embellished by social actors (individuals) from among the situations, stimuli, and events of their experience" (p. 230). Constructivist grounded theory is built on these premises, and it stresses "social contexts, interaction, sharing viewpoints, and interpretive understandings" (Charmaz, 2014, p. 14). Creswell (2013) described how the social constructivist paradigm impacts the qualitative researcher: "In social constructivism, individuals seek understanding of the world in which they live and work. They develop subjective meanings of their experiences—meanings directed toward certain objects or things" (p. 25).

Charmaz's (2014) explanation of the constructivist approach in GT "uses the inductive, comparative, emergent, and open-ended approach of Glaser and Strauss's (1967) original statement" (p. 12), but it also acknowledges subjectivity and the researcher's role in analyzing data and developing theory. Charmaz stated:

The constructivist approach perspective shreds notions of a neutral observer and value-free expert. Not only does that mean that researchers must examine rather than erase how their privileges and preconceptions may shape the analysis, but it also means that their values shape the very facts that they can identify. (p. 13)

The role of the researcher is critical in GT, and thus it is crucial to examine my own philosophical approach and experiences that situate me within the constructivist paradigm. My research questions align with the methodology, as my goal was to better understand the little-known processes in which K-2 teachers engage, and my purpose was to develop an explanatory theory rather than focus only on describing experiences. However, as an individual I had to also align with the methodology and be aware of how my values and preconceptions may have shaped the work throughout the process. I have addressed these matters with my Researcher as Instrument statement.

Researcher as Instrument

When the goal of research is understanding, Merriam and Tisdell (2016) stated "the human instrument, which is able to be immediately responsive and adaptive, would seem to be the ideal means of collecting and analyzing data" (p. 16). This is particularly true in the constructivist paradigm. According to Guba and Lincoln (1989):

Inquirers are human and cannot escape their humanness. That is, they cannot by an act of will set aside their own subjectivity, nor can they stand outside the arena of humanness created by the other persons involved. The values of the inquirer

(and of those who influence him or her, especially funders, sponsors, and professional peers) inevitably enter the inquiry in connection with the whole series of decisions involved in designing, mounting, and monitoring. (p. 88)

As the researcher in this study, it was necessary that I examine my own *humanness* and acknowledge my values and the influences that entered my inquiry. To this end, I wrote a Researcher as Instrument statement that is included in full as Appendix A. I reflected on how my “personal background, culture, and experiences hold potential for shaping [my] interpretations, such as themes [I] advance and the meaning [I] ascribe to the data” (Creswell, 2014, p. 186). My experiences living overseas for 12 years, teaching in early elementary grades for the same number of years, and as the mother of two young children have all shaped my interpretations of the world. Specifically, my work at a high-profile international school’s lower primary division in navigating the shift from a reporting system based on developmental continuums to a standards-based curriculum and assessment system highlighted some of the challenges teachers face in assessing young children.

These influences are not problematic in Charmaz’s constructivist GT, provided that I have grappled with my experiences, positions of privilege, and worldview without assuming that I can truly set it all aside (Gibbs, 2015). I do not claim that the theory I have constructed in my research is an absolute, objective truth, but acknowledge that it has been formed “through the interaction of a constructor with information, contexts, settings, situations, and other constructors ... using a process that is deeply rooted in the previous experience, belief systems, values, fears, prejudices, hopes, disappointments, and achievements of the constructor” (Guba & Lincoln, 1989, p. 143).

However, it was important throughout the work that I not only reflected on my experiences and beliefs prior to data collection and analysis, as I did in the Researcher as Instrument Statement, but also throughout the course of the study. The habit of reflexivity is important to practice in order to illuminate ways in which the researcher may shape the direction (and therefore, outcome) of the study (Creswell, 2014). For this reason, I kept a reflexive journal throughout the entire research process.

Reflexive Journal

The purpose of keeping an ongoing journal is “to engage in reflexivity and to avoid preconceiving your data” (Charmaz, 2014, p. 165). Charmaz calls this a methodological journal, as it is the place she recommends recording methodological decisions and challenges. Given the fluid nature of grounded theory, it was important to note when judgments were made to change directions or procedures why those decisions were made as well as to reflect on how my position as the researcher may have influenced those decisions (Lincoln & Guba, 1985). I chose to call my journal a reflexive journal, rather than a methodological one, in order to focus my writing on being both an audit trail of decisions and an ongoing log of personal reflections regarding my involvement in the work as a non-objective participant. Whereas the initial, static Researcher as Instrument reflection captured what I brought to the study, the reflexive journal captured *how* what I brought to the study may have influenced the co-construction of meaning throughout the process.

I kept my reflexive journal as a Word document within which I created dated entries at any point a thought or reflection occurred during the research process. I added an entry whenever decisions were made regarding the process, when new realizations about my involvement emerged, when I underwent significant deliberations regarding

next steps, or when I grappled with substantial challenges. Although I did not find it meaningful to enforce a rigid time-bound structure to the frequency of writing in my reflexive journal, I found that nearly every day I engaged in any substantive thinking work about the process, I had something to add. The reflexive journal was just over 15,000 words.

This reflexive journal is different from the memoing process, described later as part of data analysis. The primary distinction is that memos serve as ways to think about the data and the categories as they emerge, whereas the journal is a way to *think about the thinking* about the data and the process of the research. Practically speaking, it was unimportant as to whether these two processes were combined in one physical document or whether they resided in two different places. The main point was to ensure that I regularly participated in both processes, and for this reason, I chose to view my memoing and my reflexive journal as two separate things.

Conceptual Framework

In order to remain open to emerging themes from the data, GT calls for an openness to theory rather than use of a theoretical framework to shape the researcher's analysis based on a priori assumptions (Charmaz, 2014; Glaser & Strauss, 1967; Lincoln & Guba, 1985). Charmaz advised against reading theoretical literature about the research topic, yet also acknowledged that a researcher cannot undertake relevant research without having some familiarity with prior scholarship. While no theoretical framework bound this study, there are two conceptual frameworks that do inform my own perspective on the research. As they were more thoroughly addressed in Chapter 2, I only briefly mention here NAEYC's Developmentally Appropriate Practice framework and JCSEE's Classroom Assessment Standards. My familiarity with both of these frameworks provides

a lens through which I often review practices in the field. However, as my delimitations section notes, I was not trying to evaluate or measure teachers' practices, but rather to understand how they came to their approaches to assessment. Regarding the process of how teachers develop their approach, no conceptual or theoretical framework was applied to the data generated in the study.

Participants

Grounded theory methods emphasize purposeful, rather than representative, sampling of participants. As such, I sought out participants who could contribute to understanding of my research questions by initially using maximum variation sampling and then moving into theoretical sampling. Below, I describe the necessary characteristics I looked for in participants, as well as characteristics I used as initial factors to help maximize the variation in the data in order to reach a saturation point. I also describe how I recruited and incentivized participation.

Target Participants

In order to be considered for participation in this study, a teacher had to be currently teaching in a kindergarten, first grade, or second grade general education classroom in a public school setting. Furthermore, I sought out teachers who were in at least their fifth year of teaching. As there is no consistency in the literature defining what makes a teacher "experienced," I used this guideline in hopes it would allow me to gather data both about the influences prior to starting work as a teacher and the influences over time of the teaching context, including how they may interact with each other. Participants' years of experience may not have all been in an early elementary setting, provided they were teaching in a K-2 classroom at the time of the interview.

I chose to focus my research on teachers within Virginia only. Although there was an element of convenience to this decision, as I attempted to conduct as many in-person interviews as possible, it also has implications for the research findings, as teachers within one socio-political context with the same basic set of guiding policy practices impacting both their preparation and the contextual factors of their work setting may elucidate influences in a particular way that teachers in another state may not. It is important to note that Local Education Agencies (LEA) in Virginia go by the name “school division,” but the term “school district” is more widely used nationwide, so throughout this study I have chosen to use the term “district” to refer to LEA within Virginia where I conducted my research study.

Sample Size Based on Saturation

The concept of data saturation was first proposed when Glaser and Strauss (1967) formed the methods of GT, and within the methodology, it is directly tied to theoretical sampling and theoretical saturation, both discussed later in this chapter. However, the meaning and use of the term *saturation* have been applied inconsistently across a variety of qualitative methods (Aldiabat & Le Navenec, 2018; Fusch & Ness, 2015; Guest, Bunce, & Johnson, 2006; Hennink, Kaiser, & Marconi, 2017; Mason, 2010), and thus it is important to clarify the concept as it relates to participant selection and sample size here.

Several studies have attempted to address the question of when saturation is reached as it pertains to number of interviews. Guest et al. (2006) cited studies that claim the number needed to be significantly different: 6-8, 12-20, 15, 25, 35, or 36. They conducted an experiment on one of their own studies to try to answer the question of sample size to reach saturation and determined that they had reached data saturation by

12 interviews and had sufficient evidence to support their four main themes after only six interviews.

However, this study had limits when applied to GT. It used a structured interview protocol, which is advised against by grounded theorists, and was not focused on creating a theory as the outcome. As Charmaz (2014) noted, the concept of saturation depends on the purpose of the study. In a different study, Hennink et al. (2017) stated they reached *code* saturation at nine interviews, but required 24 interviews to reach *meaning* saturation.

This highlights the concerns about the confusion of the use of the term *saturation*. In grounded theory, neither *data* saturation nor *code* saturation are the benchmarks for knowing when to stop interviewing. *Theoretical* saturation is the benchmark, as theory creation is the aim. Charmaz (2014) advised, “stop when the properties of your theoretical categories are ‘saturated’ with data. In other words, your categories are robust because you have found no new properties of these categories and your established properties account for patterns in your data” (p. 213).

Although Creswell (2013) recommended 20-30 interviews for GT, the number of participants cannot be determined in advance of a grounded theory study and must be determined based on theoretical saturation as the study unfolds. Therefore, rather than setting a target number, I began my interviews using an approach of maximum variation sampling (Patton, 2002). Although the sample was homogenous in that all participants had experience teaching in a K-2 classroom, I looked for variable backgrounds and teaching experiences that could influence teachers’ development. Guba and Lincoln (1989) described this kind of sampling in a way that fits well with the constant comparative method of analysis that is a tenant of GT:

Maximum variation sampling that provides the broadest scope of information is the sampling mode of choice. ...The sample is selected *serially*, that is, no sample element is chosen until after data collection from the preceding element has already been largely accomplished. Second, the sample is selected *contingently*, both in the sense that each succeeding element is chosen to be as different as possible from preceding elements and in the sense that elements are chosen in ways that best serve the particular needs of the inquiry at that moment. (p. 178)

As themes emerged through constant comparative analysis, I moved into more theoretical sampling as the categories began to take shape for the purpose of filling out their properties and dimensions (Charmaz, 2014). It was not possible to determine in advance what categories would emerge and therefore what characteristics of participants would be salient in order to help reach maximum variation in my sample. However, the manner in which I recruited participants (described below) allowed me to create a database of willing participants that included *some* characteristics that I had projected *may* become relevant to maximizing my sample and therefore aiding in filling out the properties and dimensions of categories. Maximizing the variation was not the end goal, but a means to help reach theoretical saturation.

Recruiting Participants

In order to find potentially helpful participants, I began by selecting 12 districts varying in size, location, and contexts (urban, suburban, and rural) and started contacting school principals with a request for help identifying teachers who may be able to provide rich data. Appendix B shows the letter I used to contact school principals, including the wording for identifying potential participants. I requested that they send an attached letter directly to qualifying candidates (Appendix C), and that letter included a link to a very

simple demographic survey that interested teachers could fill out in order for me to see their interest; this survey is discussed in the next section. It became quickly apparent, however, that even with Institutional Review Board (IRB) approval from William & Mary, principals were not comfortable forwarding my invitation to participate to teachers without getting formal approval from their district's central office. Therefore, I changed strategies to gain access to teachers in various districts. This new strategy became more important as my initial analysis indicated that sampling from more districts and schools would be important to the study, as described in the later section on addressing challenges.

I shifted my focus from starting with principals to getting permission at the district level by following each targeted district's process for obtaining official research approval. Where a district's website provided information about the process, I followed their stated process. More frequently, the information was not clearly available, and I made phone calls to central office to determine who the gatekeeper was in each district. In some cases, this led to a chain of multiple calls as I followed the lead of the previous staff member, as even employees of a district did not always know who was in charge of the process. The range of gatekeepers varied greatly, all the way from the district superintendent to a human resources supervisor, but generally the role of gatekeeper fell to a research coordinator or assistant superintendent.

I prepared a summary of my research plan (Appendix D) and emailed it along with a copy of my IRB approval to the relevant district administrator if no specific information was requested, and in some cases, this was sufficient for being granted approval. In other instances, I filled out and submitted a written application, and in other cases still, a phone call and brief discussion over the phone allowed me to contact

teachers. Due to the fact that my study was entirely voluntary, required no contact with students, requested a minimal amount of teacher time with no impact on instructional time, and did not rely on collection of student, school, or district data, I generally found a favorable response except when district policy was limiting.

Once I was granted permission to a district, I inquired regarding the preferred method of contacting teachers. I realized that, when going through principals, rather than selecting specific people with whom to share the opportunity, principals were sharing my invitation to participate with all their K-2 teachers. The study was not designed or intended to evaluate or validate specific practices, so it was not necessary that participants meet a theoretical standard of best practices. Therefore, except in cases where my district contact explicitly requested that I go through principals, I looked for ways to email teachers directly and remove the request for principals to have to act on my behalf. In more than one district, I was sent a spreadsheet of K-2 teachers' names and email addresses, but for most districts, I found teachers' email addresses through public school webpages. In a few cases, the district contact forwarded my invitation to teachers on my behalf. I adapted the letter to teachers found in Appendix C slightly depending on the context and whether I was writing directly to teachers or they were receiving the letter from an administrator. Where appropriate, prior to contacting teachers I sent a letter to building principals as a courtesy, informing them of the research, although they did not have to take any action (Appendix E).

As part of receiving permission to conduct research, several districts requested that I send either a full copy of my completed dissertation or a summary of the findings. Following successfully defending my dissertation, I sent documentation to the requesting districts as well as an executive summary to all district contacts with whom I worked,

even if they did not specifically request this information, as a courtesy in recognition of their support of the work.

Building a Database of Willing Participants

As teachers received my invitation email, they could read a brief overview of my research and what I was requesting from participants: a single 45-minute interview to be conducted in person at a location of their convenience, where logistically feasible, or through video conferencing. Teachers could then indicate their interest in participating by clicking on the link in the letter to a Qualtrics survey. The survey collected only simple fill-in-the-blank information about the participant's background as a means of better predicting the ways in which the participant might be able to provide variation in the sampling process. As the factors that could be most relevant to the development of theory could not be known until the data generation and analysis processes were underway, the initial survey was not intended as a thorough way of sampling all possible variables of influence upon a teacher, but rather as one small way to possibly help maximize variation in the sample through interviewing teachers with a range of experiences and backgrounds. Appendix F lists the questions that were in this survey.

From the responses to this survey, I created a database of willing participants from which I selected interviewees. As my database grew over time, I sent an email to each new volunteer (Appendix G) to thank them and to describe what they could expect in the next stage of the research. As the interview process was spread over several months, the email explained that they may not be scheduled immediately, but they could expect to hear more from me in the coming weeks regarding scheduling. When I selected a volunteer for an interview, I used a personalized email to make arrangements; the basic template for these emails can be found in Appendix H.

Because of the time span during which interviews took place, I maintained contact with willing participants to sustain their interest in participating. Approximately every six weeks following the initial email of thanks, I sent a brief email (Appendix I) to all teachers in the database thanking them for their continued interest and letting them know that I was still collecting data and may contact them in the near future. It reminded them of the incentive drawing at the end of the study, but it also gave teachers a chance to indicate if they wished to withdraw from the study without penalty, in accordance to ethical standards. Although there were a few teachers who never responded to an email requesting to schedule an interview, no teachers requested to be removed from consideration during the process.

Incentives

In addition to the summary of research and link to the Qualtrics survey, the initial letter to teachers also stated the incentive for participation. All participants who filled out the survey as a way of expressing interest in participating were entered into a random drawing for a \$50 Amazon gift card. This helped build a large database of potential interviewees, as it was only a 5-minute survey with the potential for a relatively large incentive. In addition, every teacher who was selected for and completed an interview received a \$10 gift card from Amazon as a token of appreciation. Teachers tend to have many time demands and in this small way I could acknowledge the sacrifice of time they made without coercing anyone to participate. Only one teacher was interviewed twice, and she was given a second \$10 gift card.

The draw for the \$50 gift card happened at the end of the study after a successful defense, and the winning teacher was notified privately with the gift card sent via email. All teachers in the database of willing participants received an email at that time

notifying them that the research had been completed, thanking them for their willingness to participate, and stating that the drawing had been completed and the winning teacher had been notified (Appendix J). The executive summary of the study was sent as an attachment to this email, as many teachers expressed interest in learning about my findings. Further, I offered to send my full study after the dissertation is published in a dissertation database to any participant who expresses interest.

Addressing Challenges

Because there was no set quota for participants at the start, I needed to have a significant list of possible participants in order to do maximum variation sampling and later theoretical sampling. The emphasis on unfolding methods in grounded theory (Charmaz, 2014) required that I be flexible to respond to what the data was showing in order to decide how to select from my list of teachers and continue to build my database as needed. The reflexive journaling process helped illuminate both challenges as they arose and solutions to those challenges, while the memoing process indicated ways to proceed with purposeful sampling.

Following my initial round of contacting the first 12 targeted school districts, I was given permission to access four districts, was denied permission at one district, eliminated two districts from consideration due to their limited windows of timing for considering application, and had five pending requests. After inviting teachers in these four districts, I had a database of 35 eligible volunteers. By the end of the first 10 interviews, I had spoken with teachers in all four districts across seven different schools representing kindergarten, first, and second grades, and coming out of five different state colleges and universities and three different out-of-state institutions. Memoing during the analysis of these interviews indicated that the variable of school and school district

played a significant role, more than the other influences, so I greatly expanded the number of targeted school districts for participation.

While this was an important step to building my theory, it added a tremendous amount of effort on logistics as I continued to pursue new districts, build my database of volunteers, and interview selected participants simultaneously. Because these strands were unfolding at the same time, I was unable to pursue a purely serial sample; that is, I could not schedule, conduct, and analyze each interview then select the next participant in a strictly singular manner. Practically, I had to contact a few teachers at a time to request and schedule interviews in the future, as some interviews were scheduled as far as two weeks in advance and there was often a delay in receiving a response or, on the rare occasion, a failure for the volunteer to respond entirely. Furthermore, I could not always transcribe and analyze each interview prior to the next interview, as these were occasionally scheduled back-to-back when travel was required. Due to these pragmatic challenges, it was necessary to apply a modified serial sampling technique of clustering small groups of participants in the cycle of scheduling, interviewing, analyzing, and selecting the next round of participants based on emerging themes from the previous small group. In this way, my sampling did not strictly follow Guba and Lincoln's (1989) notions of selecting a sample serially and contingently on a one-to-one basis, but even with this modest limitation, I was able to select future participants based on emerging themes.

As I worked to expand my access to more school districts, I attempted to contact a variety of districts based on size and location. Over the full course of this process, I contacted 49 of Virginia's 138 school districts. My request to conduct research was declined at four of these districts, and another 14 districts had timelines for review of

applications that were not conducive for my study's timeline. I was granted permission to conduct research in 15 districts, and as my sample reached the point of theoretical saturation, I discontinued pursuing permission from the 16 districts with which I had pending communication but from which I had not received a timely reply.

Sample Characteristics

The 15 districts in which I was given permission to contact teachers included six rural districts with just one elementary school each, two large school districts of more than 20 elementary schools across urban and suburban contexts, and seven districts ranging in size between these extremes and including predominantly suburban areas. In total, teachers in 110 schools were invited to participate, and 84 volunteered by responding to the survey. Four volunteers were removed from the database for not meeting the requisite years of experience and another volunteer was removed for insufficiently responding to the interest survey. My preferred goal at the start of the study was to have at least 100 teachers in the database, though the 79 qualified volunteers proved to be more than sufficient to reach theoretical saturation.

As I reached out to teachers to schedule interviews, there were eight teachers who did not respond to my request, two who chose not to schedule an interview due to other commitments, and five who either did not show up at the scheduled time or who requested to postpone the interview but did not respond to a follow-up email from me to reschedule. Due to the variety available in the database, I did not need to pursue these teachers more aggressively if they did not respond to my reminder, and I assumed they were unable or uninterested in participating, so I moved on to inviting a different teacher to interview. I interviewed 35 teachers, one of whom I interviewed twice due to the rich, detailed response she was providing and our inability to extend the timing of the first

interview. How I came to reach saturation and stop interviewing is discussed later in this chapter.

There was at least one teacher from each of the 15 school districts, and in total they represented 26 different elementary schools from around the state. Teachers had a wide range of preservice experiences and came from teacher preparation and graduate degree programs completed at 25 different colleges and universities (11 out-of-state). The teaching endorsements held included preK-3, preK-6, NK-4, K-8, reading specialist, special education, English as a Second Language PK-12, and K-12 counseling. Fifteen different majors were included, with Elementary Education, Early Childhood Education, and Psychology being the most prominently represented. Teaching experience ranged from teachers in their fifth year of teaching to nearly 40 years of teaching and included work in nine other states, one foreign country, 11 additional Virginia school districts not represented by a current teacher, and multiple private schools. Current teaching assignments were spread out among kindergarten (15 teachers), first (8 teachers), and second (12 teachers) grade classrooms, with previous classroom experience including every grade from preschool through eighth grade. The sample included a large variety of additional experiences, including work in special education, school counseling, the International Baccalaureate Primary Years Programme, international schools, and dual language immersion programs as well as prior experience working in a different field and women who came to the teaching profession after raising their own children. Only two males volunteered for the study; I reached out to both, but neither one was able to schedule an interview, so all of the participants were female. For this reason, I have used a feminine pronoun to refer to participants when appropriate.

As grounded theory methodology does not make claims of generalizability and therefore does not require representative sampling, the 35 teachers I interviewed did not need to proportionally represent any particular set of characteristics or factors. Instead, the wide range of experiences included helped saturate categories of the theory as the study reached a point of theoretical saturation.

Data Generation

My research questions focused on understanding *processes, influences, and approaches*, things that cannot be readily observed or studied through collecting existing evidence. Therefore, data generation was necessary. According to Birks and Mills (2011), the relationship between the researcher and data is different in data generation versus data collection, as the researcher is considered an integral part of creating the data that is generated whereas the researcher is only connected to collected data through its analysis. Like most studies using a GT methodology, data were generated through interviews (Creswell, 2013). I was open to the possibility that, as Charmaz (2014) stated, “in the midst of the research, questions may arise that impel researchers to construct new data-gathering methods and to revise earlier ones” (p. 28). However, this did not emerge through my analysis as necessary to address my research questions.

Interviews in Constructivist Grounded Theory

Although GT methods can evolve during the course of a study, the role of the researcher in the qualitative paradigm remains relatively clear. Guba and Lincoln (1989) stated:

The major task of the constructivist investigator is to tease out the constructions that various actors in a setting hold and, so far as possible, to bring them into

conjunction—a joining—with one another and with whatever other information can be brought to bear on the issues involved. (p. 142)

Charmaz described this process as co-constructing meaning through asking and answering questions (Gibbs, 2015). She referred to this as *intensive interviewing*, which includes the characteristics of selecting experienced participants, exploring participants' experiences in-depth, relying on open-ended questions, obtaining detailed responses, emphasizing understanding the perspective of the participant, and following up on unanticipated responses or lines of inquiry (Charmaz, 2014). Rubin and Rubin (2012) described a similar set of characteristics as responsive interviewing:

It emphasizes flexibility of design and expects the interviewer to change questions in response to what he or she is learning. ...The model assumes that what people have experienced is true for them and that by sharing these experiences, the researcher can enter the interviewee's world. The researcher's role is to gather narratives, descriptions, and interpretations from an array of conversational partners and put them together in a reasoned way that re-creates a culture or describes a process or set of events in a way that participants would recognize as real. (p. 7)

Interview Structure

In order to be flexible and open, my interviews started with an unstructured format. Charmaz (2014) recommended an interview guide as a useful tool, particularly for novice researchers, as a reference of possible questions to use; she also acknowledged that you may not use it at all while conducting an interview if the respondent's story flows easily. The guide helps the interviewer by suggesting broad, open-ended questions to start and by helping the interviewer improvise smoothly. According to Charmaz:

From a grounded theory standpoint, asking a few interview questions allows the research participant to tell his or her story without the researcher preconceiving the content, or for that matter, the direction the interview will take. Such a strategy is particularly useful during early interviews but can change as the researcher moves back and forth between data collection and analysis. (p. 93)

The changing strategy is based on focusing interview questions as theoretical categories emerge. After the first 12 interviews, there were some clear but broad categories, so I shifted towards using a semi-structured interview as a more appropriate tool. Rubin and Rubin (2012) stated the primary difference between unstructured and structured interviews, “lies in the degree of control that the interviewer maintains over the interview; in semi-structured interviews, the researcher tries to focus more narrowly on the planned items that speak to the research question” (p. 31). In the case of GT, the research purpose is also crucial, as the planned items must speak to the goal of developing theory.

I initially prepared a set of questions to serve as ideas during an interview (Appendix K), but I found it to be cumbersome and unnecessary to use during an interview. Having created the questions, and by reviewing them before my early interviews, I found that I was able to easily recall questions on topics that became relevant during the interview, and that I could readily formulate questions in the moment to address emerging lines of inquiry. This became even easier over the course of 35 interviews.

Once the main categories were fairly well established, I made a much shorter interview guide of questions around those categories, but again found the interviews to be more natural and richer when I followed the direction the teacher was headed, returning

to key anchor questions around my categories when necessary. In advance of each interview I wrote down key questions or topics I wanted to be sure to address with that teacher. Continuing with this cyclical process, the interview guide looked different over time. The guide itself became less detailed as it was less necessary to have that support tool, while at the same time the interviews became slightly more structured as theoretical sampling occurred.

Stages of the Interview

Regardless of whether an interview is unstructured or semi-structured, there are a few basic principles guiding the stages of an interview. I began each interview by introducing myself and my topic in broad terms (“I am researching about how K-2 teachers develop their approaches to classroom assessment”). I then explained the informed consent document and the participant’s rights to control his or her level of participation. There was typically some small talk as the teacher and I settled in to start each interview as well, especially in cases where I was able to meet face-to-face with the participant. Particularly to the extent that I could relate as a former elementary school teacher, I established basic trust and began to build a relationship that would put the participant at ease (Rubin & Rubin, 2012).

I emailed the consent form to each teacher in advance of her interview so she had the choice to read through it in advance, but for in person interviews I brought a hard copy to sign as well. For online interviews, I asked teachers to sign and return the consent form via a scanned email. I offered to mail a hard copy with a self-addressed stamped envelope to any teacher who preferred or needed that method of returning the form. I then asked each participant if she had any questions before I began asking my questions.

Except in cases where the conversation naturally began without needing a prompting question (this occasionally occurred when I gave the brief overview of my research study), I began by asking “how would you describe your approach to classroom assessment?” This helped build context before asking follow-up questions that could be difficult or seen as more intrusive (Charmaz, 2014). The second stage of the interview was where the bulk of the questions and rich data generation occurred, starting with open-ended questions allowing the participant to share his or her story and moving towards more focused questions to build out my theory. I took notes during each interview, including jotting down questions I wanted to ask to follow up on key points as a teacher mentioned them. Using my evolving interview guide and notes in the moment, I sculpted each interview to respond to both what the participant was sharing and what my data across interviews was indicating was important.

In the closing of each interview, I inquired, “Is there anything I didn’t ask but should have asked?” (Rubin & Rubin, 2012). In most cases, teachers felt they had thoroughly addressed the important areas of influence on their approach to classroom assessment, but there were also some very valuable conversations that followed this final question, as teachers felt they had more to share or sometimes reiterated the key points of their perspective or experience. As the interview concluded, I thanked each participant and asked for permission to follow up with a second interview, if needed. As I began to reach a stage of theoretical sampling, I realized that I was not likely to need second interviews, and I dropped this question from later interviews.

I also gave teachers an overview of how the rest of my study would conclude to the extent that it was relevant to them: I would be drawing a name for the \$50 Amazon gift card incentive and notifying the winner when all of my interviewing was completed,

and I would also send all participants an executive summary of the study when it was concluded. For online interviews, I notified them that their \$10 Amazon gift card incentive would be arriving by email as soon as the interview had concluded; in-person interviewees received their card at the start of the interview. I thanked each participant for their time and support for my research. I had requested a 45-minute interview, and while a few were as short as 40 minutes, most interviews ended up being 50-55 minutes long, and a few ran past an hour when a teacher had a lot to say and was not concerned about time. The volunteers I worked with were all cooperative, helpful, and appeared to be happy to be heard.

Promptly following each interview, I used my reflexive journal to reflect on the interview process itself, making notes of what went well and what to change or adapt for the next interview. While I reserved the bulk of my memoing to occur while analyzing data for meaning, this initial reflection helped me improve the pragmatics of the interview process.

During each interview, I used an audio recording device to record the interview in addition to keeping notes by hand. Glaser (1978) actually advised *against* recording and transcribing interviews verbatim in favor of taking notes on meaningful statements and ideas, suggesting it helps lead to theoretical sensitivity. Although she agrees that notes alone may develop theory faster, Charmaz (2014) preferred the ability of full transcripts to capture and preserve rich detail that can be lost when taking just notes. For these reasons, and as a novice researcher, I both kept notes during the course of each interview and recorded and transcribed all my interviews, with participants' permission. In early interviews, my notes were more thorough, recording the bulk of what teachers said, but as I was in the stage of theoretical sampling, my notes became somewhat briefer, as I did

not emphasize recording the whole conversation and focused instead on noting the meaningful ideas that would contribute to theory development.

Logistical Issues

Though GT is necessarily open and flexible in its methods, I attempted to plan ahead for various logistical issues regarding data generation while still remaining open to problem solving as unanticipated challenges arose. The primary logistical challenges I faced in this study were related to time, due to using the key GT element of the constant comparative analysis method across 35 interviews while trying to stick to a timeline for completion of the study.

As described earlier, the process of scheduling teachers for interviews impacted my ability to select the sample in a truly individually serial manner. When I sent an invitation to interview, I offered dates typically 3-10 days ahead, and I would wait until the window of offered times had passed before assuming a teacher would not respond and moving on to contact another teacher. In cases where teachers had initially responded in a positive manner, but then for some reason did not continue to communicate, I would send a follow-up email. I did not want to coerce anyone to participate, so I stuck with a guideline of discontinuing pursuing a participant after two attempts to contact him or her to schedule or reschedule a missed interview. In cases where I was traveling to conduct interviews, I attempted to schedule two in a similar location on the same day, which added time to coordinate. Teachers not showing up or postponing interviews was a non-issue for the first half of my study, but as I did more online interviews during the second half, this challenge increased, and it led to a longer overall interview period than I anticipated and busy interview weeks at times as I fit postponed teachers into an already scheduled week.

I define a busy interview week as one with more than three interviews. While I could conduct far more interviews in a single week, the time needed to transcribe, analyze, and use the constant comparative analysis method across a growing set of interviews meant that I had to reconsider my approach to transcribing in order to keep up on busy interview weeks. My initial plan was to record and transcribe interviews using voice recognition software, but I did not find a suitable software program with accuracy that dramatically improved the time spent checking and revising transcripts as compared to transcribing them myself with the simple use of a transcription pedal. This worked at the start of the study when the interviewing pace was relatively slow. However, as I began interviewing more teachers outside of driving range, I relied more heavily on online interviews, which allowed more scheduling flexibility as well as the opportunity to schedule more interviews in a week, but to increase the interview pace without compromising the process, I had to change my approach to transcribing.

As a result, after my fourteenth interview I revised my IRB application to allow the use of an outside transcribing service. By that point in the study, I felt confident that this amendment would be acceptable to most teachers, as the content of interviews was not particularly sensitive or risky. Once given IRB approval, I revised the consent form by removing the phrase, “Only the researcher will have access to recordings of interviews” and adding the phrase “some audio files may be sent to a transcription service that uses encryption to transmit files and utilizes transcribers who have signed strict confidentiality agreements.” None of the following participants had concern with this, and many of the latter interviews were sent to an online transcription service that had a turnaround time of less than 24 hours. This crucial change in my methods allowed me to continue with the clustered method of serial sampling and data analysis.

As much as was possible, I attempted to conduct interviews face-to-face. I believe this to be a better context for connecting, building trust, and reading nonverbal communication. However, there is research to suggest that phone interviews are also effective in gathering rich data in cases that do not deal with highly sensitive or controversial topics (Rubin & Rubin, 2012). Furthermore, I was not working with vulnerable populations, so the need to establish deep trust relationships that can only be done in person was not crucial for my study. I used Zoom video conferencing with participants outside a 2-hour driving distance and, in a few cases, with teachers within that radius when scheduling challenges required it. Video conferencing still allowed me the opportunity to see facial expressions and read body language as well the chance to show my own reactions or responses in the course of conversation. I had one volunteer decline to be interviewed online, but otherwise participants were comfortable with the program and method of interviewing, and in fact several thanked me for introducing them to the technology as they envisioned new professional opportunities for them utilizing Zoom. In addition to the above description of interviews, I conducted online interviews according to Salmons' (2014) guidelines for quality online interviews. Seventeen of my 35 participants were interviewed online.

Other than time, there were minor logistical issues related to recording devices and costs of the study. Due to equipment failure, two of my face-to-face interviews were not successfully recorded. This was not a major problem for the study, as they were early interviews and so I had taken extensive notes which I could then code, and verbatim transcripts are not a mandatory component of data analysis. However, following this incident, I always carried a second device to make a back-up audio recording in case of problems with the primary recorder.

The cost of the study was another logistical issue. To help offset costs incurred for providing 35 Amazon gift cards at \$10 each, plus a single \$50 gift card incentive, driving mileage on the car, and transcribing services, I applied for and received a \$500 Student Research Grant through the Charles Center at William & Mary.

Data Analysis

One of the core principles of GT is the use of constant comparative analysis throughout the analytic process. This involves multiple phases of coding and continual memoing throughout analysis in a constant cycle rather than a linear process, and it is done with the ultimate goal of developing a theory in mind. In the 50 years since the “discovery” of grounded theory, different theorists have used different terminology to describe important analytic terms, such as stages of coding, often to reflect the particularities of their interpretation of the methodology. I use Charmaz’s terms throughout this chapter, although in places I note other theorists’ words for comparable terms in order to avoid confusion.

Data analysis in grounded theory is focused on the two concepts of fit and relevance (Glaser & Strauss, 1967). Charmaz (2014) further explained how these terms relate to the analysis process:

Your study fits the empirical world when you have constructed codes and developed them into categories that crystallize participants’ experience. It has relevance when you offer an incisive analytic framework that interprets what is happening and makes relationships between implicit processes and structures visible. (p. 133)

Constant Comparative

The zigzag process of interview-analyze-interview-analyze guides the comparative process. Although the practicalities of scheduling and conducting interviews prevented me from employing a strict one-to-one pattern of interviewing and analyzing, I was able to follow the essence of this method by clustering small groups of participants in the process of selecting, interviewing, and analyzing. As analyses guided the next group of interviews in a cyclical pattern, different types of coding were utilized over the course of the study as categories emerged and properties were developed.

Phases of Coding

Appropriate coding is the key to understanding what your data shows. According to Saldaña (2011), “Coding is a heuristic—a method of discovery—to the meanings of individual sections of data” (p. 99). Saldaña (2016) also defined coding as “a method of discovery that hopefully stimulates your *thinking* about the data you have been given and have collected” (p. 42). Charmaz (2014) described the process of coding as one in which “you *define* what is happening in the data and begin to grapple with what it means” (p. 113). As I grappled with meaning and my thinking about the data developed, the type of coding that I used to support my thinking process changed.

While different grounded theorists have slight variations in the names they use to describe the phases of coding, Charmaz (2014) used the terms *initial coding*, *focused coding*, and *theoretical coding* to describe the stages in constructivist grounded theory.

Initial coding. Called open coding by Glaser (1978) and Strauss and Corbin (1990), the initial coding stage of Charmaz (2014) calls for using gerunds to capture an action or process. She also recommends moving quickly through the data, as initial codes are rough and should trigger thinking as opposed to box in data:

Compelling codes are also provisional in the sense that you may reword them to improve their fit with data. ...Compelling codes capture the phenomenon and grab the reader. Speed and spontaneity help in initial coding. Working quickly can spark your thinking and spawn a fresh view of the data. (p. 118)

In addition to using gerunds for initial coding, I used in-vivo coding where appropriate to capture the actual language and meanings of the participants (Saldaña, 2016). This flexibility during initial coding was also reflected in the coding units I selected. Rather than forcing line-by-line coding onto transcripts, I aimed to code the smallest units of meaningful data, regardless of length, as some lines warranted more than one code and other ideas described by a single code crossed multiple lines of text. Using compelling codes in this way helped me to think broadly about the meaning in the data instead of just compartmentalizing units of data.

Focused coding. After applying initial codes to an interview, I proceeded to reflect on what initial codes meant and how they compared to each other, a process that focused on possible emerging categories. Glaser's (1978) selective coding and Strauss and Corbin's (1990) axial coding serve similar purposes, but Charmaz's focused coding emphasizes trimming away the excess and making tentative decisions regarding the value of codes. According to Charmaz (2014):

Engaging in focused coding brings you further into the comparative process. . . . Assessing your initial codes involves comparing them with data and distinguishing those codes that have greater analytic power. Thus, you also compare codes with codes and think about the ones that may be promising tentative categories. (p. 140)

Through using a focused coding process, I discovered that some interviews had sections of data that did not directly relate to the research questions and emergent categories. Where appropriate, I continued to apply focused codes to data while still remaining open to using gerunds and in-vivo codes for data not reflected in the focused codes but appearing to be relevant to my research questions. Thus, this list of focused codes changed and grew over time, and as it did, I returned to earlier interviews to look for ways new codes might be applied. The core categories were established after 12 interviews, but properties and dimensions of these categories continued to grow until much later in the interviewing; the last interviews that added new understanding to the categories were the 29th and then 33rd interview.

Theoretical coding. As the study progressed and data started to show coalescence around categories, I began to apply theoretical coding. According to Charmaz (2014), “The purpose of these codes is to help you theorize your data and focused codes. . . . Theoretical codes can help you specify possible relationships between categories you have developed” (p. 150). Theoretical coding is a process of “weaving two major threads in the fabric of grounded theory: generalizable theoretical statements that transcend specific times and places and contextual analyses of actions and events” (p. 113). This process helped lead to theory generation as described in more detail in a following section. I developed the first draft of a theory after the 22nd interview, and from that draft I developed theoretical codes to apply to subsequent interviews. The process shifted the analysis from concrete statements to sense making regarding statements and stories.

These phases of coding were applied both to individual interviews and to the process of interviewing and analyzing over time. Because the constant comparative process calls for a back and forth process of generating and analyzing data, different

coding phases were applied to later interviews than for earlier interviews, as theoretical codes emerged later in the process at a time when initial coding was minimal. In an interview with Gibbs (2015), Charmaz stated that it is not necessary to go back and re-code every interview as new categories and properties emerge. Her justification for this included the notion that “grounded theorists aim to code for possibilities suggested *by* the data rather than ensuring complete accuracy *of* the data” (Charmaz, 2014, p. 120). Five of the final six interviews were coded entirely with theoretical codes, as no new codes emerged with explanatory power.

Memoing

Whereas the reflexive journal was a way to record and reflect upon the entire research process as a co-constructor of meaning, the memoing process was done specifically to focus on thoughts about the data, concepts, and theory development throughout data generation and analysis. I used memoing after each interview and analysis to capture my thinking about codes, important quotes, links among data, connections across categories, concerns about disparate data, or other wonderings that I had during the process. I occasionally referred to Saldaña’s (2016) list of 12 things about which to write analytic memos as a means to encourage my thinking, but I found that the memoing came easily through the analysis and reflexive processes.

I did all of my coding by hand on hard copies of transcripts, a method recommended for novice researchers by Saldaña (2016) in order to immerse oneself in and better focus on the data rather than the software. I used minimal color coding on my transcripts to denote initial (and later, focused) codes, in vivo codes and key quotes, and memos or ideas about the data while coding. After coding a transcript, I made memos from the transcript and analysis in a Word document with entries labeled by participant. I

occasionally jotted memos on the pages I used to take handwritten notes during interviews as well. This allowed me to capture ideas that were spurred by a teacher's comment without disrupting the flow of the interview. Such memos were transferred into the same electronic document in order to organize all of the thinking in one place. I also gave myself the flexibility of adding on to a memo about a previous participant when the constant comparative process brought out a new idea previously unnoticed as well as the flexibility to add dated entries when I wanted to reflect on theory development, but it was not necessarily tied to a particular interview or participant's data.

Birks and Mills (2011) noted that different authors have different recommendations for memo writing, but "all agree that flexibility and freedom are essential to the process" (p. 43). They stated:

Writing is thinking—driving analysis, directing theoretical sampling, with the product a conceptually abstract, elegant and integrated grounded theory. ...For grounded theorists, writing is a fundamental act of enquiry as valuable, if not more so, as generating or collecting, reading and coding data. (p. 130)

Writing served as a very powerful way to record my thinking; by the end of 35 interviews and theory generation, my memos totaled over 45,000 words. Although some researchers emphasize the value of diagramming as a means of analysis (Birks & Mills, 2011; Corbin & Strauss, 2015), not all grounded theorists agree on the need for diagramming. I did not rely on diagramming throughout the whole process, but I found it helpful once I was generating my theory, as it made patterns more visible and understandable (Charmaz, 2014).

Theory Generation

Unlike many other forms of qualitative research that emphasize thick description as the primary outcome, the goal of grounded theory is to produce a theory to explain the data. Birks and Mills (2011) define theory in this context as “an explanatory scheme comprising a set of concepts related to each other through logical patterns of connectivity” (p. 113). While traditional grounded theorists emphasize having a theory with a core category (Glaser, 1978; Strauss, 1987), Charmaz (2014) is more open to other ways in which categories may relate to each other without a single core category, using properties and dimensions of categories to further build out a theory that provides a plausible explanation for the data. The exact form of the output, therefore, should depend on the data, and my study resulted in a visual representation of theory accompanied by written descriptions with storylining. Generating this theory relied on theoretical sampling, achieving theoretical saturation, and the use of a model to support the theory, as described below.

