Syracuse University

SURFACE

Dissertations - ALL

May 2016

Assessing Formative Assessment: An Examination of Secondary English/Language Arts Teachers' Practices

Leigh M. Tolley
Syracuse University

Recommended Citation

Abstract

This study examined how "formative assessment" (FA), a term that encompasses all of the assessment and instructional activities undertaken by teachers and their pupils to improve student learning, occurs in a secondary English/language arts (ELA) context. A sequential explanatory mixed methods approach, using a series of weekly self-report checklists about teachers' assessment practices and semi-structured interviews, was used to address the following: (a) what assessment practices secondary ELA teachers use; (b) how these teachers determine what assessment practices to use and when to use them; (c) what these teachers' perceptions of the effectiveness of the assessment practices that they use are; and (d) how teachers determine the effectiveness of their assessment practices. For the purposes of this research, there was a focus on three major underlying elements of FA: questioning, discussion, and feedback. This study was conducted with seven experienced high school ELA teachers that were instructors in a concurrent enrollment program. The results suggested that the overall use of and thinking about assessment practices in a secondary ELA context aligns with prior FA research in other content areas, although the practices reported as used in this study seemed to be more targeted to individual students and their learning needs. In addition, the findings showed that there were statistically significant differences in the practices that the participants used from week to week with their students. There were also statistically significant differences in teachers' perceptions of the effectiveness of the assessment practices that they reported using with their lower-, average-, and higher-performing students. These results suggest that the instruments and methods used for this study are a viable approach to examining FA in a secondary ELA context, but additional research is needed into the reliability and validity of these approaches with other populations.

   Keywords: formative assessment, assessment for learning, mixed methods, English/language arts, secondary education, classroom assessment, educational evaluation

ASSESSING FORMATIVE ASSESSMENT: AN EXAMINATION OF
SECONDARY ENGLISH/LANGUAGE ARTS TEACHERS' PRACTICES

by

Leigh M. Tolley

B.A., University of Pennsylvania, 2000
M.S.Ed., University of Pennsylvania, 2002
M.S., Syracuse University, 2012

Dissertation

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
Instructional Design, Development & Evaluation in the Graduate School of Syracuse University.

Syracuse University
May 2016

Acknowledgements

There have been many, many people that have supported me through this journey, helped guide me in the right direction, and enabled me to learn through myriad life experiences. I especially would like to extend my thanks to the following:

- Dr. Karen Bull, Dr. Tiffany M. Squires, and Dr. Micah Shippee, for being a safety net, resource guide, and listening ear whenever I needed you. We got that show on the road; here's to wherever our next stops on this tour will be.

- Dr. Tess Dussling, for always being willing to listen, read, and give feedback on a topic that's related to hers, but not quite.

- Dr. Glenn Wright, Dr. Peg DeFuria, Shawn Loner, Lee Pomeroy, Dina Ioannidis, and fellow Teaching Mentors, for helping make graduate school a really fun place to be.

- Fellow Dissertation Boot Campers, especially Jerry Robinson, Kathleen Joyce, and Melissa Welshans: your encouragement kept me going, and our shared goals kept me writing.

- My friends in the American Evaluation Association, especially the PreK-12 Educational Evaluation Topical Interest Group, for nurturing my love for the field of evaluation.

- Mr. Ron LaMar, whose love of teaching inspired me (and countless others) to be an educator, and more importantly, a good person.

- Mrs. Katie Hamilton, whose mentorship and guidance fostered my development as a young teacher, and whose nurturing spirit continues to influence me, no matter what, when, or where I am teaching.

- My parents, Jim and Joan Gardner, for valuing education so highly, and for teaching me so much more than I ever would have learned in a classroom.

- My family, friends, classmates, and colleagues, who have supported me through so many different stages of my life, and always made sure I came out for the better on the other side.

- Last, but certainly not least, my heartfelt thanks go to my husband, Tom Tolley, whose encouragement, support, and ability to see my potential enabled me to reach my dreams and beyond. *Grá mo chroí*, here is to wisdom and to love.

Table of Contents

List of Illustrative Materials

Chapter 1: Introduction

In 2007, I was in my fifth year of teaching high school. As a secondary English/language arts (ELA) teacher, I worked daily with over 100 students in grades 10-12 to develop their literacy skills through reading, discussing, and writing about literature. My students ranged from those with Individualized Education Programs (IEPs) and general-level learners in inclusive sections through pupils at the honors levels, including those taking Advanced Placement English Literature and Composition with me. Each year, I was able to see my students' growth and development as they read and discussed authors from Chinua Achebe to William Shakespeare. Their writing portfolios, filed away in a cabinet in my classroom, told a tale of their development and the individualized progress that had occurred. I would have been able to tell others, including other teachers, administrators, and a student's parents, about where each of my pupils currently stood in terms of their literacy, where they had been at the beginning of the year, strategies that had helped them improve, and where I hoped they would be by June. However, this knowledge did not seem to be what was important to the school district in which I taught.

Due in large part to accountability efforts to address requirements of the No Child Left Behind Act, district concerns shifted to students' performance on state-wide standardized tests and making Adequate Yearly Progress. Our students, although having made great strides individually, fell behind when compared to other students across the state in terms of proficiency on reading and math assessments. As a result, different literacy and math intervention programs were implemented each year; however, if test scores were not immediately increased, a program was cast aside for a new one. As a teacher, I became frustrated with having to focus on only one aspect of student performance, and not being able to take advantage of the knowledge I had of what was most effective in my own classroom. I chose to leave my teaching position, and within

two years, I had returned to graduate school, where I began learning about instructional design and educational evaluation. However, I kept revisiting why I had left high school teaching and thinking about ways to address the issues that my fellow teachers and I had faced. Ultimately, this developed into an interest in better understanding teachers' practices to improve student learning, as well as what they feel works and what does not in the context of their own classrooms. The study presented here evolved from that interest, and is based on and contributes to research on teachers' formative assessment practices.

Statement of Problem

The contemporary concept of accountability for teachers and their students came to the forefront of grades K-12 education with the introduction of the No Child Left Behind (NCLB) Act in 2001 in the United States (U.S. Department of Education, 2011a). Interest in educational accountability continues to grow with initiatives such as Race to the Top and the Every Student Succeeds Act in the U.S. (Executive Office of the President, 2015; U.S. Department of Education, 2011b, 2016) and similar programs worldwide (e.g., Black & Wiliam, 1998b, 2005; Crossouard, 2011). In New York State, this has developed into required annual professional performance reviews, or APPRs, which rate the effectiveness of teachers and principals (New York State Education Department, 2015). Although the aim of the APPR evaluation system is to "foster a culture of continuous professional growth for educators to grow and improve their instructional practices" (p. 7), concerns have been raised about the reliability and validity of these evaluations. Additionally, teachers' unions have stressed the need to involve teachers "as participants in, not recipients of, their own evaluations" and to use teacher knowledge and experience to drive these evaluation systems (Neira, 2015).

    With the onus on teachers' effectiveness and their potential to improve student learning,

there is a need to better understand the instructional and assessment practices that occur within

K-12 classrooms. Of the myriad approaches that have emerged in response to the current demand

for accountability, the set of practices collectively termed as "formative assessment," also

commonly referred to internationally as "assessment *for* learning" (as opposed to "assessment *of*

learning"), seems quite worthy of further study. Black and Wiliam (1998a) published a landmark

review of over 300 studies on formative assessment (FA), indicating relatively high effect sizes

(0.4 to 0.7) for its impact on student achievement. As a result of this work, FA strategies and

approaches were strongly advocated for use in educational interventions, including teacher

professional development initiatives and changes in classroom instructional practice, as a means

of addressing the gap between students' current performance and the performance required to

attain accountability standards. In recent years, the actual effect of FA has been further explored

(Bennett, 2011; Briggs, Ruiz-Primo, Furtak, Shepard, & Yin, 2012; Dunn & Mulvenon, 2009;

Kingston & Nash, 2011; McMillan, Venable, & Varier, 2013), but researchers also agree that

more study into how FA is conceptualized and how its strategies actually occur in varying

classroom contexts is warranted (Filsecker & Kerres, 2012; McMillan et al., 2013; Wiliam,

2011).

    One of the most commonly cited definitions of FA is that it is a term "encompassing all

those activities undertaken by teachers, and/or by their students, which provide information to be

used as feedback to modify the teaching and learning activities in which they are engaged"

(Black & Wiliam, 1998a, pp. 7-8). There is no single, unified definition of what FA is and how it

occurs. The instructional and assessment practices that fall under the "formative assessment"

umbrella vary depending on the context in which they take place, and may be impacted by

factors including the grade level of the students and the subject matter being taught (Hodgen & Marshall, 2005; McMillan, 2010).

FA in grades K-12 education is most often studied in the context of science and mathematics, content areas where concrete answers and methods are required as part of the instructional process. In these subjects, there is typically a specific predetermined sequence of what is learned and how it is taught (Marshall, 2007), which subsequently affects FA implementation "because feedback and instructional correctives are relatively easy to conduct for most students" (McMillan, 2010, p. 54). In contrast, there is much less empirical research into the use of FA practices in an ELA context, where questioning and feedback (and the resulting discussions) are used by teachers "in a dynamic way to enhance students' thinking and deep understanding," are initiated "based on student contributions and questions, which are often impromptu, varied, and unpredictable," and whose "outcomes are more challenging to assess" (p. 54). As a whole, the use of assessments for formative purposes in ELA allows for student engagement in authentic and purposeful reading and writing activities, provides information that teachers can use to organize, plan, and evaluate their instruction with regard to their individual students' needs, and involves students in the assessment process in order to enhance their learning (Murphy, 1997). At the secondary level, students are more capable of self-assessment and self-evaluation, and are able to play a role in selecting their own instructional correctives when needed. Therefore, it is likely that high school ELA teachers are very involved in FA practices while on the path to developing students as independent thinkers and learners.

As noted above, the label of "formative assessment" varies on what it encompasses, depending on the context in which FA occurs. It is important to focus not on the term—as any other buzzword may smell just as sweet—but on the underlying practices which it is attempting

to describe. Also important is the need to study these practices in the context in which they take place; as with studies of other educational interventions, variations may occur depending on when, where, how, and why certain strategies are used (Lendrum & Humphrey, 2012). As there is currently a paucity of empirical research into FA practices in a secondary ELA context, this study aims to address that need. In part, this was done by creating a form of standardized data collection, based on strategies that have been studied, in order to help to create a more linear connection between theory, research, classroom practices, and student outcomes (Lawrence, Rabinowitz, & Perna, 2009).

Development of This Study

This dissertation study was developed through a process of reviewing literature, conducting preliminary research, and revising and refining ideas based on these findings. To attempt to operationalize the construct of "formative assessment," I conducted a literature review of 25 studies involving professional development for teachers in grades K-12 on FA (Tolley, 2012). Although definitions and implementations of FA varied, common elements included (a) teachers setting and making explicit clear learning goals; (b) the use of questioning and feedback from teachers and students; (c) data collection, analysis, and use; (d) teacher revision/modification of instruction; (e) the development of student academic autonomy; and (f) evaluation to determine if a gap between the learning goals and current student performance exists.

These elements arose again from the results of a preliminary fieldwork study that I conducted in the spring of 2014 with administrators and teachers of the ELA branch of the

Student Preparation and Readiness through University Concurrent Enrollment (SPRUCE)[1] program. SPRUCE is a long-standing program that partners secondary schools with a private university in the northeastern United States. The Instructors are teachers in the schools that are trained by the program, and high school students can earn college credits through successful completion of SPRUCE coursework. The aim of this initial qualitative study was to understand how SPRUCE ELA Program Instructors talked about their practices to improve student learning, and this was addressed through a focus group with three administrators and semi-structured individual interviews with five teachers. The overall themes that emerged included the identification of specific instructional and assessment practices that these secondary ELA teachers used, each of which addressed one of the six elements described above in the literature on professional development about FA. This connection between research and the classroom prompted me to further study the actual practices that teachers use with their students. Other themes that emerged from the preliminary fieldwork included how these practices were learned by the teachers, how the participants defined "formative assessment," and contextual factors that impacted how efforts to improve student learning occurred. A notable finding from this research was that the term "formative assessment" was defined differently by each participant. As a result, that specific term is not used in the instrumentation for my dissertation study; instead, the decision was made to focus on underlying elements that compose FA.