Theoretical Sampling

I moved into a process of theoretical sampling to specifically refine and elaborate the emergent categories once their properties and dimensions appeared to be stable across interviews. In contrast to the maximum variation sampling method used to initially gather a wide range of experiences, the theoretical sampling process called for intentional selection of participants and focused questions as ways to try to further insights around the themes that had already shown themselves to be salient. I prioritized interviewing teachers from new schools and districts, as I knew after the twelfth interview that these contexts were important, while still looking for ways to maximize the variation in my participants based on other experiences. This allowed me to better fill out categories and

see properties or relationships that might not have been readily apparent in a more homogenous sample. For example, after talking to a teacher who described having her own children as being a key to her change in perspective on child development, I chose a participant who did not begin teaching until after her children had grown to further explore this dimension of personal development.

I did not aim to “test” my theory or categories against a representative sample, as generalizability is not the purpose of theoretical sampling (Charmaz, 2014). Charmaz described this as a mistake often made:

Perhaps researchers’ most common error results from confusing theoretical sampling with gathering data until the same patterns reoccur. This strategy differs from theoretical sampling because these researchers have not aimed their data-gathering towards explicit development of *theoretical* categories derived from analyses of their studied worlds. Instead, the patterns describe empirical *themes* in their studied worlds. (p. 199)

Theoretical sampling can happen across the course of the study and not just towards the end of data generation. According to Charmaz (2014), “theoretical sensitivity can also turn an unexpected moment during an interview into an occasion for theoretical development. Thus, opportunities for theoretical sampling may occur without being planned in advance” (p. 104). Once the main categories emerged, I was able to ask questions around these categories, leading to a credible, plausible theory from the data, with saturation as a consequence of this kind of theoretical sampling (Charmaz, 2014).

Theoretical Saturation

When Glaser and Strauss (1967) first described theoretical saturation, they emphasized that it is not when the same stories, events, or patterns are repeated, but when

no new properties of theoretical categories are emerging. The goal is not data saturation, but saturation of the theoretical categories, their properties, and the relationships among categories. Therefore, I continued to interview and analyze until new data was no longer adding to the theoretical categories. This happened with interviews 22, 25, 27, and 29, but the interviews in between provided new properties. As declaring theoretical saturation requires a judgment, I decided to continue interviewing and spoke with teachers in three previously unreached districts from the 30th through the 35th interview. Of these, only one added any relationship to the categories in the theory, and I believed my thorough, thoughtful memoing process supported my decision to stop interviewing, as there was adequate sampling to support a fully developed theory. In each of these last six interviews, I started with a few open-ended questions, and then I explained the emergent theory and asked participants to reflect on how they saw themselves in it or how they might revise the theory based on their perspective. While each participant had a slightly different set of personal experiences as they described their individual development, they unanimously supported the applicability of the explanation I provided.

Model

A well-developed theory can often be displayed in a visual that helps the reader understand the concepts presented and relationships among categories and the properties within categories. Strauss (1987) stressed the importance of a model, while Glaser (2005) stated that the thinking and narrating of relationships can suffice. Charmaz (2014) offered diagramming as a means of displaying abstract ideas in concrete images, while Birks and Mills (2011) contrasted a diagram and model in GT by stating “that the former is a strategy for analysis, while the latter is a strategy for presentation” (p. 140). Regardless of

what the visual is called, I have used a model to help explain my theory, presented along with supporting narrative, in Chapter 4.

Trustworthiness

While the post-positivist paradigm in which most quantitative research rests calls for adherence to standards of validity and reliability, Lincoln and Guba (1985) reframed these concepts for the constructivist paradigm as trustworthiness and credibility. They stated, “criteria for judging an inquiry themselves stem from the underlying paradigm” (1986, p. 76). When the naturalistic paradigm—later termed constructivist paradigm—emerged, there was a need not for verification procedures, but for standards that would guide the trustworthiness of qualitative findings. As a way of both differentiating between the values of the post-positivist and constructivist paradigms, and as a means to counter claims that qualitative research was less rigorous than the validity-tested and reliability-tested procedures of quantitative research, Lincoln and Guba (1985) proposed four criteria as alternate but parallel measures for judging the trustworthiness of constructivist research: credibility as parallel to internal validity, transferability as parallel to external validity, dependability as parallel to reliability, and confirmability as parallel to objectivity. Given these measures, Lincoln and Guba (1985) claimed, “there is no reason to believe that humans cannot approach a level of trustworthiness similar to that of ordinary standardized tests” (p. 195). This section explains how my study addressed each of these four characteristics of trustworthiness.

Credibility

Credibility has been offered as a qualitative parallel to internal validity, and Guba and Lincoln (1989) called member checks the “single most crucial technique for establishing credibility” (p. 239). However, for GT, Charmaz (2014) suggested that this

process is redundant because the constant comparison process will cause interpretations to stick if they work and to fall away if they are not emerging across data. Using focused coding, theoretical sampling, and theoretical coding bolstered credibility because the interpretations that had merit continued to emerge and develop, while those that were less reasonable did not rise to a higher level of analysis. Creswell (2014) also recommended triangulation as a way to bolster credibility, but Guba and Lincoln (1989) moved away from this term due to its implication of a positivistic stance, stating that efforts to build credibility “ought to be dedicated to verifying that the *constructions* collected are those that have been offered by respondents, while triangulation should be thought of as referring to cross-checking specific data item of a factual nature” (p. 241). The constant comparative analysis and theoretical sampling served this purpose to verify constructions, as documented through my extensive memoing and substantiated by having my final six interviewees respond directly to my theory in strong support of the constructions it presented.

Transferability

The idea of transferability is based on contextual factors. According to Guba and Lincoln (1989), “transferability is always relative and depends entirely on the degree to which salient conditions overlap or match. The major technique for establishing the degree of transferability is thick description” (p. 241). Thick description is a term coined by Geertz (1973) to describe in depth the cultural situation of behaviors and actions in an ethnography. However, the use of thick description is important in all qualitative research, as it allows the researcher to clearly explain the circumstances of their data so that the reader can determine the transferability of findings to his or her own context. Lincoln and Guba (1986) defined thick description as “narrative developed about the

context so that judgments about the degree of fit or similarity may be made by others who may wish to apply all or part of the findings elsewhere” (p. 77).

In Chapter 4, and to a lesser degree Chapter 5, I have used thick description to explain the context of my data, although the outcome of constructivist grounded theory shifts the emphasis of transferability slightly. Rather than claiming that the theory might be applicable to other contexts and providing sufficient description of the context of this study for the purpose of the reader determining transferability, my thick description is primarily for the purpose of lending credibility to the application of this theory *within its own context*. In her interview with Gibbs (2015), Charmaz stated that grounded theory is not trying to develop grand theory that applies across multiple contexts, but it is attempting to present a theory to provide a plausible explanation of the generated data within the study. Therefore, I am cautious to make claims of transferability (Lincoln & Guba, 1985).

Dependability

Much like there can be no validity without reliability, there can be no credibility without dependability (Guba, 1981). The concept of dependability within qualitative research is “whether the results are consistent with the data collected...rather than demanding that outsiders get the same results, a researcher wishes outsiders to concur that, given the data collected, the results make sense—they are consistent and dependable” (Merriam & Tisdell, 2016, p. 251). Due to the nature of constructivist grounded theory in which the theory develops over a period of data generation and analysis through the constant comparative process, it is expected that findings will evolve rather than be static, but this does not undermine dependability. According to Guba and Lincoln (1989), “methodological changes and shifts in constructions are expected

products of an emergent design dedicated to increasingly sophisticated constructions. Far from being threats to dependability, such changes and shifts are hallmarks of a maturing—and successful—inquiry” (p. 242).

Through use of memoing and my reflexive journal, I documented how my thinking and the theoretical categories emerged and developed over time. The dependability in my study does not come from the consistent nature of the data or the steady and uniform categories, but the dependability of the process through which I gathered and analyzed data. Documenting my thinking and decision making over time allowed me the opportunity to trace lines of thinking and make course corrections, when needed, and also to show a steady and consistent process of analysis and theoretical sensitivity that brought dependability to my study.

Confirmability

The use of memoing and the reflexive journal were also keys to strengthening my study’s confirmability, the constructivist’s counterpart to objectivity. As Guba and Lincoln (1989) stated, “unlike the conventional paradigm, which roots its assurances of objectivity in method...the constructivist paradigm’s assurances of integrity of the findings are rooted in the data themselves” (p. 243). Through these processes, and through the descriptions included in explaining the theory, I have worked to show my thinking, record decisions made and actions taken, and illustrate how the theory emerged from the data. Being able to show how my interpretations are grounded in the data, much like an audit trail, provides confirmability. Confirmability in this context does not mean the findings have to be confirmed by outside sources as an absolute single “right” theory, but that they can be confirmed as plausible explanations for the theoretical categories and relationships described (Gibbs, 2015).

Ethical Considerations

While the topic of this study is not particularly sensitive and was highly unlikely to cause any harm, there were still a few ethical considerations as with any research study using human subjects. Maintaining transparency in how participants were identified and how I gained access to participants was an important element in preserving high ethical standards, as was the manner in which I interacted with participants and protected their identities. When the researcher is the instrument, there is little protection from anonymity (Lincoln & Guba, 1989), but strict confidentiality can be maintained. The database of possible participants was kept on a password protected computer that remained in the researcher's possession at all times, and any other potentially identifying documents or information were also saved digitally in password-protected sources and, when in hard copy, in a locked private residence when not in the researcher's possession. Once secondary documents were created (interview transcriptions, memos, and final narrative descriptions), pseudonyms were used for participants to further protect their identities, and these documents were similarly kept in secure locations only accessible to me. Where necessary, other potentially-identifying information, such as names of school districts, have also been substituted with pseudonyms in this study.

There was no deception employed at any point in the study, and I maintained transparency with school district representatives and participants regarding the research process and how interviewees' words would be analyzed and interpreted. Participants were permitted to drop out of the study at any point, and as I had no authority over participants nor were their principals or other district administrators made aware of anyone's decision to participate or not, there was no obligation to join the study. The \$10

gift card incentive was a token of appreciation, but not substantive enough to coerce participants to join.

I received IRB approval in accordance with standard procedures for research with human subjects before starting any data collection. The study was eligible for an expedited review, and when I amended my application to permit the use of a transcription service, the revision was quickly approved as well. I communicated to all participants at the start of the interviews that the study was approved by the IRB, I explained the informed consent form to participants prior to them signing (Appendix L), and I asked if there were any questions before we began. Participants were given a consent form to save and I kept the signed copy, scanned the document to preserve an electronic version in a secured location, and kept the hard copy in a locked private residence. Thirty participants provided informed written consent and the other five participants provided oral consent after having received the written consent form in advance of an online interview and being given the opportunity to ask any questions.

Delimitations, Limitations, and Assumptions

There are a number of delimitations, limitations, and assumptions that bound this study. My primary consideration was clearly delimiting the focus of the study to teachers' reported influences on their approaches towards developmentally appropriate classroom assessment practices. The study was not intended to define or validate best practices, nor evaluate teachers' practices against any standards, although the findings have illuminated some practices that are utilized in K-2 classrooms. The focus was on strong influences and teachers' processes of conceptualizing developmentally appropriate classroom assessment within their school contexts, so it was not necessary to observe actual practices or students at work.

Furthermore, the definition of classroom assessment provides important delimitations. This study was not focused on assessment practices beyond the control of the classroom teacher (e.g., not looking at standardized testing, district-mandated assessments, screening assessments, or commercial diagnostic assessments, except perhaps in how the teacher uses or ignores these assessments.) However, as is discussed in chapters four and five, many teachers' reported conceptualizations blurred the lines between these types of assessment in the classroom. I was cautious not to constrain answers when teachers' responses showed an interpretation of classroom assessment as something different than the construct I described in Chapter 2, but I used follow-up questioning to attempt to uncover a deeper understanding regarding both classroom practices and teachers' ideas of what "assessment" means.

Other delimitations include the focus on a single state. It is not known if this delimitation had a significant impact on the findings, but I chose to focus on public school teachers within one state in order to focus on one set of licensure requirements and education policies, should those be influential.

Moreover, I did not try to validate what teachers self-reported. I trusted that what they said about their beliefs, practices, and influences was their best representation of how they view their beliefs, practices, and influences. The process of co-construction requires mutual trust, and I chose to rely on the information provided to me as trustworthy representations of teachers' emic constructions. It is also possible that by using self-reported data, the data set may not fully represent reality as viewed through other lenses.

As already discussed, there were several logistical considerations that impacted the study. These included the need to gain access to a large number of districts. The

limitation of timing for applying for permission to research impacted which districts I could target for inclusion. It is unknown if there were certain characteristics that made some districts more or less likely to participate, and if any such characteristics existed, whether those could have uniquely shaped the experiences of my participants. Similarly, because all participants were volunteers who essentially self-nominated themselves for participation, there could have been something unique about teachers who volunteered for a study about assessment in the early years, or something that led some teachers to choose not to participate.

For example, a teacher who may not feel a high level of confidence in his or her assessment skills might have been reluctant to volunteer, or a teacher who believes she faces barriers and wanted to illuminate challenges may have been more likely to volunteer. I noted in my memos that as a whole, teachers overwhelmingly spoke well of their current principals but had negative experiences with previous administrators. Since I chose to trust teachers' representations of their views, I wondered whether teachers currently working for a principal with whom they had negative experiences were less likely to participate, perhaps due to concerns that their thoughts would get back to their administrator. While I cannot state with certainty whether this occurred, I felt comfortable that the resultant theory was not dependent on the circumstance of having a strong principal or instructional leaders, as there were a few teachers who did not report positive trust relationships and the descriptions of teachers' previous experiences were also part of the analysis for theory generation.

This research approach relied on a number of assumptions. There was an assumption that teachers would be able to self-report their perspectives, experiences, and practices in meaningful ways to an unknown interviewer. There was an assumption that

my declaration of having reached theoretical saturation was a reasonable one and that, despite external time pressures, the decisions that steered the process were made based on the unfolding process and data, as documented in the memoing and reflexive journal. Finally, there is an assumption that even as a novice using this method, I was able to meaningfully employ the complex process of GT. This assumption is supported by the originator of GT who claimed, “GT is done best in the hands of the novice PhD and MA candidates” (Glaser, 2009, para. 2), as these candidates are more open to emergent theory than researchers rooted in a particular viewpoint, framework, or set of experiences that might unnecessarily limit their ability to see fresh patterns or seek relevancy with the same ambition as unformed researchers.

Summary of Methods

The constructivist grounded theory approach required flexibility to adapt methods, questions, and analysis processes in order to respond to what the data showed as being relevant and meaningful to the research questions. Not all GT studies follow the same steps, as strictly following a cookbook style recipe is counter to the goals of grounded theory (Charmaz, 2014). Nonetheless, I started with a guide that aligned the planned actions with the questions, methodology, and paradigmatic approach, and I used extensive documentation of my decisions and thinking throughout the process to show when and why I chose to make adjustments to my plan. While generating and analyzing data leading towards theory development, I kept in mind Saldaña’s (2011) statement as a guide: “Good research is not so much about good methods as it is about good thinking” (p. 119).

CHAPTER 4

FINDINGS

The purpose of this study was to address three questions regarding how teachers come to their conceptualizations regarding classroom assessment and developmentally appropriate practices. Those questions are as follows:

- (1) Through what processes and influences do experienced K-2 teachers report that they have developed their *approaches to assessment*?
- (2) Through what processes and influences do experienced K-2 teachers report that they have developed their knowledge and implementation of *developmentally appropriate practice*?
- (3) How do experienced K-2 teachers' approaches towards classroom assessment and philosophical orientations towards DAP interact with each other?

The study used a constructivist grounded theory methodology, applying a constant comparative analysis process to teacher interviews for the purpose of creating a theory to explain the influences on teachers' development. Through this analysis, many themes and implications emerged. This chapter focuses specifically on the themes that rose to the level of explanatory categories that helped produce a theory. A single basic theory can be used to address the first two research questions, and a separate discussion of relevant themes illuminates answers to the third question. The theory is presented in this chapter as the primary findings of the study, and a discussion related to the themes surrounding the third research question is found in Chapter 5.

Before describing the theory, it is important to address the overarching ideas of teachers' conceptualizations in order to properly frame the theory. The chapter begins by clarifying distinctions between teachers' perspectives and the original perspective of this study. Next, it explains teachers' perceptions of the purposes of assessment. Then, using both visual models and supporting narrative, the chapter depicts the overarching theory of influences on teachers' approaches towards classroom assessment and developmentally appropriate practice. Finally, it offers some examples of specific participants' fit within the theory. The discussion in Chapter 5 addresses these themes directly within the context of the guiding questions.

Framing Teachers' Perspectives

Over the course of 35 interviews, one of the earliest and clearest themes to emerge was that the perspective of *approach* as intended by the research and described in Chapter 2 and the interpretation of *approach* as described by teachers were different. Unless conversation flowed naturally without an initial probing question, I began every interview with the question, "how would you describe your approach to classroom assessment?" The majority of teachers responded both with a "we" statement rather a description of an individual approach, and the majority of teachers responded by describing their *practices*, not a *philosophical orientation* towards assessment. This distinction is further discussed in Chapter 5, but it is important to note before describing how teachers perceive the purposes of assessment and how they describe the influences on their *approach* towards assessment.

Limitations

It is also important to remind the reader of the limitations of this study as I present its findings. This study was bound within a single state's public school system, and some

of these themes may not be as crucial in a different context. Also, where I use the term “teachers” throughout the chapter, it is important to remember that this is used in reference to K-2 teachers, not teachers of any grade level, as this was the target for the study. Finally, recall that the intent of grounded theory is to provide an initial, plausible theory to explain the data collected within the study’s parameters, not a grand theory that claims applicability to a broader set of applications. Therefore, use of the term “teachers” refers specifically to the K-2 teachers who participated in this study and is not intended to be generalizable to all K-2 teachers.

Construct of Classroom Assessment

The definition of classroom assessment (CA) provided in Chapter 2 by McMillan (2013) is not commonly held by this study’s teachers as a distinct construct from *assessment* in general. The portion of McMillan’s (2013) description of CA that reads, “a process that teachers and students use in collecting, evaluating, and using evidence of student learning for a variety of purposes” (p. 4) holds true, although the “variety of purposes” portion is crucial. For this reason, the next section of this chapter describes the purposes teachers perceive assessment to have. Where McMillan (2013) distinguishes CA from “large-scale or standardized, whether standards-based, personality, aptitude, or benchmark- or interim-type tests” (p. 4), teachers did not make this distinction. This is also discussed in Chapter 5, but it requires specific understanding in order to make sense of the theory of teachers’ *reported* influences on their *approaches* towards *assessment*.

Rather than changing my research questions, I used the data generated to learn more about the distinction between teachers’ conceptualizations and those held by many researchers, an unanticipated finding, but one worthy of exploration.

Developmentally Appropriate Practice

While many K-2 teachers spoke about child development and things that were appropriate or inappropriate for young children, teachers did not talk about developmentally appropriate practices as a framework or set of understandings that guides their work with their students. The three core considerations of DAP, mentioned in Chapter 2, highlight what is known. There is an emphasis on knowledge along with philosophical orientation that is the foundation of practices. The way teachers spoke about this knowledge and the challenge of acting upon it in a school context reflected the research of Van Horn and Ramey (2004) that showed there can be a lack of alignment due to the nested nature of teachers' practices within school settings. For most teachers in this study, there was a level of implicit understanding of DAP, but no explicit professional discussion or focus on it. As a result, woven through teachers' descriptions of their practices are interjections of statements about developmentally appropriate practices, but it was not treated as a separate construct driving their work like assessment often was. In the following description of the theory, more emphasis is placed on the influences on assessment than on the influences of DAP because the teachers described a practice that was much more driven by assessment practices than guided by DAP considerations.

Different Vocabulary and Terms Used

In addition to having a different working definition of assessment, teachers often used educational jargon in mixed, colloquial, or even inaccurate ways. Some of these are described here for the purposes of better understanding quotes when they are provided. As noted in Chapter 3, in the state of Virginia, a Local Education Agency as referred to as a "division." Some teachers who previously worked outside the state still use the term

“district.” At other times, teachers use the term “county,” “central office,” “downtown,” or “school board” to describe the administrative structure that oversees schools. When directly citing a teacher, I use the term she used, but otherwise I use “district” throughout this study.

Another term unique to the state of Virginia is the Standards of Learning. Commonly referred to as the SOLs, the subject-based standards include minimum expectations for content knowledge and skills for students at the end of each grade level. Virginia is one of four states that never adopted the Common Core State Standards, and although the SOL program has been revised over time, it was first adopted in 1997, with SOL testing having begun in 1998. From the program’s inception, the state set goals regarding student test performance on SOL tests, but the high stakes associated with testing did not begin to amplify until after the passage of No Child Left Behind. Teachers often used the term SOL not just to describe the program of studies or specific sets of standards, but also to describe the SOL tests and sometimes to refer to a grade in which the SOLs are tested (calling third, fourth, and fifth grade “SOL grades”), even though the SOLs themselves begin in kindergarten.

A final, but perhaps more concerning difference in term usage is one that seems to have either originated at or at least been perpetuated by school districts: a misapplication of the terms *formative* and *summative* assessment. Although Popham (2008) has stated, “there is no single officially sanctified and universally accepted definition of formative assessment” (p. 3), researchers and scholars agree that the purpose or use of an assessment, not its form, determines whether it is considered a formative or summative assessment (Black & Wiliam, 2007; Brookhart, 2013; Gareis & Grant, 2015; Marzano, 2010; Moss & Brookhart, 2009; Popham, 2008). As I used McMillan’s (2007) definition

of classroom assessment in Chapter 2, I here use McMillan's definition of formative assessment: "a set of skills and activities that are undertaken by teachers to provide feedback to students to enhance their motivation and learning by designing instruction to meet student needs" (p. 1). In contrast, McMillan (2007) defined a summative assessment as one "conducted mainly to monitor and record student achievement, and is used for school accountability" (p. 1).

Teachers in a number of the 15 districts included in this study used the terms not as McMillan (2007) did, but rather to mean assessments of different weights in grading practices. When asked to clarify the difference between summative and formative assessments, Claire said the distinction was "the breadth of what we're assessing," a common usage of the term that is in contrast to the definition used by McMillan (2007), but one that aligns with how some districts manage grading expectations. I use the teachers' own terms in quotations with acknowledgment that these usages are not always precise. The Chapter 5 discussion addresses the "formative" assessment situation further.

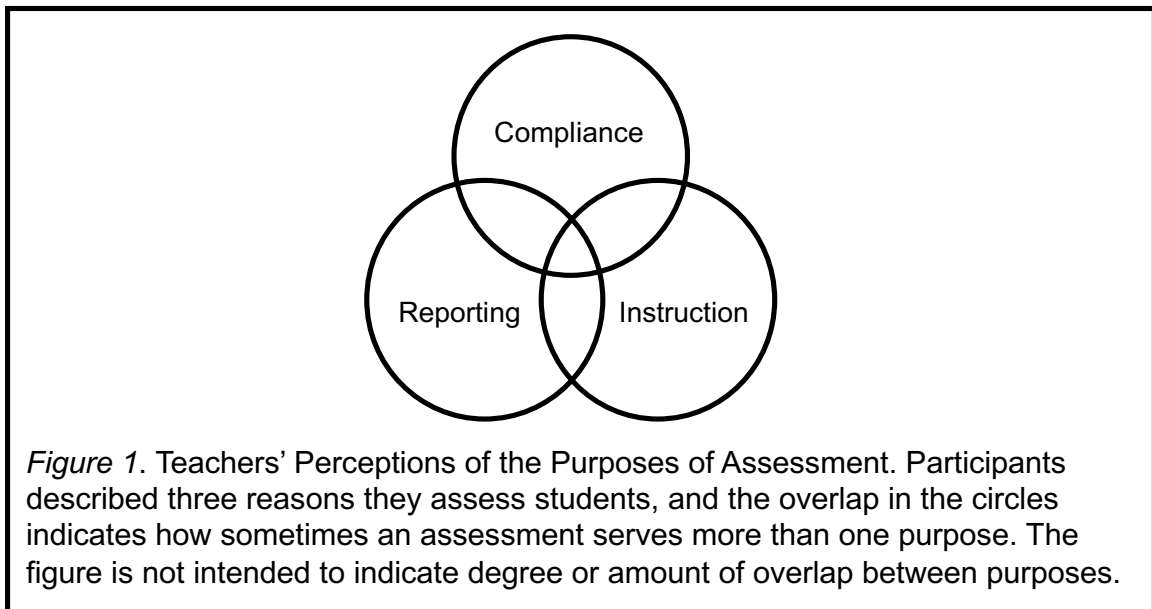
To improve clarity, where the context makes a difference, I used the terms small, mid-sized and large to distinguish school districts by using small to describe a district with a single elementary or primary school, large to describe a district with more than 20 schools, and mid-sized to include anything between these extremes. I also describe teachers' experience by using the terms "early-career" to describe a teacher with 4-9 years of experience, "mid-career" for teachers with 10-17 years of experience, and "veteran" for a teacher with 18 or more years of experience. All teachers are referred to by pseudonyms to protect their confidentiality.

Perceptions of the Purposes of Assessment

Assessment is often described in three categories based not on its format, but on its use: assessment *of* learning, assessment *for* learning, and assessment *as* learning (Volante, 2010). While these distinctions may be helpful from a theoretical perspective, teachers in this study described assessment as a practice which is undertaken for three primary purposes: assessment for *compliance*, assessment for *reporting*, and assessment for *instruction*. This section describes each of these three types and the relationships among them.

Description of Figure

While different teachers described a different balance of the three purposes of assessment, these three can best be described as circles of overlapping purposes, as shown in *Figure 1*.



Depending on each teacher's context and personal experiences, the relative weight and amount of overlap of each of these circles is different in practice, but the

meanings of them are consistent. Below I describe each of the three purposes of assessment portrayed by teachers.

Assessment for Compliance

Most teachers at some point described assessment practices that they do simply because they were required to do so. As veteran teacher Beth stated,

It's not really up to the teacher anymore. We are given so many assessments that we have to do, the thought of creating my own assessments or looking at something else is kind of like, I can't even think about it.

Cyndi, another veteran teacher, was very direct about why she does certain assessments: "We do assessments because admin needs the data. I can look in my classroom and with teacher observation, I know who's doing what, I know who needs to be grouped where."

Some teachers felt there was space for their own practices, but that they were a separate category from the required assessments: "There's what I'd like to do and then what I have to do," said first grade teacher Barbara. Although none of the participants described themselves using the word, "compliant," they collectively painted a picture of professionals who do what they are told to do. As Jill emphasized, "I hate to say it: 'I'm a little puppet and, okay, we'll get it done.'"

This perspective of assessing because they have to was often at odds with teachers' preferences of what is best for students. Early-career teacher Chantal explained,

I feel like most teachers care to some extent if something is developmentally appropriate or not, even if they're rule followers and they really are like, "Well, but this is what we have to do. This is what the county is saying."

She was one of only two teachers who specifically stated her support of accountability measures, but she also stated, "I think there's a lot of ways you could trace that and have

teachers prove that and measure that without forcing them to do everything...a certain way.” Whereas district administration may view these assessments as necessary for accountability purposes, teachers almost universally did not view these kinds of assessments through a lens of accountability, but rather as a matter of compliance, and did not place much value in this category on its own.

Beth was more direct about it when she admitted, “I tell my principal all the time, ‘I will do whatever you tell me to do, but not ‘til I say how I feel about it first.’” Bridging the gap to the next purpose of assessment described here, second grade teacher Pat also displayed a similarly compliant, but still empowered attitude as did Beth:

I’m not going to do anything behind anybody’s back, but I’ll let you know what I’m going to do because [assessment practice] is not fair. ...I still have to do what the county says I have to do, but I’m going to give them a comparable assessment because the other one’s going to tank their grade. I have to give them something that at least erases that ridiculous F.

Assessment for Reporting

The second purpose of assessment that teachers described was assessing for reporting. As Kim explained, she is prepared,

if a parent would come to me or because the people from the county will sometimes show up and they’ll come through to see what we’re working on. . . .
It’s just so we can prove what we’re doing.

Teachers express more of an understanding of the value of this purpose for assessment compared to assessment just for compliance. First grade teacher Karla has chosen to give spelling tests specifically for this purpose, as she acknowledged that some parents need to see more traditional tests and they do not understand things like checklists of skills.

“We’re still accountable to parents and the district,” she asserted. This kind of accountability seems more necessary to teachers than assessments they do simply for compliance without any other stated reason.

Anne and Janice teach kindergarten in the same district and talked about the importance of having assessment records to share with parents, although Anne’s perspective was more about connecting with parents and building relationships and Janice’s perspective was more about having evidence to support your grades. Grading practices weighed heavily on the category of assessing for reporting, and the manner in which report cards were designed often shaped how teachers collected assessment data. For example, in a school district that uses a 4-point grading scale on its report card, Claire reported using the 4-point system to design grading rubrics for her second graders’ writing.

Requirements for reporting. Requirements around reporting practices were also prevalent across most schools and many districts represented in the sample. There was no differential whether teachers taught in a small, mid-sized, or large district: many, but not all, detailed a set number of summative grades and formative grades required for different subjects for each reporting period. Jill stated that her small district requires three summative and seven formative grades each for reading and math every nine weeks. Michelle currently teaches kindergarten in a mid-sized school district and stated that she was required to have a certain number of assessments registered in an online gradebook for parents to see, but she also expressed amazement that these numbers were not consistent across schools in her district. Chantal teaches second grade in a large district and stated that, while she has some flexibility regarding “formative” assessments she uses for grades, she also is required to have nine grades per nine weeks for the core subjects.

Kim also teaches second grade in the same district as Chantal, but in a different school, and said, “The rule of thumb is—you can’t find it written anywhere, but—what the principals are looking for is you need to have two or three formative for every summative.”

In this context, teachers’ use of the terms *summative* and *formative* are not the meanings used by McMillan (2007). Even the terms *grade* and *assessment* were often used interchangeably by participants in this context. The descriptions given to summative grades or assessments were somewhat larger-scope, end-of-unit tests that cover completed learning; in essence, these mostly aligned with McMillan’s (2007) definition. Formative grades or assessments, on the other hand, were described as shorter tasks that addressed smaller portions of the content along the way to a unit’s final test. Essentially, they are what is often considered a “quiz.”

But by labeling them formative assessments, districts are either requiring grades to be derived from learning and practice activities in which risk-taking and mistake-making should be encouraged rather than penalized (J. Chappuis, 2014; Popham, 2009), or districts are devaluing true assessment for learning by giving the implication that they are doing this practice when in reality they are only requiring shorter summative assessments covering small chunks of expected learning (quizzes labeled “formative assessments.”)

Concerns with reporting practices. The purpose of this study was not to survey, document, or evaluate the range of reporting practices used across the state, although it is important to note there were many teachers who described concerns with the methods of reporting used in their school district. Only one teacher expressed concern with how the terms summative and formative assessment were used by their district, suggesting that

perhaps many teachers are unfamiliar with scholars' conceptualization of these terms and instead default to the language used in their context. But teachers did express concerns with report cards that many purported failed to adequately communicate to parents, align with assessment practices, or appropriately represent early learning goals and early learner characteristics.

Kindergarten teacher Nancy described a practice in which she mines grades from learning activities her students are already doing. She frequently spoke about "taking a grade" when she recorded data from her students' work. She described how she could "take a math grade and a history grade and a writing grade" from a single craft activity in which students recreated an American flag with construction paper and stickers by looking at their patterning in the stripes, examining their use of appropriate colors in the appropriate places of the flag, and studying how they wrote their names on their work. The manner in which she collects grades in authentic settings without adding tasks and formal assessments works well for young children, Nancy claimed. However, using the term "grading" for a practice of gathering evidence of learning points to its primary purpose as assessment to meet reporting needs rather than for guiding instruction, as in the third purpose teachers described.

Assessment for Instruction

Nearly every teacher at some point described some way in which she interprets student work to guide her instructional planning, but not all of them viewed these practices as *assessment* in the same way they viewed the more formal situations typically used in assessments for compliance and assessment for reporting purposes. Teachers described using quick, informal practices like exit tickets, observations, asking questions while students work, looking over completed papers, using whiteboards, and even

reading body language to determine in-the-moment instructional adjustments. But they also described using more formal district tests to inform their instructional planning, such as Pat did when her students performed poorly on a unit test and she adjusted her curriculum pacing accordingly to reteach some concepts.

Regardless of the format, early-career teacher Maureen described her primary purpose for assessing is for herself, stating, “I’m just doing assessment ‘cause I want to know where they are myself. I’m always like, ‘okay, what do I need to do next? What needs to change?’” Veteran teacher Annabelle spoke about the value of assessments to create and adjust differentiated instructional groups, a practice echoed by many others in the study. Veteran teacher Margaret admitted to a very informal process of noticing her students’ performance and adjusting accordingly:

Most of the time it just gets registered in my brain because I’ve tried to be that write-down-everything person and I’m not that person. But I have a mental note of who I need to be close to, who I need to watch, who needs to be sitting on the front row so I can get constantly observing. . . . It just depends on the vibes I’m getting from the kids.

Veteran teacher Betty described instructional purposes as being an important guide in determining the value of other assessments: “Any type of assessment needs to show you the progress and the understanding the child is making and...what their needs are after that and what your teaching should be.” When her small district piloted the Virginia Kindergarten Readiness Program (VKRP) this year, they determined it was not an assessment “that helped inform our teaching. . . . It was just an extra. So, we kinda threw it out for now.” Teachers gave the VKRP out of compliance because the district piloted in, but they tried to use it for instructional purposes. This cross-purpose look at

assessment illuminates the overlap of the three spheres of assessment purposes, as indicated by the Venn diagram in *Figure 1*.

Assessment for Multiple Purposes

While the figure implies that each of the three purposes teachers report for assessment are of equal weight and have only slight overlaps, the actual relation among the three distinct purposes differs depending on the contexts of district, school, and teachers. Some teachers reported maximizing the instructional value of required assessments by using the results to inform instructional groupings or by using the results of assessments given for instructional purposes to inform report card grades. The most commonly shared example of an effective overlap of purposes was the Phonological Awareness Literacy Screening (PALS). The PALS is required in Virginia public schools for K-3, and thus for the sake of compliance all teachers must give it. At the same time, many teachers spoke highly of the standardized test and its associated, typically optional “quick checks,” as they provide valuable information for forming instructional groups and identifying students in need of additional instructional support, such as is done through the Response to Intervention (RtI) framework.

Veteran teacher Maggie described a practice she began when working in a different mid-sized Virginia district out of compliance. She was required to come up with some sort of process to review skills and measure progress, so she created a daily homework sheet routine that allows students to practice previously-taught skills. Though she started using it as a requirement, Maggie found it to be such a valuable part of her instructional practice that she brought it with her to her current mid-sized district and remains adamant about going over her kindergarten students’ daily homework with them one-on-one.

Other reported practices that overlap purposes of assessment include when required assessments (such as district-wide summative unit tests) are also used for reporting or when running record results are used for both reporting students' reading levels and determining instructional level for reading groups. At times, though, the purpose can be blurred and uncertain to teachers.

Seeking clarity of purpose. Veteran teacher Claire emphasized the importance of having a formal process for gathering information on your students while also being clear as an educator why you are assessing them:

What is the purpose of the tests? That always has to be your focus. Like, why are you testing them? I mean, obviously, you did so that you can have a concrete grade so that you can say, "yes, Charlie does understand this," or "no, Charlie doesn't understand that." It's not just your feeling, like, "I don't think Charlie gets it." Although pretty much a good teacher would say, "I don't think Charlie gets it." You really don't need to assess them. But at least it gives you something.

Implied in Claire's statement is that assessment is something *different* from the process by which a "good teacher" knows her students' abilities. That process may be implicit and undocumented rather than formal and structured, and that can lead to the uncertainty regarding the purpose of a particular assessment task. Early-career teacher Pauline also asked a similar question while stating that teachers know their young students' needs without formal assessment processes:

I feel like we test so much for administration versus our kids and I think we need to make sure that the assessments we're giving are actually benefitting our kids versus just a grade to put in the gradebook. I mean, I can tell you where all my

kids are academically and which skills they need help on, but then getting that assessment—like, what is its purpose?

Problematic overlap. Sometimes an assessment that should have “crossover” value is prevented from serving dual purposes. Maya teaches at a mid-sized district and talked about how changes in reading running record practices changed the purpose of these assessments in this way: “It used to be that we do them when we felt like the kids were moving and ready. Now, we have it scheduled. The county scheduled it for us. It’s about every one and a half months or so.” Rather than being able to maximize the instructional purpose of giving a running record, by having a set timeline for compliance that is disconnected from students’ growth and learning, Maya found that the time intensive nature of the assessment and its required frequency meant that she could not be as responsive to assessing students when she determined it valuable to their learning.

Other teachers wanted to make it clear that they did not use formal assessment results for instructional purposes. After describing the required assessments (that she gives for compliance purposes), second grade teacher Bev clarified, “I really don’t drive my instruction by all this. I know we’re talking a lot about it, but I guess that’s today’s world.” It is understandable that Bev would want to ensure I did not interpret her practice as one in which her instruction was driven by high-stakes testing, as this is a concern conveyed by multiple teachers. Karla teaches in a large district and articulated the dilemma in this way: “a big part of the assessment problem is because the assessments are driving our instruction. We’re not figuring out what they actually have to learn and then how do we measure that.”

Mid-career teacher Laura illustrated this situation in her discussion of practices at her school. She stated, “We rely a lot on the PALS to form our instructional groups to see

who gets extra help from our reading specialist. . . . we work towards the PALS goals.” By setting instructional goals based on assessment results, this school has possibly inadvertently taken an assessment that has strong instructional value and made it *both* the means to and the end itself. While the highly validated PALS is an example of an assessment that aligns well with the reading skills it sets out to measure, thus mitigating issues brought on by confusion in this case, by elevating the assessment to the position of curriculum, this practice can contribute to the perceived phenomenon of teaching to the test.

Summary of Purposes of Assessment

Volante (2010) described three types of assessment: assessment *of* learning, *for* learning, and *as* learning. However, the teachers in this study did not collectively reflect a framework of these three types of assessment, and analysis of their statements instead determined a conceptualization of three purposes of assessment for *compliance*, for *reporting*, and for *instruction*. Assessment purely for compliance would fall outside of McMillan’s (2013) definition of classroom assessment, provided in Chapter 2, but assessment for reporting and for instruction could be included. Whereas the definitions of assessment *of*, *for*, and *as* learning could be considered more theoretical in nature, teachers operate with a very pragmatic perspective of assessment in their practice. As this chapter moves into describing the theory of how teachers’ approaches towards assessment are influenced, it is important to keep in mind the different ways teachers are conceptualizing *assessment*, as this theory is based on their self-reporting.

Theory of Influences

The intended outcome of a constructivist grounded theory study is the creation of a theory that provides a plausible explanation of the data (Charmaz, 2014). Here I present

the theory generated from analysis of my data. First, I present and briefly explain a basic model. Then, I display and build out each of the three main components of the model. Finally, I explain how the elements of time and resources can also influence K-2 teachers' development of their approaches towards classroom assessment.

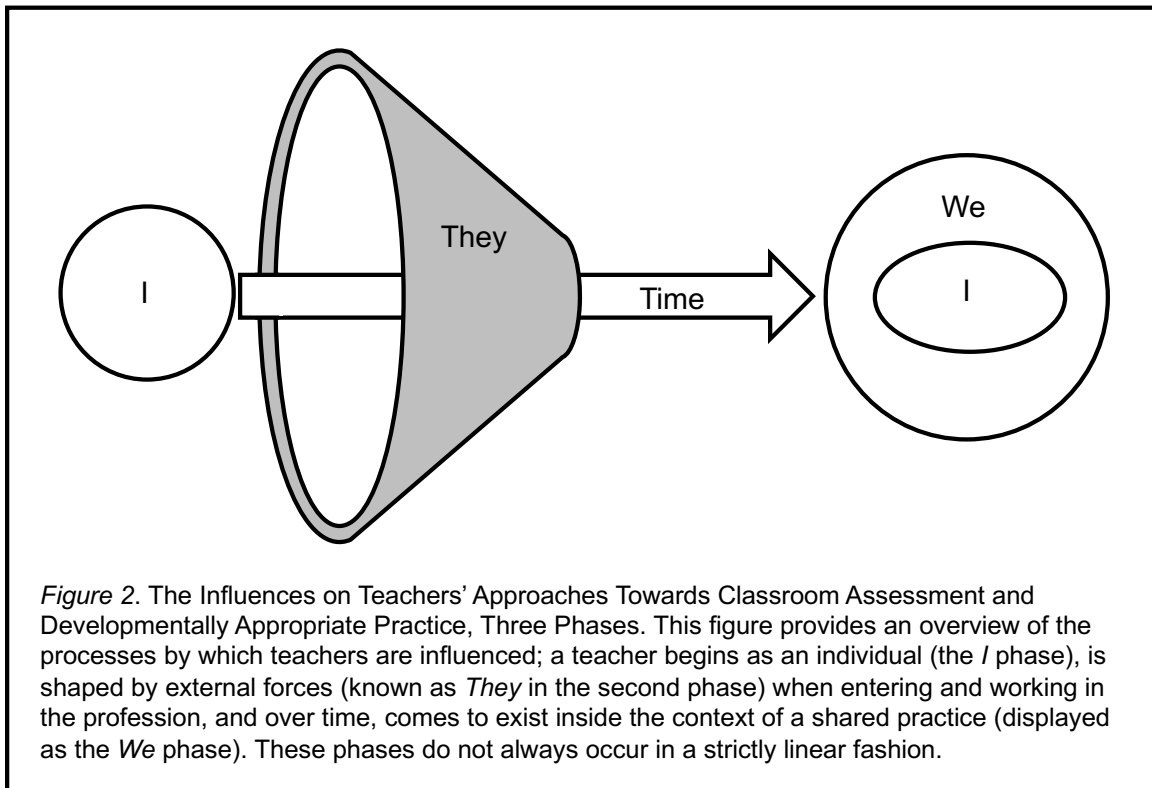
The same basic model explicates the processes and influences on K-2 teachers' knowledge and implementation of developmentally appropriate practice, albeit with different nuances and emphases. However, teachers had much more to say about assessment practices, as assessment has been a major driving force in their work and professional lives. For the teachers in this study, the framework and guiding principles of developmentally appropriate practice are not a dominant factor and existed more as a background element within their profession. Recall that DAP is a framework based on knowledge. But in teachers' pragmatic professional work, what they *do* dominates. Some teachers described close alignment between what they know (DAP) and do (in CA), but for many, their knowledge sits in the background, peeking through the mandated practices periodically.

Despite the complex relationship between knowledge and behaviors, it was possible to use a single theory to explain the influences on both these constructs. Each construct (DAP and CA) is more strongly influenced by different components in the model and I discuss each where it features more prominently. Because teachers feel that assessment drives so much of their school practices, it is a more dominant force in their professional lives and, thus, is described with more emphasis here than DAP.

Overview of Theory

When interviewing teachers, one of the themes that stood out was how teachers used the pronouns *I*, *they*, and *we*, and the meanings both explicit and implicit in using

first or third person pronouns to describe practices. While I started each interview with a question intended to prompt a reflective, individual-centered response, the majority of teachers responded with “we” statements describing collective practice. When talking about requirements and mandates (assessment for compliance, primarily), teachers used “they” statements. When mentioning beliefs or knowledge, teachers tended to use “I” statements. Analysis of the nuance of these distinctions led to the theory presented in *Figure 2*, which is built around three phases: an *I* phase, a *They* phase, and a *We* phase.



The *I* phase is marked by the individual factors and influences on the teacher, typically prior to entering the field of education. It is represented by the circle labeled *I* on the far left of the figure. These influential elements include prior work and volunteer experiences, teacher preparation programs and student teaching, raising children at home, and various personal traits and qualities that shape and frame a teacher’s philosophy and

approach heading into teaching. While some of these elements may continue to influence teachers over time (particularly raising children), they appear to have the most influence on the individual before the *I* starts teaching.

When entering the profession, a teacher moves into the *They* phase, as indicated by the arrow pointing right, moving the *I* circle through the funnel labeled *They*. A variety of entities, both identified and unrecognized, compose *They*. This phase refers to the influences that have authority or control over teachers' practices, thus heavily influencing what assessment practices they *do* and possibly shaping their approaches and attitudes towards assessment as well. These influences can be positive or negative, significant or minor, but all teachers move through the *They* phase of the context in which their teaching occurs. *They* includes state and district policymakers, central office personnel, building administrators, and often an unknown person or persons that are believed to have made decisions and created circumstances that influence teachers in a one-way manner.

The third phase, the one in which all of the teachers currently reside, is the *We* phase, as represented by the circle on the far right of the figure. As indicated by the oval inside the circle, the *I* has been shaped by the other influences over time and is now situated within the context of a collective *We*. The *I* has not been absorbed by the *We*, and may maintain a strong individual identity, but it no longer exists independent of the *We*. As shown on the arrow, Time continues to influence the individual within the community.

This basic theory is presented as a linear process of three phases, but it is important to acknowledge that individuals may navigate these three spheres in a less-strictly linear fashion. Certainly, each individual has a unique combination of personal

and contextual factors that have influenced and continue to influence their approach towards classroom assessment. In the next sections, I further describe the properties that characterize each of the three core categories of the *I*, *They*, and *We* influences on K-2 teachers.

The *I* Phase

The initial phase of the theory that explains the influences on teachers' approaches towards CA and DAP is the *I* phase, consisting of the individual elements, both chosen and incidental, that shape a teacher outside of or prior to the start of teaching, when the social contexts become more significant. The properties of this category can be broadly classified into four areas, as displayed in *Figure 3*. Each of the four properties are addressed below, but it should be noted that they are not order-dependent upon each other. For a large majority of teachers, this phase was the strongest of the three in influencing knowledge and implementation of DAP and the weakest of the three in influencing an approach towards assessment.

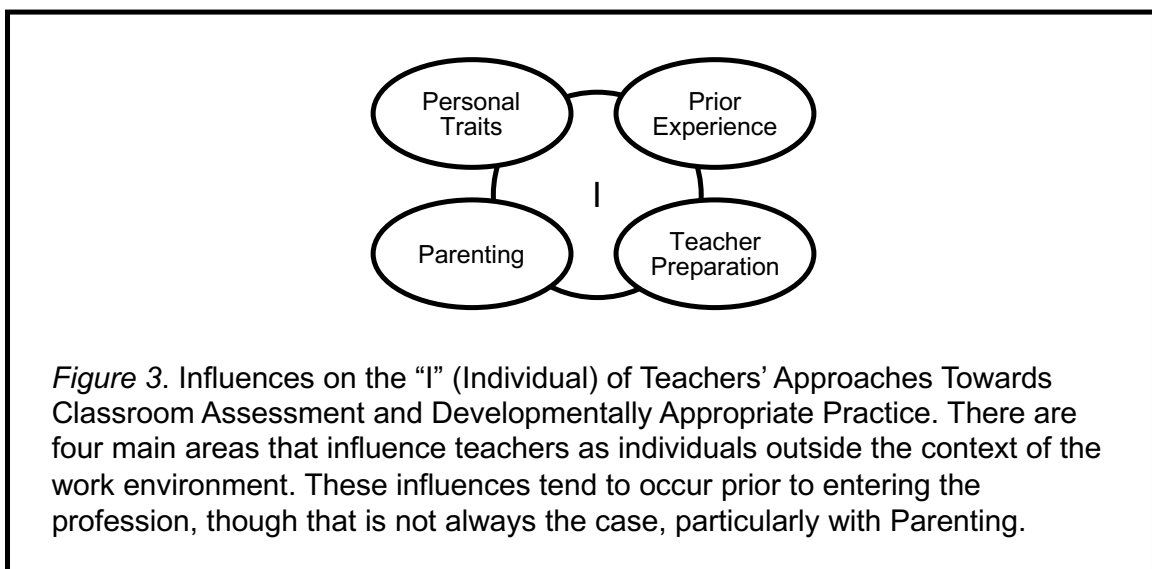


Figure 3. Influences on the “I” (Individual) of Teachers’ Approaches Towards Classroom Assessment and Developmentally Appropriate Practice. There are four main areas that influence teachers as individuals outside the context of the work environment. These influences tend to occur prior to entering the profession, though that is not always the case, particularly with Parenting.

Personal traits. The first property of the independent *I* is the personal traits that make the individual a unique human being separate from the teaching profession, traits that perhaps draw the individual towards the profession and shape how he or she approaches it. The dominant trait expressed in this sample was a penchant for working with young children. While a few teachers spoke directly about their enjoyment of a particular subject, no one described a passion towards content as a primary driver into the profession. Instead, there was a strong pattern of teachers expressing a predisposition towards enjoying being with children that influenced how they approached working with children in a classroom setting. While these traits shaped DAP more than assessment practices, a personal disposition towards honoring and nurturing the whole child over teaching and testing content did shape how a teacher approached assessment.

Prior experience. Many participants described how work and volunteer experiences influenced both their interest in kids and their philosophical orientation towards working with young children. While these influences did not have a significant impact directly on knowledge about or application of assessment practices, they often shaped how a teacher came to understand developmentally appropriate practice as teachers acquired experiences showing them the range of how children typically develop and behave at different ages.

Kindergarten teachers Janice and Michelle teach in different districts but both shared that an early influence on their approaches towards children was having worked with their own mothers in early childhood classrooms, kindergarten and preschool, respectively. Over the course of years prior to any preparation program, they each observed a seasoned teacher well before the accountability era's impact shaped early childhood practices, and they both described learning about how young children develop

by watching them at play in these contexts. Several teachers described an array of experiences, both informal and formal, volunteer and paid, that formed their understanding of developmentally appropriate practices.

Veteran teacher Margaret learned from work with kids in nurseries, daycare, and teaching Sunday School along with simply watching people, including young family members. Mid-career teacher Jill worked as a paraprofessional for years before earning her teaching license and mid-career teacher Anne worked in a private preschool before entering the public school teaching context. Mid-career teacher Gretchen collected a wide range of work experiences that shaped her understanding of young children, including being a storyteller at the library and working for a public television station's education department. Early-career teacher Maureen also described a range of experiences with children, including babysitting, running an afterschool program for latchkey kids, and serving as a "room mom" multiple times for her own children, an experience that bridges into the next property of the *I* phase.

Parenting. As some teachers became parents after starting their teaching careers and others entered the profession after having children of their own, this dimension of the *I* category is the most fluid in terms of timing. However, it is included in the *I* phase because it uniquely impacts the individual separate from the other factors included in the theory; there was no reported relationship or connection between a school administrator or teaching team and the influence of parenting on a teacher's development. Kindergarten teachers Margaret and Maureen each described how they entered the teaching profession after having children, and how their experiences as a parent have helped them understand what is typical for child development. Early-career teacher Nikki, on the other hand,

described how becoming a parent after teaching influenced and changed her approach to developmentally appropriate practice:

I was like, “I can’t send my kid to school and just have him do worksheets and stuff all day because that’s not how he learns. And that’s probably not how the kids in my classroom are learning either.” And so that really sort of changed my whole approach to teaching. . . . I don’t believe in that anymore.

As previously mentioned, first grade teacher Gretchen amassed a variety of work experiences that shaped her understanding of children, but she also intentionally sought out work experiences that fit her needs as a parent, often finding part-time work that would allow her to bring her children, such as doing support work at a child development center. As a parent, she also delved into educational TV and psychology books, and took advantage of opportunities to learn by volunteering in her children’s classrooms. She even described a particularly influential experience as a parent volunteer in which the classroom teacher allowed her to plan and carry out a craft activity that did not go well. When the teacher asked if she would still return, Gretchen replied, “Of course, I’ll come back. I’m really grateful that you allowed me to have this opportunity to fail so badly because it really made me think, ‘you know, next time, I need to tone it down some.’” Being a parent combined with other opportunities and experiences in a way that collectively helped shape Gretchen’s understanding of child development and approaches towards teaching prior to her entering the field.

Teachers often described how their approaches towards DAP were influenced by a combination of the previous three dimensions: personal traits, previous experiences, and parenting. However, these areas had little to no influence on teachers’ approaches toward assessment. In comparison, the teacher preparation program seemed to have more

of an influence on teachers' knowledge about assessment practices, although this influence was not dramatic.

Teacher preparation. Reports about the influence of teacher preparation programs or other formal schooling were widely varied. A clear trend was apparent, however, regarding the lack of a meaningful course on classroom assessment. Most teachers stated there was no course on assessment practices, or if there was one, they could not remember it. In fact, the idea seemed so incredible to her that mid-career teacher Jill laughed at the question. "Are you kidding?! I don't *think* so," she emphasized. "If it did, I don't remember it!" Early-career second grade teacher Chantal was one of a small number who remembered an influential class in assessment, and she recalled how the class addressed foundational and technical points such as unpacking the standards to ensure assessment alignment and learning about creating test items, but she also admitted that her knowledge is a bit rusty because she has often been in scenarios where she is not actually making assessments and is just administering assessments others have created. The data strongly supported the notion that the influence of teacher preparation programs on teachers' approaches towards classroom assessment was minimal at best.

Memories of coursework on developmentally appropriate practices were more diverse. While several teachers only vaguely recalled a foundational child psychology course, others described a strong training program, particularly a cluster of mid-career and veteran teachers who all had early childhood degrees or preK-3 licenses from the same state university. But even veteran teacher Melanie, who spoke well of her teacher preparation program, stated:

Those classes are so, just skimming the surface. . . . You really don't see and understand that developmental practice until you see a kid who truly just can't

read yet. And it's not because they don't know how, they're just not ready. . . .

They're going to get it when they're ready.

This idea that you cannot really be prepared for teaching until you are in the job was one of the most widely-shared understandings amongst participants. Just a few of the supporting comments are shared here:

“Were you totally prepared? No. A textbook is never going to give you what you need to know. Never.” (veteran second grade teacher Cyndi)

“They didn't teach us how to teach, they taught us information, but it's not until you're in and they're 100% [yours] that you really understand all that it takes.”

(veteran first grade teacher Marsha)

“You don't understand. I don't think that I was prepared at all for what was expected of me. . . . You know, we have textbooks and all that good stuff, but the actual experience of being in a classroom . . . I did feel very unprepared as far as that goes. . . . Sitting in the classroom won't prepare you.” (early-career

kindergarten teacher Michiko)

“I don't really feel like you can get a true sense of everything. I don't think it's going to be enough until you're living it.” (veteran kindergarten teacher Janice)

“It never prepares you for actually the class that has your name on it with your kids. . . . it doesn't give you real life experiences of figuring everything out on your own and with a colleague next door.” (mid-career kindergarten teacher

Nancy)

“I think there are a lot of things when you go through that program that do not prepare you for life in the classroom. They hit on a few things. A lot of it is theory, and then I know they put you in practicum and that gives you an

icebreaker as to what you're actually going to see in the classroom. . . . I don't think anything prepares you for today." (mid-career second grade teacher Jill)

Jill described the practicum as an "icebreaker," but a few other teachers felt that their practicum or student teaching experiences were slightly more influential. Early-career teacher Chantal described a great relationship with her mentor teacher in a kindergarten classroom when she was student teaching. "I feel like I learned how to be a teacher from her. She knew things about what kids are like and what they can handle developmentally," she recalled. Chantal continued:

I think that you learn more from that [student teaching] than you really do from classes. I feel that student teaching experience and then, like, on the job experience—it just takes time. But you could develop this proficiency in teaching from experience that's hard to get otherwise.

Second grade veteran teacher Bev frequently hosts student teachers in her own classroom and echoed this sentiment from her perspective. In talking about the student teachers she has had in recent years, she stated:

They're getting taught information, but it's not, it's not always pertinent until you're actually in the classroom and you observe it, if that makes sense. And I think sometimes we can talk and do things all day long and it'll be in a university setting, but sometimes it doesn't hit home or maybe the experience isn't as...real life, I guess, as being in the classroom doing it at that time.

Even a positive, influential relationship with a cooperating teacher during the pre-service period can only provide limited experience. This is not to say that the preservice program and student teaching period were unimportant, but that they were functionally limited on their own. In most cases, teachers stated that their preparation was a good start

towards understanding the backdrop of teaching and child development, but rarely ever did it shape how they approach assessment practices. According to veteran first grade teacher Darlene,

You have a lot of theory but you haven't really had a chance to implement it. . . . I think when you come out, as least for me coming right out of college, I did not understand any of these things. Even classroom management, any of it, any aspect of education. I thought I knew, but you get in there and you close that door and you're by yourself and you have to do it—you don't know.

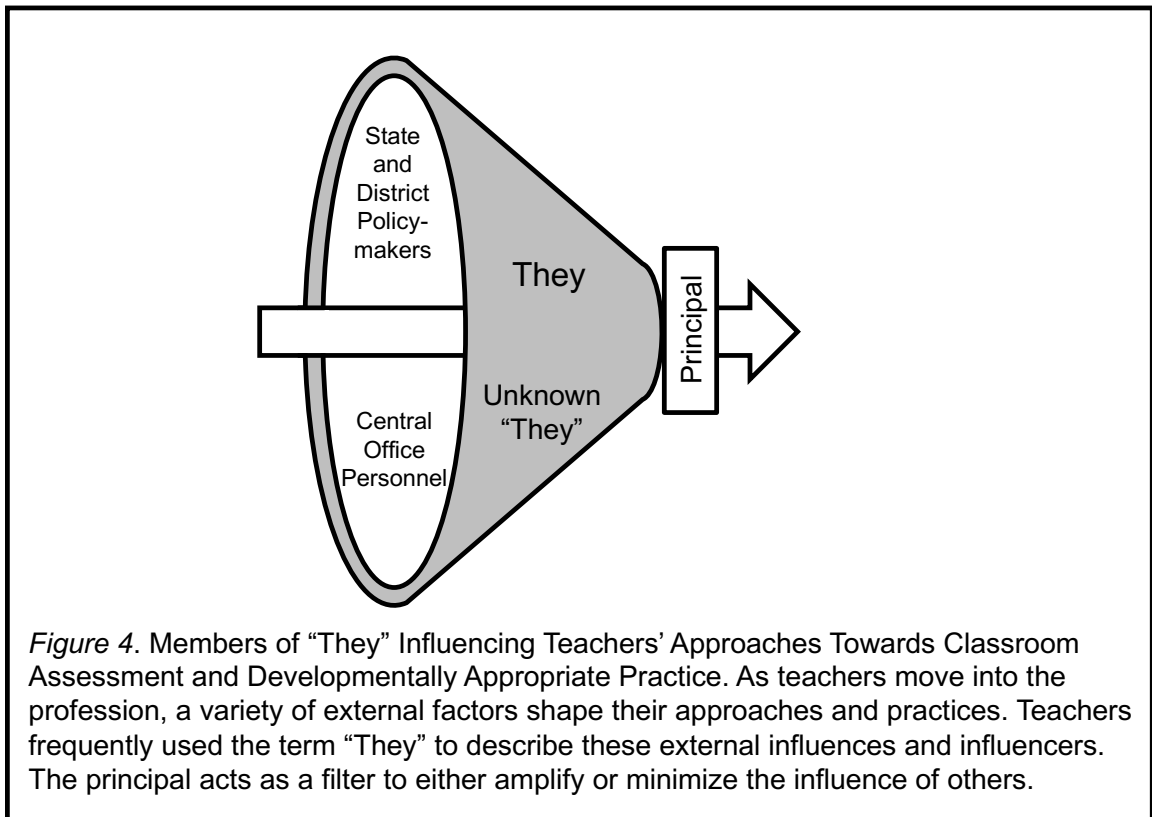
Although Darlene stated that you are by yourself in your classroom, this is vastly different than suggesting that your decisions and actions are independent and isolated from the influence of others. In fact, once teachers are in their own classroom, they enter the next phase of influence, one that is marked by a significant influence of others that can dramatically mold teaching practices, the *They* phase.

The *They* Phase

The *I* phase describes the dimensions that influence a teacher as an individual, mostly prior to entering the teaching profession. As soon as they enter a job, the *They* phase of the teaching profession begins to shape what is required and expected of teachers as well as what experiences they accumulate. Although teachers rarely spoke of *They* as influencing a philosophical approach, teachers frequently spoke of *They* as controlling or supporting their practices. The *They* can be grouped into four primary dimensions, not all of which represent actual people, but to the extent that the word “they” is third person, it applies to outside entities separate from the teacher that act in almost all cases unilaterally towards the teacher. These four dimensions are shown in

Figure 4 and described in detail later: state and district policymakers, central office personnel, the building principal, and a nebulous group of unidentified “they.”

The shape of the *They* phase is intended to show how an array of outside forces funnel and direct the teacher, the individual *I* from the first phase, and mold her practice as she moves into the context of a teaching position. In *Figure 2* and *Figure 4*, the *They* phase is situated with an arrow moving through the funnel to indicate time passing as a teacher moves from the pre-service period into the teaching context.



It was striking across the course of 35 interviews with teachers from 15 school districts and 26 schools how often the term “they” was applied both specifically and generically to authority and power held outside of the teacher and imposed on her. Just a few of the examples are provided here:

“They dictate a fair amount of it for us.” (Karla)

“This is what they make me do.” (Laurel)

“It doesn’t really match up with what they’re telling us that we have to do.” (Jan)

“I think it’s just cause that’s how they’ve always done it, to be honest.” (Pauline)

“They want us to teach developmentally, they want us to teach with manipulatives...but not assess them that way.” (Cyndi)

“They think we make the impossible happen. And it’s just not possible.”

(Michiko)

“I think they lose sight of the fact that sometimes different children need different things and they need to be evaluated in a different way.” (Kim)

“I don’t know why they won’t listen to us. It’s our report card, we teach it, we should know it.” (Nikki)

“They are very data driven in this county.” (Maggie)

The difference in this last example between saying “*they* are very data driven in this county” instead of “*we* are very data driven in this county” or “*I* am very data driven” exemplifies how teachers used the word “they” to refer to others, not including themselves in the actions or orientations. Veteran teacher Maggie reported no first-person involvement in a data driven motivation, but her practices have to support others’ inclination to be data driven. While often teachers expressed concern about the influence of *They*, stating policies, practices, and attitudes that were not perceived as being effective assessment procedures or, more frequently, not aligned with developmentally appropriate practice, it is important to note that the unilateral direction of influence outside of teachers’ control was not always described as a negative influence. In this section, I will address ways in which teachers report they are influenced by *They* in both positive and less favorable manners as this relates to each of the four dimensions below.

State and district policymakers. At the highest level of influence, the state Department of Education (DOE) and district policymakers shape the landscape of assessment practices. Because the DOE sets the SOLs and associated SOL tests, it is perceived that *they* are the primary driver of the high-stakes accountability movement and testing era. In early elementary education, there was little expressed enthusiasm for the emphasis on testing and accountability. Second grade teacher Chantal reported that teachers go into the field to help kids, but

then the system is all about standardized testing and . . . it's kind of crushing to a lot of teachers and students, and most teachers are not about it. Like, we're doing it because that's what we're supposed to be doing.

State and district policymakers weigh heavily on the category of assessment for compliance, often without a perceived benefit to student learning.

The influence from this level was not only about shaping the backdrop of teaching. Kindergarten teacher Anne reported an influence on the rigor of the standards themselves, claiming, "A lot of times, [SOLs are] made by people who have never even taught the grade." Kindergarten teacher Margaret described how an emphasis on technology-based assessment originated at this higher level: "I think a lot of that is coming from the DOE because they see bigger and better." Veteran teacher Liz came to Virginia after a long career in another state's charter school system and described the freedom she felt there in comparison, saying, "your hands aren't tied to all of the things that the state says that you will do."

A very specific way in which the state and district influence how teachers approach assessment is through the structure of accreditation. Anne works at a large

school district and described how the accreditation process impacts her school's, and thus her, practice:

Assessment is big—numbers and data. It's big at our school, being that we are, they call it "Priority Two" right now, which is, we are accredited, but we are borderline to where we would not be accredited. So, assessments are huge at our school.

Pauline teaches in a small school district, but reports an almost identical influence: "Here, it has been all data-driven, for the most part because we have been up and down with being accredited. . . . They don't want us to lose accreditation. So, we need to keep up with assessment." Maya described having previously worked in a different mid-sized district at a school that was under pressure for risk of failing accreditation. She recounted the requirement of a strict page-by-page pacing guide that usurped all of her professional decision making, and she acknowledged that the level of micromanagement was so exhausting she left the district after just one year.

While policies in regard to the SOLs and accreditation directly impact K-2 teachers, a recent shift in the state towards using performance-based assessments in lieu of some of the required SOL tests has had little impact on early elementary teachers. As students in these grades do not take SOL tests, and therefore their teachers do not have to locally create performance-based alternatives, teachers reported that this policy was not influencing their approaches to assessment. To the extent that some elementary schools were holding mandatory professional development in-service sessions to build teachers' capacity in creating performance-based assessments, a few teachers indicated that they had attended training meetings, and the influence could trickle down more over time. But as one teacher mentioned: "We've been doing this for ages." Assessing students based on

a performance of demonstrated learning has long been preferable to paper-pencil tests in the early elementary years.

No K-2 teachers reported any way that policymakers at this level impacted their knowledge of DAP, though there was some influence on their implementation of DAP. This influence was indirect in that it came about because policies the state mandated require developmentally *inappropriate* practices (such as unrealistic academic standards for all children) or because they emphasized other priorities to a degree that teachers felt they could not focus on implementing DAP. According to these teachers, DAP was simply not given a second thought at the state level, and this omission of consideration had indirect effects.

Some policies are mandated at a district level rather than at the state, but when teachers described actions at this level, they were less likely to talk about entities and more likely to talk about people as the agents of decision making. These nuances between state and district, entities and people were often blurred and at other times distinct, but at a minimum, a clear dimension of *They* emerged through the data that can be categorized as central office personnel.

Central office personnel. In over 340,000 words of transcripts, the focused code that occurred most frequently was that of “central office.” Some teachers referred to this as “the county,” “the district,” “downtown,” or even “you people upstairs” (Cyndi). The influence on teachers’ assessment practices was so critical, that multiple teachers made no distinction whatsoever between their individual approach and that of the district. When asked to describe her approach to classroom assessment, Barbara, who teaches first grade in a large district, claimed, “Our approach is required by the county and the school.” Other teachers were slightly less emphatic, but expressed a similar notion, such

as Annabelle, who also teaches in a large district, stating that the biggest influence on her approach to assessment is “the county requirement. What they want to do.” So significant was the amount teachers had to say on this topic that it demands extensive attention in the findings chapter of this study.

Emphasis on content. Virginia’s SOLs predate the rise of the accountability era nationwide, but the federal requirements for students to pass subject-based standardized tests has led to a greater focus on content area instruction and assessment. Veteran teachers reported an increase in both county- and school-based reading and math specialists or coaches. In some districts, there might be a subject-area specialist at each individual school, whereas others share specialists across schools or are guided by a single specialist at the county level, a practice more common in small districts. Some teachers lamented this shift towards subject area and away from the developmental needs of children by age, and this lament was not only from teachers who were in the field prior to the rise of high stakes testing.

As early-career teacher Chantal articulated, “We want everything to be about SOLs. So, we hire people who know about SOLs, and maybe they don’t necessarily know about, like, what it’s like to teach a younger grade and what do kids by age really know.” Second grade teacher Claire stated that these district curriculum leaders tend to know the content area well, and they possibly have sound knowledge of assessment foundations, but they do not always understand how those more technical understandings apply to children in the classroom. She argued, “They’re not in the classroom with the children and they’re not in with that particular group of children.” She felt that in giving the same assessment each year regardless of your students’ needs and interests, “I think you’re doing the kids a disservice.” Early-career kindergarten teacher Michiko teaches in

a mid-sized district, and described her experience with a content area expert lacking expertise in children's developmental needs:

We have reading specialists that work outside of our building that come and visit. Sometimes they haven't even taught in elementary school. . . . If you haven't been in the trenches and you've not experienced that, I definitely think it's harder for them to understand, you know, "oh, hey. What would be an assessment piece for them?"

The lack of consideration or understanding of students' developmental needs in favor of content expertise was reportedly rampant among central office administrators.

Experience of administrators. While this study did not collect or confirm personnel data reported by teachers, there were widespread reports of central office decision-makers who lacked knowledge and experience working with elementary school aged students. Just a few of these comments are included here to explain the situation:

Unfortunately, the school board office has become a place now where there isn't anybody with elementary experience there anymore. They're all middle school and high school. And we've expressed our concerns several times about that because there'll come times where we're like, "you guys aren't understanding this." But you know, it's like, who they hire is who they hire. We don't have control of that. . . . It's sad that they don't realize you have to have people in there that understand that level. I think they think if they're in education, [they'll] get it. But you don't, you know, you don't just understand the kindergarten world unless you teach the kindergarten world and you don't understand the elementary world unless you're in that elementary world. (veteran second grade teacher Melanie, mid-sized district)

There's an interesting sort of mentality that—and we definitely see it here—that like, high school are the important years. Like, that's where you put your best principals and then middle schools are next and then, well, you can put anybody in elementary school and that's like training ground. Like, “they might be an okay principal. We'll test them out here.” . . . [But] you can't become a senior person unless you've taught high school. They're never going to take somebody that is only in elementary and put them over curriculum, but they'll take somebody who doesn't have elementary experience, which is, these are the formative years. These are essential years. (early-career first grade teacher, Karla, large school district)

I know a lot of the people that are up there don't have a clue. If they came in and spent a day in a second grade classroom, they'd probably die. . . . Most everybody that's above us, like in the county, almost everybody are middle school or high school people. So I think that's why it doesn't translate down here. . . . [They'll say] a second grader at this time should be able to write seven to eight sentences on one topic. Yeah, their thoughts aren't seven to eight sentences on one topic. How are they going to write it on one? I'm like, “what are you people thinking?!” And sometimes I just don't, like, “oh, you aren't.” (Kim, veteran, second grade, large school district)

[Speaking about her current reading specialist:] Kindergarten's not her thing. . . . We're really feeling that right now, [what] it feels like when you don't have the reading specialist so much in your corner. [It's like], “Oh, she doesn't care what we're doing either.” (Sangita, mid-sized district)

The way Sangita expressed the impact of having a central office administrator who “doesn’t care” should not be interpreted as a concern that is a personal one. Time after time, teachers situated this concern within their students’ needs, claiming that inexperienced administrators frequently pushed developmentally inappropriate practices.

Developmentally inappropriate practices. While the purpose of this study was not to audit developmentally appropriate and inappropriate practices, it is crucial to note that many K-2 teachers reported what they considered to be developmentally inappropriate practices in the context of the assessments for compliance that were required of teachers. As veteran first grade teacher Annabelle explained, “I wish that those who made the mandated assessments were more knowledgeable and I think that they expect an awful lot out of a six- or seven-year old child.” She gave an example of a county science assessment in which first graders were asked to explain in writing why a black shirt or a white shirt would be warmer if left out in the sun. “They’re struggling to *write*, much less be able to explain something in writing,” she pointed out. “I wonder how long it’s been since they’ve been in the classroom.”

Veteran Kindergarten teacher Lyn did not wonder about administrators’ experiences. She has worked in a variety of school contexts and currently teaches in a large school district as well as goes on frequent site visits as an International Baccalaureate Primary Years Programme (PYP) trainer. Lyn reported that in all of her 27 years of teaching, she has only worked for one principal who ever taught in early elementary, and she has never met a superintendent in all of her site visits who has experience with young children. She asserted, “I’ve never had anyone who actually understood four- to six-year old children, that really gets what they really need in order to be successful or what it entails to actually give an assessment to a child that age.” She

claimed these administrators “have no idea what it is that they’re asking us to do.” As an example, Lyn described the importance of format for a kindergarten assessment:

Some people from downtown have told us, “Oh, well everyone just needs to start doing it whole group.” But in kindergarten, then you’re not going to get an accurate assessment of what the kids can do because if they miss one direction, you have to then turn it into a paper test, and then you’re not really assessing them.

Just one central office administrator lacking developmental understanding can have a significant influence over how teachers approach assessment. Veteran kindergarten teacher Sangita, who has spent most of more than 30 years in education working at the same mid-sized school district detailed:

The longer I’ve been in education, the more I see that it just takes one person, maybe in your administration or your central office, to think that this assessment tool or this way to go for teaching is the best way, and then your whole county has to do it. You’re like, “Whoa, wait a minute! How did we end up here?!”

While there was a nearly overwhelming amount of concern about developmentally inappropriate assessment practices being required by central office administrators, a few teachers expressed that their district was open to the input of teachers.

Openness to feedback. Though teachers predominantly described assessments for compliance as being mandated from a position of higher authority, a few teachers reported an openness from county administrators to receiving input on assessment practices, including required ones. First grade teacher Pauline teaches in a small district and said, “We actually develop [assessments] and the division approves them and then

they come back to us. ...Teacher input is considered, but it mostly comes down from the school board.” Her statement highlights a tension when teachers are given a heavy role of involvement in creating assessments, but the resulting assessments still feel like an obligation rather than choice because the authority to approve and require the assessments comes from others.

On the other hand, having involvement in the development of district assessments can be empowering. Veteran second grade teacher Bev works in a small school district in which teachers have worked together with central office administrators to redevelop their math curriculum and assessments. She expressed, “To know that I’ve had a hand sometimes at creating some of the assessments, that’s huge and powerful. . . . That’s powerful to know that [the district administrator] wants direct feedback from teachers as we’re using this tool.”

Unsolicited feedback, however, is not typically well-received. Karla stated that central office administrators in her large district tend to brush off teachers’ calls for more attention to developmentally appropriate practice. She claimed:

I do feel like sometimes when we bring it up, it’s like, “ugh. [Eye roll] Here we go again.” . . . [Administrators feel] that we’re “wasting time” discussing something that we can’t change. And yet, the only way to change it is for us to keep having a conversation about it.

Although some teachers reported wanting a more receptive central office administration, others expressed a desire for more support from the district, particularly to the extent that they wanted more information in order to ensure they were complying properly with required assessment practices.

Need for support. Veteran kindergarten teacher Janice works in a large district and reported some confusion regarding the district's requirements. She displayed a desire to have more direction to help prioritize her assessment practices among heavy time demands: "the county says everything is important, but it's really not." She described a requirement to assess students' sight word knowledge, but no benchmarks on what the expected rate of growth is across the year. She said, "I can tell you how many [words my students recognize], but I can't tell you what the county feels is appropriate. . . . If you're expecting us to test it, you need to do it. Give us something." Fellow large district teacher Karla also wanted more guidance from the district to meet requirements, describing it as "burdensome" to create all the demanded assessments.

Veteran teacher Lyn teaches in a different large district and detailed that clarity, not specific assessments, was most needed from her central office administration:

The people, I think, in our downtown office, they're not getting their point across when they explain things to people: "We've provided this for the teachers who need it. If you don't need it, we expect you to be on the same SOL when you're doing your shared reading as the teachers who are reading the book from [the adopted curriculum program]." People don't understand that. I guess they don't hear it when they're afraid, more like, "I'll get in trouble."

The impact of fear that Lyn mentioned was also at play in the context described by veteran teacher Darlene, who teaches at a mid-sized district. She stated there had previously been a huge emphasis on compliance with lots of benchmark testing, but a new superintendent has worked to change the culture so teachers worried less about obedience and felt more confident using professional discretion. He is constantly telling teachers, "you have permission" and is encouraging building administrators to feel

comfortable making change. Darlene declared, “[He is] less of a micromanager and more willing to let principals have some freedom in that stuff.” This is quite important, as the role of the school principal can be instrumental in moderating or amplifying the influence of central office decisions and expectations.

Principal. While the role of central office in influencing K-2 teachers’ approaches to classroom assessment is undeniable, the role of the building principal in mediating the influence of the county is also unquestionable. Although first grade teacher Barbara minimized the position of the principal (“they’re just kind of the middle person”), first grade teacher Marsha placed more significance on the ability of a building principal in regulating the pressures and requirements of the district. While talking about her current principal in the mid-sized district where she works, Marsha stated, “He’s very good at mediating. I’ve had others in the past who are just like, ‘well, that’s what the county says, that’s what you got to do.’”

Very few teachers described a current situation in which their principal was a negative influencer, and many described ways in which their principal advocated on behalf of the teachers and students in the building. At the same time, most mid-career and veteran teachers had experienced working with an administrator in the past whose perspective amplified the district’s mandates and created a negative working climate. In *Figure 4*, the principal is located at the tip of the funnel of *They*, indicating a position which can act like a filter by limiting the impact of the district or a magnifier, intensifying the impact of the district’s policies. This filter/magnifier power holds true for influencing both assessment practices and implementation of developmentally appropriate practices, though very few teachers reported ways their principal directly nurtured DAP. This tended to be more subtle, indirect, or neutral of an influence except

in cases where a principal was pushing developmentally *inappropriate* practices. In this section, I address both the perceived positive influences and negative influences of the principal as well as the impact of principal turnover on teachers. Finally, the section concludes with a description of the “trickle-down effect” that serves to illustrate how principals mediate the influence of *They* on teachers.

Positive influence. While early-career teacher Ruby acknowledged that “each building has different administrators and they each have a different perspective and opinion about how assessments should be given,” she expressed gratitude for the ways in which her building is led: “I’m really, really grateful for my administration. . . . [Administration] does as much in their power to focus on what’s really enduring and what’s really important in the classroom.” She works at a relatively new school in a mid-sized district and described how her principal has worked to establish a highly-collaborative climate that trusts teachers’ professional judgment.

Veteran teacher Melanie works at a different school but also spoke to the importance of her principal in supporting a child-centered approach, saying that her principal tells staff, “‘find purposeful ways to assess your kids. . . . We’re not here to get you. We don’t want kids to have bad grades. We want them to learn. That’s the most important thing.’” From focusing on the big picture to putting smaller things in perspective, the principal can heavily influence the comfort and confidence level of teachers in the building, even when the district has firm expectations for compliance.

Several teachers described how their principals handled the state’s and district’s required goals for teachers based on assessment data. Jill described how the principal in her single elementary school district helped take the pressure off of her when reviewing her students’ test scores by calling them “just a snapshot of a child’s progress.” Veteran

teacher Sangita described her principal as someone who “gets it. He gets the developmental part of the child.” Furthermore, he is not only open to receiving feedback, he also supports teachers’ voice in advocating for the child up to the district level. She claimed, “in our county. . . we can go to bat for the things that we really believe and think we shouldn’t have. And they absolutely listen to us.”

Mid-career teacher Pat has worked under different administrators and her experience echoed what many teachers reported:

I’ve been able to flourish under fantastic leadership that says, ‘I totally trust you and I’ve watched you teach, and your kids are thriving.’ . . . [I’ve also] learned from micromanagement to dot all i’s and cross all t’s.

Negative influence. The principals who have had this type of micromanagement style of leadership mentioned by Pat are those whose influence has reportedly been more negative. Marsha described a past principal as having five teachers sitting in a meeting telling him, “‘it’s not going to work.’ And he said they were still going to try it.” This lack of trust in their professional judgment led to teachers feeling disempowered and unmotivated, and therefore less likely to push against the status quo and work towards implementing DAP. Melanie described having worked for a high-strung principal whose focus on testing led to anxiety amongst teachers and more teaching to the test. The staff was focused on test scores, not students, she said. This both influenced how teachers approached assessment and implemented (or felt they could not implement) DAP. The negative influence of a principal was reported to increase when poor leadership was paired with frequent change.

Impact of turnover. Although she acknowledged that she has had ineffective principals in the past who needed to be changed, small district teacher Laurel said the

impact of frequent turnover on overall school climate meant their ability to make improvements was stifled. They could not focus on change because they “were always in triage.” Gretchen teaches in a different small district that has also been plagued in the past with rapid turnover, and many principals would view the position as a stepping stone to working in a larger district. Although she now has a stable principal of whom she spoke highly, Gretchen described the period in her career when the building administration changed annually:

Every year you’re try to get in a groove and then a new principal comes in and everything changes. And so, it’s always interesting to see: what is their preference going to be? What kind of things are they going to think are important? How do they view us? How do they view the kids?

This kind of extreme annual turnover was reported with some frequency at single-elementary school districts and at schools that struggle to maintain accreditation in mid-sized or large districts. Veteran teacher Kim described a period of 13 years in which she had 10 different principals. On the other end of the career spectrum, early-career teacher Maria described the instability she is experiencing in her first year at her current district as one heavily influenced by the nature of rapid change across not just the administration, but the whole staff:

I’m still learning what the division is expecting. There’s a difference between what the division expects and what my school is expecting. And not only am I new, and fifteen other teachers new, everyone that is in charge of us is new: new principal and new assistant principal. And all of the leadership people that are in our building that are helping with all this reading and writing, they’re all new to

the building as well. So, what the division might allow for wiggle room might not be what my school is allowing.

The unique relationship among the various people and positions represented in the *They* phase can be dramatically different in different contexts, but the interconnectedness among the different administrators and staff members was most frequently described with a kind of top-down power flow in which a principal might be quite limited in how much he or she can influence actions and approaches at the teacher level.

Trickle-down effect. Some of the reported change occurring in schools is “depending on what admin is hearing,” according to Cyndi. While there was widespread, but not universal, praise for current administrators under which teachers work, there was also acknowledgement that a building principal also has to respond to a higher level of administration. Second grade teacher Ruby described this structure in her mid-sized district. “We’re under some constraints. . . . There’s this top-down pressure about testing,” she said. “It’s like, you know, the administration feels like, because they’re responsible for keeping our SOL scores high and then that goes down to teachers.” Maureen described it succinctly: “You know central office is telling your principal what you need to do.”

Veteran teacher Beth teaches in a different school in the same mid-sized district as Maureen and explicated this further:

I have two wonderful administrators who respect classroom teachers and understand what’s going on in classrooms. . . . But they can’t control what’s given to them, you know. And it’s the same thing with central office. Like, I think that we have some listening ears, and we have some good things going on in our

central office, but they can't control what's given to them by the state. So, it's just this trickle-down effect.

In a similar way as to how a principal can mediate the influence of the school district on teachers, the pressure coming down from above can mediate the degree of flexibility a principal has in terms of leading teachers. "I mean, their hands get tied, too, because then their scores are...tied to that," clarified Jan. For this reason, the principal is located outside of, but directly connected to *They* in the explanatory model of the theory presented in *Figure 4*. At times, the principal makes decisions that act as influencers on teachers by either filtering or amplifying the district's focus, but the principal is also influenced by the district, and thus not entirely free to behave in an independent manner within the context of this theory. Even more convoluted is the nature of an influence that emerged as an unknown *They*.

Unknown *They*. As previously mentioned, the *They* phase encompasses a wide range of third person players who act towards the teacher in an influential manner. But sometimes, these players are unidentified, unknown, or even assumed to exist; these characters constitute the dimension of the "Unknown *They*." Teachers assigned the pronoun "they" to actors outside themselves whether they could identify them or not, as noted in Margaret's statement: "In a way it is a requirement because *they* insist upon having assessment grades. I don't know where that came from or who decided that."

Michelle teaches in a mid-sized district and also was not sure who made decisions about required assessments when asked. She acknowledged, "I don't know. I know that *they* have different people who are in charge of that and I know that the players change a good bit, too." The uncertainty around who held decision-making authority was a

frustration for Anne, who described asking her building content specialists assessment questions and not getting answers. She revealed:

I can't even tell you who they even go to if we needed more information. . . . If I need answers, they usually have to go to central office to call or to email the person who's in that department and we don't get answers. . . . I'm not being provided with any information.

Like Anne, Kim also works in a large district and describes uncertainty about the chain of command and what happens when she tried to share feedback to impact practice. She stated, "nothing ever happens," before clarifying, "I've never *seen* anything happen. I'm not saying nothing happens. *That's* kind of an unreachable person." Veteran teacher Janice works in the same large district as Kim and also wondered what happened to teachers' feedback on assessment practices. She felt her building's reading specialist listened to teachers' ideas, but she questioned whether her specialist had the ability to influence decisions. She pondered, "I don't think they do because it's coming from the county. . . . But I really feel like it's coming from those specialists in each area and not from the principals." Large district teacher Jan also wondered who was making the decisions:

Who makes these [goals] up? You know, where does this come from? And they're like, 'it basically comes from the state,' is what I was told. Because it's not from our principal, it's not from the reading specialist. It's not even from, like they say it's the county, but higher. . . . If you talked to them, I'm sure they would say that we do have, I guess, input into that. But really, we don't.