In searching for a theoretical framework to further guide research on teachers' practices that are related to FA, I examined Danielson's *Framework for Teaching* (2007, 2014), a research-based teacher evaluation tool for the purposes of improving pedagogical practice

---

[1] All program names, course titles and numbers, and individual participants' names in this dissertation have been changed to protect confidentiality.

(Danielson & McGreal, 2000) that is used in over a dozen states, including New York. The *Framework for Teaching* comprises four domains: planning and preparation, the classroom environment, instruction, and professional responsibilities. The penultimate domain, instruction, consists of five major components: (a) communicating with students, (b) using questioning and discussion techniques, (c) engaging students in learning, (d) using assessment in instruction, and (e) demonstrating flexibility and responsiveness. Although the entire instruction domain—and indeed, the majority of the framework—is well-suited for the study of FA-related practices, even this became too cumbersome for focused research. Ultimately, the decision was made to focus on assessment practices only, rather than instructional and assessment practices. This was further refined to a focus on questioning, discussion, and feedback; more about this selection is explained in the following chapters.

My frustrations that I experienced as a secondary ELA instructor about the lack of use of teachers' knowledge about effective practices to improve student learning led me to this dissertation research. I wanted to investigate the assessment practices used by secondary ELA teachers, how effective the teachers believed these practices were, and how they made these decisions. Due to the limited availability of literature addressing these questions in this particular content area and grade level, I chose to conduct this study with a select group of teachers—in particular, SPRUCE ELA Program Instructors. Through working with these expert teachers, my goal was to determine if the methods designed for this research would work as a means of better understanding secondary ELA teachers' assessment practices; if this approach was feasible with these teachers, then I would be able to apply these methods in future research with other ELA instructors that were teaching in more varied courses and contexts.

Research Questions

This study addressed the following four research questions:

1. What assessment practices do secondary ELA teachers use?

2. How do secondary ELA teachers determine what assessment practices to use and when to use them?

3. What are secondary ELA teachers' perceptions of the effectiveness of the assessment practices that they use?

4. How do secondary ELA teachers determine the effectiveness of their assessment practices?

Study Overview

This study used a sequential explanatory mixed methods design, where quantitative data are collected and analyzed, and then qualitative data are collected and analyzed (Creswell & Plano Clark, 2011; Ivankova, Creswell, & Stick, 2006). The selected population for this research was SPRUCE ELA Program Instructors who were teaching ENG 135, *Critical Reading: Issues, Theory, and Applications*, a semester-long course, in Spring 2015. This population was selected for study because SPRUCE ELA Program Instructors teach college-level courses in high schools, have applied for and hold specific credentials to be a SPRUCE Instructor, and are recognized as adjunct faculty by the SPRUCE program's partner university. The training and preparations that these individuals undergo, the context in which they teach (especially with regard to grade level and content area), their level of teaching experience, and their frequent interactions with SPRUCE staff about their instructional and assessment practices are all important qualities of this population. It was anticipated that these characteristics allowed for SPRUCE ELA Program Instructors to be more likely to be able to articulate their experiences

and perceptions of what they consider to be effective assessment than traditional classroom teachers, especially when talking about practices that can be considered as "formative assessment."

After consenting to participate in the study, SPRUCE ELA Program Instructor participants completed a brief background questionnaire to allow me to learn about them and their teaching experiences. Next, participants were asked to complete a checklist of assessment practices relating to questioning, discussion, and feedback based on their teaching of ENG 135 over the past week, and to repeat the process for the following three weeks. This administration of weekly checklists was an experience sampling method (e.g., Hektner, Schmidt, & Csikszentmihalyi, 2007; Zirkel, Garcia, & Murphy, 2015), which allows for an understanding of individuals' experiences in context as they unfold. Last, participants were asked to participate in an individual semi-structured interview (e.g., Rubin & Rubin, 2005) about their self-reported assessment practices, their reasoning behind making the decisions to use certain practices or not, their perceptions of the effectiveness of the practices they used, and how they made decisions about the effectiveness of assessment practices. Further details about the methods used and the individuals who participated in the study are presented in Chapters 3 and 4.

Broader Significance of This Study

In planning educational interventions, including teacher professional development, it is necessary to know where teachers currently are in their practice and what strategies they value. This will allow for a more effective approach to close any gaps in teachers' pedagogical knowledge, which in turn can close gaps in student knowledge. A better understanding of teachers' actual practices additionally allows for more precise tailoring of interventions to best suit both teachers' and their students' needs.

The chosen population for this study is in a range of geographic locations across several states, with varying levels of urbanicity and demographic representation of students within each state, depending on the school or school district where the teachers are employed. These contextual variations have the potential for allowing understanding of teachers' use of assessment practices in varied contexts, even if all of the instructors are teaching the same course. If the data collection methods for this study are feasible for appropriate research with this population (i.e., experienced, university-approved teachers who often discuss and reflect on their practice), these tools could be adapted for use in studying the assessment practices of other populations.

It is anticipated that the results of this study add additional insight into secondary ELA teachers' practices for improving student learning to the FA literature. Having an understanding of how SPRUCE ENG 135 Instructors implement assessment practices that fall under the aegis of "formative assessment," especially with regard to questioning, discussion, and feedback, can better inform further empirical study in this context.

Chapter 2: Literature Review

The purpose of this dissertation study was to examine the formative assessment practices used by secondary ELA teachers in their classrooms in terms of (a) what assessment practices they use, (b) how they choose which practices to use, (c) how effective they believe the practices they use to be, and (d) how they determine the effectiveness of these practices. The following literature review supports this exploratory research through an examination of assessment practices in teaching—those labeled "formative assessment" in particular—and how research on these practices in other content areas and contexts have informed this study.

The Importance of Assessment in Teaching

The main functions of classroom assessment serve formative and summative purposes, and are described using terms introduced by Scriven (1967) to the field of evaluation. Formative assessments are meant to guide and improve student learning as well as teachers' instructional practice, while summative assessments are used for reporting purposes (Black, 2013; Bonner, 2013). Any assessments that take place in a classroom are highly situated (and therefore challenging to generalize to other settings), are generally low-stakes, frequently informal, and are often quite complex and based on a variety of methods (Bonner, 2013). FA in particular focuses on the ongoing gathering and use of evidence about how students' learning is developing while instruction is taking place (Black & Wiliam, 1998b; Heritage, 2013).

Teachers as professionals using assessment. In the 20th century, educational assessment focused primarily on standardized testing for accountability (Shepard, 2000; Stiggins, 2002). Although the struggle of teachers to become valued as professionals is documented in the literature (e.g., Hoyle, 1982; Labaree, 1992), a focus on external accountability testing rather than classroom practices led to the de-skilling and de-professionalization of teachers (e.g.,

Darling-Hammond, 1988; Shepard, 2000). A flood of critiques about the U.S. education system in the 1980s prompted a variety of educational reform efforts, including endeavors to increase the quality of schools through better professional preparation for their teachers (Labaree, 1992). One of the areas addressed was a perceived lack of teachers' skills in assessing their pupils' educational progress.

The American Federation of Teachers, the National Council on Measurement in Education, and the National Education Association (1990) created a set of *Standards for Teacher Competence in Educational Assessment of Students*. These standards were an effort to refocus teacher training in assessment to include practices that occurred at the classroom level, as well as assessments at the school, district, state, and national levels. However, research by Stiggins (1999) found that teachers' competence in assessment was not required in 25 of the 50 states that were surveyed, and only 10 states required assessment course work for teachers during their training. A decade later, Brookhart (2011) acknowledged that although the 1990 *Standards* contributed a great deal to teaching and research in teacher competence, they had become dated due to their disconnect to current conceptions of FA and the teacher knowledge and skills needed in an accountability and standards-based reform context.

Recently, there has been a push for more research into classroom assessment, due to advances in standards-based education, high-stakes testing, learning and motivation theory, measurement, technology, and FA (McMillan, 2013). Continuing work in this area of teaching would serve as a means of "improving teaching as a profession, student learning, and student motivation," which can be accomplished, in part, by "conducting meaningful, well-designed studies of classroom assessment" (p. 14). Further research into FA practices as a form of classroom assessment would contribute to these efforts.

Why focus on assessment rather than instruction in FA? As previously stated, there are variations in understanding of what "formative assessment" entails, depending on the context in which its practices take place. However, in studies of the effectiveness of FA, there are four major components of the concept that emerge in the literature. The first component is questioning students appropriately in ways that engage their thinking strategies, allow for response time, and act as a means of gathering evidence to inform instruction (e.g., Black, Harrison, Lee, Marshall, & Wiliam, 2003; Shermis & Di Vesta, 2011). The second component, feedback, whether from teachers to students, students to peers, or students to themselves in the form of self-regulation, is another important piece of effective FA (e.g., Black et al., 2003; Heritage, 2010; McManus, 2008; Popham, 2008; Sadler, 1989; Shermis & Di Vesta, 2011). The third and fourth components both address assessment materials: the construction of effective assessments and the analysis and use of assessment data to drive instructional decision making (e.g., Birenbaum, Kimron, Shilton, & Shahaf-Barzilay, 2009; Griffin, 2009; Nelson & Eddy, 2008; Phelan, Choi, Vendlinski, Baker, & Herman, 2011; Yin et al., 2008). Although these components ultimately inform a teacher's instruction, the act of assessment itself (as well as its outcomes) is what guides ongoing instructional decisions and practices.

Previous Research on Assessment Practices to Improve Student Learning

A great deal of the push for FA's use in elementary and secondary education is due to Black and Wiliam's (1998a) extensive review of FA studies and their findings of reported effect sizes ranging from 0.4 to 0.7. Along with these data, however, Black and Wiliam also identified aspects of FA research that needed more attention in the literature, including "the extent to which the context of any study is artificial and the possible effects of this feature on the generalisability of the results" (p. 59). This concern was also raised by Yin et al. (2008), who noted that "the

effects of formative assessment have rarely been examined experimentally in regular education settings" (p. 336). Perhaps one reason for this is the inherent complexity of classroom environments, as there are variations at the administrative, school, classroom, and student levels (Tolley, 2011).

More recent reviews of FA research have been critical of the well-publicized effect sizes from Black and Wiliam (1998a), and have reported that approaches to the actual implementation, evaluation, and study of FA practices vary from context to context (e.g., Bennett, 2011; Dunn & Mulvenon, 2009; Kingston & Nash, 2011). Dunn and Mulvenon (2009) stressed the need for a common vocabulary to describe FA practices to facilitate communication, and called for "a sound research-validated framework for best practices in formative assessment and formative evaluation" to be established (p. 9). In a response to Dunn and Mulvenon's work, Filsecker and Kerres (2012) emphasized that the concept of FA comprises related practices that are part of the process, and that "there is no right or wrong definition of formative assessment," only the paradigms (such as educational assessment) through which its components are viewed (p. 7). Even Wiliam (2011), one of the co-authors of the article from which interest in effect size in FA stemmed, suggested that differences in FA effect size across the literature may be due in part to the sensitivity of measures used, differences in the variances of the populations studied, and, perhaps most importantly, how the concept of FA was operationalized by the researchers.