With such lack of clarity, it is unsurprising that teachers expressed a loss of a sense of agency and ability to impact assessment decisions or DAP considerations. In this

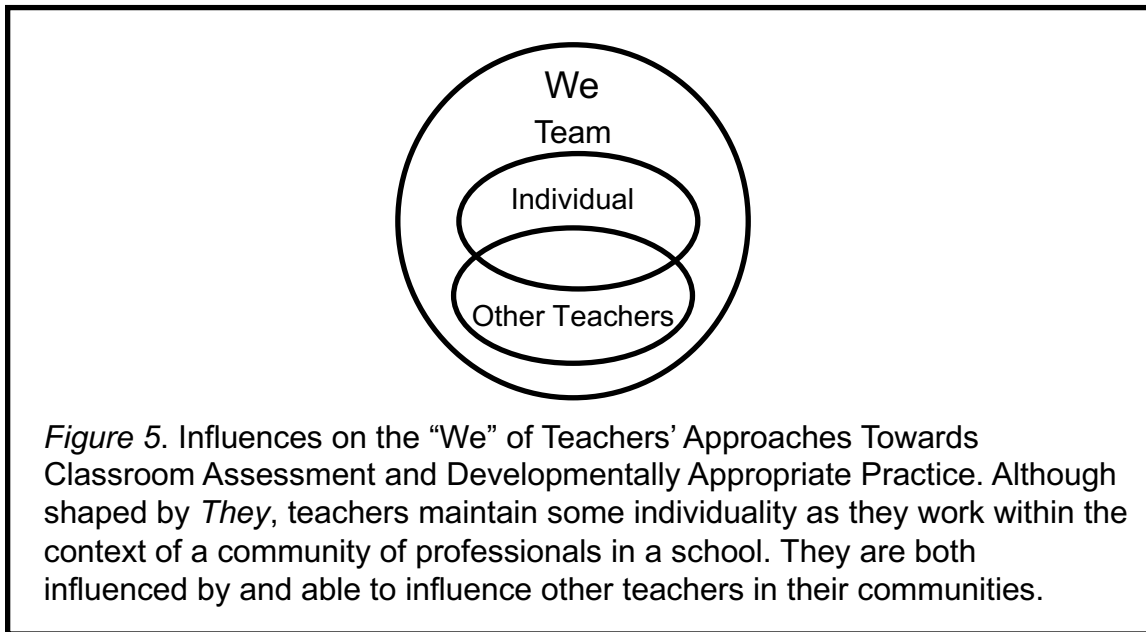
context, the term *They* was used to refer to any agent that acted unilaterally in influencing—or controlling—teachers’ approaches towards assessment. *They* might even represent a mythical entity, unknown as an actual person or position, but used as a scapegoat for acting on decisions for which teachers believed they had no input. For many teachers, *They* has come to represent any and all unwanted mandates not directly given by or clearly attributable to an individual or entity.

They was not used to describe influences on DAP, perhaps because there was a general lack of consideration about DAP from outside the *Individual*, and *They* was used to assign an action or decision to an unknown entity in authority. The vacuum of discussion about DAP meant there was no *They* doing any actions in this realm. The use of the third person *They* has a clear distinction from the first person *I* or *We*, and it has become a proxy for any outsider acting upon teachers’ practice. While this discussion presents a bleak perspective on teachers’ advocacy and agency, the third section of influence described by K-2 teachers presents more potential for action and contribution, in which teachers collectively influence each other in a shared practice, the *We*.

The *We* Phase

While teachers described a heavy influence of *They* on their assessment practices, and an indirect impact on their ability to implement developmentally appropriate practices, teachers largely described the actioning of these practices in the context of *We*. A teacher’s individual profile of *I* factors is shaped by the requirements of *They* and carried out in the setting of *We*. This is where the individual’s beliefs and knowledge regarding child development and assessment come together into practice. *Figure 5* shows how the individual remains, but her practice and approach are connected to other teachers within a team setting. The degree to which the individual influences other teachers and is

influenced by other teachers in their shared practices varies considerably, hence the overlapping ovals in the figure. But no teacher described an entirely independent practice in a school setting. The social context of the profession makes complete isolation impossible.



It is of note that teachers often changed between the pronouns *We* and *I* depending on the type of assessment practice they were describing. This distinction is encapsulated in the brief leading statement by veteran teacher Marsha: “*We* have more old-fashioned assessments, which are paper/pencil, and then *I* do a lot of informal assessments through having them come over and...making notes on index cards.” This section will further discuss how the use of pronouns reflects teachers’ conceptualizations of their practice and relationship with other educators. I will also address the shared practice of *We* in more detail before sharing about the nature of teams described by participants. After that, I explain how fellow colleagues play an instrumental role in developing strong professional practices, whether in a slightly formal mentor role or informally as

supportive teammates. Finally, I describe the “remaining I,” the degree to which the individual retains independent approaches and practices within a shared setting.

Shared practice. As stated earlier, unless discussion naturally began at the start of the interview without a question, I began by asking each teacher, “How would you describe your approach to classroom assessment?” With a few exceptions, teachers responded with *We* statements:

“We have standard assessments that we have to do throughout the year.”

(Darlene)

“Of course, we have the mandatory assessments...” (Nikki)

“We plan a unit, we start with what the assessment is, we do backwards planning.” (Betty)

“In the state of Virginia, we obviously follow the SOLs...” (Nancy)

“We look at the assessment that we, what we need to cover, and we try to, as a team, work together to come up with a variety of ways to test kids.” (Claire)

“We have the PALS assessment that we do throughout the county, and we use the VKRP...” (Maggie)

“Well, the team that I’m on, there’s four of us on the kindergarten team, and we have to do, the majority of our assessments have to be one-on-one...” (Margaret)

“We do a lot of different assessments.” (Laurel)

“You mean like what the county recommends? What we’re required to do?” (Jill)

These are just a sample of the leading *We* statements teachers used to begin a description of their approach to assessment, and each of the above teachers is from a different school and district. Regardless of school district size, grade taught, or years of experience, most teachers generally did not portray their approach as an individual

philosophy or orientation, but rather as a shared—and often required—set of practices. The final statement shared above, Jill’s question, alludes to this conceptualization: she was wondering whether a question about *approach* was intended to elicit something other than a list of required assessment tasks or activities.

For this majority of teachers who clearly situated their approach as a shared practice, they almost always followed with lists or descriptions of compulsory assessments, those that primarily fall into the assessment for *compliance* and/or assessment for *reporting* classifications. As Kim led in her interview, “We have, *obviously*, an abundance of assessments that are either state and/or locally mandated.” As the same influences demanding compliance or reporting practices (the *They*) act on all the teachers in their setting, it is logical that these practices are described as a collective, and as a collective separate from *They*.

Collaborative work. Teachers described mixed practices regarding the level of collaboration and comradery shared amongst their *We*, but for the most part, the participants in this study described congenial and hard-working teams that supported each other, even if their views on assessment practices or developmentally appropriate practices differed. First grade teacher Gretchen said about her team, “We spend a lot of time together . . . [when working on assessments] we all sit together, and we read every single question and talk about it. . . . We’re kind of really used to each other.”

Kindergarten teacher Sangita also described a strong, collaborative team:

We’re all on the same page. We’ve got our check-off lists. We’ve shared them through the years. . . . We share a lot, all of us. We share all of what we’re doing. We meet once a week. We talk all the time in the hallway. . . . We all come

together on certain things, and we all know what the other one's doing, and share, and praise each other. It's a good situation. We're lucky.

Sangita believed her situation was lucky; a few others described less-ideal situations that either may have been a result of weak team commitment or a strong group with a misguided direction.

Reluctant teammates. As previously mentioned, some schools are under pressure because of accreditation status or are heavily managed by administrators at different levels. Unsurprisingly, this can lead to teams functioning in survival mode. Veteran teacher Bev has worked on strong teams in the past, but said her current team is feeling the pressure of performance and can sometimes have narrow thinking. She described the differences:

Climate-wise, I'm in a totally different place to where I was before. It's a lot different and I have to be very careful. I don't feel I can just go out and say what's always specifically on my mind. I can't do that. Whereas before I could, but I think I grew up with all of them and they learned to live with me that way, but yet we all got along and were friends and we accepted each other for those. Like it wasn't a problem [on team at previous school].

Veteran teacher Maya also talked about a strong group, but because of the make-up the teachers on her grade level team, they often end up making decisions that are inappropriate and too rigorous for their first graders. She speculated that was because most of her teammates are teachers who have been trained in the accountability era and have previous experience in higher grades, combining to formulate a perspective that may not understand or appreciate developmentally appropriate assessment practices. Similarly, veteran teacher Lyn conveyed that her current team is more traditional and,

though all team members are agreeable in nature, they do not agree on approach:

“Everything we do in my class is hands-on and they prefer worksheets still. We don’t quite match.” She believed her teammates’ hesitation to veer from traditional methods stems from them not having learned to teach differently, perceiving it is too much work to veer from a prescriptive program, or simply feeling less brave to take risks.

A limitation of this study is that it was less likely to attract volunteers who are apathetic towards professional growth, the kind of teacher Laurel identified she has worked with at times. Whereas Laurel feels ready to take chances learning and trying new things, she described teammates who were more reluctant:

I think a lot of teachers are scared of rubrics because it is a lot of work. I mean doing performance-based assessments or anything that's not paper-pencil—it's, it is scary when you don't know [how to make them]. . . . People do not like to come out of their comfort zone. You learn what you learn, and you do what you're comfortable with and for a teacher sometimes to come out of that comfort zone—I don't know if it's a fear of failure, but I have found in [all my] years of teaching that you're going to fail. Some things are going to fail and no matter what you do, and you take it and you learn from it and I also think some of them, they just don't want to put in the work to it or they, or they don't know how.

Fear of failure is one of the elements at play when pressures of *They* create an environment in which compliance dominates the culture. The degree to which the collective team members feel they must have corresponding assessment practices can depend on a variety of factors, including how much independence the team has in functioning separately from *They*. When teachers described *We* practices, they were almost always talking about grade level teams of teachers, but continued discussion about

assessment often revealed that some decisions and actions were as part of team that included members of *They*.

Nature of the team. Understandably, the nature of teams did vary somewhat based on context. Teachers working in larger schools and districts were more likely to have content area specialists or instructional coaches housed in their building, but were less likely to feel their team was connected to those in the higher levels of district administration. Teachers who work in small systems were less likely to have dedicated, full-time subject area specialists as part of their team, but they were more likely to perceive they had approachable central office personnel, such as curriculum coordinators or assistant superintendents. This study was not designed to analyze these correlations, and the variety of team make-ups did not alter the applicability of the overall theory. There were primarily two types of teams that teachers indicated were at work when discussing assessment: grade level teams and data teams, sometimes referred to as PLCs (Professional Learning Communities).

Grade level teams. The work of grade level teams was described in more informal ways and was typically focused on instruction and assessment for instructional purposes. Amongst the tasks these teams undertake are looking at pacing guides, planning instructional units, critiquing lesson plans, and sharing resources. Grade level teams typically consist of just classroom teachers, and when a grade level was quite large, there might be two teams of teachers for easier management and to encourage more equitable contributions from all team members. Many teachers reported having more formal weekly grade level meetings for instructional planning and other meetings or informal times when grade level logistics, such as planning a field trip or upcoming holiday celebration, were organized. Due to the pragmatic nature of these team meetings, rarely

was there any general discussion about child development or emphasis on the knowledge and beliefs that guide DAP, according to K-2 teachers' reports.

Data teams. Data teams, on the other hand, consist of grade level teachers as well as other administrative and support personnel. Most teams described include the school principal (and possibly an assistant principal), a reading and/or math specialist (depending on the data to be analyzed), and a special education teacher. Where available, an enrichment specialist, instructional coaches, interventionists, Title One tutors, Reading Recovery teachers, district curriculum coordinators, and English language learner instructors are also included. At some schools, these teams meet in a dedicated data room where student assessment data is displayed in both individual and aggregate forms.

The purpose of these meetings is to analyze student data from assessments and make plans based on the data. This type of planning tends to be more mid-range and long-range planning and has more to do with determining which students need what type of support, as opposed to the finer-grained planning of grade level teams in which specific instructional strategies and materials may be discussed. Team members may evaluate and tweak assessments if they are not standardized forms (district-based benchmark tests as opposed to the PALS or Diagnostic Reading Assessment [DRA], for instance), but the driving force is a focus on student growth as measured by the results on required and common assessments. Knowledge or implementation of DAP is not at all a part of these meetings.

The perceived emphasis on this kind of data analysis, however, can be somewhat varied. "It really is a numbers game," early-career teacher Jan stated. Veteran teacher Maya implied that the intention and the outcome of the meetings may not always match. She said the purpose was to evaluate and adjust intervention groups, but she indicated,

“That’s what the purpose of all that is. And to drive instruction. That’s what they tell us, ‘to drive instruction.’ So, yes, it does that, too.” When *They* tell teachers what the purpose is rather than the *We* determining and buying into purpose, there may be mixed outcomes.

First grade teacher Karla works in a large school district and described what can happen when *They* are involved in these collaborative meetings of *We* and there is a resultant disconnect between practicality and the rigor of measuring high expectations:

We come out of a prescribed planning meeting and it . . . doesn’t always align with our practice or with our desired practice. And I always find that intriguing because we all went through teacher prep programs. We all were in the classroom and it looks like something happens when you step out of the classroom and go to the central office and there’s this whole different philosophy. But where did that come from? Like, what happens? When did, what day did that switch?

Gretchen, on the other hand, teaches in a small district, and describes a current practice that is very child-centered that was not always present under previous administration:

When I first started, I definitely felt like there was that close-your-door and mind-your-own-business attitude and through [new administrators], that philosophy really is not tolerated, and I don’t think it was done in a mean or ugly way. It’s just done in a way that in order to reach all our kids, we have to be a team and we have to be able to share our knowledge and work together in order to build strong kinds. . . . There’s a cohesiveness, there’s a trust among us.

Other teachers also described strong collaborative teams in which data was used to inform child-centered instructional practices. Teachers reported positive viewpoints of administrative involvement in data meetings when the principal respected teachers’

professional judgments, such as Betty's report of a principal who attends meetings as the secretary, alleviating teachers from the task of taking notes as well as placing herself in a support role that minimizes her perceived domination of the meeting. Whether teachers viewed their data team practices as a positive influence or not seemed to depend in part on the nature of the relationship among teachers and the *They* who joined their data team meetings.

Collegial learning. Data teams and grade level teams were almost unanimously described as formal groups that shared a common purpose and practice, but teachers placed greater learning value on the importance of informal relationships with colleagues. While on occasion, this was done through a formal mentoring program for new teachers, the organic, grassroots-support of other teachers was particularly instrumental for teachers early in their careers as they navigated the context of their work with limited experience. Some of the veteran teachers explained how they have transitioned into being a mentor, and while the early familiarization with the realities of the profession was the most critical period within which fellow teachers influenced each other, most teachers describe an ongoing orientation towards learning from and with each other, regardless of the stage of their career and regardless of whether the discussion was about the construct of CA or DAP.

Mentors. A number of teachers began their careers in a district that had a formal mentoring program, but teachers frequently spoke about mentors in unofficial ways, not as a title, but as a role someone assumed. Ruby was one of the newest teachers interviewed, and she expressed her appreciation for those who helped her acclimatize to her teaching role. She said,

I'm really, really grateful for the mentors I've had and the systems that I've been a part of that have helped me to be as intentional as I can with the demands that we have surrounding us these days in schools.

Mid-career teacher Jill described how, despite already having a decade of experience, she relied on a mentor when moving from kindergarten to a second grade classroom:

“Fortunately, next door to me was a veteran teacher who had taught second for 20 plus years. She took me under her wing.”

First grade teacher Gretchen explained the role an expert teacher on her grade level has for all her team members, as well as the collective role they share in supporting each other:

There is a teacher in our grade level who's been there I think close to 30 years and I mean, she's like, The Knowledge. We go to her and ask those questions that you ask to people that have been around a long time. . . . We have each other. We're not alone on an island, so if somebody's having an issue, we can take it to the group. . . . It's like a village. We need to help each other out in order to make our kids successful, and at the end of the day, that's what it's about.

Supporting colleagues. Sometimes a relationship has a clear mentor, like “The Knowledge” illustrated by Gretchen. But just as often, it is the casual, undefined relationships marked by openness and sharing that teachers described as influencing their practice the most. Veteran teacher Maggie claimed:

I would say 80% of what I've learned, I've learned from friends that teach. . . .

Teachers have lots and lots and lots of ideas and you know, I'm not the tell all and they aren't either and when you can share, it's just wonderful . . . and that's how I

have gotten through the years, how I have gotten a lot of my lessons and a lot of my supplies is just from talking to other teachers.

While learning from other teachers applied to matters of management and instruction, teachers also said it specifically influenced their approaches to assessment. Veteran teacher Annabelle said she learned how to assess “from colleagues . . . watching what they do.” Veteran Darlene also described learning about assessment from older teachers who were willing to explain things or let her watch in their classrooms. Early-career teacher Michiko describes this as part of her active practice of learning (“I kinda just watch what they’re doing”) and described how she is part of a line of knowledge and experience that is getting “passed down from previous teachers.” Veteran teacher Darlene was an English as a Second Language teacher for a period of time, and she described the opportunity to be in every other teacher’s classroom supporting English language learners as an incredible learning opportunity as she could observe many styles, practices, and philosophies in action.

Figure 5 suggests how the individual and other teachers interact with each other in the context of the shared practice of *We*. But the individual is not relinquished to the group, and in some cases, the remaining *I* can be strong enough to completely retain its individual characteristics and perspectives despite the collective context.

Remaining *I*. Despite the heavy emphasis on the influence of *We*, many of the teachers in this study described ways in which they exerted their professional autonomy, sometimes even when they felt their approach was at odds with the other members of the *We* team or even the requirements of *They*. The oval of the *Individual* in *Figure 5* shows that it is embedded in a team context, but it may not always behave in common ways with other teachers. The examples and explanations provided here illustrate some of the ways

in which a teacher preserves elements of the individual *I* when approaching assessment or DAP.

Although there was no attempt to correlate factors such as experience, it is noteworthy that the teachers who described highly independent practices tended to be more experienced practitioners who were confident in their knowledge and/or philosophical orientation towards teaching children. A strong remaining *I* was also more frequently portrayed by teachers who were describing their focus on developmentally appropriate practices, and where *Individual* assessment practices were described, it tended to be in light of teachers making their own adaptations in order to better address their students' developmental needs in assessment decisions.

Learning orientation. As previously mentioned, because of the voluntary nature of this study, it was more likely to attract participants who are motivated to grow and assume responsibility for their own development and learning as a professional. Many of the teachers described themselves as self-motivated learners (Laurel claimed, "I am a sponge") and avid consumers of professional resources ("I read a boatload of books," said Pat.) Despite having taught for over three decades, Maggie described how she had sometimes found working on big teams to be difficult because she admitted she is, "kind of more set in my ways." Now, she is the only kindergarten teacher at her school, and she stated her motivation to learn more about assessment:

I don't want to get stale in what I'm teaching so I try to keep pretty current with that type of thing. So, I think a person that is the only teacher [in a grade level], you have to be a pretty strong person to be able to kind of see that about yourself.

While many teachers talked about their individual learning in response to the needs they encountered while in the field, there was a mix of experiences in relation to

whether that learning was an influence on others or whether the new knowledge stayed with the individual.

Closed doors. Laurel, the self-proclaimed “sponge” of a learner, has found some resistance when she has tried to share new ideas with some of her teammates. But when she has encountered opposition, she focuses on the learners in her classroom over collective unity just for unity’s sake. “This is really bad,” she confessed. “These are my four walls and that door closes. As long as I know I’m doing right by my children...” She was not the only teacher who talked about her door.

“I like to just close your door,” veteran teacher Darlene unapologetically proclaimed. “You do your thing with your students.” Fellow veteran Maya, who works in a different mid-sized district than Darlene does, also talked about closing her door:

That’s exactly what I do. When it comes to assessments, I do try to follow it. I mean, especially if it’s county assessments. [But] I do things on my own all the time, it’s just how I am. I have to be comfortable in my own room. . . . The saying is, ‘she does her own thing.’

Pat is not as covert when she defies the shared practice decided by the team. When she gives a formal, common assessment and sees that a student does not perform as well as she knows from her informal assessment the student could perform, she takes initiative: “Then I create five other assessments that will highlight what they can do. . . . And it makes my teammates crazy because I will over-grade kids. I will grade everything [to ensure a poor test result doesn’t misrepresent a child.]” This kind of autonomous practice is described in various ways as teachers navigate best practices in assessment.

Autonomous practice. The amount of oversight and dominance exerted over teachers from *They* can influence how much individuality remains in a teacher's approach towards assessment or DAP, but other individual traits of confidence and conviction also seem to be influential. Veteran teacher Kim described her uniqueness in the following way:

I am a little bit of an anomaly in the fact that I am comfortable going outside of the box. I'm not uncomfortable explaining why I'm doing [something different]. So, a lot of times mine is different. Our team sits down and says, "These are the assessments we're going to do." I present to them, typically: "Well, I understand why you're doing it that way. I'm going to try to assess it this way. If anybody's interested, I can give you more information." So, I offer it to them also.

While Kim stated she still had good relationships despite the times when she does things a different way, Michelle acknowledged that probably some of her colleagues were not so fond of her because she is strong-willed and her orientation does not align with theirs. She admitted:

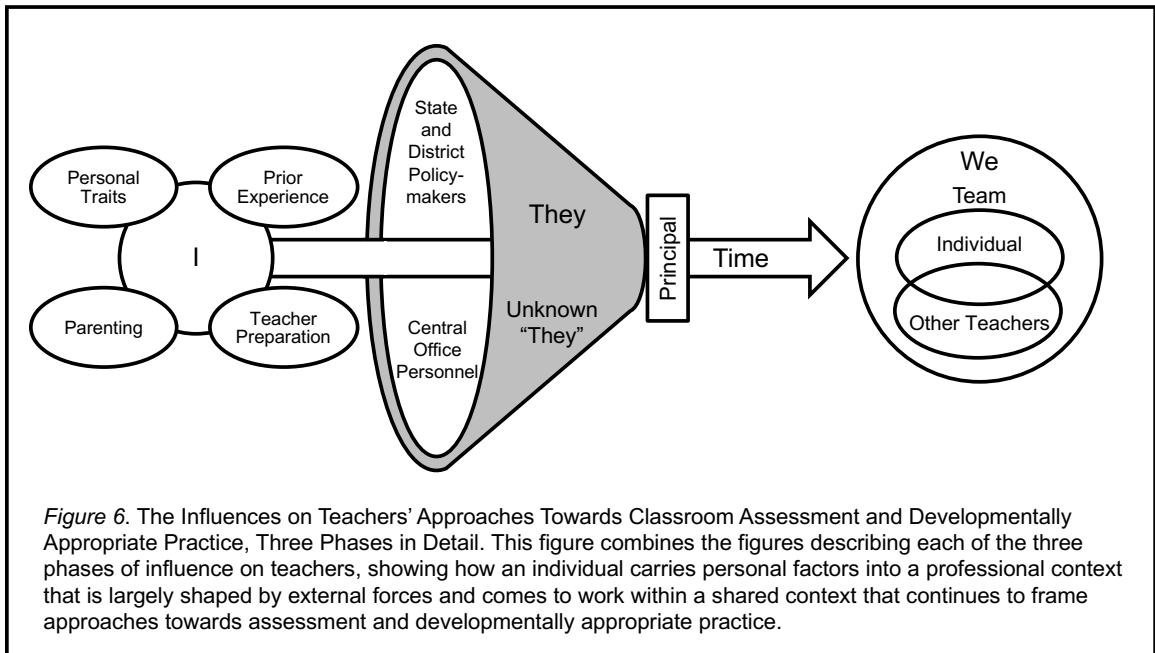
There's some that weren't really fond of me in my teaching because I was so child-centered and not as hardcore as they wanted to be. I always just roll back to that idea that they're five, six, and seven years old. You think about the fact they've only been on the earth since 2011. And my group, 2013, some of them, you know, '12. ...but I just have to roll with that. And luckily, I'm old enough that I, you just pick your battles and I don't need to be your friend and I don't need to do what you do. I need to do what's good for me.

Michelle lightheartedly attributed some of her confidence to being "old enough," but there is certainly an element of time that influences teachers' approaches toward

assessment and implementation of DAP. In the next section, I show how time connected the *I*, *They*, and *We* while shaping teachers.

Time

Time is the variable that moves teachers through the three phases of influence. *Figure 6* illustrates how the pre-service *I* moves through the *They* as she enters the field, being shaped by context filtered through a principal. As time passes, the teacher settles into a *We* context in which her practice is enacted. This practice is not static, and can continue to evolve as other factors change, such as if the teacher shifts to a different district, works under a new principal, experiences turnover on her teaching team, becomes a parent for the first time, or simply continues teaching over more years.



Although the arrow points in one direction, it is important to remember that not all teachers experience these phases in a strictly linear fashion. But virtually every teacher identified time and experience as crucial influences on her knowledge of DAP and

approach towards assessment of young children. In this section I describe the properties of time as experiential time, elapsed time, and the cycle of time.

Experiential time. As already discussed in the earlier section on teacher preparation, the theoretical knowledge acquired in the pre-service classroom can lay a solid foundation, but alone is insufficient preparation for teaching. Some types of learning can only be acquired through practical, on-the-job training. As early-career teacher Nikki stated, “with teaching, you just don’t know until you get in there. It’s just one of those things, you have to get in there and get your feet wet, see what’s happening.”

Veteran teacher Bev disclosed that while she knew she had a course in assessment, she could no longer remember it, and her assessment knowledge has been learned through trial and error as well as working in an intense setting when her school was under focus for accreditation standards, which she described as a stressful time but one that led to powerful professional growth, the kind that can only be accomplished by *doing*.

Veteran teacher Janice agreed that understanding assessment has required action over studies or training. She confessed that she did not develop assessment skills, “until I got in the trenches and was actually on my own having to do it.” Given that she was having to navigate the trenches on her own, it is not hard to see how well-intended teachers could end up developing assessment practices that are not valid or reasonable for young children. This highlights a potential danger in experiential learning, one which Karla revealed while admitting her understanding has come largely from experience: “you have to be really careful there because you know, you’re only operating within your own experiences.” While learning by experience requires time, this kind of time has a

slightly different nuance than the *elapsed time* reportedly needed for understanding child development.

Elapsed time. Veteran teacher Betty praised her early childhood degree program, but still said it took time to know what children at different stages are really like: “That was a good foundation, I guess, but most of my experience has just come from doing it and just gradually over the years coming to understand this age.” This kind of understanding requires an accumulation of experiences across time. Liz has taught kindergarten for over 30 years and credited her longevity with giving her a depth of insight into what is developmentally appropriate for her students: “I’ve lived in a five-year old’s world for a long time.”

Other teachers expressed similar sentiments about the importance of passing time, including veteran first grade teacher Darlene who admitted, “I just think watching that process over and over helps you understand where they should be. . . . It comes down to experience.” Mid-career kindergarten teacher Nancy revealed that she also needed a lot of time and practice to understand her students:

It took a lot for me to figure out what “developmentally appropriate” really meant because what I was expecting in my mind was not practical. . . . I had to figure out over many years to meet the students where they were.

While elapsed time appeared to be particularly influential on teachers’ understanding and implementation of developmentally appropriate practice, it also influenced teachers’ approaches toward assessment. Veteran teacher Maggie’s statement demonstrates this:

As I've taught year after year and I just feel a little bit more competent in my abilities and just from looking at test scores and things like that . . . I feel a little bit more empowered. . . . I feel like I can do it.

Passing time can build up teachers' confidence and sense of empowerment, as Maggie described. But other teachers described how the passage of time also can allow resistant teachers to ride out new movements in education because they believe things will change again in time.

Cyclical time. Many teachers acknowledged the pendulum or cycle at work in education systems, and they often classified assessment practices as changing trends. It was not only veteran teachers who had endured changes who identified this phenomenon. Early-career teacher Karla interpreted the issue:

Teachers can—not all, but—teachers can definitely get locked into this way of thinking and be resistant to change. . . . You do see where that cycles. Like, every two years we're starting something new. And then the teacher who's been here 30 years is like, "I'm gonna keep doing my thing 'cause it's always worked."

Mid-career teacher Michelle also spoke about this in regard to other teachers, claiming they have "learnt that in their experience that it swings, the pendulum swings, it goes this way and then it comes back," and thus they lack a sense of urgency to embrace new practices and methods of assessing. Veteran teacher Maya placed herself squarely in this mindset when she asserted:

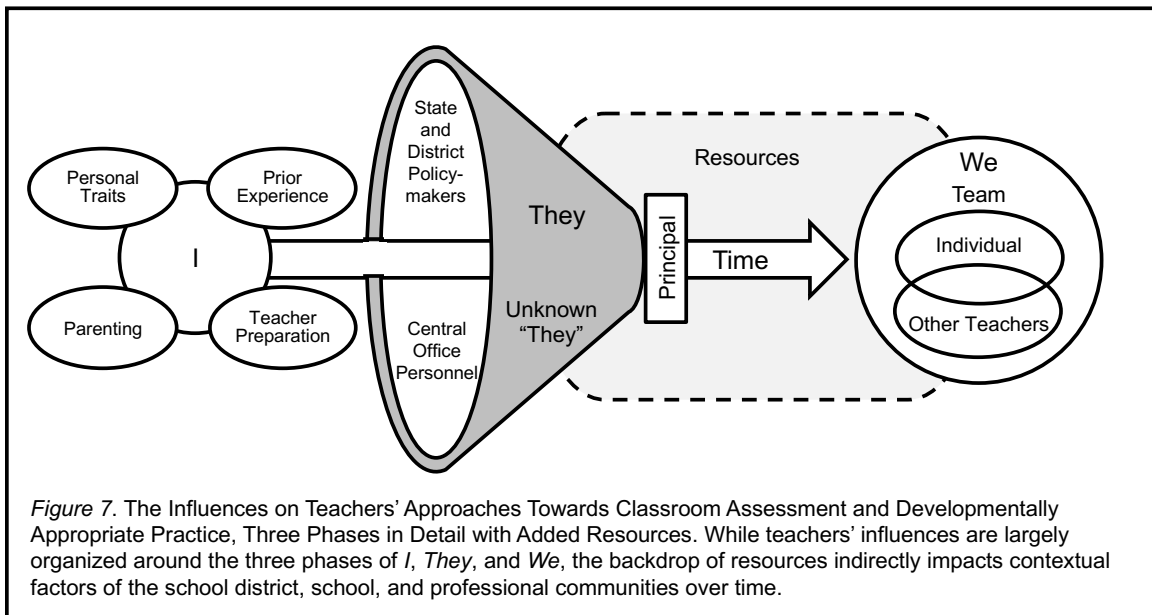
You have to understand that it's all going to comeback around again. So, they're going to realize, in my experience this is what I've realized is that if they go too far one way, or too far the other way, they tend to go back. So, right now they're all into the data, and then they're going to realize, "wow, what are we doing to our

kids?” and they're going to go back towards, I'm not saying all the way back towards developmentally appropriate where it used to be, but hopefully there will be a balance.

Maya's use of the pronoun *they* highlights the distinction between *others'* decision to emphasize assessment data as opposed to her involvement in and embrace of the decision. She has enough experience to wait out the cycle, and she is hopeful a new balance will be achieved between a data-orientation and developmentally appropriate practice, two ideas which she positions on opposing sides of the pendulum. This description elucidates change as a long process, but the final figure of this study's theory shows that impacts can also occur through relatively short and discreet events and resources.

Resources

The final element in the model explaining the theory of how teachers develop their approaches towards assessment and DAP is broadly labeled “Resources.” In *Figure 7*, it is displayed as something in the background, that can affect *They* and *We* as well as be an ever-present backdrop across the passage of time, but it does not serve as a primary driver or major influence. The dotted line around the figure is to indicate that this category is somewhat more nebulous than the others, and it includes both tangible and intangible assets, materials, and opportunities. While unique resources could influence teachers in particular contexts, the main properties of this category are teaching assistants, curriculum guides, supplies, technology, and professional development.



Teaching assistants. There was no consistency across the teachers interviewed regarding whether they had teaching assistants in their classroom, but teachers who had other adults in their classroom reported being able to do more one-on-one assessments with their students. Teachers related the difficulty of focusing on a single student with 20 other children in the room, many of whom have not yet learned to work independently and are naturally attention-seeking. When teachers had the support of an assistant, the assistant could manage the rest of the class while the teacher did individual assessment work. Some teachers reported having volunteers who could do basic assessments of discrete skills with young children, such as a high school student who listened to kindergarteners count and recorded the highest number each child reached before making a mistake. Teachers who did not have additional adult assistance found it more difficult to manage a classroom and give individual assessments, and therefore may not have been able to assess students one-on-one as frequently. Teaching assistants could indirectly impact assessment practices or developmentally appropriate practices because the additional support opened up opportunities for different classroom strategies and formats,

but there were no reports of assistants influencing knowledge or beliefs about either construct.

Curriculum and pacing guides. While curriculum and assessment are inextricably linked, there were specific attributes of curriculum that teachers reported as influences. Scripted programs were frequently critiqued, as they do not allow instruction geared to students' needs and therefore associated assessment practices yielded little useful information about what students can do. As veteran teacher Lyn stated:

In order to meet the needs of where each child is and move each child forward, you have to deviate from pacing guides or from scripted programs. It just doesn't work otherwise. Even if it's a good quality scripted program, it's not going to meet everyone.

The type of curricular program was an issue for some teachers, as were the intended outcomes set by *They*, including districts' pacing guides and benchmark assessments. Veteran first grade teacher Maya recalled encountering poorly leveled assessments required by her district when she moved down from teaching third grade to first because she said some of the same benchmark math questions within the online PowerSchool system were used in both grades. When people asked how she liked first grade, she responded, "Well, I'm teaching third grade to six-year olds." Beth described how assessments are not always properly sequenced with the SOLs, such as a kindergarten assessment in which students had to show they could count the number of dots on a page, an appropriate outcome for kindergarten, by circling the matching number word, despite reading number words being a first grade SOL.

Supplies. Closely tied to curriculum materials as a potential influence are the educational supplies available for teachers to use. Many teachers described the various

supplies beyond standard paper and pencils that they used to have students demonstrate their learning, such as using fake coins as a manipulative for a unit test in math or white boards and markers to informally assess early writing and phonics skills. Several teachers described using resources they found in teaching manuals or curriculum-based kits or from internet-based resources, such as Teachers Pay Teachers or Pinterest, when they had freedom of choice. Sometimes teachers were compelled to use certain assessments because their district had purchased them, and a few teachers reported a lack of supplies impacting their practice. For instance, Ruby teaches Spanish in a dual language immersion classroom and said her district did not pay for a subscription to PALS Español, so she had to resort to other methods of assessing her students' language and literacy skills in Spanish. In Ruby's case, the impact of *They* on her assessment practice was significant, but in general, supplies were less influential than the impact of technology.

Technology. While K-2 teachers reported using the internet to find resources for their assessments, technology emerged as a significant factor in how teachers assess and, in conjunction, a major source of required developmentally *inappropriate* practices. There was an overwhelming sense that tech-based assessment formats were inappropriate for students in early elementary grades and result in invalid information on student learning. For the sake of focus on the theory regarding influences on teachers' approaches, I do not address the plethora of ways teachers found these formats ill-suited for gathering meaningful information about young children's learning here, and I instead focus on the two main ways technology influenced teachers' assessment practices.

Requirements to use technology. Teachers widely reported assessment for compliance measures that required the use of computers. In some cases, the school or

district had purchased a program that would allow teachers to create assessments and track student results for ease of data collection. Even when teachers did not want to use such a format, they reported having to do so. Second grade teacher Melanie works in a mid-sized school district and stated,

We have to do it on the computer, and they try to prepare the students—talk about not developmentally appropriate—they want us to try to prepare the students for the SOLs in third grade, but they want them to see those TEI items.

TEI (Technology Enhanced Items) were widely panned as inappropriate for young learners. TEI include questions asking students to click and drag images, rotate and manipulate shapes, or select all answers that apply. Though TEI are promoted as a way to gather information that is difficult to collect through traditional test question formats, most K-2 teachers found that students did not have the manual dexterity, abstract reasoning, or patience to use the computer programs to show their learning. Furthermore, K-2 teachers preferred to utilize demonstrations of learning that are more authentic than traditional test questions anyway. Second grade teacher Pat expressed her concern regarding the influence over practices that these test requirements have, and how they impact her ability to make accurate conclusions about student learning:

We don't want them to fail because they don't understand the, the device, the delivery. We want them to succeed because they understand the content. That's where the disconnect can happen. If a kid doesn't understand how to use the medium, then it doesn't matter if he can tell you what the answer is.

Not all teachers felt all forms of technology were impediments in assessing students. First grade teacher Marsha appreciates the opportunities technology affords her students in being able to capture and record their oral language, such as through the iPad

application Chatterpix. However, the way this technology influences how she can assess is a more informal use as opposed to scenarios when *They* require assessments to be done on a computer, and it requires access to working technology.

Access to technology. Marsha's school district has implemented a 1-to-1 initiative, so her students have access to individual iPads. Teachers who reported only having a few computers in their classroom, a single class set of computers shared across a grade level, or limited time in a computer lab stated that they struggled to use technology to assess in a whole class setting. Furthermore, even in contexts where machines are available for every student, bandwidth support could be an issue. Multiple teachers reported nightmare scenarios where assessment programs froze or crashed when an entire class of students attempted to go online at the same time. "You just pray the technology is going to work the day you have to use it," stated veteran teacher Margaret, who works in a rural school. When the technology was unreliable, teachers expressed frustration with required assessments and tended to entirely avoid using technology to assess in order to avoid the frustration and wasting time.

Professional development. Another resource with promising opportunity, but sometimes considered by teachers to be a waste of time if not functionally applicable, was professional development. Several teachers interviewed described themselves as proactive about seeking individual professional development (PD) opportunities. According to early-career first-grade teacher Pauline, "I am one to, where if I'm struggling with something, I do seek PD. So, I have some kind of source to help me with that issue." Mid-career second-grade teacher Pat professed, "I'm kind of a weirdo that I like going to professional development, learning how other people do things and what works for them."

Although Pat described herself in a self-deprecating manner, implying she is unique, in the particular sample of teachers who volunteered for this study, there were quite a few who portrayed themselves as self-motivated in continuing to learn and grow in their careers through PD, personal reading materials, or attending conferences. However, this type of learning resided in the *Individual* and rarely influenced the *We*. The most common type of PD described was in-service sessions offered in schools or districts for all teachers on staff, the whole context of *We*.

Foundations. Although very few teachers could recall an assessment course in their preservice training, several teachers reported having valuable professional development on the foundations of assessment provided through their school. Beth, Laura, and Nancy teach in different districts, but all reported having professional development delivered by non-school personnel within their school context. This PD included the basic concepts of validity and reliability, aligning curriculum and assessment using Bloom's taxonomy, using a table of specifications to create assessments, and evaluating what makes good test items. While those who experienced this type of PD valued the learning, Beth indicated that it applied differently to earlier grades than older grades and other teachers commented that the impact was minimal if given in a one-time workshop style of delivery. More teachers stated they had not been offered such professional development and cannot recall anyone addressing the fundamentals of assessment. No teachers reported any PD addressing DAP in any manner.

Training. On the other hand, many teachers did report receiving training on how to give and score specific assessments. Although this type of training was labeled PD by teachers, I describe it separately as *training* here, because the goals and outcomes are different. Unlike PD that is intended to build professional capacities within teachers that

can be applied using judgment to a variety of matters, training is designed to increase the fidelity with which teachers carry out practices largely removed from professional decision-making. PD is more likely to influence assessment for *instruction* whereas training is needed for teachers to properly carry out assessment for *compliance*. These can have a heavy overlay; several teachers reported that they received training to give the PALS, a requirement, but then they used those PALS results to inform their instruction using their professional judgment.

Beyond the PALS, teachers reported receiving training to give assessments using Wilson Phonics, the DRA, and the Interactive Achievement program through PowerSchool, among others. Veteran teacher Kim works in a large school district, and she speculated that administrators emphasized training over PD because of an assumption that the assessments were effective and sufficient. “I don’t think they think we need to know why these questions they feel are valid are, but I think sometimes that’s why there’s a disassociation and I don’t think some of the questions are valid,” she pondered.

Other teachers questioned the value of formal PD or training, as did veteran kindergarten teacher Sangita:

I think for as many workshops and conferences and junk that you can go to, the best way [to learn] is to observe your peers. ... You learn from others. You have to, the people that know what they’re doing. Especially in kindergarten.

Kindergarten. The distinction of kindergarten is one that spanned multiple issues, including professional development. Sangita continued to discuss this issue by claiming that PD sessions she has found helpful are “the ones that say ‘Kindergarten.’ And that’s it. Not K-2. Not K-1. ... professional development needs to be geared to just me.” Early-career kindergarten teacher Michiko stated that unless something is geared towards K or

K-1, it is not really helpful, and veteran kindergarten teacher Janice has learned to be more assertive in expressing her need for kindergarten-specific work:

I'm sorry, but any session I sit in that's K-5, it's not going to be for me. Even K-2, I have to really push to get stuff out. I prefer if I'm going to do training, I want K training.

Summary of influence of resources. These discreet resources can shape teachers' assessment practices, but they can also give a false sense of understanding assessment. With the possible exception of PD, these resources are likely to permit, prohibit, or improve certain practices, but they do not inherently influence how teachers view assessment as a construct or how they approach assessment as a philosophy of gathering data on student learning. When resources ease teachers' workload with administering, checking, and tracking assessment practices, they can imply that facility equates with validity. This category of Resources as an influencer should thus be interpreted with some caution. It should be noted as something that does not dominate the landscape, but rests in the background for teachers' practices, especially in the area of DAP, which was not reported as ever being addressed in PD and only very indirectly impacted by available material resources.

Examples of Teachers in the Theory

As the purpose of grounded theory is to produce a theory that explains a process or relationships in a way that makes "implicit processes and structures visible" (Charmaz, 2014, p. 133), the model presented in *Figure 7* serves as a kind of summary of the study. However, in order to better understand how it works to explain the influences on a teacher's approach to assessment, I present a few examples of teachers with different

combinations of experiences in the *I*, *They*, and *We* phases of development, but all of whom have a journey that can be explained using the theory.

Veteran and Early-career Teachers

The heavy emphasis teachers placed on the value of experience and time to learn about both assessment practices and developmentally appropriate practices suggests that there could be a difference between veteran teachers and early-career teachers regarding how they approach assessing young children. While this study did not endeavor to measure, compare, or make correlations between such variables, a look at the difference between an early-career teacher and a veteran teacher can help explicate the theory of influences on teachers.

Veteran teacher. Liz has taught kindergarten for over 30 years, and she has recently begun teaching in Virginia after spending most of her career in a different state, much of that in a charter school system. While her recollection of her teacher preparation program and other pre-service experiences is limited, she expressed the importance of experience over time in shaping her individual approach to assessments. Early in her career she observed colleagues to better understand working with kindergarteners. Liz recalled, “watching how they dealt with developmental issues and things like that when I was a young teacher I think helped. And then it's just getting in there and doing it and trying something.”

She describes her collaborative approach to being part of a team and learning from other teachers:

I have been one that I have always enjoyed sharing ideas with other teachers. And when I first started teaching, I always was asked to go and be on teams to look at new innovative ideas, year-round school or looping or transitional classes. And

so, I had the opportunity of traveling around the state . . . and visiting lots of different schools. And in all that visiting classrooms, I saw lots of different teaching styles and methods from other teachers. I've always said watching someone else do something will teach you a lot.

From her many years of living in a five-year old's world, Liz has nurtured a very developmentally appropriate approach towards children, and she lamented the changes in kindergarten and the pressure placed on young children to perform before they are ready. She contended:

I think we have gotten away from developmentally appropriate practices in kindergarten. Kindergarten is what I would consider first grade from years ago. Everybody forgets the little things have only been breathing for five years. . . . But I've always said to parents, "You don't stick a newborn on a bicycle and expect them to pedal." . . . There's nothing you can do except let that little person mature and give them that gift of time.

Liz brought her strong commitment to early childhood with her to Virginia, but now she is navigating a new *They* with different requirements and allowances. She questioned, "Who's mandating that we do this? Where did this come from? And what are we going to do with all the information that we're gathering? What are we doing with it all?" Although she is working to determine the tangle of required assessments and uses of the collected data, she has maintained her child-centered approach developed over so many years of experience.

Early-career teacher. Early-career teacher Pauline has worked in the same small district for her whole career and even completed her student teaching in the same school. As a result, all of her experience has come in the same context of *They*. She has taught

kindergarten, third grade, and now first grade, and she admitted that she does not have a specific background in early childhood, and that may influence her assessment practices:

I think it does for me more now that I have taught third grade reading and I know their expectations, and I know what they need to know. . . . Even now I'll try to use those question stems just so they're, they're hearing 'em and maybe they know how to use them. But I do, me personally, I hate that I think that way because it, I mean that's not what should matter but we are in an area where we struggle with being accredited sometimes and right now we are [accredited], but of course at any time we may not be, so we've got, we're like a little roller coaster.

In comparison to Liz, Pauline does not have years and years of developing her individual approach to fall back upon, and so her approach to assessment more closely resembles that of her district's requirements. Her *I* and *We* practices are very similar, and *They* strongly shapes both. She described a collaborative team and a lot of shared methods regarding assessments. Pauline stated that her team often creates the district assessments, as they are the only elementary school in the district. They have had training sessions for using specific assessments, but her team largely relies on teachers' individual's strengths and backgrounds for understanding how to create, deliver, and interpret assessments. As such, she has been able to influence the literacy assessments in her grade level, but she acknowledged that the SOLs drive their collective practice more than child development:

I think we push so much for the standards that we may not consider whether this is developmentally appropriate or not. And I'm guilty of that as well. Where [we might] give them something that I know, this really isn't okay to be giving you right now. But I hate to say it, like, I'm a little puppet and okay, we'll get it done.

But yeah, because it's a lot we have to get done and a lot of it just means, we forget their age, I think.

Pauline's experiences differ in part from Liz's because she has not had as much experience to build up an *I* approach separate from her *We* context, heavily shaped by a *They* focused on accreditation. This study did not endeavor to measure distinctions such as these, but it is noteworthy that Liz described an ability to maintain a stronger developmental approach to assessment than could Pauline in her more directed context.

Two Teachers, One School

Because the context of *They* can significantly influence how K-2 teachers approach assessment, it is helpful to look at how two teachers in the same *They* context have developed as individuals and how different grade levels could influence their approaches. This section looks at how Melanie and Michiko, who teach at the same school in a mid-sized district described their influences in approaching assessment. It should be noted that there is not control over other significant variables of influence; Melanie is a veteran teacher and Michiko is an early-career teacher, and the distinction here could also be reflected in a comparison.

Kindergarten teacher. Early-career kindergarten teacher Michiko entered the field with a degree in early elementary education and previous work experience at a preschool, so she felt she brought a good foundational understanding of child development into the field. Yet, "if I could do it all over again, I would beg to at least have two years of student teaching. You don't understand. I don't think that I was prepared at all for what was expected of me," she acknowledged.

Michiko taught in fourth grade but quickly moved down to kindergarten and has learned a lot from her colleagues. "I didn't have anyone to rely on in fourth grade," she

confessed. “I definitely like the age of kindergarteners better . . . and I was able to have somebody else to rely on when I moved down.” As her preparation program had no course in assessment and she has had training in specific test administration but not professional development in assessment foundations, Michiko admitted that most of her assessment knowledge has come from her colleagues. When she started teaching kindergarten, one of her teammates had been teaching for over 20 years and the other for over 10. She revealed:

I really used what they had done in the past. They definitely knew what they’re talking about. Um, I think with them, too, it was kind of like trial and error and, I’m learning from the next one, the next person who was here before me, um, and kind of just passed down.

Michiko talked about requirements from the county (PALS testing three times a year, for instance), but did not describe an overbearing *They*. She called her current principal helpful, willing, and supportive, but did acknowledge that he did not have experience teaching early elementary, and therefore she sometimes finds that a challenge. She claimed, “a lot of people don’t understand what it’s like to teach K through 2, K through 3, and how different it really is from the upper grades, even though we are all in the same elementary school.” She appreciates the flexibility and freedom in her context, but would also like some additional support and guidance. She commented:

It would be nice to have something where I could just go and, you know, this follows the second nine weeks report card or whatever it is. . . . instead of kind of just like, okay, well I guess this is what, um, you know, I should use. I’m just kind of second guessing myself basically.

Michiko retold of a cumulative test that used to be required in kindergarten math that was, “page upon page upon page. And even just the fact of having a kindergartener turn a page and go on to the back is a lot of work. Doing this multiple times was stressing them out, stressing us out.” When teachers complained about it, she found “*they* listened.” Though her principal may not have a thorough understanding of early childhood practices, she believes he is open to feedback and so the *We* has some control over their assessment practices, and they were able to get rid of the test that she thought was “a true waste of time.” But this pushback and response only applied to challenging specific practices or actions; there was not a proactive advocacy effort made to advance to DAP.

Second grade teacher. Veteran teacher Melanie currently teaches second grade at the same school as Michiko, and she expressed a similar perspective about having a supportive principal who empowers teachers with flexibility in their teaching. She reported that he tells faculty, ““find purposeful ways to assess your kids. . . . we’re not here to get you. We don’t want kids to have bad grades. We want them to learn, that’s the most important thing.”” But the influence of *They* still shapes practice and can change with different leadership. Melanie stated, “it depends on the support of your administration and your school board office and how they understand developmentally [appropriate practice] and then, you know, then you fall into the standards and what you have to have there.” The hierarchy of authority exists even if it is not overpowering.

The hierarchy creates the trickle-down effect of state expectations that have a stronger impact on her second grade approach than when she previously taught in kindergarten. Melanie claimed, “some administrators really see that and understand it, but so often we're told, ‘oh yeah, we know, but, we have to get them to this point.’ And that's

where I think teachers feel that frustration is.” In second grade, teachers are required to give computer-based math assessments so the students are exposed to TEI prior to SOL testing. “It’s stressful for the kids and we hate it,” she succinctly stated, but she follows the directive and gives those assessments for compliance.

While Melanie reports that her *We* team all dislikes what they view as developmentally inappropriate computer-based assessment practice, she also points out that, like Michiko, her team would still appreciate more guidance in some areas of assessment. “Teachers don’t want to always be told what to do, but then we want to be given some guidelines as to what we’re supposed to do as well,” she confessed. Teachers just hope those guidelines will include developmentally appropriate assessment practices, something which is de-emphasized in second grade in favor of preparation for the third grade SOL testing year.

One School District, Two Teachers

The next pair of examples is a comparison of two teachers who work in different schools in the same large school district, Lyn and Maria. Although the district and central office personnel are dominant forces in their shared contexts of *They*, the distinctions of how the principal has served as a mediator in their settings has had a heavy influence over their current practices. Furthermore, the past experiences of the *Individual*, and the differences in their teaching teams has also shaped these two teachers in different ways. Each of them affirmed during interviews the theory as having explanatory power in their situation, even though their paths and major influences are very different.

Lyn’s story. Lyn is a veteran kindergarten teacher who has taught in several contexts, including in another state and overseas. Her *I* has been largely shaped by two teaching experiences: her first year and her time overseas. Lyn described her first year of

teaching as one heavily influenced by resources; she worked in a low-income school district with few materials, and teachers would wait hours to make copies on a single ditto machine. Along with another new teacher, Lyn realized that she could avoid the late nights waiting her turn to make copies *and* create authentic learning experiences for her students that were more effective than using worksheets. So, she started teaching with more hands-on methods and has continued doing so ever since. She learned early on to be creative and has maintained a child-centered, hands-on approach to teaching and assessing for over 25 years. While teaching overseas, she was in a professional context that provided her freedom and flexibility while the school adopted the PYP curriculum. The supplies and curriculum shaped her practices and approaches to teaching, and in turn, to assessment.

For over a decade now, Lyn has worked in a large school district in Virginia. Initially, she was at a PYP school where she was encouraged to continue teaching in her child-centered way. Her fellow grade level teachers do not share her approach. “Everything that we do in my class is hands-on and they prefer worksheets still,” she stated. Lyn continued:

I mean the three of us get along really well and we're a really a good team and we all work towards the same goals, but we just don't share the same ideas or practices in actual . . . I guess, when it comes to our pedagogical style.

Their pedagogical styles are about both how they teach and how they assess. Lyn compared her approach to assessment with that of her teammates, saying,

I like to do just something quick, easy, something that doesn't always feel like a test and something that is actually connected to the way I teach. I'm not one of those people that likes to give the little kids a multiple choice test when that's not

actually how we're actually testing them just so that they can learn how to do the third grade test. . . . We're teaching them the alphabet. That's not a multiple choice test. They actually have to know all the letters and the sounds.

Lyn has been allowed to work in her own way within the context of a team that has different practices because her principal's filter has kept out some of the district's heavy management for compliance and allowed her to continue teaching in her own style. According to Lyn, when her school lost PYP funding several years ago, her then-principal told her, "'Just shut the door. Make sure your SOL matches the pacing guide and go about your life.' That's what I've done." Although she no longer has the same principal, subsequent administrators have similarly permitted her to continue to do things her way, as long as she completes the minimum required district assessments.

After describing the theory of influences, Lyn confirmed that it explained her situation, wherein she had a strong *Individual* approach, but the shaping of *They* into a common *We* was minimal in her case because her principal's filter keeps top-down pressures at bay and has open spaces for confident, capable teachers to pass through, retaining their individualism. "Thank goodness that I get to go through those holes as you explained," she professed.

Maria's story. Maria is an early-career teacher who works at a different school than Lyn in the same large district. This year is her first year in the district after having taught in a private school since the start of her career. She states that her teacher preparation program helped, but the biggest influence on her understanding of child development was having children of her own. In her private school, the *They* influence was negligible. There were neither required or even provided assessments, and with only seven children in her classroom, she was able to assess constantly in informal ways. "I

knew my kids,” she claimed. “Like I could tell by the look on their face if they were getting it or not.”

Now, she teaches second grade and the influence of *They* is overbearing. She does not get to choose her assessments. She confirmed that she would not need to have any assessment literacy, as all the decisions are made for her and assessments provided from the district. “I feel like all I do is test and data,” she critiqued. Administrators require adherence to a strict pacing guide set by the district because the school has struggled with its accreditation status in recent years. Maria claimed:

They don't allow any wiggle room in the curriculum whatsoever. And it might be more of a wiggle room if our building was doing better on SOLs, and we weren't under academic whatever. But, um, the state is really coming through, like as of last week, they were really taking a look and they're walking room to room to room.

According to Maria, her principal does not buffer any of the pressure from *They*, and she may in fact amplify it, as the principal is in her first year at the school and responsible for test scores. In response to hearing the theory of influence and how the principal can serve as a filter, she admitted:

Honestly, I felt like this year I haven't been able to get through that filter as much as I would like to where last year, I mean I didn't have a filter and you know, it was pretty much just go and swim and it was fine. But now, it's like, brick wall, a brick wall, a brick wall.

Because her individual practice is being blocked, her practice of *We* is limited. The other members of her team are also required to comply to the district's requirements, and while they do have to carry out common practices of assessment, all of her grade

level teammates are also in their first year at the school. The team is in survival mode, and she does not predict they will also survive long enough to build a collaborative team practice. Maria speculated:

As far as making it to the *We*, I would love to be able to make it to the *We*, the problem is there is no *We* yet, because we're all new and they're not going to stay. So, I'm going to be the only *We* probably left. Um, and I'm okay with that. I'm, I'm hoping that next year then I would be able to help develop some more *We*, um, with other colleagues

Maria would prefer to use more informal observations and quick checks of student learning in the moment so she can adapt her instruction, but with a strict pacing guide, she has to find time to fit in adaptations. “But now I'm starting to see some wiggle room in my own personal teaching, which makes me happy because then I don't have to feel like I'm in a box,” she said. “Like I can be in my classroom box, but I'm not having to be in their box with everything.” She even described techniques for trying to get around the control of *They*: “It takes a little creativity, like if some big wig is about to walk in—because we keep our doors locked—um, I've gotten really good at knowing what my curriculum is so I can always pull it out.”

Maria even described ways that other teachers at her school are working to assert their individual professional judgment despite the context of heavy management:

[Other teachers] are like, “it's called Sneak and Teach.” And I'm like, “what?!” And they're like, “you sneak it and teach it and you do what you need to. . . . I've been Sneak 'n' Teaching and I know that's wrong, but at some point, if I know my kids don't get it, I need to reteach it.

Maria's story shows how the influence of *They* can be so strong as to nearly overwhelm the ability of the *I* or a professional *We* to have any decision in a teacher's approach toward assessment, but when contrasted with Lyn's story, it also shows how a principal can significantly mediate the impacts, and a change in contexts of *They* can potentially allow the same individual to flourish.

Unique Teacher Perspectives

The theory displayed in *Figure 7* is one that explains the patterns of influence of K-2 teachers' approaches to assessment and DAP; even when teachers had unusual circumstances or contextual factors, this theory applied to all of the stories shared in this study. In this section I briefly address some of the more unique circumstances and how the theory illuminated the major influences on each of these teachers.

Influence of parenting. Nikki is an early-career kindergarten teacher in a mid-sized school district, and she portrayed a trajectory across her development that was similar to many other teachers in this study: her pre-service experiences were insufficient to really prepare her, most of her early learning about assessment came on the job from other teachers, her practice is very much embedded in a *We* team orientation, and *They* is composed of people without an understanding of early childhood. A major distinction in Nikki's development is that becoming a parent after she had begun teaching had a dramatic impact on her approach to early childhood and, therefore, her approach towards assessment in early childhood. Nikki retold how she started her career by replicating her teammates' practices:

I mean, I was just doing what everyone else is doing. So, I was like, "I'm just giving out worksheets. I guess this is what I'm doing." I was new to teaching then and I was like, "I guess this is what I'm supposed to do. This is what my team

leaders are telling me to do so it must be right, and it must be working." But then the more I started reading and seeing my son and thinking about it, and I was like, "I don't think this is working at all. And I'm not . . . I could be a lot more effective in my teaching."

When she was home on maternity leave, she educated herself by reading a lot of articles on the internet "about play-based learning and just how important it is for their executive functioning skills and things like that. And studies are proving that this is the way that kids learn." Her independent learning, rather than a professional development event or the ongoing influence of time and her coworkers, shaped her approach to assessment. While other teachers described how being a parent influenced the *I*, the *Individual*, Nikki's description of this influence happening dramatically and after she had begun teaching, was unique. She talked about how the change in her approach is slowly influencing others:

My principal and I kind of worked together and at our school . . . the direction we took things is a more hands-on approach to learning. I feel like my class learns better that way and I seem to have had good testing results that show that my kids are growing even though we are doing some more hands on play.

According to Nikki, her principal is advocating for more developmentally appropriate practices. "He tries hard for us. But it's sort of met with I think, closed ears," she speculated. "Everybody at central [office], I don't think any of them ever taught—I don't even know if they taught at the elementary level." Nonetheless, because the district is not heavy-handed, she and her principal are able to shift the direction of their practice. "I don't want to say leave us alone, but they sort of let us do our thing," Nikki said. "I

mean, of course they implement their mandated stuff like the PALS and the math VKRP and the running records, but other than that they really leave it sort of up to each school.”

As to whether she has been able to influence the rest of her *We* team, Nikki stated: I'm slowly bringing them on board. . . . And one of the girls on my team has kids the same age as mine. So, she is much more on board with me and the other two who are close to retirement are much more sitting at the desk learning. So, I kind of just had to force some of the hands-on stuff, but they're coming around slowly.

Dual language immersion teacher. Early-career teacher Ruby was one of two teachers interviewed who works in a school with a dual language immersion program. She did not just respond to the interview’s opening question with a *We* statement; in anticipation of questions about her approach to assessment, she had prepared a response that “a high priority for assessment is doing team assessments.” The context of the dual language program called for high collaboration and strong teamwork, which Ruby described she has had in her district since she began teaching.

“I started off my career with—which I think is a very rich way to start—being mentored heavily by my teammates, and being very team based with our assessment,” Ruby testified. The team commitment comes “from the top down with our administration, with them coming to our PLCs, our team meetings.” She appreciated the strong leadership, a very positive influence of *They*, that is not always present in every school:

Administration does as much in their power to focus on what’s really enduring and what’s really important in the classroom. . . . Each building has different administrators and they each have a different perspective and opinion about how assessments should be given. . . . I’m really, really grateful for my administration, and my dual team because they have been so focused.

Ruby is also very grateful for a strong *We*. While she acknowledged that she did not have a course or other professional development on the foundations of assessment, she contended that her team is able to share their expertise and work together to understand the whole child. There are a variety of teams that meet: English and Spanish language arts teachers across a grade, Spanish language arts teachers across grades, whole teams of grade level teachers, and other configurations of educators working together to understand the diverse needs of students with a whole-child emphasis.

“Our principal just said at our last dual language meeting, ‘It’s the staff that makes the program,’” Ruby recounted. Her principal and other administrators have shown respect for the staff and their ability to collectively make sound educational decisions around assessment. “That’s one thing I appreciate about my district, well, my school, is that there is a certain level of validity that’s given to that professional, like *mmph*,” Ruby proclaimed. *They* do not have a highly controlling sway and instead entrust teachers to work together to develop a strong *We* practice. Indeed, Ruby expressed very little of the *Individual* in her approach to assessment, but as she felt the *We* had a strong, child-centered emphasis that took DAP into consideration, there was no concern over this.

Teacher with strong philosophical bent. Earlier in this chapter, I posited that teachers categorized the three reasons they do assessment into the three primary purposes as for compliance, for reporting, and for instruction (*Figure 1*). But there was one teacher who presented a very strong philosophical approach towards assessment *as learning*. Gretchen expressed an individual perspective of assessment as part of a process of learning and supporting child growth that was unique among all participants. She began her interview with the declaration, “I feel very strongly that students, especially younger

students, need to have the opportunities to practice before giving an assessment.” This was not just as a means to ensure they were successful on an assessment, but an extension of a strong emphasis on growth mindset as framing her work with first graders.

Gretchen experienced the same influences in varying degrees as other teachers (“you name it, we’re testing it,” she admitted), but she talked about a process of using assessment to build students’ resiliency, not just measure their performance. She likes to give a pre-assessment before introducing a new concept, provide lots of opportunities for students to practice with the concept, then give students a post-assessment and have the children think about how their brains grew and stretched and all they learned. Rather than looking at errors as evidence of failure, she frames mistakes as opportunities to learn. Gretchen functions from a perspective of acceptance that the testing era exists, and rather than trying to fight against it, she works to harness the power of assessments and redirect that power from being punitive to being positive learning.

“As we know later, that test anxiety's going to hit and if they've had a positive experience testing in kindergarten, first, and second, then hopefully that anxiety will not be so great when they get to third grade,” Gretchen reasoned. Her approach towards assessment as a tool for teaching soft skills like flexibility, tenacity, and risk-taking was unique in the study, but the influences on her development of this approach can still be understood through the model presented in *Figure 7*.

Summary of Theory

As this section has shown, there are a wide range of ways that the theory of the influences on teachers’ approaches towards classroom assessment and DAP can be applied. The illustrations given of specific teachers’ stories are not intended to have thoroughly presented each of these individual’s narratives, nor to suggest the full array of

possibilities for combinations of influences on teachers. In the use of this methodology, I have not claimed to have reached *data* saturation, but having attained *theoretical* saturation, the explanatory power of the theory presented can illuminate and give understanding to each of the accounts presented by the 35 teachers in this study.

Summary of Findings

In the fourth chapter, I have presented the findings of this grounded theory study as a framework to interpret what is happening as teachers develop their approaches to assessment and DAP and to demonstrate relationships of different influences on teachers (Charmaz, 2014). First, I described the ways in which teachers perceived the purposes of assessments they give in their classroom. Then, I explained the theory of influences on teachers, including details about the three phases of development, the *I* phase, the *They* phase, and the *We* phase. Finally, I provided a few examples of teachers in relation to the theory to help explicate how it applies in different situations. Chapter 5 takes these themes and discusses them in direct response to the three research questions used to guide this study.

CHAPTER 5

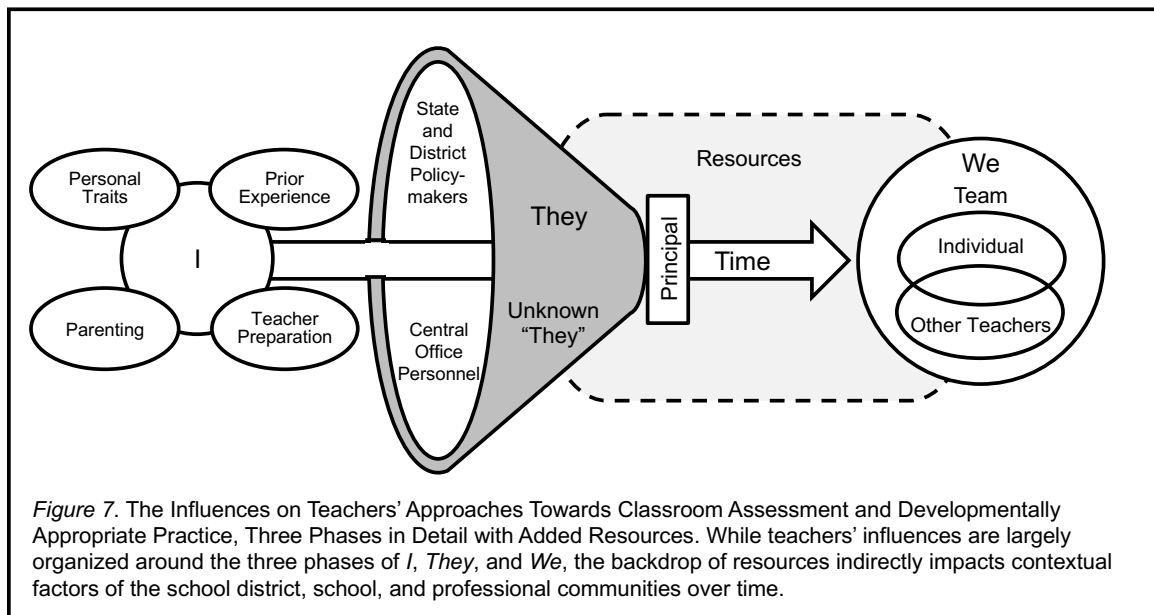
DISCUSSION

As stated in the first chapter, the overarching question driving this study was *How do K-2 teachers come to their conceptualizations regarding developmentally appropriate practices and strong classroom assessment practices?* Data were generated and analyzed with the purpose of developing a theory to explain this process using three guiding research questions. While an extensive number of notable themes emerged from this process, the key categories around which the theory was created are the ones addressed in this chapter. As previously described, grounded theory calls for a delay of a full literature review; now that key themes have emerged, I return to literature in a discussion of the most significant themes here.

The implications for practice and recommendations for research will be addressed in the final chapter. The discussion in this chapter is framed by each of the three research questions:

- (1) Through what processes and influences do experienced K-2 teachers report that they have developed their *approaches to assessment*?
- (2) Through what processes and influences do experienced K-2 teachers report that they have developed their knowledge and implementation of *developmentally appropriate practice*?
- (3) How do experienced K-2 teachers' approaches towards classroom assessment and philosophical orientations towards DAP interact with each other?

In this section, the discussion focuses around the major themes in response to each of the three stated research questions. While this discussion is not exhaustive, it does emphasize the key findings in light of relevant literature. The findings are organized by research question, but as they frequently refer to the theory of influence described in Chapter 4, the model of the theory is presented again here for reference (*Figure 7*). Note that, as in the previous chapter, the names of the key phases of *I*, *They*, and *We* are used in this chapter’s narrative to refer to the phases of the theory and relevant policy actors and practice-drivers within the phases.



Influences on Approaches to Assessment

The first research question was: Through what processes and influences do experienced K-2 teachers report that they have developed their *approaches to assessment*? The distinction of *approach* to assessment is one made relatively recently in the field in order to present teachers’ theoretical leanings, perspectives, and practices towards assessment as a reconceptualization of traditional notions of assessment *literacy* that are based on knowledge and skills (Barnes et al., 2017; Coombs et al., 2018; DeLuca,

LaPointe-McEwan, & Luhanga, 2016; Deneen & Boud, 2014). As opposed to a set of professional competencies, an *approach* to assessment suggests that a teacher's philosophical orientation toward assessment is an important factor in determining how assessment is (or is not) used effectively in support of student learning.

I chose to examine K-2 teachers' approaches towards assessment rather than assessment literacy (AL) based on a presumption that an analysis of philosophical approach would be more appropriate than an analysis of technical knowledge for early elementary teachers. After all, developmentally appropriate practice is a research-based framework, but it is also enacted as a philosophical orientation towards teaching young children; therefore, early childhood teachers function on a philosophical continuum of developmentally appropriate practice (Copple & Bredekamp, 2009). My assumption that teachers would also conceptualize their approach towards assessment as a set of theoretical underpinnings did not hold true for participants in this study. In their self-reporting, the dominant conceptualization for teachers regarding assessment was a description of their *behaviors*, not their beliefs or knowledge. This could be attributed to the overwhelming influence of others (*They*) on teachers' assessment practices, most significantly shaped by the school district.

The School District

Mid-sized district teacher Maya directly placed the school district at the top of the influences on her "approach" to classroom assessment:

The hierarchy is: the county tells you what you have to do, in first grade, because we don't have SOL tests in first grade. The county is going to tell you what assessments you need to do three times a year. Those are the big assessments.

Then from there the principal and the county lets us know what assessment

options are available. We have something called PowerSchool, and we can make tests through that. That is a way to make your end-of-unit tests and things like that. That would be a grade level choice. The grade levels make a lot of choices, and decisions about how to assess each unit that they go through. . . . After that, then you can decide what you want to do, the little stuff.

Maya described a context in which her individual approach does not get considered until after meeting the district's requirements, the principal's suggestions, and the grade level decisions. By the time the individual's decisions are in play, it is only over "the little stuff." As Chapter 4 elucidated, teachers feel *They* determine the core of assessment practices. Although *They* may not directly prohibit or limit teachers' individual approaches from making classroom assessment decisions, the sheer amount and the accountability to complete requirements leave most teachers feeling they have limited classroom time (or professional energy) to supplement with their own approaches to assessment. In this assessment context, the central office personnel who shape assessment policy and decisions appear to be a greater shaper of the student experience than the classroom teacher is.

Separation of assessment and instruction. While the participation of district level administrators in systems of assessment is not inherently a negative matter, the more directed assessment requirements come from sources outside the classroom, the more likely teachers are to see assessment as an outside thing done *to* them and their students (Gareis & Grant, 2015). This raises concerns about the value of assessment in supporting student learning. As Black and Wiliam (2018) stated, "assessment is an essential aspect of effective education because the relationship between the instruction students receive and what they learn as a result of that instruction is complex" (p. 570). If

teachers believe that assessment practices from the district do not connect with their instructional decisions in the classroom, the effectiveness of students' education is undermined.

This issue is well documented in literature. According to J. Chappuis (2014), "when teachers are required to give predetermined common assessments at a predetermined time, accountability for covering material has superseded instructional use" (p. 25). Black and Wiliam (1998) also described a misplaced primacy when "the collection of marks to fill in records is given higher priority than the analysis of pupils' work to discern learning needs" (p. 142). Frequently, assessment is viewed as an outside requirement, not an integrated part of teachers' work in the classroom. As Datnow and Hubbard (2016) indicated, "as long as data use is associated with external accountability, teachers will distrust it and experience it as yet another top-down initiative" (p. 24).

This separation of assessment from outside the classroom and instruction from within results in what Stiggins (2014) called a "two-silo division of responsibility" (p. 69). He described how assessment professionals take charge of testing, stating, "we have those who teach and those who assess, and never the twain shall meet" (Stiggins, 2004, p. 26). While not all teachers in this study would agree that assessment *professionals* were making their tests (that is, they questioned the assessment expertise of assessment-makers), many did describe a situation in which county assessments were not integrated into their instruction. Essentially, these were not *classroom* assessments.

This practice was described more frequently in larger districts where a county-wide curriculum director, subject coordinator, or instructional coaches designed assessments without teacher input. In the training section of the last chapter, Kim, who works in a large school district, was quoted as saying she does not think her

administrators believe teachers need to have the technical knowledge to understand how to develop strong assessment practices. Stiggins (2004) described how some administrations handle this:

Rather than providing teachers with the professional development they need to manage the assessment process effectively, some districts try to circumvent the problem by providing teachers with the tests they need. It's just that those assessments are often developed in the absence of quality control and so can be inaccurate. (p. 26)

According to Stiggins (2014), this separate view of assessment removed from the daily life of the classroom has happened because "those driving federal, state, and local school policy and practice don't understand the differences between sound and unsound testing practices" (p. 67). This mistaken understanding is perhaps most easily seen in an examination of the use of the term "formative" assessment.

"Formative" assessment. As mentioned in Chapter 4, many school districts influence how K-2 teachers approach assessment through the way the terms they use shape shared understanding and practices. While it may seem like a relatively benign error of semantics, misusing the term *formative assessment* in a way that implies it is a smaller, informal, less-comprehensive, lighter-weighted, or lower-stakes assessment than *summative assessment* can inadvertently suggest that it is just a mini-version of the "more important" summative assessment, reducing its power as a potent instructional tool in favor of more consequential means of measuring student learning. Districts may be unintentionally "corrupting the meaning of the term formative assessment, thereby diminishing the potentially positive effect of such assessments on student learning" (Popham, 2006, p. 86).

Black, Harrison, Lee, Marshall, and Wiliam (2004) contended that “there may be a need to review current school policies because such policies can actually constrain the use of formative assessment” (p. 21). Districts that require a certain number of “formative” assessments to be given—with recorded scores—over a set amount of time make it more difficult for teachers to harness the in-the-moment teaching and learning adaptations a well-designed formative assessment process can offer. According to Moss and Brookhart (2009):

Misconceptions are the inevitable result of misunderstanding and often cause teachers to question the formative assessment process. Clearly these misconceptions can dilute the effectiveness of formative assessment and block its consistent use in the classroom. School leaders can take an active role in helping teachers build accurate understandings of what formative assessment is, and, perhaps most important, what it *is not*. (p. 13)

Definitions of formative assessment. The current challenge, it seems, is that school leaders themselves do not have accurate understandings of what formative assessment is and can do. Gareis and Grant (2015) claimed, “a common misconception among many educators and policymakers is that formative assessments are a type of assessment in and of themselves. This is not the case. Rather, what makes an assessment formative is the way it is used” (p. 144). Popham (2008) also recognized this in stating, “one of the most difficult tasks for educators who accept this conception of formative assessment is to grasp the overarching idea that it is a process rather than a test” (p. 8). Chapter 4 introduced McMillan’s (2007) definition of formative assessment: “a set of skills and activities that are undertaken by teachers to provide feedback to students to

enhance their motivation and learning by designing instruction to meet student needs” (p. 1).

Other leaders in assessment work provide similar definitions. Popham’s 2008 book *Transformative Assessment* described formative assessment as “a planned process in which teachers or students use assessment-based evidence to adjust what they’re currently doing” (p. 6) and Moss and Brookhart’s (2009) definition is similar: “an active and intentional learning process that partners the teacher and the students to continuously and systematically gather evidence of learning with the express goal of improving student achievement” (p. 6).

Rather than focus on process, Black and Wiliam (2007) emphasized the importance of purpose:

What distinguishes formative assessment is that the purpose of the assessment is the generation and use of the information to assist students’ learning, rather than simply to record it or to make decisions such as what kinds of education or employment should follow. (p. 4)

Marzano (2006) highlighted the timing of assessments as a way to help distinguish the meanings: “formative assessment was defined as occurring while knowledge is being learned. Summative assessment was defined as occurring at the end of a learning episode” (p. 8). Brookhart (2013) stressed the outcome as evidence of the type of assessment, claiming “formative assessment is about forming learning—that is, it is assessment that gives information that moves students forward. If no further learning occurred, then whatever the intention, an assessment was not formative” (p. 102). None of these definitions stayed with Darlene, who mentioned a “vague memory” of having district-level training on the difference between formative and summative assessment,

though she could not recall the distinction from the one-time training, which serves as an example of the impact from training sessions when districts did not adequately support accurate learning.

Effects of misunderstanding formative assessment. Despite the various nuances of definitions given here, the colloquial use of “formative assessment” as a smaller, intermediary checkpoint along the way to a summative assessment is at odds with any of the features described that make formative assessment a pedagogical tool. Recall that Claire defined formative assessment by the “breadth of what we’re assessing” because that is how her district uses the term. When a school district incorrectly frames formative assessment, it is risking undermining the effectiveness of teachers’ assessment use. In their seminal 1998 work on formative assessment, Black and Wiliam identified that, “teachers clearly face difficult problems in reconciling their formative and summative roles, and confusion in teachers’ minds between these roles can impede the improvement of practice” (p. 148). But, they do not believe it is helpful to have a strict separation of formative and summative purposes, and in a return to the topic 20 years later, Black and Wiliam (2018) acknowledged there could be a “helpful overlap” (p. 554). Being intentional as a teacher about the aim of one’s assessment practices is more important than absolute distinction in definition. They contended:

The formative-summative distinction rests in the kinds of inferences that are drawn on the basis of assessment outcomes, and although an assessment can be designed to serve both purposes, when priority is be given [*sic*] to one particular purpose, this is likely to influence the design of the assessment. (p 562)

Huebner (2009) addressed assessment design at the system level, asserting, “the challenge for schools is designing a balanced assessment system using the strengths of

summative, interim, and formative assessments to address instructional, accountability, and learning needs” (p. 85). Maria, Jill, Annabelle, Laura, and Barbara were among the teachers who described their districts requiring a certain number of formative and summative assessment scores to be reported, but they stated there was no discussion regarding balancing the uses and purposes of these, or other, assessment types. This could be because many schools and districts may lack the AL to design such balanced systems. But in describing the “trickle-down effect,” several teachers also acknowledge that external pressures may be influencing the district’s choice of assessment practices as well. As Stiggins (2014) reported, “a societal blind spot has been created by the common belief that standardized test results are the only truly acceptable evidence of student achievement” (p. 68).

Grading practices. While schools and districts are accountable to society in the form of test results, teachers are accountable to parents and families in the form of report card grades. Gareis and Grant (2015) defined grading as “the translation of student performance on an assessment into a system of relative numbers or symbols to communicate the teacher’s judgment about the nature and degree of student learning” (p. 152). The key purpose of *communicating* a judgement is crucial in any grading and reporting scheme, and thus clarity is necessary. However, Muñoz and Guskey (2015) have pointed out that the idiosyncratic ways in which grades are formulated makes their meaning nearly impossible to interpret. They asserted, “to make grades more meaningful, we need to address both the purpose of grades and the format used to report them” (Muñoz & Guskey, 2015, p. 65).

Teachers in this study described a variety of reporting systems with different grading mechanisms behind the report cards, supporting Muñoz and Guskey’s claim of

idiosyncratic grading methods. Jill described a county-level system in her small district requiring a certain number of “summative” and “formative” grades to be entered each nine weeks. She would assign percentage grades to student products including worksheets and writing samples, then the computer system would calculate these scores into a 4-3-2-1 leveled reporting system. Problems with a 100-point, percentage score to communicate meaningful information are well-documented (Brookhart, 2013; Marzano, 2006, 2010; O’Connor, 2007) but teachers still reported having to provide a numerical score to work for the district’s reporting requirements.

Cyndi works at a different small district than Jill and complained about having to capture learning as percentage scores despite their lack of value in communicating meaning: “What does that mean? Okay, yes, 71. Next nine weeks, you go to 79. Great—how’d they grow? You don’t know that. I told you they just grew eight points. You have no clue how they grew.” As we discussed in the interview, trying to put a six-year old’s reading growth into percentage scores is like grading a toddler’s walking with percentage scores. “How do you judge that?” she rhetorically questioned.

While standards-referenced rubrics tend to be better at capturing the nature and degree of learning, they are not always appropriate for all types of content. Nancy’s ability to capture multiple pieces of evidence from a single activity, such as students creating an American flag, was documented in Chapter 4. But as she pointed out, how can a student score a 4 (*exceeds expectations*) on a skill like recognizing the flag? Attainment of this sort of objective is better tracked with a yes/no checklist, she stated, but checklists are even more problematic than rubrics for translating into percentage grades. In her small district, once her kindergartener’s summative percentage grades were calculated, they would then be assigned a letter grade based on a 10-point grading scale

where each letter had a meaning. The exact letters have been withheld to ensure confidentiality, but an example of this would be 90-100% is E for Excellent, 80-89% is G for Good, 70-79% is S for Satisfactory, and 69% or less is N for Needs Improvement.

Perhaps the most puzzling grading practice was the one Claire described in her large district. She conveyed how she would use a rubric to gather useful information about a second grade student's writing sample, but then had to assign numbers to levels of the rubric to calculate a percentage grade, a practice that leaves valuable information lost in translation (Brookhart, 2013). Most ironically, the computerized grading system takes all of Claire's converted 100-point scores across the nine weeks and applies a (secret-to-her) algorithm to the balance of summative and formative grades to come up with a report card grade—based on the same 4-point scale as her original rubric.

In a review of 100 years of research on grading, Brookhart et al. (2016) reduced the overarching inquiry: “the central question underlying all of these studies is, ‘what do grades mean?’” (p. 804). Even without the back-and-forth coding of information regarding student performance in Claire's example, it was clear that school districts in this study were also grappling with how to answer the question “what do grades mean?” Not only were district-mandated reporting systems struggling to represent learning in ways that can meaningful communicate about student learning, the systems were also influencing how teachers viewed their classroom assessment practices. Karla confessed, “I'm not sure which is the chicken and which is the egg—whether it's the change in the grading structure drove the change in assessments or whether this change in assessments drives that grading structure.” Either way, she acknowledged, it was the school district driving the changes.

Although a thorough analysis of reporting and grading practices was outside the scope of this study, as an influence on teachers' assessment practices, grading schemes and systems that come from the district must be included in this discussion.

The Principal

While the district sets the policies, the principal sets the tone. In the theory of influences on teachers' approaches to assessment, the principal served as a filter that could moderate or amplify pressures coming from the central office level of *They*. A principal's influence, thus, was powerful but indirect. This finding aligns with other research about the impact of a building principal (Leithwood, Harris, & Hopkins, 2008; Louis & Robinson, 2012). Despite being indirect, it was reported by most teachers as a critical element in establishing a positive work climate. Beth stated, "definitely having been in my building through four administrators, I have seen what a difference quality leadership can make. I mean it's huge. It's huge."

According to Leithwood et al. (2008), "School leadership is second only to classroom teaching as an influence on pupil learning" (p. 28). The researchers continued to describe how principals make such an impact: "School leaders improve teaching and learning indirectly and most powerfully through their influence on staff motivation, commitment and working conditions" (p. 32). By establishing a climate of trust, teachers are motivated and empowered to support strong student learning environments (Louis & Wahlstrom, 2011; Tschannen-Moran, 2004).

Sensemaking of policy. Louis and Robinson (2012) described how a principal works as a filter or a lens through which sense is made of policy: "Policies themselves comprise complex ideas that are often ambiguous and open to multiple interpretations" (p. 631). They continued by stating, "unmediated policy interventions will have less

chance of motivating change than those that are mediated in skillful interactions designed to engage the views of implementing agents” (p. 632). Although teachers are not the only implementing agents of district assessment policies, they are greatly influenced by how principals interpret these policies, as it is principals’ “interpretations of policy, rather than an invariant and objective policy idea, that is implemented” (p. 631).

Assessment leadership. Principals are powerful influencers as mediators between the district and teachers (Levin & Datnow, 2012), but they also can have a prominent role in guiding assessment practices, as assessment leadership is an integral facet of instructional leadership (Stiggins & Duke, 2008). In addition to nurturing a climate of trust and collaboration, “Instructional leadership also requires an understanding of the role of sound assessment in efforts to improve teaching and learning. The well-prepared principal is ready to ensure that assessments are of high quality and used effectively” (Stiggins & Duke, 2008, p. 286).

While many teachers in this study described positive views of their current principal, there was less overall confidence in the AL of elementary school principals as a group, a fundamental necessity for strong assessment leadership (Stiggins & Duke, 2008). Darlene spoke well of her principal, saying that as a leader he provides stability, supports a collaborative work environment, and mediates the impact from the mid-sized district’s central office. But she said nothing about his skills or knowledge with assessment. Perspectives similar to this one were very common, suggesting that to teachers in this study, AL was not as important a leadership quality as other traits their principals had and roles they filled.

It is also possible that these teachers were unaware of the ways their principals were acting behind the scenes to shape district decisions and work with different faculty members. According to Stiggins and Duke (2008):

School leaders must ensure that their assessment systems provide a wide variety of decision makers with a variety of different kinds of information in different forms at different times to support or to verify student learning, depending on the context. (p. 286)

This type of support—ensuring a variety of stakeholders’ assessment needs are met in a variety of ways at a variety of times—was reportedly inconsistent across teachers’ careers.

Leadership of data teams. Strong instructional leaders play important roles in influencing how teachers collaborate to align curriculum, instruction, and assessment as well as how they use assessment data to inform decisions (Lunenburg, 2013). Research supports the importance of principals being directly involved in accountability measures, such as data teams (Crum, Sherman, & Myran, 2009). While technical knowledge about assessment data is needed, it may be more critical that “principals help to set the tone for data use among teachers” (Datnow & Hubbard, 2016, p. 17). Principal participation in data teams or PLCs was widely reported and primarily considered a positive influence, although some teachers’ past experiences supported Datnow and Hubbard’s (2016) assertion that data leadership could influence teams in different ways: “Some leaders may promote an accountability-focused culture in which data is used in a short time frame to identify problems and monitor compliance, whereas others may promote the use of data for continuous improvement” (p. 17).