In order to understand more about and attempt to clarify the actual effects of FA on student achievement in K-12 education, research in varying grade levels and subject areas has been conducted. For example, Kingston and Nash (2011) orchestrated a meta-analysis of over 300 studies reporting FA use, but reported that only 13 studies provided sufficient information to calculate effect sizes. Based on 42 independent effect sizes from these studies, moderator

analyses of FA effect sizes were estimated at 0.17 and 0.09 for mathematics and science, respectively, yet were found to be most promising for ELA, with an estimated effect size of 0.32. However, the methods for inclusion of studies in this meta-analysis were criticized by other researchers (Briggs et al., 2012; McMillan et al., 2013) as being too restrictive, and were charged with contributing to the uncertainty of the effect that FA practices have on student achievement.

A review of literature on FA practices and their relationship to student achievement shows a predominance of science and mathematics as the content focus, yet, given the effect sizes reported by Kingston and Nash (2011), whether completely accurate or not, perhaps more should be explored about the use of FA practices in ELA. How *is* FA evaluated in content areas that are more subjective, such as advanced ELA, where students are encouraged to think in non-linear ways, interpret texts and their meaning, and support their hypotheses in written and oral arguments? What measures have been used to capture teachers' assessment practices, and would they be suitable for use in a secondary ELA context?

Presented in this section are brief overviews of research on assessment at the secondary level, typically ranging from grades 7-12 (and at times including 6th grade), or studies indicated in the research as taking place in at least a "middle school" or "secondary" context. First is a brief summary of research in a general secondary context, followed by research specific to a secondary ELA context. Finally, an overview of existing measures of teachers' assessment practices is given.

Research in a secondary context. Several studies have been done comparing FA implementation (or similar forms of assessment) and its potential impacts on instruction and student achievement in multiple disciplines. For example, McMillan (2001) investigated the actual classroom assessment and grading practices of teachers of science, social studies, math,

and English in grades 6-12. Of the almost 1500 respondents, 27 percent were English teachers. An analysis of the types of assessments reported as being used by the teachers resulted in the findings that the English teachers showed more emphasis on constructed-response assessments, teacher-developed assessments, and major exams, and less use of recall items, objective assessments, and quizzes. In addition, the English teachers showed a greater reliance on higher-order thinking skills from their students than did the science and social studies teachers.

Hodgen and Marshall (2005) observed lessons in English and math in secondary schools in the United Kingdom to learn more about the use of assessment for instructional practices in these content areas. The investigation was intended to clarify how these contrasting contexts might impact FA in the classroom. Surprisingly, the researchers discovered that the lessons had fundamental similarities. Both the English and math lessons engaged students in problem-solving tasks, involved teachers using interventions to extend students' thinking through feedback, and the scaffolding of the lessons allowed for peer interaction and feedback to help students progress in their learning. In this sense, these particular lessons demonstrated that there may be parallels in how the overall implementation of FA occurs across subjects, and that what is learned about FA in one content area may inform the study of FA in another discipline. However, Hodgen and Marshall also noted that the observed lessons were greatly impacted by teachers' subject knowledge and their application of this knowledge to their practice, both of which remained vastly different between the two content areas. As a result of these findings, the researchers identified a need for adaptation of the FA guidelines outlined by Black and Wiliam (1998a) to the specific demands of each school subject. In addition, Hodgen and Marshall called for "teachers and researchers [to] work together to transform research knowledge into practice" (p. 172).

Research specific to a secondary ELA context. As noted earlier in this chapter, McMillan et al. (2013) critiqued Kingston and Nash's (2011) meta-analysis of FA effect sizes, which showed a much greater effect of FA on student achievement in an ELA context compared to science and math effect sizes. Although the overall moderator effect size for ELA was estimated by Kingston and Nash as 0.32, their restrictive inclusion criteria only allowed for two of the 13 studies to involve ELA classrooms: Brookhart, Moss, and Long's (2009) study involving kindergarten and first grade students; and Rich, Harrington, and Kim's (2008) study of seventh and tenth graders. For the latter study, when researching the effect of FA on classroom instruction, or even how FA was implemented, McMillan et al. (2013) noted that there were only "scant, anecdotal indications that instructional correctives were made" (p. 13). Unfortunately, lack of empirical data to support claims about FA's effectiveness in secondary ELA education appears to be a common problem in the literature, although this seeming lack of information may be due to the nature of the subject matter.

McMillan (2010) noted that in contrast to the more "planned and structured" use of FA in math and science, ELA and other humanities teachers conduct much of the assessment of their students "'at the moment,' without preplanned feedback and responses" (p. 54). FA-related practices in ELA classrooms, such as classroom talk and questioning, feedback, sharing learning intentions and success criteria with the learners, and peer and self-assessment (Marshall, 2007) are still likely to occur, but the FA progression itself is likely to be "a more meandering, organic affair" (p. 137).

In the secondary ELA classroom, students should be guided by their instructors into "guild knowledge" that allows them to be able to assess the work that they and others produce with relation to the subject's core concepts and processes (Marshall, 2007; Sadler, 1989). Guskey

(2003) noted that writing teachers in particular emphasize the importance of the process of learning, and that gaining knowledge from one's own mistakes and their correction helps develop students' ability to learn on their own. Perhaps due to the increasing emphasis on student self-reliance, there is not a great deal of empirical literature on the actual classroom assessment practices of secondary teachers and how these data are used to inform instruction for improvement (McMillan, 2001).

One exception to the dearth of empirical research in secondary teachers' assessment practices is an ELA-oriented qualitative study conducted by Fisher, Lapp, Flood, and Moore (2006), which used observations and interviews as part of a professional development initiative to help secondary school teachers in grades 9-12 use assessment data to guide their instruction. The 25 teachers in the study participated in a three-course program that included study of (a) literacy and language, (b) literacy assessment, and (c) clinical field experience, which involved one-on-one tutoring with their students that needed additional help (10 students per teacher, for a total of 250 students). The qualitative data were triangulated with the students' results on two assessments of student achievement, obtained from school records: their scores on the Gates-MacGinitie reading assessment, and the state content standards test. Teachers participating in this assessment-to-instruction approach reported growth in their assessment knowledge, skills, and dispositions, and the students who received extra help from the teachers learning to link assessment with their instruction gained an additional half-year of reading growth over the academic year, when compared with their peers.

A second example of an empirical approach to the study of secondary teachers' assessment practices is the research by Frey and Fisher (2013) on the creation of an error analysis tool. The need for this tool was prompted by survey response data from almost 550 high

school students about what they felt would benefit them most during the developmental stage of the writing process. The resultant tool was based on a specified purpose for learning within a writing context, helped students and teachers identify student errors (what they do not know), and prompted teachers to change their instruction based on what their students did and did not know. The use of this tool then led to guided instruction on the part of the teacher, typically involving three key scaffolds: questions to check for understanding, prompts to apply specific cognitive and metacognitive resources, and cues when the learner needs their attention shifted more overtly to improve their learning.

A third example is the work of Hunter, Mayenga, and Gambell (2006), who used pan-Canadian survey data from over 4,000 secondary English teachers of 13- and 16-year-olds and the teachers' self-reported assessment practices to describe the teachers' preferred evaluation tools, analyze each of these elements for potential underlying factors, and investigate possible relationships between teaching experience and qualifications and assessment practices in the secondary ELA classroom. The researchers found that the use of examples of students' own writing along with teacher-constructed open-response and essay tests were the primary classroom assessment tools among the respondents, and that homework, student improvement, participation, and student effort were also highly valued. Hunter et al. also found that the use of these tools was primarily for the support of one-on-one interactions with students, rather than in small groups, and that the more qualified and experienced teachers were more likely to rely on examples of student writing for assessment of students' individual progress. Surprisingly, teachers identified as being at the midpoint in their careers (6-20 years of experience) were most likely to use examples of student writing and portfolios as assessment tools. From these results, the researchers stressed the need for more research in North America into "notions of [teacher]

expertise, comparing novice and expert practitioners in their interpretative approaches to writing assessment information and its derivation from tool use" (p. 62). This call for further study ties back to McMillan's (2001) observations about the lack of empirical literature on the actual classroom assessment practices of secondary teachers and how these data are used to inform instruction for improvement.

The previous three summaries of studies on assessment use in an ELA context were empirical in nature, but this was not a common occurrence in the literature review conducted for this dissertation study. Much more common were articles written from the perspective of the ELA teacher about strategies that they used in the classroom and how they felt these practices (whether termed as FA or not) helped to improve student learning.

Although a great deal of extant literature on secondary ELA teachers' FA practices is anecdotal in nature, it still draws upon prior FA research on these practices and their implications for use. For example, Gorlewski (2008) noted that the use of FA "provide[s] an exceptional opportunity for teachers to collect, analyze, and use data in meaningful ways" (p. 97). It was through Gorlewski's use of FA, in particular, a "ticket out the door" activity, that she was able to realize that her perception of her lesson varied from her students' actual understanding. As a result of her experience, she encouraged other ELA teachers to "require students to generate evidence of what they have learned" as a means of assessment to gauge student learning and to help them achieve learning objectives (p. 97). The use of similar strategies, such as "exit slips" and informal communication with students to check for understanding, were described by Pappageorge (2013). Additional research has emphasized the usefulness of student writing portfolios to involve students in the assessment process and to make their thinking about writing and reading more explicit (Murphy, 1997; Murphy & Smith, 2013).

Other anecdotal research into the use of practices to improve student learning in the secondary ELA classroom, although not addressing FA specifically, does relate to some of FA's main principles or components. This body of work includes teacher support of writers during drafting and revising, and supporting students and intervening when appropriate in the writing process. These practices include the use of conferences, response logs, and participation in literary discussion groups (VanDeWeghe, 2006), or the use of student writing portfolios to assist students in the self-assessment process and to allow for personal instructional or "program" evaluation by the teacher (Gorlewski, 2010). Reading strategies reported as successful by secondary ELA teachers include setting learning objectives based on students' individual comprehension abilities or levels, guiding students' response writing based on these levels, and encouraging student small group discussion of each other's responses (FitzPatrick, 2008). Other reading strategies include allowing students to "play a participating and active role in the evaluation of their learning" (Noskin, 2013, p. 73) by establishing and sharing learning objectives with them, and involving students in the assessment process through the writing and revisiting of reading response journals, group discussions with passage explication, and written responses to the text, and using the results of this process to adjust teaching and learning (Noskin, 2013).

Existing measures of teachers' assessment practices. There are several existing self-report measures of teachers' classroom assessment practices and beliefs that are relevant to the research problem presented in Chapter 1. Although similar constructs are explored, it is important to note that most of the instruments below focus on frequency of use of practices, and overall do not delve into the perceived effectiveness of these practices nor the reasoning behind their use. These self-report instruments indirectly measure assessment practice; currently there

are no instruments available that directly measure the assessment practices of teachers in the classroom (Randel & Clark, 2013).

Assessment Practices Inventory (API). The API is a 67-item self-report instrument that lists assessment practices ranging across constructing tests, communicating assessment results, and using assessment results. Respondents use two response scales for each item: one is a frequency of use, ranging from 1 = *not at all used* to 5 = *used very often*; the other measures teachers' self-perceived assessment skills from 1 = *not at all skilled* to 5 = *very skilled* (Zhang & Burry-Stock, 1994, as cited in Randel & Clark, 2013).

School Achievement Indicators Program (SAIP) Writing Assessment III Teacher Questionnaire. Hunter et al. (2006) reported a selection of findings from the 2002 administration of the SAIP Writing Assessment III Teacher Questionnaire, a supplemental survey given to teachers of students who participated in the SAIP in Writing assessments in Canada. The items chosen by Hunter et al. were Likert-type, and asked about the district/provincial context, the out-of-school context, the school context, student characteristics, program design, teacher characteristics, and classroom instruction and climate. Items specifically about teachers' assessment practices were rated on a 4-point scale, using either a general scale of 1 = *rarely or never*, 2 = *a few times a month*, 3 = *a few times a week*, and 4 = *almost every class*, or a reporting of classroom practices scale, with 1 = *rarely or never*, 2 = *once or twice a class*, 3 = *several times a class*, and 4 = *many times a class*. Teachers were additionally asked to rate their level of agreement on a 4-point scale from 1= *strongly disagree* to 4 = *strongly agree* about their perceptions regarding the teaching of writing. The findings of this instrument from use with secondary teachers are presented in a previous section of this chapter, "Research specific to a secondary ELA context."