As assessment can serve multiple purposes, including compliance and improvement on a larger scale and daily instruction on a smaller, though consequential, scale of student learning, the influential role of principal as an assessment leader is crucial. According to Louis and Robinson (2012):

Given this layered accountability context, school leaders' responses to federal accountability mandates are likely to reflect a complex interaction between their perception of state policies and support, the specific district contexts in which those policies are situated, including ongoing district reform initiatives, and their own leadership beliefs and practices. (p. 630)

Having a building principal who can navigate multiple demands using strong AL and empowering leadership skills can shape a teacher's approach towards assessment.

Ironically, the only teacher to specifically praise her principal's expertise in working with assessments works in one of the districts that utilizes the previously mentioned inaccurate uses of "formative assessment" to refer to activities covering a small range of content from which a mandatory grade is derived and reported, demonstrating the "complex interaction" Louis and Robinson (2012) noted exists between principal and district.

Formal Learning

While all teachers work in a nested context where school district and building administrators influence their assessment practices, not all teachers experienced formal learning opportunities to develop their approach to assessment. In fact, this is one reason why principals need AL. According to Stiggins and Duke (2008), "the stronger the assessment literacy background for new and practicing school leaders, the more able they will be to develop or arrange for the professional development their colleagues need to find remedies to their problems" (p. 290). The problems teachers have stem from a lack

of assessment knowledge and skills, and the three primary sources for developing these are in the teacher preparation program, through structured in-service professional development, and in collaborative Professional Learning Communities (PLCs).

Teacher preparation. Very few teachers reported having any background training in assessment during the pre-service period or through a course on assessment foundations during her preparation program. Recall from Chapter 4 that Jill responded with laughter at the preposterous notion of having an assessment class in her licensure program. The literature widely supports the idea that classroom assessment is under-emphasized at best in teacher preparation programs (DeLuca et al., 2013; Fives et al., 2016; Popham, 2011; Stiggins & Duke, 2008).

DeLuca et al. (2013) found that explicit opportunities to learn about assessment improve both competency and confidence. Presuming that a student teaching internship will successfully transmit such learning is unwise. According to Stiggins and Duke (2008):

Expecting students to learn about assessment practice during internships can be risky, since it cannot be assumed that field-based supervisors are prepared to provide instruction in this area. Until more school leaders develop appropriate levels of assessment literacy, internships will not provide a solution. (p. 290)

Popham (2009) also noted a need for improved AL across the field, stating, “until preservice teacher education programs begin producing assessment literate teachers, professional developers must continue to rectify this omission in educators’ professional capabilities” (p. 4).

Professional development. The role of professional development (PD) to improve AL was not an overwhelming influence on teachers in this study. Some reported

that one-off PD sessions were not helpful, but the minority of teachers who experienced ongoing, job-embedded professional development described it as a positive influence in building their knowledge and skills for using classroom assessment. Guskey (2014) acknowledged that there has been weak evidence of the effectiveness of PD in part because of the lack of well-planned development that stated a clear purpose and cohesive direction in its goals. Stiggins and Duke (2008) contended that it is a key function of the principal to be able to select and secure valuable PD, the kind that is well-planned and focused not just on training for a specific assessment, but on building teachers' AL, like the kind Laura described in her interview.

Gareis and Grant (2015) posited that administrators are not the only ones who can support meaningful PD, claiming, "teacher leaders can constructively influence the professional practice of other teachers by advocating for and, indeed, by providing professional development opportunities in assessment" (p. 174). This kind of shared influence and learning is possible through formal PD, but it was more widely reported through collaborative learning communities. This finding is supported in the specific context within AL of data literacy by Datnow and Hubbard (2016), who stated, "teachers' capacity to use data and their beliefs about data use are shaped within their professional communities, in training sessions, and in their interactions with coaches, consultants, and principals" (p. 7).

Professional learning communities. Interactions with coaches, consultants, and principals were widely reported through structured team meetings that were planned and held regularly, but somewhat less formal than professional development programs. Whereas well-planned PD may have as its intended outcomes the growth of teachers' AL through an applied practice in the school context, PLCs were more likely to have as their

outcome improved student learning through the analysis of assessment data and actions taken from such analyses. Growth of teachers' AL tends to be a secondary intended outcome, or even by-product, of PLCs. According to Popham (2008), both PD and PLCs can be useful in developing teachers' skills. There are subtle distinctions between the two, however.

A PLC emphasizes long-term growth and collective focus of school priorities. Louis and Wahlstrom (2011) described the type of PLC that leads to positive outcomes:

Teachers need to be part of a professional community that breaks down the isolation and short-term focus of the traditional school culture. Professional communities motivate teachers to take collective responsibility for ensuring that all students learn, and they strengthen teachers' resolve to break out of old practices. (p. 54)

Data teams. Like Louis and Wahlstrom, Boudett and Moody (2005) emphasized the role of a professional community in shaping school culture, but they focused on "data culture" as being key. Boudett and Moody (2005) found that the three activities that support a data culture are "creating and guiding a data team; enabling collaborative work among faculty; and planning productive meetings" (p. 12). Datnow and Hubbard (2016) also address data teams as a specific type of PLC:

In some cases, teacher collaboration for data use also involves principals, instructional coaches, university researchers, or consultants who serve as facilitators. When these individuals frame data use as part of a cycle of inquiry and careful reflection (rather than simply being about accountability) and are able to contribute deep knowledge of instruction and data use themselves, it is much

more likely that their involvement has a positive impact on teacher capacity building. (p. 23)

Almost all K-2 teachers in this study described the influence of a data team over their assessment practice, though the degree of influence as well as its nature varied. Where Gretchen described the strong collaborative nature of her team's focus on analyzing student data for improving instruction, Jan portrayed her team's emphasis to be more of a numbers game. The attention placed on purpose was one way these teams varied; another variant was the purported level of data literacy among team members.

Data literacy is not the same as, but shares commonalities with, AL, which, as previously discussed, is relatively weak among educators. Datnow and Hubbard (2016) described data literacy in the following manner:

Data literacy for teaching is the ability to transform information into actionable instructional knowledge and practices by collecting, analyzing, and interpreting all types of data (assessment, school climate, behavioral, snapshot, longitudinal, moment-to-moment, and so on) to help determine instructional steps. It combines an understanding of data with standards, disciplinary knowledge and practices, curricular knowledge, pedagogical content knowledge, and an understanding of how children learn. (p. 2)

According to participants, this type of literacy is not taught in preparation programs, and teachers in the study described how they learned to analyze data by doing it on the job. This result is mirrored in Datnow and Hubbard's (2016) finding that "almost all efforts to build teachers' capacity to use data rely on the teacher collaboration model. Structured training in how to use data is less common" (p. 14). By analyzing data, such as

PALS results or district math benchmark tests, teachers reported learning how to analyze data in these PLCs, which in turn influenced how they approached assessment.

Popham (2009) reserved the name PLC for a group when an administrator is involved and used the term *teacher-learning community* to describe a group composed of just teachers. In Chapter 4, the use of “data team” meets the definition of a PLC, whereas the category of “grade level team” meets the conceptual notion of a teacher-learning community, frequently referred to as an informal network or team that functions as a *community of practice*.

Community of Practice

Although there was one teacher in the study who was the sole teacher on her grade level, all other teachers described some sort of grade level team, whether informal or formal, that had the potential to influence their assessment practices. In Sangita’s case, she has a strong team of teachers at her grade level that meet regularly as well as connect through casual hallway conversations about their teaching. Laura described a team of mixed input where her teaching partner likes to share and work together, but others are less interested in anything perceived as additional work. Gretchen has experienced both the “closed door” type of community as well as a supportive network of professionals.

Whereas the lack of a strong community of practice did not appear to directly impact teachers negatively (other than the loss of a potential learning opportunity), teachers who did work with collaborative teammates described learning from each other. Black et al. (2004) wrote that this type of community is crucial when making changes in assessment practices and approaches:

Collaboration with a group trying out similar innovations is almost essential.

Mutual observation and the sharing of ideas and experiences about the progress of

action plans can provide the necessary support both with the tactics and at a strategic level. Support for colleagues is particularly important in overcoming those initial uncertainties when engaging in the risky business of changing the culture and expectations in the classroom. (p. 20)

This type of collegial learning may be “essential,” but it brings with it risks, such as untrained teachers developing inaccurate assessment practices and perpetuating them throughout a team (Stiggins, 2004). Other threats to a community of practice include the climate in which they function. According to Deneen and Brown (2016), “assessment literacy may be at risk if teachers fear assessment and testing, have negative perceptions of assessment, lack adequate training, or face strong pressure to place accountability purposes over improvement purposes” (p. 12). Due to its unstructured nature without a clear leader and a composition of members with diverse skills and attitudes, a community of practice may be one of the most organic—and least controllable—influences over a teacher’s approach towards assessment.

Summary of Influences on Approach to Assessment

K-2 teachers in this study described a variety on influences on their assessment practices that mirror those described by Black and Wiliam (2007): “Teachers learn from one another, in informal groups within schools, within the formal structures of their school, in in-service professional development groups, and in professional associations” (p. 45). While this kaleidoscope of inputs results in a wide range of outcomes regarding how teachers approach assessment in their classroom, the heavy sway of *They* is always present in teachers’ individual practices. As McMillan (2003) described, assessment decision making is “a process in which teachers balance the demands of external factors

and constraints with their own beliefs and values to determine classroom assessment practices” (p. 42).

Influences on Developmentally Appropriate Practice

The second research question that shaped this study was: Through what processes and influences do experienced K-2 teachers report that they have developed their knowledge and implementation of *developmentally appropriate practice*? Unlike teachers’ descriptions of assessment practices, which were driving much of their work in schools, teachers’ descriptions of ways in which developmentally appropriate practice (DAP) was emphasized in their work were scant. Many teachers spoke about some of the underlying concepts about child development as a backdrop to their personal approach towards teaching young children, but DAP was an unfamiliar framework and in the context of their schools, it was not an explicit guiding factor in any way. Recalling the literature from Chapter 2, it is not a surprise that the accountability era has steered early childhood education away from a foundation of developmentally appropriate practice (Bassok et al., 2016; Hatch, 2002; Miller & Almon, 2009; Minicozzi, 2016).

In absence of a dominant force, like the policy mandates of *They* overseeing assessment practices, a range of teacher understandings and philosophies regarding teaching in the early elementary years can propagate. Vartuli (2005) stated:

Diversity of teacher beliefs about how young children learn can be attributed to personal experiences, early schooling and instruction, teacher preparation (certification or degree), variations in program/school sponsorship, the personal nature of teaching, student teaching experience, or isolation of teachers from colleagues. (p. 80)

In this section, I discuss the major influences represented in the theory: individual factors, time, colleagues, administration, and curriculum.

In discussion about the findings of this particular study, it is important to remember both a limitation and a delimitation: it is likely that the teachers who volunteered are ones with a philosophical inclination towards DAP due to its mention in the recruiting email, and the study did not attempt to evaluate whether reported practices or attitudes were truly reflective of NAEYC's (2009) framework of DAP. As developmentally appropriate practices exist on a continuum (Buchanan et al., 1998), the reported influences described in this section could impact practices across the spectrum.

Individual Factors

Because there is currently a lack of collective emphasis on developmental considerations in elementary schools, individual factors can have a substantial influence over teachers' attitudes, knowledge, and behaviors. "Teachers learn to rely on themselves as a potent source of information about the needs, interests, uniqueness, and diversity of the children they live with in the classroom," claimed Cohen, Stern, Balaban, and Gropper (2016, p. 1). Personality characteristics can influence the use of DAP (Parker & Neuharth-Pritchett, 2006), and several teachers described how they have been oriented towards work with children for most of their lives. Many of these teachers chose to study child psychology or elementary education, and teachers who earned preK-3 teaching endorsements typically reported that their preparation program had a substantive influence over their philosophical orientation.

Although this study did not endeavor to directly compare preK-3 and preK-6 endorsement holders, Vartuli's (2005) study determined that early childhood majors have higher DAP scores than those with general elementary education backgrounds. Vartuli

(2005) also found that, “as grade level increases, teachers’ scores on surveys of developmentally appropriate practice beliefs decrease” (p. 81). Teachers in this study who taught in kindergarten described more of an orientation towards developmental practices than those in second grade, but it is unclear how much of this can be attributed to the contextual factor of looming third grade SOL tests shaping second grade teachers’ practices versus personal characteristics that make some teachers more inclined to seek positions in kindergarten.

A common refrain heard in interviews was an expression of desire not to teach in higher grades. Kindergarten teacher Janice directly confessed, “First scares me. It’s too much. I mean if there’s this much pressure in kindergarten, I don’t want to know what’s happening in these first grade rooms right now.” Parker and Neuharth-Pritchett (2006) stated that “teachers who endorsed a more child-centered perspective noted that they felt more pressure” (p. 65). Jill, Annabelle, Karla, Michiko, and Anne all admitted to trying to avoid being moved to an upper elementary classroom due to the testing pressure and the climate in third through fifth grades. Implicit in Nancy’s explanation for this is an acknowledgement of children’s individual rates of growth: “I don’t want to be responsible if a student doesn’t know something because they may not be ready to know it.”

Self-confidence. Statements of fear or aversion towards the upper grades should not be confused with a lack of professional confidence. In fact, many teachers expressed a strong belief in their ability to understand children’s developmental and learning needs. Anne downplayed the challenge of grasping child development, suggesting a skilled teacher can “just look at the child! They’re only five.” Maggie described the flexibility her principal gives her teaching team because, “we know our kids and we know our

students and we know what they are capable of.” Jan had a similar sentiment about her ability to understand her students’ development:

We get pretty good at being able to pick stuff out . . . and know that it’s developmental or there’s something else not right. . . . You can pick it out. You know, and you can’t say, like I’m not qualified to say what it is, but I know it’s something. It’s not just normal immaturity or, you know, even just not ready to be here. It’s something else.

In her research, Vartuli (2005) found that high teacher self-efficacy is associated with high developmentally appropriate belief scores. If a teacher believes she has the ability to shape and inform instructional decisions, she chooses strategies to increase learning whereas teachers with low self-efficacy just cover curriculum (Vartuli, 2005). This personal self-confidence was expressed almost universally by veteran teachers and most mid-career teachers, such as Michelle, who talked about her laid-back approach to assessing young children:

I don’t lose sleep about it. It really goes back to the, back in the day when you would just say, “I know my kids.” You know, I sit with them every single day in different groups all day long, and I know I could tell you that this little boy still doesn’t know his sounds.

“Old school.” The concept of “back in the day” was frequently employed to conjure up a past era of teaching that was less technical and more intuitive. It was also used almost as an excuse for being more child-centered and less curriculum-driven. Veteran teacher Bev hedged her confident description of her approach to teaching with “I might be old-fashioned. I don’t know.” Veteran teacher Margaret explained, “I’m weird. Old fashioned.”

While recognizing that being “old fashioned” may not be lauded in the current education environment, the idea of being “old school” bolstered some teachers’ self-confidence in their practices. Having seen the pendulum of education swing back and forth, many experienced teachers felt comfortable relying on what their time in the profession has shown them works with kids. As veteran teacher Janice rationalized:

Just because the county says it doesn’t mean necessarily that that’s the way it’s gonna roll in this room. . . . So, there’s some of that old school that’s still in me that just closes the door and does it the way I feel is best for my kids.

Janice specifically situated her “old school” perspective as being more developmental and child-centered whereas newer teachers are more curriculum- and assessment-driven, she claimed. Veteran teacher Sangita also used the same term to compare formalized assessment tasks with child-focused observation procedures, saying, “I’m not a big one for assessing. We assess one-on-one, old school, watch what they’re doing. . . . I’m more of a one-on-one, good old-fashioned checklist assessment kind of girl.”

The idea of “old school” was seen in two factors: the era during which the teacher did her training and the amount of time the teacher had been working in a classroom with children. The latter was reported as a particularly substantial influence over K-2 teachers’ understanding of developmentally appropriate practices. This issue has not been well-addressed in the current research base, which is why it is mentioned again in Chapter 6.

Time

The importance of both experiential time and elapsed time in shaping a teacher’s understanding of DAP was described in Chapter 4. As veteran teacher Melanie asserted:

You really don't understand teaching until you're into it. You really don't see and understand that developmental practice until you see a kid who truly just can't read yet. And it's not because they don't know how, they're just not ready.

“Old school” teachers have had more time to observe children and learn from actual experience, not textbook knowledge, what children are like. The importance of learning by observing is supported by Cohen et al. (2016) who confirmed, “the technique of studying one child in detail leads to deeper understanding of the one child and broader knowledge of all children” (p. 7). During the early years of one's career, before having an accumulation of observations and experiences, new teachers rely on older teachers' accumulated experiences and lean on their colleagues to better understand children's developmental needs and practices that will best support them.

This can lead to practices across the spectrum of developmental appropriateness. As is found in Chapter 4, Nikki initially followed her team's lead without question, though it was dominated by teachers who were content with relying on didactic teaching methods such as worksheets and direct instruction. In reflecting on how student teachers she has had in her room behaved, veteran teacher Maya speculated, “they were just so scared they wanted to do what was right, they didn't think about developmentally appropriate practices.” If newer teachers believe that what those around them are doing is “right,” they are likely to be heavily influenced by those practices, regardless of where they fall on the continuum of developmental appropriateness.

Administration

While teachers respected and followed their colleagues' experiences working in early childhood, very few teachers worked with a principal who had any early elementary teaching experience, and many did not even have experience with elementary aged

children prior to becoming the building principal. While some principals without elementary teaching experience were still described as strong instructional leaders, neither principals nor district administrators in this study were described as positive influences on teachers' knowledge of developmentally *appropriate* practices, though some did influence (or even require) the use of developmentally *inappropriate* practices.

This finding aligns with a study of kindergarten teachers' work in navigating the conflicting paradigms of developmentally appropriate teaching and curriculum-centered academic teaching completed by Minicozzi (2016): "It is fair to say that all of the teachers felt a real lack of support from the central office or reported passive support from their district superintendent at best" (p. 307). Passive support *at best* meant that administrators deferred to faculty experienced in early childhood regarding developmental questions, and they did not try to mandate inappropriate practices. This was seen in the kind of advice Lyn reported receiving from a previous principal: "Just shut the door. Make sure your SOL matches the pacing guide and go about your life."

Second grade teacher Melanie described less of a neutral force in administrators and more of a barrier in implementing developmentally appropriate practices, as she related:

So often we're told, "oh yeah, we know, but we have to get them to this point."

And that's where I think teachers feel that frustration is, we know. . . sometimes we are pushing these kids so far beyond what their brains can accept and then we get frustrated because they can't do it.

The results of teachers feeling this kind of frustration regarding their lack of autonomy to teach in appropriate ways include use of developmentally inappropriate teaching methods, according to a study by Parker and Neuharth-Pritchett (2006):

The majority of teachers using didactic practices perceived themselves as having little control over their instructional decisions, whereas nearly all of the child-centered teachers noted that the decision to use instructional practices was entirely theirs to make. One can conclude that when teachers perceive they have the professional freedom to make instructional decisions, they will use child-centered, developmentally appropriate strategies more frequently. (p. 75)

For teachers in this study who felt they were granted professional freedom and respect, many credited their principal for nurturing an empowering climate of trust. In a review of literature on school leadership, Leithwood et al. (2008) found that, “school leaders improve teaching and learning indirectly and most powerfully through their influence on staff motivation, commitment and working conditions” (p. 32). Though none of these teachers described a principal who led from a position of prioritizing child development and developmentally appropriate practice, many spoke of how their principal motivated staff and fostered a positive working environment, like the findings of Leithwood et al.

Other leadership studies have found important behaviors of school leaders for influencing student learning, but a leader’s knowledge of or commitment to developmentally appropriate practice has not been adequately studied. In Marzano’s (2005) prominent meta-analysis of leadership behaviors, 21 responsibilities were shown to impact student learning, but none of them addressed knowledge of child or human development, much less developmentally appropriate practice. It is unclear if this was because this factor did not have statistical significance or has simply not been studied closely. Either way, the influence of school leaders on developmentally appropriate practice appears to be a neutral one at best.

Curriculum

Although the curriculum is largely shaped by district administrators, it stands as its own influence as it was so consistently described by teachers as expecting too much, too soon from young children. This finding is widely supported by research literature that describes how the accountability movement's focus on standards has pushed down through all the elementary years to kindergarten (Bassok et al., 2016; Gallant, 2009; Graue, 2009; Gullo & Hughes, 2011a; Hatch, 2002; Minicozzi, 2016; Nicolopoulou, 2010; Parker & Neuharth-Pritchett, 2006; Popham, 2011).

In their widely-cited report *Is kindergarten the new first grade?* Bassok et al. (2016) described how developmentally appropriate practices “centered on play, exploration, and social interactions have been replaced with highly-prescriptive curricula, test preparation, and an explicit focus on academic skill-building” (p. 1). This trickle-down effect of academic standards into kindergarten (Minicozzi, 2016) is even more concerning when considering that nearly 20 years before the report, Shepard (1994) stated that the trickle-down of academic curriculum starting in the 1980s had already resulted in first grade demands being placed on kindergarteners. Bassok et al.'s report reflects another generation, or second wave, of academic “shovedown” (Hatch, 2002, p. 457).

According to Minicozzi (2016), “the accountability pressures have influenced the focus of curriculum in kindergarten to such an extent that it remains difficult to distinguish between curriculum and methodology in classrooms for young children and those of later elementary grades” (p. 304). This sentiment was frequently reported by teachers in this study, such as first grade teacher Pauline who lamented, “I think we push so much for the standards that we may not consider whether this is developmentally

appropriate or not. . . . we forget their age. . . so little.” Kindergarten teacher Michelle was more direct about what happens when children enter kindergarten: “child development goes to the side.” Teachers described pacing guides that did not allow time to respond to students’ needs or abilities and benchmark expectations from the district that grew beyond what students could reasonably achieve. One district’s kindergarten sight word list inexplicably doubled from the previous year, and some teachers reported being evaluated based on the growth their students made in reading this unfeasibly-long list of sight words. With such high stakes, it is no surprise that the curricular demands shape the degree to which teachers feel they can use developmentally appropriate practices.

While much of the research around the increasing academic nature of early elementary has been focused on analyzing practices in kindergarten, it is reasonable to assume that the trickle-down has affected first and second grades on its way down to kindergarten. According to Nicolopoulou (2010), “this emphasis on more didactic, academic, and content-based approaches to [early childhood] education comes at the expense of more child-centered, play-oriented, and constructivist approaches, which are dismissed as obsolete or simply crowded out” (p. 1). Maria, Annabelle, Janice, Pauline, Beth, Sangita, and Jan were among the teachers who lamented the loss of play and authentic, hands-on learning and social opportunities for their students.

Despite the increased rigor in the standards, Bassok et al. (2016) claimed this does not have to come at the expense of child-centered teaching. They wrote, “Teaching academic content need not be at odds with ‘play’ and other types of pedagogical approaches considered developmentally appropriate in early childhood” (p. 22). Nonetheless, teachers in this study widely reported that the curriculum limited, or at least

heavily influenced, their use of developmentally appropriate practice. If there is a way highly-academic standards can be taught using developmentally appropriate pedagogy, this ideal has not yet fully materialized for K-2 teachers around the state.

Summary of Influences on Developmentally Appropriate Practice

According to NAEYC (2009), decisions regarding the education of young children are made by a variety of stakeholders. Because of this, “classroom environments are directly affected by outside forces, such as accountability, high-stakes testing, parental involvement, parental pressure about children’s success, financial resources, and administrative policies and procedures” (Parker & Neuharth-Pritchett, 2006, p. 69). This section described the major forces that K-2 teachers in this study recounted as being influential over their knowledge and implementation of developmentally appropriate practices in their classrooms.

While the literature base regarding DAP is substantive in the preschool years, there is a dearth of research addressing DAP in first and second grades. The research on DAP in kindergarten has become increasingly focused on concerns about kindergarten becoming “too much, too soon, too fast” (Gallant, 2009, p. 201) over the last 20 years as the accountability movement has taken hold. Though Bassok et al. (2016) posited that the goals of the accountability movement do not need to be at odds with a child-centered pedagogy, McMillan (2003) acknowledged the tension created when dominant assessment practices from the high-stakes testing movement do not easily integrate with constructivist theories of learning. He stated, “assessment decision making is characterized by tension between the internal beliefs and values of the teachers and external influences that are imposed on them” (p. 35). The next section discusses how these two competing interests—a value system based on developmentally appropriate

practice and the externally-required practices of assessment procedures—interact with each other.

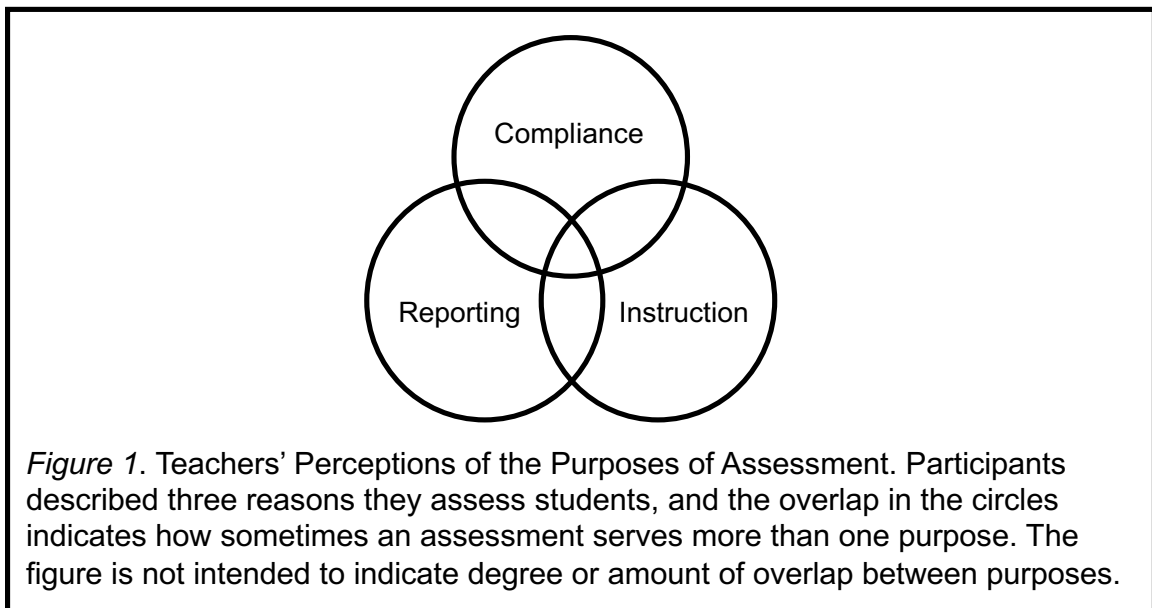
Interaction of Approaches to Assessment and DAP

The third research question this study aimed to illuminate was: How do experienced K-2 teachers' approaches towards classroom assessment and philosophical orientations towards DAP interact with each other? There is research evidence that teachers can navigate these seemingly conflicting paradigms (Pyle & DeLuca, 2013), and some teachers in the study described ways in which they approached their profession with a “both/and” mindset supported by intentional teaching decisions (Copples & Bredekamp, 2009). Yet, other teachers struggled with integrating developmentally appropriate practices and more technical aspects of required assessment practices. This section discusses how teachers conceptualized assessment, how they viewed the purposes of assessment in their early elementary contexts, and the impact of assessment practices on their pedagogy.

Conceptualizations of Assessment

NAEYC authors McAfee, Leong, and Bodrova (2004) defined assessment as “the process of gathering information about children from several forms of evidence, then organizing and interpreting that information” (p. 3). The dominant conceptualization expressed by teachers in this study, however, was not one that defined assessment as a process, but rather one that described it as a collection of pre-determined tasks or activities that were used to provide a score, grade, or level for audiences or purposes outside of the classroom. The gathering and interpreting of information about children for classroom instructional purposes certainly happened, but it was not lumped in with these other forms of assessment and labeled “assessment.”

Essentially, teachers described formal assessment procedures that originated from outside their individual control and had an audience beyond the teacher and child as “assessment,” and other informal practices within the classroom that lead to instructional decisions as “the little stuff,” according to first grade teacher Maya. It would not be reasonable to expect K-2 teachers to hold a single common working definition of the term “assessment,” given that it is a complex construct that varies depending on its purpose, scope, use, audience, time, and form. As scholars have developed different conceptual frameworks for understanding assessment, it is only normal that teachers will also have developed different understandings based on the combination of contextual factors in their professional experiences. Next, I present a few conceptualizations of assessment as provided in literature and describe how they compare to conceptualizations held by K-2 teachers in this study. For reference, I repeat here *Figure 1*, which shows how the study’s teachers perceived the purposes of assessment.



Assessment as, for, and of learning. As previously described in the section on formative assessment, some educators prefer to categorize assessment based on how it is

used: assessment *of* learning tends to be summative assessment to determine what learning has been completed, assessment *for* learning tends to be formative assessment used to adapt instructional plans in the short term, and assessment *as* learning tends to be dynamic assessment involving student self-reflection and modifications to one's own learning approach.

Teachers in this study who described a strong developmental approach to teaching heavily favored practices that would fall into the assessment *for* learning category, describing ways they are constantly responding to students' needs in the moment. Furthermore, these teachers were not inclined to use traditional assessment of learning forms like tests at the end of units, because their instructional approach—when not directed by the district—was unlikely to be segmented into discreet content-based learning units. These teachers were also unlikely to value a single point-in-time assessment as being representative of students' true capabilities.

Only one of the teachers provided a clear description of assessment *as* learning practices. This could be because the fluid nature of instruction and informal assessment allowed for teachers to seamlessly use self-reflection on performance, but these informal practices were mentally considered teaching tools rather than more formal assessment practices. It could also be because teachers did not believe their students were capable of participating in deep meta-cognitive reflective practices, as first grade teacher Karla commented: "That's abstract reasoning. They don't have that skill yet!" She provided the example of the responses she got when asking her students *how* they found an answer. "They don't know. And that's a legitimate answer. They don't!"

Fundamental roles. While the resulting three categories are similar to those just described, Gareis and Grant (2015) used the *timing* of when assessments occur relative to when instruction occurs to present another way of distinguishing the roles of assessment:

1. Pre-assessment is the assessment of student learning prior to teaching.
2. Formative assessment is the assessment of student learning integrated into the act of teaching.
3. Summative assessment is the assessment of student learning at the end of some period of instruction. (p. 5)

It is important to note that though the timing of when these tests are given is used to describe them, the purposes are also distinguished from each other. Preassessment can serve as both a way to plan instruction and measure growth and, for some students, as a means to reflect upon and spark learning. Gretchen was the only teacher to describe the use of a preassessment as an important learning tool in her classroom, as she would have her first graders reflect back on their learning in order to help develop a growth mindset.

Developmentally appropriate practices lend themselves well to assessment practices integrated into the act of teaching, as one of NAEYC's (2003) indicators of effectiveness for assessment is, "Assessment evidence is gathered from realistic settings and situations that reflect children's actual performance" (p. 3). Assessment that occurs "at the end of some period of instruction" may be less frequently used by teachers with a strong developmental philosophy, as children's development occurs along continuums of relatively predictable sequences, but at varying rates (Copple & Bredekamp, 2009). Growth does not occur in neat, isolated chunks of learning, and it certainly does not occur in neat chunks for all children at the same pace (NAEYC, 2009). Some teachers in this study espoused a developmentally appropriate position like this one about children's

growth, but because of the structures of school and ways in which the SOLs and their curriculum are organized, they follow through with practices that do group common learning targets for all students during the same period of time, and thus, they are able—or required—to utilize common summative assessment practices at the end of learning units, even if their philosophical preferences may be at odds with these requirements.

Continuum of conceptions. In a study of K-12 teachers, Barnes et al., (2017) conducted an exploratory factor analysis that led to a continuum of conceptions from extreme pedagogical on one side to an extreme accountability conception on the other. On the extreme pedagogical side, the purpose of assessment is to advance learning, whereas on the extreme accountability side, the purpose is for teacher and/or school accountability. Between these extremes falls a purpose of student accountability, and outside of the continuum rests a notion of assessment as irrelevant.

None of the teachers in this study described assessment holistically as being irrelevant, though specific assessment tasks were sometimes deemed as useless, such as when Betty described her school’s decision to discard the VKRP after piloting the test because it was “just an extra” and did not serve any purpose. The K-2 teachers’ personal inclinations were to be more philosophically oriented towards pedagogical uses of assessment, but they described practices required from the dominant accountability paradigm. As such, most teachers lived in tension of the two ends of the continuum, rather than primarily resting in a particular location along the spectrum.

The three purposes of assessment as described by Barnes et al. (2017) have similarities to the three purposes perceived by teachers in this study. Although not identical in nuance, “assessment to advance learning” is similar to the instructional purpose, “assessment for student accountability” is similar to the reporting purpose, and

“assessment for teacher/school accountability” most closely correlates with teachers’ description of assessment for compliance.

Four frames. In a study with K-12 teachers in New York City, Deneen and Brown (2016) used a framework that aggregates assessment into four major purposes: “informing educational improvement, evaluating students, evaluating schools and teachers, and irrelevance” (p. 4). In their conceptualization:

Improvement focuses on assessment informing teachers and students regarding what students need to learn next and how this may guide learning. Evaluation has to do with appraising student performance against standards, assigning scores/grades and awarding qualifications. School evaluation has to do with using assessment results to determine the performance of teachers and schools.

Irrelevance has to do with either rejecting assessment as having meaningful connection to learning or believing it to be bad for students. (p. 4)

Black and Wiliam’s (2007) frames omitted “irrelevance,” but otherwise described three very similar purposes: “the support of learning, with certification, i.e. with reporting the achievements of individuals, and with satisfying demands for public accountability” (p. 4). Like Deneen and Brown, Black and Wiliam (2007) do not relate these purposes to each other using a continuum, but essentially, both of these pairs of researchers’ categories readily map onto those of Barnes et al. (2017) with only slight variations that are primarily semantic rather than functional.

Early childhood perspective. While the above conceptual frameworks categorizing assessment are intended to be applicable to K-12 settings, there are unique considerations in assessing children in early childhood, thus suggesting a framework designed specifically for young children may be more informative to this study. McAfee

et al. (2004) provided the following reasons for assessing young children: “to monitor children’s development and learning, to guide our planning and decision making, to identify children who might benefit from special services, to report to and communicate with others” (p. 6-8).

The first purpose adds a nuance not seen in the other conceptualizations: monitoring children’s *development*. Implied in this is that it is important to attend to students’ development in domains outside the dominant academic learning paradigm represented in previous frameworks. This would include development in the social, emotional, and physical domains. A few of the teachers in this study expressed a value of acknowledging children’s growth in non-cognitive domains, but this was typically done in the context of lamenting a loss of focus on areas of growth beyond the academic-cognitive domain. For instance, Claire noted how the kindergarten curriculum pace no longer allows time for fine motor work with sand or clay, and she can see the difference in her second graders’ difficulties with the physical aspects of writing. Schools are not formally monitoring students’ development outside of the cognitive domain, and so only a few teachers in this sample reported keeping informal observational records for these purposes. Costanza and Falkenstein (2018) reported that this imbalance has led to a system that does not adequately educate the whole child, a finding portrayed by teachers in this study.

The second reason for assessing young children reported by McAfee et al. (2004), to guide planning and decision making, addresses the assessment for *instruction* purpose described by K-2 teachers, though it also includes the idea of mid- to long-term program planning. It was very rare for a teacher in this study to mention anything about using assessment for program evaluation, such as making broader decisions about curriculum

spacing or materials. What is described as assessment for *compliance* by teachers could, from the perspective of central office administration, be used to inform these broader planning purposes. However, teachers may feel these decisions are outside their realm of influence, and therefore, irrelevant in their perceptions about the purposes of assessment.

The third reason for assessing young children is one not directly addressed in any of the previous frameworks: to identify children who might benefit from special services. Although this purpose could be implied in other conceptualizations, it is curious that only the early childhood frame explicitly states this reason, and at the same time, early childhood teachers interviewed did not identify this purpose for assessment. A few teachers described taking assessment results to a child study meeting, but this was not viewed as a reason for giving assessment and rather treated as an additional use of assessment results to support or prove the teacher's own judgment or concerns about a child's development (essentially, an assessment for *reporting* purpose). Screening assessments, which are prevalent in early childhood, were rarely discussed, perhaps for the same reason program planning purposes were not addressed: these decisions tend to be made, and were often carried out, by other professionals and are not primarily the realm of the classroom teacher.

The fourth reason for assessing children aligns with teachers' conceptualization of the purpose of *assessment* for reporting.

The Approaches to Classroom Assessment Inventory (ACAI). The final conceptualization presented here is one that shows more complexity in teachers' approaches to assessment. Building off of research that indicates a teacher's approach to assessment can be as important as the teacher's AL, researchers developed an instrument to measure this construct, the ACAI (DeLuca, LaPointe-McEwan, & Luhanga, 2016;

DeLuca, Valiquette et al., 2016). At its core, it presents three main approaches to assessment:

Approach A focuses on summative approaches to assessment that emphasize reliable assessment design and the standardized administration of assessments of learning across groups. Approach B focuses on formative approaches to assessment that emphasize validity through emphasis on use and scoring of assessments of learning and equitable treatment of students. Approach C focuses on individualized approaches to assessment that emphasize both reliability and validity concerns through communication and a differentiated approach to assessments through assessments as learning. (DeLuca, LaPointe-McEwan, & Luhanga, 2016, p. 252)

The instrument codes teachers' responses to scenarios in respect to four dimensions of assessment: assessment purposes, assessment processes, assessment fairness, and measurement theory. As a result, a teacher's profile can be rather varied across the three approaches and four dimensions. This multi-faceted view of assessment may be a more realistic look at how teachers actually approach assessment when compared with simpler frameworks. Certainly, teachers in this study expressed different combinations of value and priority across the three perceived purposes of assessment; contextual factors and the level of autonomy teachers had in different assessment decisions shaped the degree to which the three areas overlapped, as modeled in *Figure 1*. While the framework beneath the ACAI might best represent the dynamic experiences of K-2 teachers in this study, a closer look at the ACAI as an instrument highlights the problem in the field of not taking into consideration the unique needs of early learners.

The ACAI uses scenarios to prompt teacher responses, and the scenarios are not all applicable in an early childhood classroom. For example, the first scenario reads, “You give your class a paper-pencil summative test with accommodations and modifications for identified learners. Sixteen of the 24 students fail. As a teacher, your ideal priority would be to...” (DeLuca, LaPointe-McEwan, & Luhanga, 2016, p. 252). Kindergarten teachers in particular explicitly shunned the use of traditional paper-pencil assessments entirely for their students, and summative tests like this one are rarely used in first grade either. In situations where they are required by the district, teachers often criticized their validity, even in second grade. Moreover, none of the study’s teachers spoke of students “failing” a test. The second scenario deals with the teacher’s response to a student plagiarizing some of an assignment. One could argue that all emergent writing is “plagiarized,” as it is often directly copied or explicitly modeled after exemplar writing. This is how early writers learn to produce words and sentences, and thus, it is an expected and celebrated part of the learning process for beginning writers (Coker & Ritchey, 2015).

Given that at least two of the five scenarios deal directly with situations an early elementary teacher would not face in her classroom, the validity of the ACAI for measuring K-2 teachers’ approaches to assessment could be questionable. This is not intended to be an evaluation of the ACAI; the conceptual framework behind the instrument shows a complex and contextualized matrix of approaches that could be used for understanding the intricacies of teachers’ varied approaches towards assessment, including the approaches of early elementary teachers. Rather, this is to demonstrate that even the most recent and well-conceptualized instruments are being developed within the

context of a profession that insufficiently recognizes and considers the importance of DAP.

Summary of conceptualizations of assessment. The importance of better understanding teachers' conceptualization of assessment rests in the fact that their conceptualization influences their implementation of assessment practices (Barnes, Fives, & Dacey, 2015; Brookhart, 2011; Deneen & Brown, 2016). While this influence may not override the influence of *They* in all contexts, where teachers feel they have freedom of choice, their view on the purposes and priorities of assessment will influence the practices they select. According to Deneen and Brown (2016), "Teacher assessment practices will not be consistently or reliably enhanced if conceptions of assessment are not also enhanced" (p. 11). This view is held by other scholars, including Black and Wiliam (2018), who added to this idea the importance of integrating conceptions: "as a result of our work with teachers, we have become convinced that any approach to the improvement of classroom practice that is focused on assessment must deal with *all* aspects of assessment in an integrated way" (p. 552). For this reason, the next section discusses how teachers integrate the various roles and purposes of assessment.

Integrating Purposes of Assessment

Although six different frameworks for conceptualizing assessment were presented in the previous section, it is important to acknowledge that teachers do not neatly classify all of their beliefs and practices about assessment in clear cut categories that work in isolation, and they are able to hold multiple beliefs simultaneously (Barnes et al., 2017). As Black and Wiliam (2007) stated, "There are many possible cases of overlap, tension or synergy, between the different purposes" (p. 7). However, it is not desirable to always multi-task assessment practices. "As noted by Pellegrino, Chudowsky, and Glaser (2001):

‘Often a single assessment is used for multiple purposes; in general, however, the more purposes a single assessment aims to serve, the more each purpose will be compromised’ (p. 2)” (Marzano, 2010, p. 27).

Purposes are also compromised when they are not appropriately aligned. As addressed in the section of Chapter 4 on assessment for multiple purposes, K-2 teachers often struggle with carrying out assessments when they feel the design, emphasis, and purposes do not align. A teacher’s priority needs to be clear, as indicated in Claire’s guiding question: “What is the purpose of the test? That always has to be your focus.” This view aligns directly with a description of strong teaching in the NAEYC position statement on DAP:

A hallmark of developmentally appropriate teaching is intentionality. Good teachers are intentional in everything they do—setting up the classroom, planning curriculum, making use of various teaching strategies, assessing children, interacting with them, and working with their families. Intentional teachers are purposeful and thoughtful about the actions they take, and they direct their teaching toward the goals the program is trying to help children reach. (NAEYC, 2009, p. 10)

In addition to being intentional about approaching all aspects of teaching, including assessment, it is also important to be intentional about recognizing and addressing the possible detrimental “side effects” of educational decisions (Zhao, 2017). Black and Wiliam (2018) expressed concerns with how accountability purposes can unintentionally undermine the objectives their assessments are trying to achieve. They provided an example which, though not contextualized for an early childhood classroom, illustrates this phenomenon:

When science achievement is measured only through written assessments, particularly in a “high-stakes” testing environment, then teachers may focus on increasing their students’ achievements in the most efficient way possible, and this may well involve doing less practical science. What is important here is that, at least for Messick, the validity of the science assessment *would* then be jeopardized. The adverse social consequence—students doing less practical work in science—is directly traceable back to the fact that the assessment underrepresents the construct of science, at least as defined in the English national curriculum. The important issue here is not that the adverse consequences are bad; it is that they are directly traceable to deficiencies in the assessment. (p. 565)

An example from early childhood would be if a first grade literacy assessment focused only on reading comprehension at a surface level. Teachers could plan their instruction in ways to specifically raise performance on this skill, but the side effect could be that less instructional time is spent supporting oral language development, which is a foundation for strong reading skills. The deficiencies of the assessment would undermine progress towards the learning outcomes the assessment purports to measure.

This same situation can happen with deficiencies in assessment *systems* that misalign the purposes for which assessments were created and their use. As an example from the study, Jan described how her professional goal this year, which contributes 40% of her annual evaluation score, set by an unknown *They* in her large district, is to raise the performance of her students on the Concept of Word (COW) portion of the PALS test. Based on her perception, a county administrator read a research study that correlated kindergarten COW assessment performance with third grade reading levels, and so they are using the PALS’ measure of COW to evaluate teacher effectiveness. The COW word

list is 10 words read in isolation from context but having been taken from a short text students are to have orally memorized. The fall score (number of words read accurately) is compared to the spring score on a list of 10 different words, and Jan is evaluated based on her students' improvement on a score out of 10.

The layers of concerns this system raises are substantial. But focusing here on the student learning outcomes, it is conceivable that Jan and her colleagues would, due to the high-stakes nature of the test's outcomes for their own professional evaluation, focus their instruction on improving students' performance on this one PALS subtest at the cost of teaching the full range of skills needed to actually become successful readers in third grade, the target implied by the administrator's use of a research study using that as a target outcome.

Integrating purposes of assessment is important to having a balanced and effective system of assessing students' learning, but clear understanding of the limits of different assessments is equally critical. For teachers and administrators weak in AL, these may be difficult distinctions to appreciate. To the extent that many of these system decisions are outside the control of the classroom teacher, it may be more useful for teachers to focus primarily on the value assessment can add to a classroom instructional program. According to Stiggins (2014), "We must teach teachers and their supervisors to use assessment as a teaching tool—not merely as a grading tool" (p. 72).

As a teaching tool, assessment needs to be understood within the context of pedagogy (Black & Wiliam, 2018), a conception rarely described, though occasionally implied, by teachers in this study. K-2 teachers could talk about their preferred pedagogical approach and methods, but for the most part, teachers viewed assessment as something separate from their pedagogy. Many teachers described a pedagogy that

embraced developmentally appropriate practices—to the extent permitted by their contextual setting—but as assessment was a separate thing to their pedagogy, it was difficult to integrate a developmentally appropriate *assessment* practice into their teaching.

DAP and Assessment as Pedagogical Approaches

Over 20 years ago, Black and Wiliam’s (1998) influential work on formative assessment practices set the stage for viewing assessment as a tool of learning. In this work, published as the standards movement was just beginning to ramp up, they stated, “a focus on standards and accountability that ignores the processes of teaching and learning in classrooms will not provide the direction that teachers need in their quest to improve” (p. 140). Based on the reports of teachers in this study, it would appear that Black and Wiliam’s concerns came to fruition—an accountability system that has ignored the processes of teaching and learning.

Twenty years later, to address the commentary that their initial meta-analysis had served a pragmatic, but not theoretical, purpose, Black and Wiliam (2018) wrote:

We propose that formative assessment cannot be fully understood except within the context of a theory of pedagogy. As evidence that the business is, indeed, unfinished, we would point to the fact that much, and perhaps most, of the scholarly literature in the area of pedagogy says little or nothing about assessment.

In part, this may be because most writers on pedagogy have been concerned to focus on the social, cultural and political context within which schooling in general, and the curriculum in particular, are framed. In part, it may also be the result of a distinction drawn by some between *pedagogy* and *instruction*, in which

only the latter is concerned with the day-to-day work of teachers in classrooms.

(p. 552)

To help explicate how they chose to frame pedagogy, Black and Wiliam (2018) use another scholar's meaning:

We adopt Alexander's (2008) definition: *pedagogy is the act of teaching together with its attendant discourse of educational theories, values, evidence and justifications. It is what one needs to know, and the skills one needs to command, in order to make and justify the many different kinds of decisions of which teaching is constituted* [original emphasis]. Curriculum is just one of its domains, albeit a central one. (Black & Wiliam, 2018, p. 555)

This definition demands that pedagogy is treated as a broad construct that includes multiple dimensions. Using this definition, *developmentally appropriate practice* could be understood as more than just a research- and knowledge-based framework of practice. It could be interpreted as a theory of pedagogy of early childhood: *developmentally appropriate practice is what one needs to know, and the skills one needs to command, in order to make and justify the many different kinds of decision of which teaching is constituted.*

If developmentally appropriate practice is viewed as a set of practices or actions, it can sit in isolation of the set of practices and actions of assessment, as conceived by teachers in this study. They did not describe an *approach* towards assessment, they described behaviors that occur separate to instruction. It is no wonder that teachers expressed frustration over assessment practices required by *They* which are in conflict with the underlying theoretical approach to learning held by *I* or *We*. Perhaps the only

way to justify the misalignment of instruction and assessment is to hold them as separate things and not attempt to integrate them.

This cognitive dissonance (Festinger, 1957) is not a new phenomenon. In fact, nearly 20 years ago Shepard (2000) described how this kind of dissonance occurs because:

Instruction (at least in its ideal form) is drawn from the emergent paradigm, while testing is held over from the past. . . . dominant theories of the past continue to operate as the default framework affecting and driving current practices and perspectives. Belief systems of teachers, parents, and policymakers derive from these old theories. (p. 4)

Shepard described how an emergent social-constructivist paradigm guides instructional decisions in the classroom, while the older social efficiency/behaviorist learning theory paradigm still guides systems' decisions that shape assessment practices. Her article holds true today as it did at the turn of the century.

Given the perspective of developmentally appropriate practice as a theory of pedagogy, the conflicting paradigms become more complex. NAEYC first issued its position statement on developmentally appropriate practice in 1987. Although a full discussion of the ways in which it emerged from a social-constructivist paradigm is beyond the scope of this paper, the following excerpt from the most recent position statement highlights its paradigmatic roots:

Young children construct their knowledge and understanding of the world in the course of their own experiences, as well as from teachers, family members, peers and older children, and from books and other media. . . . Children take all this input and work out their own understandings and hypotheses about the world.

They try these out through interactions with adults and other children, physical manipulation, play, and their own thought processes—observing what happens, reflecting on their findings, imagining possibilities, asking questions, and formulating answers. When children make knowledge their own in these ways, their understanding is deeper, and they can better transfer and apply their learning in new contexts. (NAEYC, 2009, p. 14)

As a theory of pedagogy, developmentally appropriate practice predates the standards and accountability movement. But Shepard (2000) describes constructivism as the emergent paradigm. In her description, teachers are moving towards constructivism while holding on to testing practices from the past behavioristic paradigm. (She may not have been able to foresee how the measurement theory of the “past” would become the dominant force in the accountability movement.) For proponents of DAP, however, they feel they are being forced into a *new* paradigm of testing and measurement that does not align to their understanding of how learning and growth occur in young children. While the rest of the K-12 world is attempting to embrace formative assessment and performance-based assessment practices, developed from a more constructivist approach to learning, DAP-centered experienced K-2 teachers described being forced to *abandon* those types of practices in order to make room for the mandated assessments from behaviorism. Rather than being dragged behind the upper grades’ movements through educational paradigms, it would seem early childhood should be leading the way.

Summary of the Interaction Between Approaches to Assessment and DAP

An important finding of this study was a confirmation of the tension between external pressures shaping assessment practices and internal beliefs driving pedagogy (Black & Wiliam, 2007; Gullo & Hughes, 2011a; McMillan, 2003; Pyle & DeLuca,

2013; Shepard, 2000). But this study goes further in illuminating the specific ways in which this tension impacts K-2 teachers. While teachers across the K-12 spectrum may experience the tension between policies and beliefs, K-2 teachers have to cope with the ways their pedagogy of developmentally appropriate practice, uniquely shaped by the age and development of their young students, sits in stark opposition to the demands of the testing and measurement paradigm within which they must carry out their profession. This has implications for both practice and research, as described in the final chapter.

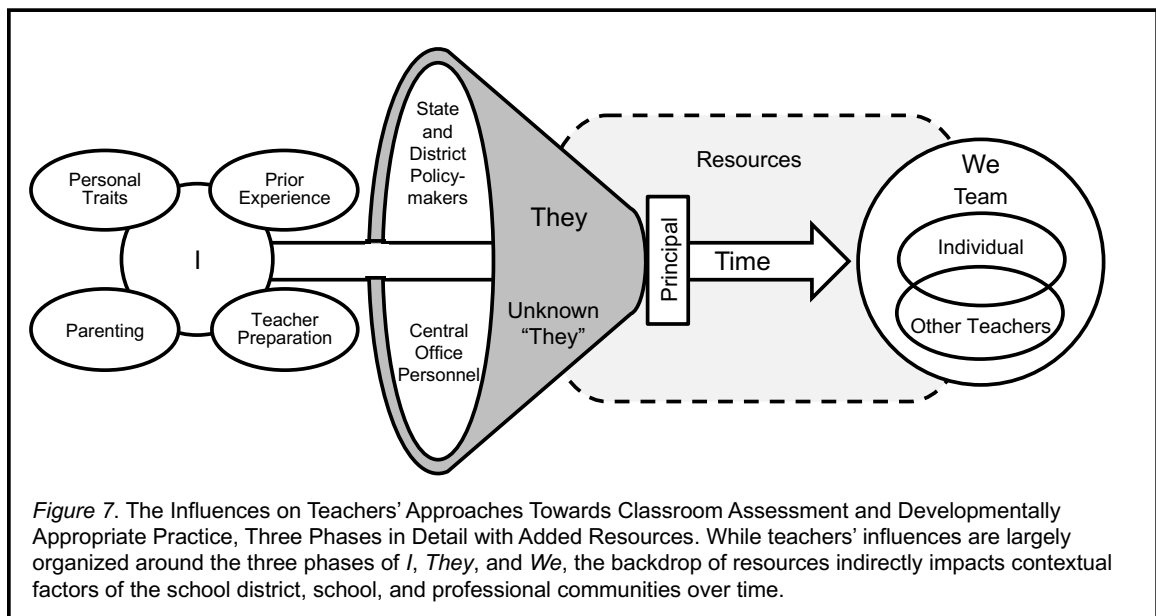
Summary of Discussion

Following the chapter on findings, this chapter has served the purpose of explaining the study's findings in light of existing research. The theory generated by the study allowed for deeper understanding of how teachers describe the influences on their approaches to assessment, the influences on their knowledge and implementation of developmentally appropriate practices, and the way the two concepts of approach to assessment and DAP interact with each other. The sixth and final chapter describes the implications for practice and recommendations for research that can be extrapolated from this discussion of emergent themes.

CHAPTER 6

IMPLICATIONS FOR PRACTICE AND RECOMMENDATIONS FOR RESEARCH

Through the use of a constructivist grounded theory methodology, this study has elucidated the influences on experienced K-2 teachers' approaches to assessment and developmentally appropriate practice. The outcome, a theory to explain the influences on teachers, is shown in *Figure 7*, presented again here for reference. Chapter 5 served to discuss the three research questions of the study in light of the major themes presented from the theory. Chapter 6 serves to suggest ways this study's findings may impact the future in the form of implications for practice and recommendations for future research.



Implications for Practice

The theory of influences on teachers' approaches towards classroom assessment is presented in three phases, the *I* phase, the *They* phase, and the *We* phase. Although teachers do not experience these three phases in a strict, linear fashion, the three spheres of influence are presented in this order to serve as a framework for sharing the implications for practice that appeared to be most important within each sphere. This presentation format is not intended to imply that implications are purely isolated to a single sphere or that each implication addressed has an equal influence on a teacher's development. Rather, it is meant to indicate that different sources and actors may have the ability to make changes in different spheres, so suggestions are made across the different major categories of *I*, *They*, and *We*. A brief mention of implications for Time and Resources is made at the end of this section as well.

Implications for *I*

The first sphere is the *I* phase, where the dimensions of influence that are individual factors are grouped. Because of the isolated, individual nature of many of these influences, the ability to make recommendations for systemic change in these areas is rather limited. However, the area of teacher preparation and the pre-service period is one worthy of attention.

Emphasize early childhood in licensure programs. Teachers whose pre-service formal educational experiences included an emphasis on child development and/or early childhood reported a better understanding of developmentally appropriate practices than teachers who entered the field with a non-education based bachelor's degree and a preK-6 or other teaching license endorsement covering a wide range of grades. These teachers trained in early childhood education spoke confidently of their ability to work with young

children and understand their needs as well as to advocate for early childhood practices. According to Bornfreund (2012), preparation programs tend to emphasize *either* child development and pedagogy *or* content and pedagogy. However, early elementary teachers need pedagogy based on an understanding of child development *and* pedagogical content knowledge.

As long as the state of Virginia has two elementary endorsements (preK-3 and preK-6), this distinction may continue. While consideration of splitting the endorsements, such as having a preK-3 elementary endorsement and a grade 4-8 middle years endorsement, would be a worthwhile inquiry for the state to consider, a full look at state licensure policy procedures was outside the scope of this study. Nonetheless, the findings do support the importance of teacher preparation programs laying a foundational understanding of child development, and therefore schools of education should consider how they are addressing development in early childhood for all their students earning endorsements for teaching children in the early elementary years. This suggestion is supported by other research (Buchanan et al., 1998; File & Gullo, 2002; Parker & Neuharth-Pritchett, 2006; Vartuli, 2005).

Provide preservice assessment instruction. The weakness of AL instruction in teacher preparation programs has been documented in literature earlier in this study and appeared as a finding as well, as very few teachers had a course on assessment. Teacher preparation programs should give attention to this important professional dimension and not relegate assessment to be touched on in other methods courses (McMillan, 2003) by instructors who themselves may not have strong assessment knowledge (Stiggins, 2014). According to Popham (2011), “what’s needed is something far more akin to a regular, honest-to-goodness assessment course in which the professor relies on a legitimate

assessment textbook” (p. 272). The present study supports this recommendation, but also suggests that a class alone is inadequate. Unless the course considers assessment *in the context of early learners*, it runs the risk of teaching theory and techniques with little practical application for early elementary teachers. The core concepts of assessment, such as validity and reliability, are applicable in all assessment situations, but the conditions of assessment decisions can dramatically change which practical decisions are valid and reliable, and if teachers are not able to translate their theoretical understanding into practice, that knowledge is useless at best, damaging at worst.

Increase experiences. At risk of deriving an implication from the findings that is difficult to leverage for change, it must be noted that teachers need an accumulation of personal experiences to understand children’s needs and development. Time and again teachers stated that they learned about teaching, assessing, and especially what children are like at different ages, through experience. While there may be little practical guidance for ways to increase personal backgrounds such as babysitting as a teen, volunteering with a youth group during college, or having children of one’s own, teacher preparation programs can and must emphasize extensive field placement experiences. There is some research to suggest that teachers with longer student teaching experiences have a stronger sense of self-efficacy (Dorel, Kearney, & Garza, 2016), but little has been done to examine how the grade level of student teaching experience influences teachers’ understanding of child development. Many teachers only experienced one classroom across practicum and student teaching assignments, and viewed this experience as insufficient, particularly if they were placed in an upper grades classroom and their first teaching job came in the early years.

Results of this study indicated that time working with young children was more important than the expertise of the cooperating teacher, although the potential impact of a strong child-centered cooperating teacher being able to frame a pre-service teacher's experience with young children was also noted (as in Chantal's case.) Although it would require a dramatic shift in the profession, considering ways for teachers to intern for longer periods of time or begin work with stronger systems of mentoring (such as is done in the medical profession), would be a great benefit to helping teachers gain more experience in their pre-service and early in-service years.

Implications for *They*

Research supports the importance of the pre-service time period in shaping teachers, as that is “a time when teachers' conceptions and approaches to teaching have yet to be fully formed. Once teachers' conceptions and approaches are established, they are resistant to change” (Vanlommel, Van Gasse, Vanhoof, & Van Petegem, 2018, p. 117). However, what the findings in this study show is that the authority and power of *They* is so substantial that it can outweigh any individual factors brought into the job by a teacher. Teachers' conceptions and approaches may not change, but *They* can impact or even control teachers' behaviors. As *They* set the mandates and steer the priorities, it is critical to look at ways to improve practice in this sphere.

Acknowledge that early childhood is different. The simplicity of this suggestion belies the complexity of actually creating change in this area. But frequently, teachers in this study stated how “it's different,” especially in kindergarten, and administrators failed to acknowledge that early childhood needed special consideration and not just scaled-down versions of upper-grades practices. The first needed step in improving outcomes for early learners is simply for *They* to acknowledge that early

childhood is a unique time period in children's growth and development (Copple & Bredekamp, 2009).

Improve “early childhood literacy.” As important as that first step is, it is not enough. This study highlighted the crucial need of central office and building administrators increasing their understanding of early childhood. Spinning off from the phrase “assessment literacy,” I propose here that all educators who make decisions impacting young children also need “early childhood literacy.” I borrow from Popham's (2011) definition of AL presented in Chapter 2 in forming a definition that states: early childhood literacy consists of an individual's understandings of the fundamental concepts of early childhood growth and development and the procedures deemed likely to positively influence educational outcomes for young children.

Using NAEYC's 2009 position statement, the fundamental concepts of early childhood would include the three core considerations (areas of knowledge to consider in making decisions), the 12 principles of child development and learning that inform practice, and five guidelines for developmentally appropriate practice. It is not unreasonable to expect that elementary principals would have *at a minimum* a working knowledge of these areas as a professional competency. Central office administrators or content area specialists, such as literacy coaches and curriculum directors, who make decisions that impact children up through age eight should also have this knowledge as a professional competency. Knowledge about content area or pedagogy separate from the centrality of an understanding of the *learner* is not sufficient.

Given the importance teachers placed on experience as the best (or in some cases, the only) way they learned about what is developmentally appropriate for young children, it is logical that teachers who would be wary of the understanding of administrators who

have not taught in the early years and particularly of administrators who have no experience teaching elementary school. Because teachers' experience so heavily shaped their own knowledge, they reason that a lack of experience in others would likely equate with a lack of understanding. This held true in the experiences teachers described, as most central office personnel and principals did not demonstrate an appreciation for the unique nature of teaching young children.

McDaniels, Issac, and Hatch (2005) wrote of a "general mismatch between the philosophy teachers bring to their work and the expectations of other adults who influence the teaching environment" (p. 21). One way to address this mismatch is by using professional development for administrators to increase their understanding of early childhood (Minicozzi, 2016). It would also serve the field well to look for ways to recruit more experienced early childhood teachers into administrative roles as well as consider how knowledge of child development should be reflected in administrative licensure coursework or proficiency tests. While this issue is most acute in the early years, surely improving the understanding of children's growth across all stages of development would be helpful for all district and building administrators.

Develop assessment literacy. While there is slightly greater awareness of the need for AL than early childhood literacy, this study also supports the crucial importance of increasing a contextualized knowledge of assessment practices for all administrators. According to Stiggins (2014), "very few practicing teachers and almost no practicing school leaders have been trained to develop quality assessments or to use them in effective ways regardless of the purpose" (p. 69). The teachers in this study rarely identified and labeled deficiencies in their administrators' decisions as a lack of overall AL, though many commented on the inappropriateness of the application of assessment

decisions to their classroom context with young children. Teachers *indirectly* questioned their administrators' competency in creating, delivering, and analyzing assessments effectively by describing practices they felt were ineffective or counter-productive. While Popham (2009) stated that assessment literacy PD for school administrators should be slightly different than that for teachers, the recommendation for improving AL through well-designed PD is supported both by the body of literature and the findings in this study. The gross misuse of the term formative assessment and the potential impacts of this error serve to highlight just one reason it is critical for administrators to improve their AL.

Value teacher expertise. The social and political landscape has not been kind to teachers in recent decades, characterized by distrust of teachers and a perception that external accountability measures are necessary (Black & Wiliam, 1998). According to Stiggins (2004), "our collective assessment actions over the past 60 years reveal a fundamental lack of trust in teachers and school leaders to accurately assess the achievement of their students" (p. 26). As Hatch (2002) described, "the movement toward standards is a movement away from teacher responsibility and agency" (p. 459). While shifting an entire paradigm is no simple matter, there are smaller ways that *They* can work from within the system to empower teachers and trust their professional judgment.

Vartuli (2005) stated, "to increase teachers' self-efficacy, administrators must give teachers opportunities to influence the schools' decisions related to curriculum content, selection of instructional materials and activities, professional development, [and other areas]" (p. 82). This study would further support Vartuli's recommendation, as many teachers expressed frustration with not feeling heard or respected, such as Chantal,

who stated, “I wish there were more flexibility for teachers to make decisions that they, professional decisions, that they know are best for their individual group of students. . . . teachers aren’t often treated as professionals as they really are.”

The study also provided evidence of how teachers respond when they *are* respected. According to Darlene:

I feel more valued as a professional that I can make some of those decisions on my own, and they don’t feel that they need to come down and, you know, with a heavy hand say, “this is what you must do at this time or this is the way you must do it.”

Granting teachers autonomy based on trust in their professionalism must be done from both the levels of district and building administration, but teachers also should advocate for themselves by proving their competency. So many teachers used variations on “you just know” to explain how they can determine their students’ learning performance and needs. Of course, they do not *just* know; they know by applying their understanding and experience to make decisions based on tacit procedures and processes that are not visible to others and often so second-nature to the teachers themselves that they are not even consciously aware of the processes they are employing. Recall that many teachers did not label as *assessment* the small, ongoing, formative processes they perform in the name of adapting instruction; perhaps these processes become so innate with experience that teachers are not able to isolate and identify them with an outside label. They are not viewed as formal procedures to plan and carry out; these types of assessments are simply a second-nature part of what they intuitively do as teachers.

As part of improving the value placed on teacher expertise, teachers must learn to use the vocabulary of assessment and educational practices to explain, authenticate, and

defend the decisions they make on a daily basis. At the same time, administrators must provide opportunities for sharing professional responsibilities and trusting their teachers' judgment.

Share learning and leadership. One of the key findings of this study is the importance of the mediating role of the principal in supporting teachers and buffering the effects of heavy-handed or misguided district management, in cases where it existed. This finding suggests that principals can and should take a strong lead in supporting an environment of shared learning and respectful teamwork. Principals can build teacher confidence by providing useful and effective PD opportunities and providing ongoing support for collaboration (Black & Wiliam, 2018; Huebner, 2009). According to Louis and Wahlstrom:

Teachers and administrators need to engage in deeper organizational learning—learning that uses all of the knowledge and resources that can be brought to bear on the core problems of practice in their particular setting. . . . it also includes teachers' knowledge, including what they already know but may not have shared and emerging knowledge that they create through action research. (2011, p. 54)

While some teachers had experienced meaningful PD in building AL, no teachers reported having PD in developmentally appropriate practices. Knowledge of DAP could be the “teachers' knowledge” about which Louis and Wahlstrom wrote. Assessment experts may be able to provide deep theoretical learning experiences in strengthening AL and teachers can provide their experiential expertise in early childhood to together collaborate on improving assessment practices and systems for use with young children.

While valuing teachers' expertise and sharing decision-making responsibilities with teachers is crucial, principals must be responsive to teachers and ready to take the

lead in communication when needed. Finding a balance between supporting teachers and providing autonomy can be a challenge, and administrators should be ready to adapt as circumstances change. According to Melanie, “teachers don’t want to always be told what to do, but then we want to be given some guidelines as to what we’re supposed to do as well.” Where there is space for autonomy, a principal’s role in mediating should be used to empower teachers, but when the district has not permitted flexibility, principals must communicate clearly about expectations and advocate for the needs of students as expressed through teachers’ expertise.

Implications for *We*

The third sphere of influence is the one in which teachers are most likely to not just be influenced, but also be an influencer. In the context of *We*, teachers should take individual responsibility within the context of the group to act at two levels for improving the use of developmentally appropriate assessment practices.

Broaden the understanding of assessments. The way teachers view what is—and is not—assessment was an important finding of this study. While *assessment* includes the formal, required computer-based tests and district-made, paper-and-pencil, common unit tests, it also includes a full spectrum of other behaviors and activities that occur in early elementary classrooms on a daily basis. Gareis and Grant (2015) provided a list of assessments across a continuum of formality that range from project-based assessments to reading facial expressions. While many teachers in this study described actions they took or procedures they used in their classrooms to adjust instruction—such as a quick thumbs-up, a turn-and-talk routine, writing on whiteboards, making anecdotal notes about students’ work while they are practicing, and asking children to explain their thinking—teachers rarely categorized these practices as *assessment*.

If teachers began to acknowledge and accept all the ways assessment is at work in their classrooms, they would be able to better support their professional decisions and actions and, with improved confidence, advocate for practices that substantiate reasonable judgments about student learning that do not depend on developmentally *inappropriate* methods. Due to the emergent nature of young children's literacy skills, so many of the forms of assessment that yield sensible inferences are based on oral language or demonstrations of learning, which cannot be as easily captured on paper or with traditional assessment methods. If teachers continue to narrow their understanding of assessment to only those traditional methods, they undermine the legitimacy of the judgements they make through more appropriate methods.

Build shared assessment literacy. In addition to broadening what they consider qualifies as *assessment*, it is also imperative that teachers continue to build their AL. Due to the initial literature review that situated the research questions, this finding was not surprising. However, this study adds nuance to why it is particularly important for early elementary teachers to improve their AL. Foundational understandings, such as the concepts of validity and reliability, are crucial. But early childhood teachers must also be able to justify and defend non-traditional assessment methods for use with their young students. Administrators and fellow teachers who may hold a more social efficiency/behaviorist learning theory view of testing will likely need to be convinced of the validity of unconventional or informal methods of assessing, and they may need support in shifting their approach to assessment to support a more pedagogical stance. Teachers must have stronger AL in order to better communicate and promote for effective assessment practices that benefit their students' learning journeys.

Because of the strong shared practices of *We* that teachers described, this may well be a case of the rising tide lifting all ships. According to Gareis and Grant (2015), “a teacher can influence the professional practice of teachers and instructional leaders in a school first by developing the competency of assessment within him- or herself and then by becoming a teacher-leader in the area of classroom assessment practices” (p. 167). Teachers should actively work to break down the “two-silo division of responsibility” described by Stiggins (2014) by working collaboratively to develop professional assessment competencies through processes like a book study (Popham, 2004) and sharing examples of what works (Black & Wiliam, 1998). Due to the strong nature of collegial learning and a community of practice, early childhood teachers have the opportunity to spread early childhood assessment knowledge and bolster the use of effective practices even without the direct support of *They*.

Focus on the child. According to Hatch (2002), “by training and disposition, teachers of young children are concerned with understanding and teaching the whole child” (p. 458). However, this study indicated that the outside forces of *They*—and sometimes limitations of the *I* or influences of the *We*—largely shaped these teachers’ practices, which therefore did not always align with their philosophical approach of teaching the whole child. It is therefore important that, as part of their shared learning opportunities, teachers look not only at assessment, but also at improving their understanding of child development at large and focusing discussion and priorities on the *child* over the curriculum or the structure of schools.

Brookhart (2015) stated that professional judgment needs to be “informed by clear criteria and clear understandings of students and content” (p. 6), and in the case of understanding students, this means having knowledge of typical childhood development

(Shepard, 1994). Teaching teams should use developmental continuums to support their learning about what expected behaviors and abilities are for the age of students in their classrooms. Developmental continuums can also serve as anchors for some assessment practices. In recent years in the United States, there has been a growing curiosity around the pedagogical methods of the Reggio-Emilia approach to early childhood education. While the full range of implications of such an approach are well outside the bounds of this current discussion, the key tenant of documentation of student learning is one that is worthy of study for a teacher team (Krechevsky, Mardell, Rivard, & Wilson, 2013).

The potential for using developmental continuums to anchor observational notes taken while watching children has multiple benefits: it would enhance the depth of teachers' understanding of their students as individuals, it would strengthen teachers' knowledge about typical behaviors at different developmental stages, and it would provide assessment evidence of students' abilities in appropriate formats that can lead to valid inferences about student learning. Similar to the methods used for lesson studies, teachers should work as teams to practice their observation, documentation, and inference-making skills, using developmental continuums as frameworks for professional conversations about student growth and learning.

Advocate for early childhood. If teachers are able to gather evidence that can lead to more sound educational decisions, they will be in a better position to collectively advocate for their students' needs and, thus, for the needs of young children in general. Many teachers expressed concerns about the impacts of putting pressure on young children to perform and to reach benchmarks that are outside reasonable expectations for their development. But teachers must have confidence to express these concerns to *They*, to the decision makers who can support change initiatives. This kind of confidence comes

through increased knowledge and experience, such as through some of the previously-suggested ideas for practitioners. Maxwell et al. (2001) also had specific advice about the nature of the message teachers should try to share:

By marketing the concept of developmentally appropriate practice as one that applies only to young children, supporters of developmentally appropriate practice may inadvertently be doing more harm than good. People may dismiss the ideas as relevant only for very young children when, in fact, many of the principles apply to children and adults of all ages. (p. 446)

While proponents of DAP may recognize the unique developmental needs of young children, their message is less likely to be received by *They* if *They* are not already inclined towards early childhood. Based on the reports of teachers, central office personnel and administrators are highly unlikely to have backgrounds in early childhood, and therefore less likely to be open towards readily embracing a new way of interpreting education based around a set of principles with which they are not familiar. Hatch (2002) commented about the ways society's elite shapes curriculum, and his comments apply here as well to how some administrators may subconsciously think about educational practices: "It's easy for those with cultural power to reach consensus about which knowledge is of the most worth: it's the knowledge they already have" (p. 460).

Very few central office administrators have knowledge about early childhood, according to participants, and so they are likely to be biased against knowledge of DAP as less important or less relevant than other concepts in education. In order to push against the dominant guiding structures, teachers have to find ways to make their message seem relevant to all, and to connect it to the knowledge that those with power already have. Although it is counter to the prevailing standards-based, high-stakes

accountability paradigm, if early childhood educators can advocate for ways to center education around the *learner at all* stages of development, they are more likely to find receptive listeners who can appreciate and support a child-centered practice.

Implications for Time and Resources

Change, especially at the institutional level, takes time. According to Deneen and Brown (2016), “changes in teachers’ beliefs or attitudes usually take longer than changes in knowledge or skill” (p. 4). It is important that teachers advocate with a sense of urgency, but support learning and change initiatives with patience and recognition of the time needed to shift thinking. Administrators should strategically plan to support new programs or undertakings with the time needed to sustain change. For instance, if a grade level team decides to do a book study to improve their knowledge of the foundations of assessment, the principal should not add additional responsibilities or other new projects during the same time period. If resources can be put towards hosting ongoing professional development sessions on AL, the district needs to set aside specific time across a span of time to nurture long-term learning. If teachers wish to work with each other in informal ways to practice new assessment strategies or strengthen their understanding of child development through child observation techniques, the principal should protect their time by providing assistance from a teacher’s aide or substitute to release the classroom teacher to devote time to her learning.

Time can be seen as a valuable resource in this context, but material resources may also be needed. Popham (2004) recommended a book study as a viable means to improving AL. At present there are no textbooks that adequately address the considerations of assessing early elementary learners; most K-12 textbooks do not even mention the impact of child development on assessment practices and judgments (Shy,

2018). While there are a number of strong professional books that support growth in AL, this study's findings suggest that materials that do not adequately acknowledge young learners' needs may prove to be more of a frustration than a support. There are a number of books targeted specifically to early childhood assessment knowledge and practice; while these tend not to be marketed to the average elementary school teacher—perhaps because of the structural difficulties in carrying out some of these practices in the elementary school system—I would recommend carefully considering one of these resources as better able to support the needs of early elementary school teachers as they look for alternative, appropriate assessment practices.

Finally, as an implication of the resource of curriculum, the reports of K-2 teachers suggest that the state standards and districts' associated pacing guides need to be reconsidered in light of the developmental needs of early learners and the pedagogy that best supports their learning. Rather than using a push down model of looking at the end of the K-12 journey and backwards-mapping grade level expectations to reach the long-term ends based on content outcomes, creators of curriculum should integrate a start-up consideration for what students bring at the beginning of their journey. Based on the skills, experiences, and developmental abilities of five-year olds, what is a reasonable amount and the appropriate nature of learning we should expect each year? How can we focus on the *learner* rather than content-covering? This need not neglect important content-based learning, but it centers schooling differently. Policy leaders need to address what Hatch (2002) called the “curriculum shovedown” (p. 457) and consider that early childhood education's

standards presume that young children are complex human beings who learn best when they are guided, nurtured, and cared for, not lifeless commodities that must

meet standards of production. Our accountability comes from an ethical commitment to do what is right for every child, not from measuring productivity according to an arbitrary set of narrowly defined outcomes. (p. 461)

Summary of Implications for Practice

While this study highlighted many categories and dimensions of influence over teachers' approaches to assessment and developmentally appropriate practice, this section shared a few of the more important implications for practice that can be derived from the findings and discussion about the most noteworthy influences. The next section addresses implications for future research.

Recommendations for Research

Because the goal of grounded theory is to explore a process or experience that is not well-studied or understood, it serves to uncover variables rather than prove them (Charmaz, 2014). Each newly-recognized variable, therefore opens a new area for exploration, and the possibilities for future research are numerous. Here, I recommend a few areas that the study suggests would be high priority for next steps in research. This study could be replicated and, due to the constructivist nature of grounded theory, a different researcher or different participants could illuminate different dimensions or properties of the questions involved, even in the same context. However, as grounded theory does not claim generalizability, it does not necessitate replications to confirm its findings.

Knowledge and Experience in Central Office

The backgrounds, experiences, and knowledge of central office administrators regarding the education of young children is relevant to the practice of early childhood teachers, but it has not been studied. Research on demographics, knowledge, and

philosophical orientations towards early childhood would better illuminate the specific needs and strengths of district leaders as well as highlight ways to improve practice and support growth in meaningful ways. While there is a growing amount of literature on administrators' AL, the dearth of research on administrators' approaches towards DAP rises to a higher research priority based on this study's findings. In fact, one of the participants posed this inquiry in her interview: "I'd be really curious to know if there's like a profile or a profile could be developed of what kind of person makes that choice to go from classroom into central office. Like, if there's a distinct type."

In addition to better understanding the current status of central office administrators and specialists' approaches towards developmentally appropriate assessment practices, theoretical and practical research that would support training programs for administrators to increase their "early childhood literacy" would work well in tandem with research that highlights administrators' needs to improve practice.

Historical Look at Teachers' Development

The methods for this study were not designed to support a strong comparative analysis of teachers across different stages of their careers: veteran teachers, mid-career teachers, early-career teachers, and novices, who were not included in this study. But the descriptive findings suggest that how teachers respond to challenges of assessing in early childhood may vary based on years of experience as well as the paradigm within which the teacher was trained and began her career. An exploration of K-2 teachers' meaning with the language of "old school" teaching could be illuminating, as could a qualitative study analyzing the differences in new and veteran teachers' philosophical approaches towards as well as practical knowledge of AL. A historical analysis of preparation programs would also shed light on the current situation. Such an analysis could look at

how preparation programs have changed over time in how they have taught assessment knowledge and skills and the degree to which these programs emphasized early childhood development. These findings could support a better understanding of how degree programs, elapsed time, and the changing paradigms of education mingle and interact as they influence early childhood teachers.

K-2 Teachers' Approaches to Early Childhood Pedagogy and Assessment

This study shed further light on the complexities of how philosophical orientations, practical knowledge and skills, and structural obstacles intermingle to shape practices, but more research is needed to better understand the intricacies of these relationships. Further study regarding the interaction of an orientation towards DAP and assessment practices would add to the literature base, and a theoretical consideration of the role assessment plays as pedagogy in the early years would also contribute to the body of research. One recommendation for research is to look specifically at early elementary teachers' self-efficacy for their pedagogical approach towards assessment. Examining how knowledge about child development plays into pedagogical content knowledge and practical assessment knowledge of teachers would also elucidate the multi-faceted constructs of teachers' beliefs, pedagogical orientations, and conceptualizations of the purposes and uses of assessment.

Teacher Preparation Programs

Teachers in this study rarely expressed confidence in the quality of their licensure programs for preparing them to effectively develop and utilize strong classroom assessment practices. Because so few teachers had a strong assessment course, research looking at the effects of such a course on the students fortunate enough to have one and comparing those effects to well-designed in-service PD would be a useful contribution

towards understanding ways to effectively raise teachers' AL and shape their approaches towards assessment.

Reporting Systems in Early Elementary

As this study was focused on assessment approaches, given that one of three emergent conceptualizations of assessment was that it served a purpose of reporting, it is logical that issues of reporting would arise through the analysis. These issues suggest the need for more research on reporting practices in early childhood. A large-scale study of research on reporting by Brookhart et al. (2016) included over 100 studies from the previous 100 years, *none* of which focused specifically on early elementary reporting practices. Several studies reported looking at elementary school, but further examination showed at least two of these (McMillan, Myron, & Workman, 2002; Welsh, D'Agostino, & Kanislan, 2013) claimed to look at elementary practices, but only included Grades 3-5, suggesting that the handful of other studies that claimed to look at elementary practices may not have included the early grades either.

Most teachers reported that their district's report cards changed somehow from kindergarten to third grade, but none of the teachers expressed a strong understanding for the base of theory, practice, or knowledge that supported their report card methods. While issues in older grades around reporting may be more about grading schemes, this study indicated that grading practices, scales, and symbols used in the early years may not adequately track growth or effectively communicate about students' abilities. A content analysis of report card elements and components would reveal the range of practices in place, such as the use of different letter grade symbols, point scales, and comment sections. Also, consideration of how different content areas are organized by strands, concepts, and objectives as well as how teachers decided grades for these different areas

and how parents interpreted the grades would be beneficial. Further research on the effectiveness of various practices to communicate intended meanings could help reporting systems better capture and share meaning about student growth and performance. Research on the impact of reporting systems on assessment practices would also be a valuable addition to the field.

What Works in Early Elementary Assessment Practices

Another outcome the research questions were not specifically designed to lead towards was an audit of assessment practices in the early elementary grades, yet in the course of responding to questions, most teachers described practices that they believed were not valid or reliable for their young learners. But there is a lack of evidence regarding the impacts of various forms of assessment on early learners' abilities to adequately demonstrate their knowledge and skills. Specifically, evidence that analyzes the use of forms of non-traditional, that is, not paper-and-pencil, methods of assessing young children would help teachers make more effective assessment decisions. Almost all research on assessment methods focus on methods that rely on basic literacy skills, and for students who are not yet readers and writers, most evidence of method effectiveness is anecdotal or experiential.

Several teachers commented on wanting to have guidelines to help them create assessments that are developmentally appropriate. Research-based recommendations on formats, length of assessments, and language considerations in developing assessment activities would be helpful, they stated. Furthermore, teachers expressed a desire to better understand how issues of validity and reliability operate with children who, developmentally, are unreliable by nature. Do the constructs of validity and reliability

take on different dimensions when it cannot be assumed that the learner is able to produce his or her “best” efforts on assessments?

Furthermore, there is a need to explore ways to improve the degree of validity given to teacher judgments about student learning. According to McMillan (2003), “assessment, like instruction, involves informed professional inquiry and judgment. From this perspective, assessment is much more subjective and even intuitive than is implied by the use of technical measurement concepts” (p. 39). In the present paradigm, technical measurement theory dominates, but it cannot be carried out with confidence for young children to the same degree as it can with older learners. Thus, educators cannot place the same level of confidence in the technical merits of the construction of an assessment without considering its application with the learner. But, can they have confidence in their inferences and judgments when they involve intuition?

In the context of conducting qualitative research in the constructivist paradigm, Lincoln and Guba (1985) claimed, “there is no reason to believe that humans cannot approach a level of trustworthiness similar to that of ordinary standardized tests” (p. 195). This statement merits deep consideration—and further research—in the context of whether teachers can approach a level of trustworthiness similar to, or perhaps even above, standardized tests.

Summary of Implications for Research

As DeLuca and Hughes (2014) stated, the research in support of teachers integrating assessment in the early years is “alarmingly sparse” (p. 442). In this section, I mentioned six areas that demand closer analysis, and I made recommendations for further research based on questions that arose from this study. With most good research, attempts to explore questions often end up creating more questions than they answer. This

grounded theory study was no different, and I highlighted just a few of the most promising or urgent research needs suggested by the study's findings.

Summary of Study

The forces of conflicting educational paradigms tug at the daily practice of kindergarten, first, and second grade teachers. With the dominant accountability and testing era's emphasis on standardized measurement, the topic of assessment has become one of the most prominent areas of tension. Bodrova and Leong (2018) addressed this tension:

Although *assessment* has become a loaded word in the field of early education—often associated with bias, push-down expectations that are not developmentally appropriate, and pressure on children and teachers to perform—appropriate, authentic assessment lies at the heart of learning and teaching. (p. 15)

How teachers handle inappropriate expectations while optimizing the pedagogical benefits of appropriate assessment practices is a question that has been largely ignored in educational research. This study aimed to help fill the gap in our understanding about how different forces influence and shape K-2 teachers' assessment approaches and practices, given the unique developmental needs of their young learners.

The study was guided by three research questions and employed a constructivist grounded theory methodology to address the questions (Charmaz, 2014). I interviewed 35 K-2 teachers in the state of Virginia using an unstructured interview process that evolved as I reached theoretical sampling and theoretical saturation. The results of the study are a theory that explains the influences on teachers' approaches to assessment and developmentally appropriate practices. The theory presents three phases to show how teachers bring individual beliefs and experiences into the field (*I* phase), how teachers'

practices are shaped by the context and district policies—mediated by building principals—within which they are situated (*They* phase), and then how teachers enact the combination of their beliefs and practices in a shared context (*We* phase.) The findings also illuminated additional understandings, such as how teachers conceptualize assessment.