Assessment Practices Survey (APS). The APS asks teachers to reflect on their FA practices, which are divided into eight categories: written feedback, revisions, developing questions, selecting students during questioning, criteria for success, collaborative learning, peer assessment, and self-assessment. Respondents are asked to report on the frequency of those practices using a 5-point scale: *never*, *quarterly*, *monthly*, *weekly*, or *daily* (Lyon, Miller, & Wylie, 2011, as cited in Randel & Clark, 2013).

Classroom Assessment and Grading Practices survey. Unlike the measures described above, one instrument that explored decision making in assessment (as opposed to frequencies of practice) is the Classroom Assessment and Grading Practices survey, which has been used to describe the factors involved in elementary and secondary teachers' math and ELA practices in determining final semester grades (McMillan, 2001; Randel & Clark, 2013). This 34-item self-report survey includes items about "(a) factors that teachers consider when assigning grades, such as student effort, improvement, and academic performance, (b) types of assessments used, and (c) the cognitive level of the assessments (e.g., recall knowledge, understanding, application, and reasoning)" (McMillan, 2001, p. 23). The items were introduced by the stem "'To what extent were final first semester grades of students in your single class described above based on…'" (p. 23), and were rated on a 6-point scale "ranging from *not at all* to *completely*" (p. 22). The results of the use of this instrument with secondary teachers are presented in a previous section of this chapter, "Research in a secondary context." Of note, however, is the remark by McMillan that "these findings suggest the continued need to take subject matter into consideration when examining and understanding assessment and grading practices" (p. 31).

Appreciating and Using Teacher Knowledge

Of course, another major factor to be considered when attempting to understand teachers'

assessment practices is the knowledge of the teachers themselves. The kinds of knowledge that

teachers have and use within a particular context are critical to their actions. Below is a summary

of research on teacher knowledge, how it is supported, and how we as researchers may learn

from this knowledge. Such understanding is critical to this dissertation study, as teachers'

decision making about their assessment practices is grounded in their implicit knowledge, and

this research aimed to capture this knowledge and make it explicit.

What does teacher knowledge entail? According to Shulman (1987), teacher

knowledge, *at a minimum* [emphasis mine], includes the following kinds of knowledge: (a)

content knowledge; (b) general pedagogical knowledge, including classroom management and

organization; (c) curriculum knowledge; (d) pedagogical content knowledge; (e) knowledge of

learners and their characteristics; (f) knowledge of educational contexts, from classroom through

culture; and (g) knowledge of educational purposes and values and their philosophical and

historical grounds. Of particular note is pedagogical content knowledge, "that special amalgam

of content and pedagogy that is uniquely the province of teachers, their own special form of

professional understanding" (p. 8). Shulman asserted that teachers' knowledge is developed from

at least four major sources, which he identified as scholarship in content areas, the materials and

structure of teaching, ongoing research in the field, and the wisdom that comes from conducting

the practice itself.

What helps to support teacher knowledge? There are several factors that support the

development and use of teacher knowledge. Three are briefly presented below: (a) communities

of practice, (b) effective professional development, and (c) ongoing reflection in practice.

Communities of practice. A supportive context where teachers can use their knowledge in their daily work is very important (Day & Gu, 2007; Kelleher, 2003). The development and maintenance of a professional learning climate through a teacher's colleagues (Hill, Hawk, & Taylor, 2002) and the school administration (Day & Gu, 2007; Webster-Wright, 2009) encourage teacher learning. If a teacher is involved in a professional community that is "nurtured and developed both within and outside the school, then the effects […] can bring about significant and lasting change" (Lieberman, 1995, p. 596).

Effective professional development. Teachers bring a wealth of experiences from the classroom, their prior learning, and their lives into development of their knowledge, and their values and perceptions of themselves and the world should be included in professional development. Effective professional development allows for the interchange of ideas and the sharing of teachers' individual experiences with each other, which allows for practices to become more generalizable to others and therefore more applicable in different contexts (Beavers, 2009).

Teacher learning through professional development has been reported to be most effective when teachers are given the opportunity to engage with the new knowledge and apply it to their classrooms (Kelleher, 2003; Lieberman, 1995; Opfer & Pedder, 2011). Actively working on genuine problems relevant to their specific situations, experiencing the processes involved, and perhaps even being actively involved about the decisions made on what is offered through professional development is more engaging and effective to teachers-as-learners (Lieberman, 1995; Webster-Wright, 2009).

Ongoing reflection in practice. Teachers are better able to engage in transformative learning if they are able to critically examine their practice and increase their understanding of themselves and others (Beavers, 2009). As Kelleher (2003) noted, "good teachers are by nature

reflective learners," and they can use their own reflexivity to develop ideas to share with their colleagues and "work together through many professional development activities" with that clearer understanding of the self and their own practices in mind (pp. 755-756). Discussing that reflection with others is also an important part of the process: Lieberman (1995) remarked that "people learn best through active involvement and through thinking about and becoming articulate about what they have learned" (p. 592).

How do we learn from this knowledge? Expert teachers, or those with strong practical knowledge, have a certain set of features that is similar to experts in other fields. In part, these features include excelling within particular contexts, being more opportunistic and flexible in their teaching than novices, and perceiving more meaningful patterns in the domain in which they are experienced (Berliner, 2001). Although the time for the development of expertise may vary, "a reasonable estimate for expertise to develop in teaching, if it ever does, appears to be 5 or more years" (p. 479). Along with the use of written questionnaires, more in-depth information can be gleaned from teachers in narrative form, such as through interviews where teachers reflect on their practice (e.g., Clemente & Ramírez, 2008; McMillan, 2003) to learn about their decision-making processes. This combination of approaches to learn more from the expert teachers that are SPRUCE ELA Program ENG 135 Instructors about their assessment practices is what was used for this dissertation study.

Chapter 3: Methods

The purpose of this mixed methods study was to examine secondary ELA teachers'

formative assessment practices through exploration of the assessment practices that they used,

how they selected these practices, how effective they believed these practices to be, and how

they determined the effectiveness of these practices. This was achieved through the

implementation of both quantitative and qualitative research methods with a sample of secondary

ELA teachers that were all teaching the same course within a concurrent enrollment program in

the northeastern United States.

Study Design

This study used a sequential explanatory mixed methods design (Creswell & Plano Clark,

2011; Ivankova et al., 2006). In this design, the qualitative phase builds upon the quantitative

phase and helps to explain the findings (Ivankova et al., 2006). For the purposes of this research,

equal priority was given to the phases, rather than ascribing more weight to the quantitative

phase, as is typically done with this research design (Creswell & Plano Clark, 2011; Ivankova et

al., 2006). The integration of both phases was done in the interpretation and discussion of the

findings, which are presented in Chapter 5.

For the initial quantitative phase, participants completed background questionnaires and

checklists about their assessment practices, and these responses were analyzed. These data and

findings then informed the qualitative phase of the study, where the participants were

individually interviewed using a semi-structured protocol. The interview data were analyzed to

determine emergent themes. The findings of both phases of the study are reported in Chapter 4.

Study Participants

The target population for this study was SPRUCE ELA Program Instructors teaching ENG 135 (*Critical Reading: Issues, Theory, and Applications*) in Spring 2015. ENG 135 was the only ELA course being offered through SPRUCE during that semester; the same course content was covered by the teachers in similarly-structured ways across their respective schools. The other SPRUCE ELA course, ENG 121 (*Academic Writing Processes*), had been taught by many of these Instructors in Fall 2014. At the time of recruiting for this study, there were 49 individuals in this population, and the invitation to volunteer to participate (see Appendix A) was extended to all of them.

SPRUCE ELA Program Instructors overview. SPRUCE ELA Program Instructors are a unique population of secondary ELA teachers. To become a SPRUCE Instructor, teachers apply for the position. Requirements include several years (preferably at least five) of prior teaching experience, a master's degree in their subject area or a related field, and participation in required SPRUCE training. Certified SPRUCE Instructors teach concurrent enrollment college-level courses in high school, and are considered to be adjunct faculty in the program's partnership university. As professionals who elected to be a part of this program, SPRUCE Instructors are motivated teachers who are experienced in their field, and are better able to determine what instructional and assessment strategies work for them, as well as how to best adapt their practices to meet the needs of their students.

All SPRUCE teachers go through extensive professional development, including summer workshops, on-site training, and continuing check-ins with SPRUCE staff. They also frequently collaborate with other SPRUCE Instructors, both within their own school and in other schools offering SPRUCE courses. These strong professional networks, which involve a focus on

common interests and objectives, a variety of engaging activities, a discourse community that encourages exchanges among members, and leadership opportunities for teachers outside of their own schools, benefit both teacher learning and the potential for educational change (Lieberman & McLaughlin, 1992). By agreeing to teach in the SPRUCE ELA Program, the teachers lose some autonomy over the major content to be covered (which is described in the next section), but are permitted flexibility in the supplemental texts that they use with their students to meet the course goals. Due to ongoing communication and collaboration, these Instructors have a shared vocabulary to discuss the teaching practices they use in these courses. In addition, they are more likely than typical classroom teachers to be able to make their implicit teaching knowledge explicit because of their frequent sharing of ideas and pedagogical practices.

Access to SPRUCE ELA Instructors for this study was granted through the SPRUCE Director and ELA Program administrator. Additional support was provided by the SPRUCE Associate Director, who oversees research and evaluation activities within the program.

SPRUCE ENG 135 overview. SPRUCE ENG 135 is a semester-long writing-intensive critical reading course in which students study the influence of culture on texts and their meaning in many forms, including literature, advertisements, film, and images. The course is centered around three major themes of contemporary criticism (subjectivity, ideology, and agency), which students learn about, engage with, and apply to the texts that they are exploring. Students participate in frequent writing activities, and are involved in ongoing dialogue and the development of their ideas through individual writing assignments and in-class discussions with their teacher and their peers to share, refine, revise, and expand upon their work.

There are four course goals for ENG 135, according to the syllabi for the course provided to me by the teacher participants. These goals involve students: (a) gaining an expanded

understanding of textuality, or examining the world as text; (b) learning to apply the language and methods of textual criticism; (c) developing a working knowledge of strategies and genres of cultural analysis and argument; and (d) gaining a sense of how context, and in particular culture, shapes meaning and influences the reception and interpretation of texts by their readers.

To achieve the course goals, all ENG 135 students are expected to complete the following assignments, with the final grade percentage in parentheses for each: a subjectivity unit essay (20%); an ideology unit essay (20%); an agency unit multimedia project/portfolio (20%); at least six think/response papers (30%); and informal writing and class participation (10%). The unit essays are expected to be six to eight pages in length, and are developed from students' think/responses papers for that unit, which are more informal writing pieces that are usually one to two pages in length. Another example of informal writing for the course is students keeping a dialogic journal, where they note summaries of excerpts of the texts they are reading on one side of a page, and document their interpretation or analysis to correspond with these notes on the other side. Like the think/response papers, the dialogic journals inform and contribute to students' development of their unit essays.

Instrumentation

There were three instruments used in this study, all of which were created for the purposes of this research. They are (a) a background questionnaire, (b) the English/Language Arts Teachers' Experiences (ELATE) checklist, and (c) a semi-structured interview protocol. Details about each of these instruments are presented below. Instructions for participation in the study (Appendix B), along with the informed consent document, a background questionnaire, and an initial ELATE checklist, were sent to SPRUCE ENG 135 Instructors who expressed interest in participating in this research.

Background questionnaire. The background questionnaire (Appendix C) included a total of eight items, and was anticipated to take approximately 5-10 minutes to complete. The first three items asked for participants' gender and years of teaching experience, as of Spring 2015, both generally and in the SPRUCE ELA Program. The remaining items asked about participants' current teaching load in Spring 2015: (a) for the current academic year [2014-2015], the number of sections of SPRUCE courses that the Instructor is teaching, the average number of SPRUCE students per section, and the total number of all students taught; (b) the current number of students in a participant's ENG 135 section(s); and (c) other courses taught this academic year [2014-2015], including course name, number of sections, and grade level(s) of the students taking those courses. The background questionnaire was emailed to participants as a PDF form, which they completed, saved, and emailed back to me.

English/Language Arts Teachers' Experiences (ELATE) checklist. The ELATE checklist (Appendix D) contained 25 items—a list of assessment practices related to questioning, discussion, and feedback—as well as space for Instructors to indicate other assessment practices that they used which were not included in the list. Each ELATE checklist was anticipated to take approximately 15-20 minutes to complete. Participants were asked to complete one checklist per week, documenting their assessment practices from Friday through Thursday, for four weeks. These days were chosen instead of Monday through Friday so as to not overburden teachers at the end of the work week. Requesting the submission of the ELATE checklists on Thursday also budgeted an additional work day to follow-up with teachers who had not yet sent in their checklists, if needed.

Instructors were asked to indicate how many times over the past week they used each practice with their SPRUCE ENG 135 students. For the practices that they used, they were asked

to rate their perceived effectiveness of that practice to improve student learning with their lower-performing, average-performing, and higher-performing ENG 135 students using a 6-point Likert-type scale (0 = *extremely ineffective*, 1 = *very ineffective*, 2 = *moderately ineffective*, 3 = *moderately effective*, 4 = *very effective*, and 5 = *extremely effective*). Each ELATE checklist was emailed to participants as a PDF form, which was completed, saved, and emailed back to me. In some cases, due to technical difficulties, the forms were also printed on paper by the participants, completed by hand and then scanned, and the images were emailed back to me.

The checklists, taken together, are a form of experience sampling methods (e.g., Hektner et al., 2007; Zirkel et al., 2015). This approach allows for an understanding of participants' experiences as they are happening, which reduces retrospective and other biases and also allows insight into practices that are occurring within a particular context (Zirkel et al., 2015). By completing a series of four ELATE checklists, SPRUCE ELA Program ENG 135 Instructor participants were able to document their assessment practices on a weekly basis, while also reflecting on what they were doing and how well these approaches worked with their students.

The items for the ELATE checklist were based on research indicating the importance of questioning, discussion, and feedback to FA practice. As noted in Chapter 1, a previously conducted literature review of 25 studies involving professional development for teachers in grades K-12 on FA (Tolley, 2012) resulted in six common elements of the construct, including the frequent mention of the use of questioning and feedback from teachers and students (e.g., Aschbacher & Alonzo, 2006; Blanchard, 2008; Brookhart et al., 2009; Dixon & Williams, 2003; Herman, Osmundson, & Silver, 2010; Hollingworth, 2012; McGatha, Bush, & Rakes, 2009; Olson, Olson, Slovin, Gilbert, & Gilbert, 2010; Phelan et al., 2011; Sato, Wei, & Darling-Hammond, 2008; Sondergeld, Bell, & Leusner, 2010; Wiliam, Lee, Harrison, & Black, 2004). A

second contributor to these items is Danielson's *Framework for Teaching* (2007, 2014), particularly the "using questioning and discussion techniques" and "using assessment in instruction" components of the instruction domain. The former includes elements of quality of questions/prompts, discussion techniques, and student participation, and the latter includes elements of assessment criteria, monitoring of student learning, feedback to students, and student self-assessment and monitoring of progress (Danielson, 2014). Item wording was adapted from the questioning techniques section of the SAIP Writing Assessment III Teacher Questionnaire (Hunter et al., 2006), the examples in practice for the selected *Framework for Teaching* components (Danielson, 2014; Danielson et al., 2009), and the preliminary fieldwork study I conducted prior to this research, which involved interviews with SPRUCE ELA Program administrators and Instructors about the Instructors' instructional and assessment practices.

The content validity of the ELATE checklist was first reviewed by the SPRUCE Director and ELA Program administrator, who recommended slight changes to the wording of items to better reflect the language used by SPRUCE Instructors. Additional review of content validity was provided by the SPRUCE Associate Director after these revisions were made, and this individual also determined that the ELATE checklist was appropriate for use with SPRUCE ELA Instructors. Following these reviews, the ELATE checklist was converted from a Microsoft Word document into an Adobe PDF form, and was distributed to six individuals who were either current or former teachers who were not part of the study population. These individuals reviewed the checklist and its accompanying instructions for face validity, gave feedback as appropriate, and completed and returned the form to me via email to ensure that the PDF form functioned as expected.

Semi-structured interview protocol. An individual semi-structured interview (e.g., Rubin & Rubin, 2005; see Appendix E for the full interview protocol), which took approximately 45-60 minutes to administer, was conducted with participants after their completion of the weekly checklists. The interview protocol was finalized after the background questionnaire and ELATE checklists data were analyzed, and was created to include individualized questions specific to each Instructor. The participant-specific questions added to the general protocol were intended to check the validity of ELATE checklist responses and request clarification of why certain assessment practices were used (or not) or deemed effective (or ineffective). In addition, questions were asked of participants to help them reflect on their perceptions of assessment in their teaching after having completed the series of ELATE checklists and thinking about their own practices.

Procedures

This study was conducted at the end of the 2014-2015 academic year, with the background questionnaires and ELATE checklists completed in May and June by the teacher participants, and the interviews conducted in June and July. Research activities commenced immediately after approvals from both the Institutional Review Board (IRB) and my dissertation committee were granted, which impacted the timing of the study. Conducting most of the study in May and June may have affected its findings, due to end-of-the-year course and school schedules not being representative of the rest of the semester. However, I chose to collect data within this timeframe because the SPRUCE ELA Program was offering ENG 135 (*Critical Reading: Issues, Theory, and Applications*) in the Spring 2015 semester for the final time, and was planning on offering a new replacement course the following year. Unlike the writing-intensive ENG 121 (*Academic Writing Processes*), which is offered by the SPRUCE ELA

Program in the fall semester, I reasoned that ENG 135, which focused on critical reading of texts and theory application in addition to writing development, would be more representative of the kinds of activities (in particular, questioning, discussion, and feedback) that may occur in a more traditional ELA course.

Following IRB approval of this research, potential study participants—SPRUCE ELA Program ENG 135 Instructors teaching in the Spring 2015 semester ($N = 49$)—were identified with the assistance of the SPRUCE Associate Director, who shared their contact information with me. A recruitment email, describing this study and inviting these Instructors to participate (see Appendix A) was sent to each person individually (April 29, 2015), with two follow-up reminder emails sent as needed at one week (May 6, 2015) and two-and-a-half weeks (May 18, 2015) after the original message. A total of 24 of the 49 ENG 135 Instructors responded to these invitations to participate in the study; the remaining 25 did not reply to any of the three recruiting emails. Of the 24 respondents, 10 replied that they were interested in participating, 13 declined to participate, and one recommended another individual in her/his school as a potential participant, but did not indicate her/his own interest in participating, even after follow-up emails were sent to her/him. Eleven of the 13 declining ENG 135 Instructors cited lack of time due to personal and professional obligations, especially those due to the end of the school year, as the main reason they were unable to participate in the study.

The ten ENG 135 Instructors that expressed interest in participating in this dissertation research were sent an email that included the following: (a) instructions for participation in the study (Appendix B); (b) the IRB-approved informed consent document for the study, both as a PDF attachment and in text form pasted below the email message; (c) the background questionnaire as a PDF form (Appendix C); and (d) the first blank copy of the ELATE checklist

as a PDF form (Appendix D). Potential participants were asked to provide their written consent

to participate in the study via email, and also to indicate whether or not they agreed to be audio

recorded during the interview phase of the study.

Of the ten individuals that originally expressed interest in participating in the study, seven

ultimately consented to and took part in this research. The characteristics of these individuals and

the contexts in which they were teaching at the time of this study are presented in Chapter 4.

Each of these SPRUCE ELA Program ENG 135 Instructors was asked to return her/his

background questionnaire and first ELATE checklist about the previous week when each form

was complete. An email was sent to each Instructor acknowledging receipt of these documents,

and clarifying questions were asked of the participants as needed. In addition, a blank ELATE

checklist was attached to each of these messages for the next week; confirmation emails were

also sent when these completed forms were received. Two more ELATE checklist copies were

sent over the next two weeks, and their receipt was acknowledged each time a participant

returned a completed form. Reminder emails were sent as needed to ensure that the checklists

were returned in a timely manner. Data were recorded, coded, cleaned, and analyzed

quantitatively for potential trends as they were received. Overall, this phase of data collection

spanned a six-week period from May 1st through June 11th, 2015, as each teacher began

participating in the study as appropriate for her/his schedule.

After the ELATE checklist data were collected for each participant, an interview was

scheduled with that Instructor. A semi-structured interview protocol was developed based on

preliminary analysis of the quantitative (i.e., checklist) data. Of the seven teachers who

participated in the ELATE checklists phase of this study, six were interviewed; the seventh

Instructor, Tony, did not respond to multiple requests for an interview. The interviews were

conducted via Skype videoconferencing to allow for visual/nonverbal interaction (i.e., the use of

body language) between me and each participant. At the start of each interview, the Instructor's

willingness to participate was confirmed, and her verbal permission to participate in the

interview was obtained (see the interview protocol, Appendix E). Participants' consent to be

audio recorded was reconfirmed and audio recorded; all interviewed participants did grant their

permission to be recorded. The interviews, averaging approximately 52 minutes each in recorded

length, took place between June 25th and July 17th, 2015, as Instructors' schedules permitted.

Data Analysis

As a former secondary ELA teacher, it was important for me to understand my own

positionality prior to conducting this study and, more importantly, performing analyses of the

data. In order to better understand my role as a researcher instead of as a practitioner, I

conducted preliminary fieldwork with SPRUCE ELA Program Instructors and administrators, as

described in Chapter 1. From that fieldwork, I learned that terms and concepts that may be

perceived as having universal definitions, such as "formative assessment," may vary

significantly from teacher to teacher. As I result, I was aware of potential differences in the use

of terms for this research, and attempted to mitigate potential issues in understanding by asking

teachers for clarification when they described practices that they reported using with their

students. For example, although a "Round Robin" discussion may have been a strategy I had

remembered from my own teacher training, when a participant for this study mentioned it in her

interview, I asked her to explain what she meant by those words. By being cautious about

making assumptions and asking teachers for further explanation, even if we ultimately had

shared understanding of a phrase, concept, or strategy, I was able to better learn about the teacher

participants' assessment practices and was more likely to set aside my own biases.

Data from the quantitative phase of this study (background questionnaires and ELATE checklists) were entered into Microsoft Excel and IBM SPSS Statistics, v. 22. Descriptive statistics (e.g., frequencies, means, and standard deviations) were calculated, as were chi-square tests, an adjusted one-way repeated measures analysis of variance (ANOVA) with post hoc analysis, and Pearson correlations for the ELATE checklist items (Leech, Barrett, & Morgan, 2011; Morgan, Leech, Gloeckner, & Barrett, 2013; Sprinthall, 2007). The results of these analyses are presented in Chapter 4 and discussed in Chapter 5. Participants' responses to open-ended items of this phase were reviewed and coded for emergent themes, and are reported in this dissertation as appropriate.