The findings led to numerous implications for practice and recommendations for research, while the overall impact of the study is the way in which it shines a spotlight on the unique challenges early elementary teachers face in balancing their beliefs about pedagogy for young learners with the constrictions and demands for practice on their assessment behaviors. Black and Wiliam (2018) stated:

Assessment cannot be understood without a consideration of the wider context within which that assessment takes place. Teachers and schools are constrained, at least in the short term, by the cultural traditions, the political and public expectations of education, and the norms of the various institutions within which they operate. (p. 570)

This study adds “consideration of the wider context within which that assessment takes place” by suggesting that the context of *early childhood* is unique and crucial to understand when making decisions about assessment. Deneen and Brown (2016) also emphasized the importance of context, stating, “changes to theory and teacher education are not enough if tensions in practice remain unresolvable. ... To be an assessment literate teacher requires knowledge, skill, and beliefs appropriate to the context in which assessment is deployed” (p. 12). The findings of this research can help practitioners and researchers further examine the knowledge, skills, and beliefs appropriate for assessing children in the early elementary grades.

APPENDIX A

RESEARCHER AS INSTRUMENT

As the researcher is a key part of any qualitative research study in the constructivist paradigm, I must address my own assumptions, experiences, and perspectives that may influence the construction of the study process and its findings. I do not claim to be able to fully bracket all my biases and positions of privilege, nor does Charmaz (2014) believe this to be possible or even necessary. But I reflect in hopes of raising my own awareness of these issues and being cognizant of how they may influence the research. Here I highlight four major life experiences that have influenced that way I view social reality.

First, I spent twelve years as an adult living in Asia. While living in vastly different cultures, I developed both a need and an ability to question my own behaviors, thoughts, perspectives, and assumptions as well as question those of others without making value judgments based on behaviors alone. I recognize that there are unseen but powerful forces of culture at many levels, and those lead to multiple ways of seeing the world rather than one objective “right way” of understanding people and ideas.

Second, I have taught in early elementary grades for twelve years in three very different schools. My own teacher training was weak in both AL and in addressing developmentally appropriate practices, and my growth in understanding each of these constructs came out of struggle over the years as I grappled with different implications of these areas in different contexts in my career. As a teacher, I am comfortable with ambiguity in plans (or research methods) to the extent that flexibility allows one to respond to the needs presented at the time, and I think the balance of having solid plans

and knowing when and why it is necessary to change them in the course of action is a skill that will carry over to grounded theory work well.

Third, while teaching, I earned my master's degree in school administration and completed a lengthy internship with a principal in a preK-grade 2 division of a large international school with exceptionally high professional standards. The principal I worked with lead from a deep understanding of and appreciation for developmentally appropriate practice, and she instilled in me a greater appreciation for child development and its implications in the classroom setting. She also entrusted me to lead the school's multi-year effort towards shifting reporting and assessing to a new standards-based curriculum, a process that truly opened my eyes to the challenges of navigating both of these paradigms simultaneously.

Finally, and perhaps most importantly, I am the mother of two young children and see the miracle of child development magically unfolding in front of my eyes daily. Now that my son is in kindergarten, I have added the lens of *parent* to the lenses of *teacher* and *administrator* that I already carried with me when considering practices in school settings.

While there are certainly other significant influences on my life, given the context of this research study, I believe these four areas shape my view of the issues the most, and therefore, it will be important for me to continue to utilize reflexive journaling throughout the course of the study to identify ways that my experiences are present in the co-constructing of reality with my participants.

APPENDIX B

EMAIL TO PRINCIPALS SEEKING RECOMMENDATIONS

Dear _____,

I am a Ph.D. candidate at the William & Mary School of Education studying Educational Policy, Planning, and Leadership and conducting my dissertation research under the guidance of Dr. Chris Gareis. My research will examine how kindergarten through 2nd grade teachers develop their approaches towards classroom assessment. I am hoping that you can help me identify teachers in your school who may be able to help me.

I am looking for teachers with at least four years of teaching experience (at any grade) who are currently teaching in a kindergarten, first, or second grade classroom. I will interview each teacher once or twice for approximately 45 minutes each time. All participants selected for interviews will receive a small incentive for participating, and any and all data will be kept strictly confidential.

I will not be conducting any observations in classrooms of students or teachers; my research is strictly interview-based. I am not looking to validate specific practices or to evaluate teachers' perspectives or behaviors. I am seeking to better understand the factors that influence teachers' development and perhaps also learn more about the current state of classroom assessment in K-2 classrooms.

If you have a few teachers in your building who you think would be good candidates for this study and who might be able to provide rich interview data on this topic, I ask that you forward this email to them with a simple request to see the attached letter and copy me on the email to your teachers (lkshy@email.wm.edu).

If you have questions of any kind, please do not hesitate to ask. Thank you for your support of this research and its goal to further the knowledge regarding assessment practices in the early elementary years.

Kind regards,

Leah Shy

Ph.D. Candidate in Education Policy, Planning, and Leadership

Adjunct Instructor in Education

William & Mary School of Education

Dissertation research conducted under the guidance of Dr. Chris Gareis

APPENDIX C

**LETTER TO TEACHERS SEEKING PARTICIPANTS (ATTACHED TO EMAIL
TO PRINCIPALS)**

Dear Teachers,

I am a Ph.D. candidate at the William & Mary School of Education studying Educational Policy, Planning, and Leadership and conducting my dissertation research under the guidance of Dr. Chris Gareis. In my research, I am examining how kindergarten through 2nd grade teachers develop their approaches towards classroom assessment. Your school principal identified you as a teacher who might be a good source for my work.

I am reaching out to see if you would be willing to be interviewed as part of my research study. Participants will be asked to meet at a mutually convenient location (it does not have to be your school setting) for one or possibly two interviews. Each interview will last approximately 45 minutes, and the second interview may not be necessary for all participants or could be conducted virtually. There will not be any observations of your teaching practice. Your recommending administrator will not know if you have chosen to participate or not, and your identity will be kept strictly confidential throughout the research process.

If you would like to be considered for an interview, please use [this link](#) to fill out a brief background survey (it should only take about 5 minutes to complete.) I taught in early elementary school classrooms for twelve years, and I know how precious your time is. As a small token of my appreciation, every teacher who fills out the interest survey will be entered into a draw for a \$50 Amazon gift card. In addition, all teachers who are selected for interviews will receive a \$10 gift card to Amazon.

If you have questions for me about the study or your participation, please do not hesitate to ask. Thank you for your support of this research and its goal to further the knowledge regarding assessment practices in the early elementary years.

Kind regards,

Leah Shy

Ph.D. Candidate in Education Policy, Planning, and Leadership

Adjunct Instructor in Education

William & Mary School of Education

Dissertation research conducted under the guidance of Dr. Chris Gareis

APPENDIX D

SUMMARY OF RESEARCH PLAN

Context of Research Topic

Teachers of children in kindergarten through second grade face a unique problem as they attempt to navigate conflicting educational paradigms. With the developmentally appropriate practices of early childhood pulling in one direction and the standardized practices and measurements of the accountability movement pulling in another, K-2 teachers often have to navigate a tricky balancing act in order to best support their students and honor policy demands of the school and state. One area where this tension is particularly pronounced is in creating, delivering, and interpreting developmentally appropriate classroom assessments.

Research Questions

At present, there is research that shows K-2 teachers tend to be relatively weak in their understanding of developmentally appropriate practices and similarly weak in their classroom AL. However, there is little that discusses the overlay of these two constructs from two different paradigms. In particular, there is no research that explains how teachers arrive at their conceptualization of developmentally appropriate classroom assessment practices. To address this gap, my research will focus on three questions:

1. Through what processes and influences do experienced K-2 teachers report that they have developed their approaches to assessment?

2. Through what processes and influences do experienced K-2 teachers report that they have developed their knowledge and implementation of developmentally appropriate practice?
3. How do experienced K-2 teachers' approaches towards classroom assessment and philosophical orientations towards DAP interact with each other?

Methodology

To answer these questions, I am using a constructivist grounded theory approach. I will interview teachers using an interview guide rather than a strictly structured protocol so that questions can be responsive to the emerging data and themes. Each teacher will be interviewed once (or on the rare occasion, twice), and I will remain open to the possibility that the emergent themes may suggest another form of data generation or data collection might help inform the process.

I will use a constant comparative process to analyze and generate data by analyzing each interview before conducting another interview, then cycling back to analyze previous interviews compared to emergent themes or categories from the new data. I will undertake initial coding using gerunds, then move to focused coding to see which codes suggest the most analytic power, and finally move into theoretical coding as a way of helping develop a theory to explain how K-2 teachers develop their approaches to classroom assessment.

Throughout the process of data generation and analysis, I will use memo writing as a way to record thinking about the data and reflexive journaling as a way to record thoughts about my participation in and influence over the process. Constructivist grounded theory does not call for a strict bracketing of the researcher out of the process

as a means of validating findings; rather, it relies more on trustworthiness to support findings. I will focus on Lincoln and Guba's (1985) categories of credibility, transferability, dependability, and confirmability through the use of member checking, the constant comparative process, thick description of contexts, and the memoing and reflexive journaling processes.

Participants

Participants for this study will be general education teachers currently teaching in a kindergarten, first grade, or second grade classroom. Participants must have completed at least 4 full years of teaching, although not all of their experience must have been in a K-2 class. Participants must be working in a public school in the state of Virginia, but they will have a range of backgrounds on other matters, such as the university program where they undertook their preparation program, the endorsements held (preK-grade 3, preK-grade 6, etc.), the background of the school division (urban, suburban, rural), and other possibly salient factors.

Participants will be recruited by following division protocols for contacting teachers; in some cases, this will be by direct email invitation and in others it may be through a referring principal, but the least disruptive method for schools will be used. Teachers will have the option of responding to the invitation without any division administrators knowing of their participation.

Participants will receive a \$10 Amazon gift card for their participation. Teachers who are asked to complete a second interview (and do so) will receive a second \$10 card. This amount for this population is small enough to avoid coercion to participate, and there is no foreseeable harm or risks of involvement to participants.

I will continue to interview teachers through theoretical sampling until I reach a point of theoretical saturation, which is impossible to predict prior to the study. However, researchers often suggest a minimum of 20-30 participants for grounded theory (Creswell, 2013).

Privacy and Confidentiality

Although the nature of the data to be collected does not indicate a significant risk in regard to matters of security, all necessary precautions will be taken to protect the privacy of participants and the confidentiality of their responses. Records of this study will be kept strictly confidential. All electronic information will be coded and secured on the hard drive of a personal laptop that is password-protected and always remains in the investigator's possession. A back-up copy of interview transcripts will be saved on the cloud in a password-protected account and without attached identifying information. Any paper records will be kept locked in a file cabinet.

Once transcribed, all recordings will be electronically erased. Participants will be assigned a pseudonym for the transcriptions of interviews, and pseudonyms will be used in any published material. Names of elementary schools, school divisions, universities, or other personal identifying information will also be assigned pseudonyms before publishing, and no information will be included in any report or other written material that would make it possible to identify participants.

Impacts of Participation

This study is considered to have minimal impact on teachers, schools, divisions, and most importantly students and their learning opportunities. It requires no observation

of teaching, no contact with students, and no access to student, school, or division data. As the study is focused on teachers' individual development and not evaluating particular practices, it can be conducted entirely outside the work environment in teachers' personal time and will have no impact on instructional time. While it is likely that some teachers will comment on influences from the context of the school or division, the study is looking at teachers from across the state, and not examining or reporting on any particular school setting or set of division policies.

Research Outcomes

The outcome of this study will be a theory, most likely in the form of a visual representation, that explains how teachers come to their conceptualization of developmentally appropriate classroom assessment. It is hoped that in addition to discovering variables that may be key influences over teachers' development, the study will also provide a frame for future discussions in the field and support for future research opportunities.

APPENDIX E
COURTESY EMAIL TO PRINCIPALS ABOUT STUDY

Dear Principals,

I am writing to inform you that it is possible some of your teachers may volunteer to participate in a study I am conducting for my dissertation research through William & Mary. The study is looking at the influences on K-2 general education teachers' development of their approaches towards classroom assessment. It is not focused on any particular school or even the division as a whole, but rather involves teachers from divisions throughout the state.

This research has been approved by [name of person/committee in the district], and it will not involve any teacher observations, contact with students or parents, or collection of student, school, or division data. It is entirely based on teacher interviews conducted during their off-work hours. I will be sending an invitation to teachers today. There is nothing you need to do for this project, I just wanted to let you know about the study in case you hear word in the hallways about it. If you have any questions, please do not hesitate to contact me.

Kind regards,

Leah Shy

Ph.D. student in Educational Policy, Planning, and Leadership

Adjunct Instructor in Education

William & Mary School of Education

Dissertation conducted under the guidance of Dr. Chris Gareis

APPENDIX F

QUESTIONS FOR QUALTRICS SURVEY

Thank you for your interest in participating in this research study. Please fill out this brief survey of background questions to show your interest. These will be used only for the purpose of selecting interviewees from a range of background experiences. If you are selected for an interview, identifying information will be associated with your interview data, but your participation will remain strictly confidential and any personally identifying information will be removed from any publications. You may skip any question you do not wish to answer, but in order to be entered for the \$50 Amazon gift card draw, you must enter your name and email address.

- What is your preferred name?
- What is your email address?
- At what school do you currently work?
- How would you describe your current school?
Urban, suburban, rural, other
- What grade(s) do you currently teach?
Kindergarten, First grade, Second grade, None of the above
- What other grades have you taught (if any)? (check all that apply)
preK, Kindergarten, 1st grade, 2nd grade, 3rd grade, 4th or 5th grade, middle school, high school, none (all experience in current grade level), other context (please describe)
- How many years of teaching have you completed? (Do not count the current year.)
Less than 4 full years, 4-9 years, 10-17 years, 18 or more years
- Have you had a career in a field other than education?
Yes, No [If yes, in what career field other than education have you worked?]
- What teaching endorsement(s) do you hold? (Select all that apply.)

preK-grade 3, preK-grade 6, special education, other (please list)

- What is your undergraduate major?
- At what college or university did you complete your teacher preparation program?
- Which of the following best describes your student teaching experience?

Assigned to same classroom for student teaching and practica experiences, full-time student teaching in more than one classroom, practicum and student teaching were completed in different classrooms, experienced a variety of classroom placements across practica and student teaching, other (please describe)

- Do you have or have you had children living in your home?

Yes, No

- Is there anything unique about your teaching background that you would like to add?

APPENDIX G

RESPONSE EMAIL TO TEACHERS AFTER SUBMITTING SURVEY

Dear _____,

Thank you for your willingness to participate in my research study. I recognize this would be a sacrifice of your time, and appreciate your desire to further our understanding about issues surrounding assessment in early elementary classrooms. You have now been entered in a drawing for a \$50 Amazon gift card to be held at the end of the research study.

I would like to explain the current stage of my work so you know what to expect. I will be conducting interviews with teachers over the next three to five months. Each interview will be approximately 45 minutes, and some teachers may be asked for a second interview. As my research progresses, I will be selecting teachers from my list of willing participants and contacting them to schedule interviews. As much as possible, I would like to interview in person, though I recognize that an online interview may be necessary, particularly for any follow-up interviews that may be held. All teachers who are interviewed will receive a \$10 gift card to Amazon.

It is possible that I will have more willing participants than I am able to interview; I will stay in regular contact with you regarding timing for scheduling an interview, or in the possible event that your participation is no longer needed, to let you know I have completed the interview stage of the research. You will still be eligible for the \$50 gift card draw.

As a reminder, all interview data will be kept confidential, and your referring administrator will not know if you have chosen to participate or not. If you have any questions, please feel free to contact me at any time. Thank you!

Kind regards,

Leah Shy

Ph.D. Candidate, Education Policy, Planning, and Leadership

William & Mary School of Education

APPENDIX H

TEMPLATE FOR EMAIL TO REQUEST AN INTERVIEW

Dear _____,

Thank you for volunteering to participate in my dissertation research on how K-2 teachers develop their approaches towards classroom assessment. I would like to arrange a time to meet with you for an interview, if it is possible to do in person. As a reminder, all responses provided in the interview as well as the initial interest survey and the status of your participation will be kept strictly confidential. The interview will be recorded in order to make a transcription, but I will be the only one to hear the interview and it will be erased once transcribed.

I will not be observing any teaching practices or meeting with any students, and it is not necessary to interview in your school building. If, however, that is the easiest place for you to meet, that is possible. Otherwise, I can suggest a quiet coffee shop or library near your school, but if there is a place you can recommend for your convenience and timing, please let me know.

Here are possible dates and times that I could meet for an interview; please let me know what would work best for you: [Fill in appropriate dates/times.]

Having taught elementary school for twelve years, I recognize how precious your time is, and in gratitude for your sacrifice, you will receive a \$10 gift card to Amazon.

Thank you again, and I look forward to meeting you soon!

Kind regards,

Leah Shy

Ph.D. Candidate, Education Policy, Planning, and Leadership

William & Mary School of Education

APPENDIX I

EMAIL TO TEACHERS IN DATABASE TO MAINTAIN INTEREST

Dear Teachers,

Thank you for volunteering to be interviewed as part of my dissertation research on how K-2 teachers develop their approaches towards classroom assessment. I am still in the process of conducting interviews, and I may contact you in the near future about scheduling a time to meet. If your interest or availability in the study has changed and you would like to be removed from further consideration, please let me know. As a reminder, you have already been entered into a draw for a \$50 Amazon gift card (to be held at the completion of all interviews), and all teachers who are interviewed also receive a \$10 gift card.

Thanks for your continued interest, and I will be in touch again soon!

Kind regards,

Leah Shy

Ph.D. Candidate, Education Policy, Planning, and Leadership

William & Mary School of Education

APPENDIX J

EMAIL TO ALL TEACHERS WHEN RESEARCH IS DONE

Dear Teachers,

Thank you for volunteering to be interviewed for my research on K-2 teachers' approaches to classroom assessment. I have now completed all of the necessary interviews for my study, and thus I will not be needing to schedule an interview with you. All teachers who filled out the initial background survey were entered in a draw to win a \$50 Amazon gift card. Due to confidentiality procedures, I cannot announce the name of the winner, but the winner has been notified.

I appreciate your willingness to participate in this study. If you have any further questions about this research, please feel free to contact me. Best wishes for the remainder of your school year!

Kind regards,

Leah Shy

Ph.D. Candidate, Education Policy, Planning, and Leadership

William & Mary School of Education

APPENDIX K
INTERVIEW GUIDE

- How would you describe your approach to classroom assessment?
- Tell me about how you developed your approach to classroom assessment? Please include all the influences you can think of as you tell me your story.
- How has your approach to assessment changed over time?
- Tell me about how your approach to classroom assessment is impacted by your current school context?
- How would you describe your philosophical orientation towards teaching young children?
- Describe what “developmentally appropriate practice” means to you?
- In what ways has “developmentally appropriate practice” influenced your teaching?
- How do you see classroom assessment and developmentally appropriate practice interacting with each other?
- Tell me about your background in teaching?

Possible more specific questions:

- Why did you choose to become a teacher?
- How would you describe your current school context, including administration/leadership style?
- Did you want to teach in the early elementary grades? If so, why? If not, how did you come to teach this age level?
- Tell me about your experience in your teacher preparation program.
- How did your teacher preparation program address classroom assessment? (separate course, embedded in methods, through student teaching, etc.)

- If you had the opportunity to learn more about classroom assessment, what topic(s) would you want to address?
- Describe the extent to which you feel you can make your own decisions regarding classroom assessment practices. What factors support or hinder your decision making?
- If you could change or improve your assessment practices in some way, what change would you make?
- Have you received any formal training in DAP? How/Was it addressed in your preparation?
- What kinds of practices do you engage in that fit DAP? Are there practices you engage in that do not fit the DAP framework? Why is this?
- DAP is often considered a continuum. What parts of your teacher preparation program influenced your development along this continuum? What parts of your school experience influenced you?
- How would you describe your school's relationship with DAP?
- What factors in your current or previous school setting have shaped your practice of DAP? What factors have supported it? What factors have limited your practices of DAP?
- Is there anything about DAP that you would like to understand better or receive support to implement?
- Is there a particular philosophical approach to assessing young children that you believe is important or necessary?
- What do you consider to be critical knowledge for teachers assessing young children?
- If you had the ability to change anything in your school's assessment practices, would you change anything? What? Why or why not?
- What challenges exist in assessing children in your classroom?

- Is there anything else I should have asked that would help me understand ways in which K-2 teachers develop and integrate their approaches to classroom assessment and DAP?

APPENDIX L

INFORMED CONSENT

Consent to Participate in a Research Study

Title of study: Navigating conflicting paradigms: How K-2 teachers develop their approaches towards classroom assessment

Investigator: Leah Shy **E-mail:** lkshy@email.wm.edu **Phone:** 757-512-4930

Introduction

You are being asked to be in a qualitative research study of the ways K-2 teachers approach classroom assessment. You were invited as a possible participant because you have at least four years of experience teaching and are currently teaching in a K-2 general education classroom in a Virginia public school. Please read this form and ask any questions that you may have before agreeing to be in the study.

Purpose of Study

The purpose of the study is to develop a theory that explains how K-2 teachers navigate the conflicting demands of the standards-based era and early childhood as it pertains to classroom assessment. This is a dissertation study, and the work may also be presented at professional conferences, published in an academic journal, or included in part of a book.

Description of the Study Procedures

If you agree to be in this study, you will be asked to do one or two interviews of no more than one hour each. You may be asked to share other documents or data, if it becomes relevant to the study. Interviews will be recorded, transcribed, and coded for research purposes.

Confidentiality

Your participation is strictly confidential. The records of this study will also be kept strictly confidential. Research records will be kept in electronic form on a password protected personal computer and in paper form at a locked personal dwelling. Only the researcher will have access to recordings of interviews. Once transcribed, all recordings will be electronically erased. Your identity will not be disclosed in the transcriptions or in any material that is published, and no information will be included in any report that would make it possible to identify you. A pseudonym will be used in interviews, transcriptions, and published material.

Incentive for Participation in the Study

All participants selected for interviews will receive a \$10 gift card from their choice of Starbucks, Amazon, or iTunes as an incentive and show of appreciation for their time.

Risks of Participation in the Study

There are no reasonable, foreseeable (or expected) risks involved. There may be unknown risks.

Duration of Participation

Interviews for the study will be conducted over several months, but any one participant will only be interviewed up to twice.

Right to Refuse or Withdraw

The decision to participate in this study is entirely up to you. You may decline to take part in the study at any time. You have the right not to answer any single question, as well as to withdraw completely at any point during the process.

Right to Ask Questions and Report Concerns

You have the right to have any questions about the study answered by the investigator before, during, or after the research, at lkshy@email.wm.edu. To report any concerns about the study, contact Professor Tom Ward, chair of the School of Education Internal Review Committee, at tward@wm.edu.

Consent

Your signature below indicates that you are at least 18 years of age, that you have decided to volunteer as a research participant for this study, and that you have read and understood the information provided above. You will be given a signed and dated copy of this form to keep.

Participant's name (print) _____

Participant's signature _____ **Date** _____

Investigator's signature _____ **Date** _____

REFERENCES

- Aldiabat, K. M., & Le Navenec, C. (2018). Data saturation: The mysterious step in grounded theory method. *The Qualitative Report*, 23(1), 245-261.
- American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students*. Retrieved from <https://buros.org/standards-teacher-competence-educational-assessment-students>
- Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teacher beliefs* (pp. 284–300). New York, NY: Routledge.
- Barnes, N., Fives, H., & Dacey, C. M. (2017). U.S. teachers' conceptions of the purpose of assessment. *Teaching and Teacher Education*, 65, 107-116.
- Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade? *AERA Open*, 1(4), 1-31.
- Birks, M., & Mills, J. (2011). *Grounded theory: A practical guide*. Los Angeles, CA: Sage.
- Black, P. (2013). Formative and summative aspects of assessment: Theoretical and research foundations in the context of pedagogy. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 167-178). Thousand Oaks, CA: Sage.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004, September). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan* 86(1), 9-21.

- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- Black, P., & Wiliam, D. (2007). Large-scale assessment systems: Design principles drawn from international comparisons. *Measurement*, 5(1), 1–53.
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principals, Policy & Practice*, 25(6), 551-575.
- Bodrova, E., & Leong, D. J. (2018). Common assessment terms and how to use them: A glossary for early childhood educators. In H. Bohart & R. Procopio (Eds.), *Spotlight on young children: Observation and assessment* (pp. 15-20). Washington, DC: NAEYC.
- Bornfreund, L. (2012). Preparing teachers for the early grades. *Educational Leadership*, 69(8), 36-40.
- Boudett, K. P., & Moody, L. (2005). Organizing for collaborative work. In K. P. Boudett, E. A. City, & R. J. Murname (Eds.), *Datawise: A step by step guide to using assessment results to improve teaching and learning* (pp. 1-28). Cambridge, MA: Harvard Education Press.
- Brookhart, S. M. (2001). *The Standards and classroom assessment research*. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education, Dallas, TX. (ERIC Document Reproduction Service No. ED451189)
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12.
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria, VA: ASCD.

- Brookhart, S. M. (2015). *Performance assessment: Showing what students know and can do*. West Palm Beach, FL: Learning Sciences International.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803-848.
- Buchanan, T. K., Burts, D. C., Bidner, J., & White, V. F. (1998). Predictors of the developmental appropriateness of the beliefs and practices of first, second, and third grade teachers. *Early Childhood Research Quarterly*, 13(3), 459-483.
- Campbell, C. M., & Evans, J. (2000). Investigation of preservice teachers' classroom assessment practices during student teaching. *Journal of Educational Research*, 93(6), 350–355.
- Campbell, C. M., Murphy, J. A., & Holt, J. K. (2002, October). *Psychometric analysis of an assessment literacy instrument: Applicability to preservice teachers*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, OH.
- Chappuis, J. (2014). Thoughtful assessment with the learner in mind. *Educational Leadership*, 71(6), 20-26.
- Chappuis, S., Chappuis, J., & Stiggins, R. (2009). The quest for quality. *Educational Leadership*, 67(3), 14-19.
- Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.
- Cohen, D. H., Stern, V., Balaban, N., & Gropper, N. (2016). *Observing and recording the behavior of young children* (6th ed.). New York, NY: Teachers College Press.

- Coker, D. L., & Ritchey, K. D. (2015). *Teaching beginning writers*. New York, NY: Guilford.
- Coombs, A., DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education, 71*, 134-144.
- Cooper, P. M. (2009). *The classrooms all young children need: Lessons in teaching from Vivian Paley*. Chicago, IL: University of Chicago Press.
- Copple, C., & Bredekamp, S. (Eds.). (2009). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8* (3rd ed.). Washington, DC: NAEYC.
- Corbin, J., & Strauss, A. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Thousand Oaks, CA: Sage.
- Costanza, V. J., & Falkenstein, R. (2018). Combating assessment fatigue in K-3 classrooms. In H. Bohart & R. Procopio (Eds.), *Spotlight on young children: Observing and assessing* (pp. 98-106). Washington, DC: NAEYC.
- Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches* (3rd ed.). Los Angeles, CA: Sage.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles, CA: Sage.
- Crum, K. S., Sherman, W. H., & Myran, S. (2009). Best practices of successful elementary school leaders. *Journal of Educational Administration, 48*(1), 48-63.
- Datnow, A., & Hubbard, L. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of Educational Change, 17*, 7-28.

- 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.
- DeLuca, C., Chavez, T., Bellara, A., & Cao, C. (2013). Pedagogies for preservice assessment education: Supporting teacher candidates' assessment literacy development. *The Teacher Educator, 48*(2), 128-142.
- DeLuca, C., & Hughes, S. (2014). Assessment in early primary education: An empirical study of five school contexts. *Journal of Research in Childhood Education, 28*(4), 441-460.
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice, 17*(4), 419-438.
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Approaches to classroom assessment inventory: A new instrument to support teacher assessment literacy. *Educational Assessment, 21*(4), 248-266.
- DeLuca, C., Ogden, H., & Pero, E. (2015). Reconceptualizing elementary preservice teacher education: Examining an integrated-curriculum approach. *The New Educator, 11*(3), 227-250.
- DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, & Luhanga, U. (2016). Teachers' approaches to classroom assessment: A large-scale survey. *Assessment in Education: Principles, Policy, & Practice, 1-21*.
- Deneen, C., & Boud, D. (2014). Patterns of resistance in managing assessment change. *Assessment & Evaluation in Higher Education, 39*, 577-591.

- Deneen, C. C., & Brown, G. T. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education*, 3(1), 1-14.
- Dorel, T. G., Kearney, W. S., & Garza, E. (2016). Ready from day one? The relationship between length of pre-service teacher field residency and teacher efficacy. *Critical Questions in Education*, 7(1), 38-52.
- Duncan, A. (2013, December 3). *The threat of educational stagnation and complacency*. Retrieved from <https://www.ed.gov/news/speeches/threat-educational-stagnation-and-complacency>
- Earl, L. (2003). *Assessment as learning*. Thousand Oaks, CA: Corwin.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson, and Company.
- File, N., & Gullo, D. F. (2002). A comparison of early childhood and elementary education students' beliefs about primary classroom teaching practices. *Early Childhood Research Quarterly*, 17(1), pp. 126-137.
- Fives, H., Barnes, N., Dacey, C., & Gillis, A. (2016). Assessing assessment texts: Where is planning? *The Teacher Educator*, 51(1), 70-89.
- Fusch, P I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report*, 20(9), 1408-1416.
- Gallant, P. A. (2009). Kindergarten teachers speak out: "Too much, too soon, too fast!" *Reading Horizons*, 49(3), 201-220.
- Gareis, C. R., & Grant, L. W. (2015). *Teacher-made assessments: How to connect curriculum, instruction, and student learning* (2nd ed.). New York, NY: Routledge.

- Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In Clifford Geertz (Ed.), *The interpretation of culture*. New York, NY: Basic Books.
- Gibbs, G. R. (2015, February 4). *A discussion with Professor Kathy Charmaz on grounded theory* [YouTube video]. Retrieved from <https://www.youtube.com/watch?v=D5AHmHQS6WQ>
- Glaser, B. G. (1978) *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.
- Glaser, B. G. (2005). Basic social processes. *Grounded Theory Review: An International Journal*, 4(3). Retrieved from <http://groundedtheoryreview.com/2005/06/22/1533/>
- Glaser, B. G. (2009). The novice GT researcher. *Grounded Theory Review*, 2, 8. Retrieved from <http://groundedtheoryreview.com/2009/06/30/914/>
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago, IL: Aldine.
- Goldstein, L. (2007). Beyond the DAP versus standards dilemma: Examining the unforgiving complexity of kindergarten teaching in the United States. *Early Childhood Research Quarterly*, 22(1), 39-54.
- Graue, M. E. (2009). Reimagining kindergarten: Restoring a developmental approach when accountability demands are pushing formal instruction on the youngest learners. *School Administrator*, 10(66), 10-15.
- Groen, M. (2012). NCLB-the educational accountability paradigm in historical perspective. *American Educational History Journal*, 39(1), 1-14.
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology Journal*, 29, 75-92.

- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
- Gullo, D. F., & Hughes, K. (2011a). Reclaiming kindergarten: Part I. Questions about theory and practice. *Early Childhood Education Journal*, 38(5), 323–328.
- Gullo, D. F., & Hughes, K. (2011b). Reclaiming kindergarten: Part II. Questions about policy. *Early Childhood Education Journal*, 38(6), 393-397.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough?: An experiment with data saturation and variability. *Field Methods*, 18(1), 59-82.
- Guskey, T. R. (2014, May). Planning professional learning. *Educational Leadership*, 71(8), 10-16.
- Guskey, T. R., & Jung, L. A. (2013). *Answer to essential questions about standards, assessments, grading, and reporting*. Thousand Oaks, CA: Corwin.
- Hatch, J. A. (2002). Accountability shovedown: Resisting the standards movement in early childhood education. *Phi Delta Kappa*, 83(6), 457-462.
- Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code saturation versus meaning saturation: How many interviews are enough? *Qualitative Health Research*, 27(4), 591-608.
- Huebner, T. A. (2009). Balanced assessment. *Educational Leadership*, 67(3), 85-86.
- Jeynes, W. H. (2006). Standardized tests and Froebel's original kindergarten model. *Teachers College Record*, 108(10), 1937–1959.
- Klinger, D., McDivitt, P., Howard, B., Rogers, T., Muñoz, M. & Wylie, C. (2015). *The classroom assessment standards for preK-12 teachers* [E-book]. Joint Committee on Standards for Educational Evaluation.

- Krechevsky, M., Mardell, B., Rivard, M., & Wilson, D. (2013). *Visible learners: Promoting Reggio-inspired approaches in all schools*. San Francisco, CA: Jossey-Bass.
- Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership and Management, 28*(1), 27-42.
- Levin, J. A., & Datnow, A. (2012). The principal role in data-driven decision making: Using case study data to develop multi-mediator models of educational reform. *School Effectiveness and School Improvement, 23*(2), 179-201.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.
- Lincoln, Y. S., & Guba, E. G. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. In D. D. Williams (Ed.), *Naturalistic evaluation*. San Francisco, CA: Jossey-Bass.
- Lincoln, Y. S., & Guba, E. G. (1989). Ethics: The failure of positivist science. *Review of Higher Education, 12*(3), 221-241.
- Louis, K. S., & Robinson, V. M. (2012). External mandates and instructional leadership: School leaders as mediating agents. *Journal of Educational Administration, 50*(5), 629-665.
- Louis, K. S., & Wahlstrom, K. (2011). Principals as cultural leaders. *Phi Delta Kappan, 92*(5), 52-56.
- Lunenburg, F. C. (2013). The principal as an instructional leader. *National Forum of Educational Administration and Supervision Journal, 30*(2), 30-40.
- MacLellan, E. (2004). Initial knowledge states about assessment: Novice teachers' conceptualizations. *Teaching and Teacher Education, 20*(5), 523-535.

- Manna, P. (2011). *Collision course: Federal education policy meets state and local realities*. Washington, DC: CQ Press.
- Marzano, R. J. (2005). *School leadership that works: From research to results*. Alexandria, VA: ASCD.
- Marzano, R. J. (2006). *Classroom assessment and grading that work*. Alexandria, VA: ASCD.
- Marzano, R. J. (2010). *Formative assessment and standards-based grading: Classroom strategies that work*. Bloomington, IN: Marzano.
- Mason, M. (2010). Sample size and saturation in PhD studies using qualitative interviews. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 11(3), Art. 8. Retrieved from <http://nbn-resolving.de/urn:nbn:de:0114-fqs100387>
- Maxwell, K. L., McWilliam, R. A., Hemmeter, M. L., Ault, M. J., & Schuster, J. W. (2001). Predictors of developmentally appropriate classroom practices in kindergarten through third grade. *Early Childhood Research Quarterly*, 16(4), 431-452.
- McAfee, O., Leong, D. J., & Bodrova, E. (2004). *Basics of assessment: A primer for early childhood educators*. Washington, DC: NAEYC.
- McDaniels, G., Issac, M., & Hatch, J. A. (2005). Confronting K-3 teaching challenges in an era of accountability. *Young Children*, 60(2), 20-26.
- McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, 22(4), 34-43.

- McMillan, J. H. (Ed.). (2007). *Formative classroom assessment: Theory into practice*. New York, NY: Teachers College Press.
- McMillan, J. H. (Ed.). (2013). *SAGE handbook of research on classroom assessment*. Thousand Oaks, CA: Sage.
- McMillan, J. H., Myron, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research*, 95, 203-213.
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). San Francisco, CA: Jossey-Bass.
- Mertler, C. A. (2003). *Classroom assessment: A practical guide for educators*. Los Angeles, CA: Pyrczak.
- Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, 33(1), 49–64.
- Mertler, C. A., & Campbell, C. S. (2005, April). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the Assessment Literacy Inventory*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, QC.
- Michnick-Golinkoff, R., Hirsh-Pasek, K., & Singer, D. G. (2006). Why play=learning: A challenge for parents and educators. In D. G. Singer, R. Michnick Golinkoff, & K. Hirsh-Pasek (Eds.), *Play=learning: How play motivates and enhances children's cognitive and social-emotional growth* (pp. 3–12). New York, NY: Oxford University Press.
- Miller, E., & Almon, J. (2009). *Crisis in the kindergarten: Why children need to play in school*. College Park, MD: Alliance for Childhood.

- Minicozzi, L. L. (2016). The garden is thorny: Teaching kindergarten in the age of accountability. *Global Studies of Childhood, 6*(3), 299-310.
- Moss, C. M., & Brookhart, S. M. (2009). *Advancing formative assessment in every classroom*. Alexandria, VA: ASCD.
- Muñoz, M. A., & Guskey, T. R. (2015). Standards-based grading and reporting will improve education. *Phi Delta Kappan, 96*(7), 64-68.
- National Association for the Education of Young Children. (2003). *Early childhood curriculum, assessment, and program evaluation* [Position statement]. Washington, DC: Author.
- National Association for the Education of Young Children. (2009). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8* [Position statement]. Washington, DC: Author.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Author.
- Neill, M. (2006). The case against high-stakes testing. *Principal, 85*(4), 28-32.
- Neill, M., & Medina, N. J. (1989). Standardized testing: Harmful to educational health. *Phi Delta Kappan, 70*(9), 688-697.
- Nicolopoulou, A. (2010). The alarming disappearance of play from early childhood education. *Human Development, 53*(1), 1-4.
- O'Connor, K. (2007). The last frontier: Tackling the grading dilemma. In D. Reeves (Ed.), *Ahead of the curve: The power of assessment to transform teaching and learning* (pp. 126-145). Bloomington, IN: Solution Tree Press.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332.

- Parker, A., & Neuharth-Pritchett, S. (2006). Developmentally appropriate practice in kindergarten: Factors affecting teacher beliefs and practices. *Journal of Research in Childhood Education, 21*(1), 65-78.
- Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Popham, W. J. (2004). *America's failing schools: How parents and teachers can cope with No Child Left Behind*. New York, NY: Routledge.
- Popham, W. J. (2006). Phony assessments: Buyer beware. *Educational Leadership, 64*(3), 86-87.
- Popham, W. J. (2008). *Transformative assessment*. Alexandria, VA: ASCD.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice, 48*, 4-11.
- Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator, 46*(4), 265-273.
- Pyle, A., & DeLuca, C. (2013). Assessment in the kindergarten classroom: An empirical study of teachers' assessment approaches. *Early Childhood Education Journal, 41*(5), 373-380.
- Pyle, A., & DeLuca, C. (2017). Assessment in play-based kindergarten classrooms: An empirical study of teacher perspectives and practices. *The Journal of Educational Research, 110*(5), 457-466.
- Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York, NY: Basic Books.
- Ravitch, D. (2013). *Reign of error: The hoax of the privatization movement and the danger to America's public schools*. New York, NY: Random House.

- Rubin, H., & Rubin, I. (2012). *The art of hearing data* (3rd ed.). Thousand Oaks, CA: Sage.
- Saldaña, J. (2011). *Fundamentals of qualitative research*. Oxford, UK: Oxford University Press.
- Saldaña, J. (2016). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage.
- Salmons, J. (2014). *Qualitative online interview: Strategies, designs, and skills* (2nd ed.). Thousand Oaks, CA: Sage.
- Schmoker, M. (2006). *Results now: How we can achieve unprecedented results in teaching and learning*. Alexandria, VA: ASCD.
- Shepard, L. A. (1994). The challenges of assessing young children appropriately. *Phi Delta Kappan*, 76(3), 206-12.
- Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Shy, L. K. (2018, October). *Developmentally appropriate practice in classroom assessment textbooks: A content analysis*. Paper presented at the meeting of CREATE Conference, Williamsburg, VA.
- Stiggins, R. J. (1999). Are you assessment literate? *The High School Journal*, 6(5), 20-23.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86(1), 22-27.
- Stiggins, R. J. (2014). Improve assessment literacy outside of schools too. *Phi Delta Kappan*, 96(2), 67-72.
- Stiggins, R., & Duke, D. (2008). Effective instructional leadership requires assessment leadership. *Phi Delta Kappan*, 90(4), 285-291.

- Stipek, D. (2004). Teaching practices in kindergarten and first grade: Different strokes for different folks. *Early Childhood Research Quarterly, 19*(4), 548-568.
- Strauss, A. (1987). *Qualitative analysis for social scientists*. Cambridge, England: Cambridge University Press.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). Newbury Park, CA: Sage.
- Tschannen-Moran, M. (2004). *Trust matters: Leadership for successful schools*. San Francisco, CA: Jossey-Bass.
- Van Horn, M. L., & Ramey, S. L. (2004). A new measure for assessing developmentally appropriate practices in early elementary school, a developmentally appropriate practice template. *Early Childhood Research Quarterly, 19*(4), 569-587.
- Vanlommel, K., Van Gasse, R. V., Vanhoof, J., & Van Petegem, P. (2018). Teachers' high-stakes decision making: How teaching approaches affect rational and intuitive data collection. *Teaching and Teacher Education, 71*, 108-119.
- Vartuli, S. (1999). How early childhood teacher beliefs vary across grade level. *Early Childhood Research Quarterly, 14*(4), 489-514.
- Vartuli, S. (2005). Beliefs, the heart of teaching. *Young Children, 60*(5), 76-85.
- Volante, L. (2010). Assessment of, for, and as learning within schools: Implications for transforming classroom practice. *Action in Teacher Education, 31*(4), 66-75.
- Welsh, M. E., D'Agostino, J. V., & Kaniskan, R. (2013). Grading as a reform effort: Do standards-based grades converge with test scores? *Educational Measurement: Issues and Practice, 32*(2), 26-36.
- Zhao, Y. (2009). *Catching up or leading the way: American education in the age of globalization*. Alexandria, VA: ASCD.

Zhao, Y. (2017). What works may hurt: Side effects in education. *Journal of Educational Change*, 18, 1-19.

VITA

Leah Kathleen Shy

- Education:** 2014-2019 Doctor of Philosophy
Educational Policy, Planning, and Leadership
William & Mary School of Education
- 2009-2011 Master of Arts
Organization and Leadership
University of San Francisco
- 1995-1999 Bachelor of Arts
Humanities
Pepperdine University
- Experience:** 2015-2018 Adjunct Instructor in Education
Elementary Reading and Language Arts Methods
Language Development and English Language
Learners
Supervised Teaching in Elementary Education
Elementary Science Methods
William & Mary School of Education
- 2015-2018 Director/Lead Instructor of the Graduate Business English
Language and Culture Program
William & Mary Reves Center for International Studies
- 2014-2018 Graduate Assistant
William & Mary School of Education
- 2013-2014 Educational Consultant
Hong Kong International School
- 2008-2013 Classroom Teacher, 2nd grade
Hong Kong International School
- 2003-2008 Classroom Teacher, Year 2 (British system)
St. Michael's International School
Kobe, Japan
- 2001-2003 English Language Teacher/Christian Missionary
Osaka, Japan
- 1999-2001 Classroom Teacher, 1st and 2nd grades
Juan Cabrillo Elementary School
Malibu, CA