Reliability statistics (Cronbach's alpha and inter-item correlations) were attempted with these quantitative checklist data. However, they were not able to be calculated due to the design of the ELATE checklist. For example, if an ENG 135 Instructor did not use a practice that week, they were instructed to not rate the effectiveness of that practice, so those data were missing by design. Another example were the responses from two of the teachers, who stated that their students were all higher-performing, and therefore they only rated the effectiveness for the practices they reported as using for this performance level and not for average- and lower-performing ENG 135 students. Additionally, teachers' reported practices varied from week to week, even for the same individual, which also impacted the ability to analyze test-retest reliability. To try to correct for these issues, multiple imputation of missing data was attempted, but SPSS calculation errors still prevented these analyses from being conducted.

For the qualitative phase of the study, all interviews were audio recorded with the participants' permission, and were transcribed verbatim into Microsoft Word format with the aid of Dragon NaturallySpeaking, v. 13. I chose to transcribe the interviews myself, since I had

conducted them, had first-hand knowledge of what was being discussed and expertise in the interview subject, and had the advantage of interacting verbally and nonverbally with the participants, benefits noted by Halcomb and Davidson (2006) for researchers transcribing their own data.

Once transcribed, the qualitative data were reviewed and coded for emergent themes (Creswell, 2008; Stake, 2010), which are summarized and reported in Chapter 4 and discussed in Chapter 5. The interviews were analyzed in the order in which they were conducted, after all interviews and transcriptions had been completed. This allowed for more time between when the discussions occurred and when each conversation was revisited, and therefore permitted a more careful and less biased reading of each transcript due to distance from my initial impressions and reactions to what was said. An open coding approach in which I reviewed each transcript line-by-line was used; although some of the codes were associated with particular interview questions (e.g., "ELATE represent?" for the fourth question on the protocol, which asked "In thinking about the assessment practices you reported using across all of the checklists, how representative are they of what typically takes place over an ENG 135 semester?"), or were influenced by my own knowledge of the relevant literature (e.g., "FA for modifying practice"), the majority of the codes emerged from the transcript data. All of the codes were documented in a Microsoft Excel spreadsheet, and an explanation for each code was recorded in this codebook.

After the first interview was coded, the emergent themes in the codebook were organized by similar concepts into categories, which were either labeled with an existing code, or the group of codes was given a new label. These categories were flexible, and as coding continued, the codes were redistributed and/or the category names were revised, as needed. As the analysis of each transcript occurred, the developing codebook was used as a reference to ensure that the

codes being used were consistent in their meaning and to minimize the creation of redundant

codes. At the end of the analysis of the transcripts, a total of 367 codes had emerged, and they

were sorted into 32 categories. In order to determine which codes were appropriate for inclusion

in this dissertation, the research questions for this study were revisited, and the relevant

categories and their respective codes were sorted by research question in the codebook. Finally,

the interview data that had been identified with these codes were re-examined to ensure that the

coding was appropriate and that the content from the transcripts was relevant.

The findings from the second, qualitative phase of the study helped to explain the initial

quantitative results in more depth. Although the qualitative phase was originally intended to

address the second and fourth research questions (i.e., how teachers made decisions about the use

and determined the effectiveness of their assessment practices), these data were also used to

supplement the quantitative data that addressed the other two research questions in order to

provide a more valid and appropriate representation of the assessment practices of the teachers

that participated in this study. Chapter 4 presents the relevant study findings, organized by

research question.

Chapter 4: Results

This research was conducted to explore (a) the assessment practices used by secondary ELA teachers, (b) why they choose to use these practices, (c) their perceived effectiveness of these practices, and (d) how they determine the effectiveness of these practices. Results from the quantitative and qualitative analyses, conducted with the ELATE checklist and interview data, respectively, are presented in this chapter along with the research questions they were intended to address, and discussion of these mixed methods findings are examined in Chapter 5. Graphs summarizing teachers' reported use of the ELATE checklist items, their perceptions of the effectiveness of these assessment practices, and comparing use and perceived effectiveness of these items are presented as Appendices F, G, and H, respectively. In addition, snapshots or summaries of these findings for each teacher, their reported use of assessment practices, and their thinking behind the use of these practices are included as Appendix I.

Characteristics of Participants and Their Contexts

Seven SPRUCE ELA Instructors who were teaching ENG 135 in Spring 2015 participated in this study. These individuals were secondary ELA teachers in high schools across one U.S. state, and each worked in a unique context, which may have impacted their self-reported use and perceptions of the effectiveness of their assessment practices. The characteristics of each teacher and the school and school district in which she/he was teaching in Spring 2015 and data about the student population at these schools were primarily collected via the Teachers' Background Questionnaire (Appendix C) and from the National Center for Education Statistics (NCES). Additional information about the context in which these teachers served as ENG 135 Instructors was learned from email exchanges and interviews with the teachers.

Teacher characteristics. Detailed information about the teacher participants'

characteristics, as of Spring 2015, are presented in Table 1. Of the seven teacher participants, six

were female and one was male. Their total years of teaching experience ranged from 10 to 27

years ($M = 18.14$, $SD = 6.23$), with between 3 and 15 years ($M = 8.14$, $SD = 3.77$) of teaching in

the SPRUCE program. Four of the teachers were Instructors for one SPRUCE ELA section in

2014-2015 academic year, one taught two sections, and two taught three sections. The number of

SPRUCE sections impacted the total number of ENG 135 students for which each teacher was

responsible. Although the overall average number of SPRUCE students per section was

approximately 16, the total number of ENG 135 students for each teacher ranged from 14 to 51

($M = 28.14$, $SD = 16.39$).

Table 1

Characteristics of Teachers, Spring 2015

Teacher	Gender	Total Years Teaching	Years Teaching in SPRUCE ELA Program	SPRUCE Sections Taught in 2014-15	Average Number of SPRUCE Students/ Section	Number of Students in SPRUCE ENG 135, Spring 2015	Total Number of Students in 2014-15
1. Mia	Female	17	7	3	18	50	75
2. Helen	Female	27	15	3	16	51	102
3. Allison	Female	25	10	1	17-20[a]	17	95
4. Charlotte	Female	20	6	1	17	17	101
5. Rachel	Female	15	3	1	12	16	76
6. Krystal	Female	13	9	2	17	32	110
7. Tony	Male	10	7	1	16	14	105

[a] Values indicated by Allison on Teachers' Background Questionnaire.

All of the teachers were responsible for instructing other courses besides ENG 135 at their respective schools in the 2014-2015 academic year. These courses included those for 9th grade students (e.g., English 9, Academic Intervention Services), writing courses for 11th and 12th graders, Advanced Placement (AP) offerings, and electives (e.g., Theatre Arts, Film Studies, and Mythology). The number of sections of other courses taught by these teachers ranged from 2 to 6, with an average of 4 sections taught in addition to ENG 135.

School district characteristics. The school districts in which the teacher participants worked varied in size in terms of numbers of schools and students they served, as shown in Table 2. Student/teacher ratios ranged from 10.73 to 1.0 in Helen's district to 14.88 to 1.0 in Tony's district ($M = 12.86$, $SD = 1.27$). The number of students in each district for whom English was not a first language also varied: Charlotte's district had the highest percentage of these students at 3.20%, and Rachel's district had the lowest at 0.07%. The percentage of students receiving specialized instruction and related services, as indicated by the number of students with Individualized Education Programs (IEPs), ranged from 10.98% (Allison's district) to 20.30% (Mia's district).

Table 2

Characteristics of Teachers' School Districts

Teacher	Total Schools in District	Total Students in District	Classroom Teachers (FTE)	Student/ Teacher Ratio	ELL Students (% of Total)	Students with IEPs (% of Total)
1. Mia	9	4,832	374.35	12.91:1.0	22 (0.46)	981 (20.30)
2. Helen	10	6,404	597.06	10.73:1.0	116 (1.81)	780 (12.18)
3. Allison	6	4,236	319.74	13.25:1.0	24 (0.57)	465 (10.98)
4. Charlotte	5	3,092	257.40	12.01:1.0	99 (3.20)	607 (19.63)
5. Rachel	4	1,452	112.55	12.90:1.0	1 (0.07)	235 (16.18)
6. Krystal	13	7,227	542.27	13.33:1.0	120 (1.66)	1,333 (18.44)
7. Tony	6	4,761	320.04	14.88:1.0	19 (0.40)	658 (13.82)

Notes. FTE = full-time-equivalent. ELL = English language learners. IEP = individualized education program. Data from Common Core of Data (CCD) Public school data, 2013-2014 school year (National Center for Education Statistics, 2016).

School characteristics. The teachers who participated in this study were employed in seven different schools across New York State. As shown in Table 3, the locale (i.e., urbanicity) and number of students at each school varied, although it should be noted that urban schools were not represented in this sample. Excepting Rachel's school, which served students in Grades 7-12, all schools where the teacher participants were employed were at the high school level, or served Grades 9-12.

The proportion of students eligible for free or reduced-price lunch provides a proxy measure for the concentration of low-income students within a school (National Center for Education Statistics, 2015). By this indicator, Rachel's school had the greatest percentage of low-income students, with almost 38% eligible for free or reduced-priced lunches; students at

Krystal's school were at 25.41%. In contrast, students at Helen's school were the most affluent

according to this measure, with only 2.47% of students eligible for free or reduced-price lunches.

Table 3

Characteristics of Teachers' Schools

Teacher	School Locale	Grade Levels of School	Total Students in School	Enrollment by Gender		Students Eligible for Free/Reduced-price Lunch (% of Total)
				Male	Female	
1. Mia	Suburb: Large	9-12	1,659	858	801	124 (7.47)
2. Helen	Suburb: Large	9-12	2,186	1,149	1,037	54 (2.47)
3. Allison	Rural: Fringe	9-12	1,526	763	763	125 (8.19)
4. Charlotte	Suburb: Large	9-12	1,029	542	487	152 (14.77)
5. Rachel	Rural: Fringe	7-12	716	377	339	272 (37.99)
6. Krystal	Suburb: Large	9-12	2,263	1,163	1,100	575[a] (25.41)
7. Tony	Suburb: Large	9-12	1,600	788	812	99 (6.19)

Note. Data from CCD Public school data, 2013-2014 school year (National Center for Education Statistics, 2016). [a] Data are from 2012-2013 school year, as 2013-2014 data did not meet NCES quality standards.

Student characteristics. As shown in Table 4, the ethnic background of students in the

teachers' schools was predominately White, non-Hispanic (82.71% overall). In Helen's school,

over one quarter of the students (580, or 26.53%) identified as Asian/Pacific Islander. While the

majority of students in Charlotte's school (69.48%) identified as White, non-Hispanic, 13.12%

identified as Hispanic, and 10.40% identified as Asian/Pacific Islander. Krystal's school had the

largest Black, non-Hispanic student population (212, or 9.37%). However, she commented in our

interview that the students in her ENG 135 courses were "not representative of the high school as

a whole, which is something that frustrates me. […] I would like to see better representation, um, than we currently have" (interview dated July 9, 2015). According to the six Instructors that were interviewed, the students in their SPRUCE ENG 135 classes were predominantly White.

Table 4

Enrollment by Race/Ethnicity of Students in Teachers' Schools

Teacher	Total Students in School	Enrollment by Race/Ethnicity (% of Total Students)					
		Amer. Indian/ Alaska Native	Asian/ Pacific Islander	Hispanic	Black, non-Hispanic	White, non-Hispanic	Two or More Races
1. Mia	1,659	1 (0.06)	24 (1.45)	63 (3.80)	7 (0.42)	1,545 (93.13)	19 (1.15)
2. Helen	2,186	2 (0.09)	580 (26.53)	49 (2.24)	13 (0.59)	1,535 (70.22)	7 (0.32)
3. Allison	1,526	4 (0.26)	116 (7.60)	35 (2.29)	53 (3.47)	1,309 (85.78)	9 (0.59)
4. Charlotte	1,029	4 (0.39)	107 (10.40)	135 (13.12)	53 (5.15)	715 (69.48)	15 (1.46)
5. Rachel	716	3 (0.42)	2 (0.28)	21 (2.93)	7 (0.98)	648 (90.50)	35 (4.89)
6. Krystal	2,263	9 (0.40)	112 (4.95)	63 (2.78)	212 (9.37)	1,831 (80.91)	36 (1.59)
7. Tony	1,600	5 (0.31)	41 (2.56)	25 (1.56)	18 (1.13)	1,498 (93.63)	13 (0.81)

Note. Data from CCD Public school data, 2013-2014 school year (National Center for Education Statistics, 2016).

All of the ENG 135 students taught by the teacher participants in this study were in 12th grade. The prerequisites for the students to be in this course varied by school. For example, Mia said of her students, "some of them come from the honors track, some of them come from the AP track, and some of them do come from our regular Regents track. So, um, you do get a mix, in, in the classroom." She added that all students were required to have a "minimum 85 [average] to get in to the [ENG 135] class" (interview dated July 17, 2015). Helen reported similar

requirements at her school, and shared via email that to be in the ENG 135 class, students "need an A average in English for the first three years of high school" (email dated May 7, 2015). In Allison's school, ENG 121—the SPRUCE course on academic writing—was the prerequisite for students who wanted to take ENG 135, but she also noted in our interview that "most of these students have been in an honors track all the way through" (interview dated July 1, 2015). In contrast, Charlotte stated that her school had an open enrollment policy, so any student who was interested in taking ENG 135 was permitted to do so (interview dated June 29, 2015).

Students' reasons for taking the SPRUCE ENG 135 course also varied. Mia indicated that the concurrent enrollment factor was valued by students: "generally the college credit is very attractive to them, and the fact that they can get college credit without taking a, one major exam, not having everything hang on that one AP exam is, uh, a pretty popular choice because of that" (interview dated July 17, 2015). Helen recounted that community, district, and parental expectations for student success were very high, and that "by the time we get to senior year, pretty consistently, half the senior class is in AP and [SPRUCE]," and that her students in ENG 135 are "interested in good grades, and transferring credits, and all that sort of stuff" (interview dated June 26, 2015). Allison also brought up differences between students who chose to take SPRUCE courses rather than those with the AP designation:

> a lot of students who take AP […] are often taking it because they want to see AP on their transcript, and anticipate that they're applying to schools that will recognize it — the AP brand, and that might not recognize the [SPRUCE] brand. […] on the other hand, I think a lot of the kids who end up in the [SPRUCE] courses rather than the AP courses have been recommended by teachers who saw that they were interested in ideas, and, um, and that they are going to be going

into fields where they really need to have, um, good academic research skills, and

good critical reading skills in a much different way than you would have if you

were taking a literature course. (interview dated July 1, 2015)

Allison added that a difference between AP and ENG 135 students was that the latter

are much more willing to push back, which is like, sort of more what I prefer, and

kids who are willing to question, and, I don't know — engage with ideas in a

different way. They're, they're less willing to be, sort of, passive consumers of

information, and more interested in engaging with, fighting with, you know,

grappling with ideas. (interview dated July 1, 2015)

Students' willingness to push themselves by taking SPRUCE ELA Program courses rather than

other senior-level English courses was also described by Charlotte, who noted that students make

comments such as "'I know it's heavy writing, and I want to become a better writer'" (interview

dated June 29, 2015).

Most of the interviewed teachers noted differences between their ENG 135 students and

the other students in their school, especially in terms of work ethic. For example, Mia shared that

these students "are motivated to work, for the most part" and that "they understand the

importance of what they're doing," including applying knowledge and skills gained from this

course to other subjects. She continued by remarking: "they become a different type of student,

they become — become a different type of person than the kids that are taking a regular English

course that just focuses on, you know, symbolism in literature, and that type of thing" (interview

dated July 17, 2015). Rachel noted that students in SPRUCE ELA Program courses tended to

have more drive and initiative than their peers, and that "they were always the kids who were

receiving awards, and high honors and — so, you know…school and education mattered to them" (interview dated June 25, 2015).

The ENG 135 Instructors' perceptions of their students' performance levels varied. Both Mia and Helen said that their students were all higher-performing due to prerequisites for the ENG 135 course in their school. As a result, their completed ELATE checklists only rated the effectiveness of the assessment practices that they reported as using for this student performance level. It is important to note that several of the teacher participants taught their students in previous courses over the years, such as in an English 11 Honors course prior to SPRUCE ENG 135, or in SPRUCE ENG 121 in the prior fall semester. For instance, Krystal remarked that her familiarity with students she had taught before allowed her to "get a sense of what's effective for particular kids" in terms of her assessment practices (interview dated July 9, 2015), and Allison said that she was able to remind students of what they had learned in their previous courses with her (interview dated July 1, 2015).

Participants' Use of Assessment Practices

The first research question for this study, *What assessment practices do secondary ELA teachers use?*, was addressed primarily through statistical analyses of data collected from teacher participants' ELATE checklists. Supporting data were culled from interviews and email exchanges with the teachers.

Completion of ELATE checklists. It was originally anticipated that the SPRUCE ELA Program ENG 135 Instructor participants would complete an ELATE checklist (Appendix D) once a week for four consecutive weeks. However, due to varying dates when participants joined the study, the timing of the study at the end of the academic year, and students' events such as AP and Regents exams and senior celebrations, response numbers and times fluctuated. Table 5

shows the dates for which the checklists were completed by the seven teachers who participated in this study. Mia, Helen, Allison, Charlotte, and Rachel all began checklists in the beginning of May, 2015. Krystal began a week later, and Tony joined the study at the end of the semester. Both Allison and Charlotte completed five checklists; Mia, Rachel, and Krystal completed four; and Helen and Tony completed three. Overall, participants completed a total of 28 checklists over the span of six weeks, or an average of four ELATE checklists each.

Table 5

Weeks ELATE Checklists Were Completed by Teachers, Spring 2015

Teacher	Week 1 May 1st- May 7th	Week 2 May 8th- May 14th	Week 3 May 15th- May 21st	Week 4 May 22nd- May 28th	Week 5 May 29th- June 4th	Week 6 June 5th- June 11th
1. Mia	▓	▓	▓	▓		
2. Helen	▓			▓	▓	
3. Allison	▓	▓	▓	▓	▓	
4. Charlotte	▓	▓	▓	▓	▓	
5. Rachel	▓	▓	▓	▓		
6. Krystal		▓	▓	▓		▓
7. Tony				▓	▓	▓

Number of times each practice was reported as used. As presented in Table 6, the number of times each practice on the ELATE checklist was reported as used varied each week from the beginning of May through mid-June. (A graphical representation of the use of each practice by item is included as Appendix F.) For example, Week 1 (May 1-7) had the largest number of reported uses of questioning and discussion assessment practices. Weeks 1 and 4 (May 22-28) were when the majority of feedback on students' writing was given. According to the interviewed teachers through discussion of their practices, ENG 135 course syllabi that several of the teachers shared with me, and the description of the course from the university

collaborating with SPRUCE, all of the study participants were teaching a similarly-structured version (i.e., the same content, with SPRUCE-approved variations on supplemental texts used) of the ENG 135 course. The shifts in practice reported through the ELATE checklists correspond to teachers' instruction, as concepts were introduced and talked about, students wrote brief papers about their thinking on the topic (called "think papers" or "response papers"), and these ideas were developed into more formal work. In the case of SPRUCE ENG 135, the course ended with students creating a multimedia project rather than a terminal written paper.

Teacher participants were asked to write in and rate the effectiveness of any additional questioning, discussion, or feedback assessment practices that they used with their students in the "Other Practices" section of the ELATE checklist. The particular practices that were submitted were specific to each teacher and their lessons, and the responses given often related more to instructional rather than assessment practices. For example, one teacher included "had students read aloud from a play; involved all class members," while another wrote "Visual presentation of all unit ideas' relationship to each other." All teacher-provided practices were only reported as used for one week, and were indicated as being used four times or fewer. Of the 18 total responses in this section of the ELATE checklists across the six weeks, 14 were from the same teacher. For the purposes of this dissertation, only the 25 assessment practices that were included as items on the ELATE checklist are presented, analyzed, and discussed, as the write-in responses were not related to the questioning, discussion, and feedback assessment practice domains that were being examined in this study.

Table 6

Number of Times Each Item Was Reported as Used by Week

Item	In teaching ENG 135 this week (last Friday through Thursday), I…	Number of Times Item Reported as Used by Week						Total Times Used
		Week 1 5/1-5/7	Week 2 5/8-5/14	Week 3 5/15-5/21	Week 4 5/22-5/28	Week 5 5/29-6/4	Week 6 6/5-6/11	
	Questioning and Discussion							
1	Asked students questions to determine how well they understood a concept, idea, or strategy.	57	25	50	44	15	20	**211**
2	Asked questions of individual students by name.	123	53	18	27	35	35	**291**
3	Asked questions of the class as a whole.	90	36	36	63	28	32	**285**
4	Asked questions specifically of students I felt were not paying attention.	9	3	10	15	6	5	**48**
5	Asked questions of students I thought would be more likely to respond well.	24	11	10	32	15	22	**114**
6	Asked questions of reticent students to help improve their participation.	17	10	22	16	5	11	**81**
7	Asked questions requiring brief responses (e.g., a word or phrase).	93	13	35	30	20	35	**226**
8	Asked questions requiring more elaborated responses (e.g., a few sentences).	60	50	31	73	37	40	**291**
9	Asked questions intended to stimulate a general discussion.	151	47	29	63	13	25	**328**
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	7	5	5	9	2	0	**28**
11	Used large group (5 or more students) discussion to determine how well students understood a concept, idea, or strategy.	45	8	29	9	0	2	**93**
12	Used whole-class discussion to determine how well students understood a concept, idea, or strategy.	57	27	19	21	4	1	**129**

Table 6 Continued

Item	In teaching ENG 135 this week (last Friday through Thursday), I…	Number of Times Item Reported as Used by Week						Total Times Used
		Week 1 5/1-5/7	Week 2 5/8-5/14	Week 3 5/15-5/21	Week 4 5/22-5/28	Week 5 5/29-6/4	Week 6 6/5-6/11	
	Feedback							
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	106	4	5	31	2	4	**152**
14	Reviewed and gave feedback (oral or written) on students' dialogic journals.	11	1	0	16	0	0	**28**
15	Reviewed and gave feedback (oral or written) on students' portfolios of their writing.	4	0	0	16	0	0	**20**
16	Reviewed and gave feedback (oral or written) on students' *draft* versions of a writing assignment.	10	1	2	71	0	0	**84**
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	10	3	5	36	1	0	**55**
18	Used a rubric or rubrics for *informal* feedback to students on their writing.	2	2	2	0	1	0	**7**
19	Used a rubric or rubrics for *formal* feedback to students on their writing.	2	3	5	37	0	0	**47**
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	6	2	8	76	3	0	**95**
21	Gave students *general or holistic* written feedback about their writing, such as comments at the end of an assignment about how they did overall.	6	3	5	87	13	4	**118**
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	4	2	55	58	0	2	**121**
23	Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing.	6	3	55	76	1	0	**141**
24	Gave oral feedback to multiple students at once about their writing (e.g., discussing strategies with groups or the entire class).	10	2	1	10	2	0	**25**
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	6	3	5	15	5	3	**37**

Across all six weeks, there were great differences in the total number of times that each assessment practice on the ELATE checklist was reported as used by teacher participants. These variances were largely a function of the different numbers of respondents each week. As shown in Table 5, the number of teachers, as well as the individuals who were submitting ELATE checklists each week varied: there were five participants in Weeks 1-3, seven participants in Week 4, four participants in Week 5, and two participants in Week 6. In order to examine the total number of practices that were reported as used, aggregated reports across all six weeks are presented in Table 7, and are sorted from most times used to least times used.

Most frequently used practices. The six most frequently used assessment practices were related to questioning and discussion; the practice that was reported as most used by teachers (328 times) was item 9, "Asked questions intended to stimulate a general discussion." When asked why she might use these particular practices so often, Helen replied that class discussion encourages students to ask questions, learn from each other, and make connections. She further explained:

> So the classroom discussion is — I think for a lot of students, it generates ideas that they have, um, that they didn't even perhaps know that they possess, but when they hear it from another student, or when somebody says, 'But what about this? Is that…?' And you — and then the student can say, 'Oh yeah, that is,' you know. So it's empowering, and it — it opens minds, and, um — […] so many times when you think of your reaction to something, or your idea about something, you think you're the only one, or, you know, [you might say] 'I don't know what other people think, but…', you know. And then you find out, 'Oh,

you're not alone,' you know. 'Other people do think this way.' (interview dated

June 26, 2015)

A similar sentiment about the value of group discussion was raised by Rachel. She remarked

about her experiences with students in class discussions, "When they heard from their peers, it

was a lot easier sometimes for them than coming from the instructor […] it means more, it makes

more sense coming from someone who's in their own age group" (interview dated June 25,

2015).

Allison was another teacher who reported using discussion very often in her ENG 135

class. She stated that she used guiding questions to help the students both stay focused and

wrestle with challenging course material. In addition, she used the information she learned from

asking students questions to assess student learning and change her teaching:

very often, it gives me […] a pretty good idea of whether or not they understood

[*laughs a little*] basic concepts. And, […] it helps me figure out what gaps I need

to fill in. And sometimes, though, I'll find that they're — they've really handled it

well, and then we can move on to something else. But, sometimes, you know, on

Bloom's taxonomy, we're pretty low if we're dealing with something really

challenging. [Based on students' understanding,] I might adjust what I'm planning

to do for the next class period, or what I expect them to do for homework for the

next class period, or what I have them talking about on [the online discussion

boards]. (interview dated July 1, 2015)

Other teachers also described using questioning to quickly assess students' knowledge. For

example, Charlotte stated, "it helps me see how much they are thinking about the work"

(interview dated June 29, 2015), while Krystal recounted how she used observation and

questioning of groups to determine the status of multiple students at once, and individual

questioning to assess the learning of a particular student (interview dated July 9, 2015). These

narratives aligned with teachers' frequent reported use (211 times) of item 1, "Asked students

questions to determine how well they understood a concept, idea, or strategy."

Table 7

Total Number of Times Each Item Was Reported as Used, from Most to Least

Item	In teaching ENG 135 this week (last Friday through Thursday), I…	Item Type	Total Times Used
9	Asked questions intended to stimulate a general discussion.	Q & D	328
2	Asked questions of individual students by name.	Q & D	291
8	Asked questions requiring more elaborated responses (e.g., a few sentences).	Q & D	291
3	Asked questions of the class as a whole.	Q & D	285
7	Asked questions requiring brief responses (e.g., a word or phrase).	Q & D	226
1	Asked students questions to determine how well they understood a concept, idea, or strategy.	Q & D	211
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	F	152
23	Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing.	F	141
12	Used whole-class discussion to determine how well students understood a concept, idea, or strategy.	Q & D	129
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	F	121
21	Gave students *general or holistic* written feedback about their writing, such as comments at the end of an assignment about how they did overall.	F	118
5	Asked questions of students I thought would be more likely to respond well.	Q & D	114
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	F	95
11	Used large group (5 or more students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	93
16	Reviewed and gave feedback (oral or written) on students' *draft* versions of a writing assignment.	F	84
6	Asked questions of reticent students to help improve their participation.	Q & D	81
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	F	55
4	Asked questions specifically of students I felt were not paying attention.	Q & D	48
19	Used a rubric or rubrics for *formal* feedback to students on their writing.	F	47
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	F	37
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	28
14	Reviewed and gave feedback (oral or written) on students' dialogic journals.	F	28
24	Gave oral feedback to multiple students at once about their writing (e.g., discussing strategies with groups or the entire class).	F	25
15	Reviewed and gave feedback (oral or written) on students' portfolios of their writing.	F	20
18	Used a rubric or rubrics for *informal* feedback to students on their writing.	F	7

Notes. Q & D = Questioning and Discussion. F = Feedback.

Least frequently used practices. As shown in Table 7, four of the five assessment practices that teachers reported as using the least were all related to giving feedback to students on their writing. Item 18, "Used a rubric or rubrics for **informal** feedback to students on their writing," was only reported as used seven times. In our interview, Charlotte commented that she preferred to use checklists of grading criteria instead of rubrics for assessing students' work, because "I feel like it's more accurate measurement, and it gives students a little bit more pointed feedback, because it's just not development" (interview dated June 29, 2015). Krystal shared that she uses rubrics for individual feedback to students on their final papers, but that she always gives written comments as well (interview dated July 9, 2015). As previously noted, however, since students were completing multimedia projects for the end of the course rather than writing final papers, teachers' assessment practices likely varied in accordance with the types of assignments they were evaluating. This would also account for the low numbers of reports of teachers' reviews of and providing feedback on students' portfolios of their writing (item 15).

Other practices that were reported as less frequently used related to giving students feedback about their writing (item 24) and determining students' understanding of a concept, idea, or strategy (item 10) in situations where multiple students were involved (25 times and 28 times of reported use, respectively). However, these findings make sense, as all six of the interviewed teachers reported preferring the use of individual conferencing with students to discuss their writing. Several participants said that they required students to schedule appointments with them to review feedback that they had provided to students on draft and final versions. Others also mentioned meeting as needed with students while they were thinking about and developing their ideas prior to putting them on paper.

Representativeness of practices. Because the study occurred over several weeks at the end of the semester, all six interviewed teachers were asked if the assessment practices that they reported using through the ELATE checklists were representative of what typically occurs throughout the SPRUCE ENG 135 course. Participants affirmed that these self-reported practices were typical of what would normally be done in this course, with several adding the caveat that there was more of a focus on discussion than writing at the end of the semester due to the final multimedia project. Additionally, all interviewed teachers stated that the questioning, discussion, and feedback practices that they reported using through the ELATE checklists with their ENG 135 students carried over into use with the other courses that they taught. The extent to which the teachers used similar practices depended primarily on the level of the students and the amount of control the teachers had over the curricula. For example, teachers tended to report similar or identical use of assessment practices in AP and honors courses, especially with 12th-grade students, and less flexibility with assessment practices for courses for younger students that were driven by the Common Core State Standards.

Chi-square analysis of use. To determine if there were statistically significant differences between actual and expected (i.e., average of the data)[2] use of the ELATE checklist items over time, a 1 x *k* (goodness-of-fit) chi-square was computed that compared the frequency of occurrence of teachers' reported use for each item across the six weeks in which the data were reported.

---

[2] In order to conduct a chi-square test, an even distribution across cases, called "expected use," must be assumed. For example, a total of 150 reports of use of a practice across Weeks 1-6 would result in an "expected use" each week of 25 times. Although this assumption of "expected" values is statistically necessary to conduct the analysis, it is important to note that this does not mean that these averages, or equal frequencies, are the values anticipated by the researcher to be found in the actual use of these practices.

Table 8 presents the frequencies (mean, standard deviation, and range) of reported use by all teachers, the chi-square value, and the p value for each item. These data show the amount of difference between the actual and expected reports of use of each individual item across the six weeks. For example, item 1, "Asked students questions to determine how well they understood a concept, idea, or strategy," was reported as used a mean of 35.17 times across the six weeks, with a minimum of 15 times and a maximum of 57 times. Given a χ^2 value of 43.08 with 5 degrees of freedom, the difference between observed and expected values had a p value of 3.57E-08, which is significant at the $p < .0001$ level.

A significant difference was found between the observed and expected values for almost all items; the exception was item 18, "Used a rubric or rubrics for ***informal*** feedback to students on their writing" ($\chi^2(5) = 4.14$, $p = .5290$). This finding is congruent with the low number of reports of use of rubrics by teacher participants, as described earlier in the section titled "Least frequently used practices." As rubrics were seldom used, there was little variance across the weeks, so expected and actual values were much closer than for other items.

Table 8

Total Item Use Frequency Analysis: Descriptives and Chi-square Tests

Item	*M*	*SD*	Range (min, max)	Chi-square (*df* = 5)	*p* value
1	35.17	17.41	42 (15, 57)	43.08	3.57E-08****
2	48.50	38.28	105 (18, 123)	151.08	7.85E-31****
3	47.50	24.21	62 (28, 90)	61.72	5.37E-12****
4	8.00	4.29	7 (3, 10)	11.50	.0423*
5	19.00	8.53	22 (10, 32)	19.16	.0018**
6	13.50	6.02	17 (5, 22)	13.44	.0196*
7	37.67	28.48	80 (13, 93)	107.66	1.28E-21****
8	48.50	15.78	42 (31, 73)	25.68	.0001***
9	54.67	50.37	138 (13, 151)	232.01	3.96E-48****
10	4.67	3.27	9 (0, 9)	11.43	.0435*
11	15.50	17.74	45 (0, 45)	101.52	2.53E-20****
12	21.50	20.12	56 (1, 57)	94.12	9.16E-19****
13	25.33	41.01	104 (2, 106)	331.87	1.40E-69****
14	4.67	7.03	16 (0, 16)	53.00	3.36E-10****
15	3.33	6.41	16 (0, 16)	61.60	5.67E-12****
16	14.00	28.18	71 (0, 71)	283.57	3.40E-59****
17	9.17	13.61	36 (0, 36)	101.11	3.09E-20****
18	1.17	0.98	2 (0, 2)	4.14	.5290
19	7.83	14.41	37 (0, 37)	132.62	6.62E-27****
20	15.83	29.61	76 (0, 76)	276.94	9.06E-58****
21	19.67	33.18	84 (3, 87)	279.83	2.16E-58****
22	20.17	28.19	58 (0, 58)	197.00	1.24E-40****
23	23.50	33.27	76 (0, 76)	235.47	7.19E-49****
24	4.17	4.58	10 (0, 10)	25.16	.0001***
25	6.17	4.49	12 (3, 15)	16.35	.0059**

Notes. df = degrees of freedom. †*p* < .1. **p* < .05. ***p* < .01. ****p* < .001. *****p* < .0001.

Chi-square analysis of use by function. To further analyze teachers' reported use for the assessment practices listed in the ELATE checklist, the items were categorized by their function. These categories were as follows: *questioning* (items 1-9), *discussion* (items 10-12), *writing process feedback* (items 13-17), *rubric feedback* (items 18-19), *written feedback* (items 20-23), and *conferencing feedback* (items 24-25). A 1 x *k* (goodness-of-fit) chi-square was computed comparing the frequency of occurrence of teachers' reported use for the items in each category within each week for which data were reported.

Table 9 presents the weekly frequencies (mean, standard deviation, and range) of reported use by all teachers, the chi-square value, and the *p* value for each function. These data show the amount of difference between the actual and expected reports of use of each of these groups of items within each week. For example, in Week 3, the items related to *written feedback* (#s 20-23, involving giving students targeted written feedback, giving students general or holistic written feedback, writing questions when giving feedback to students to help prompt their thinking, and correcting students' spelling, grammar, and mechanical errors, respectively) were reported as used a mean of 30.75 times that week, with a minimum of 5 times and a maximum of 55 times. Given a χ^2 value of 76.64 with 3 degrees of freedom, the difference between observed and expected values for this group of items had a *p* value of 2.31E-13, which is significant at the $p < .0001$ level.

For all six weeks, a significant difference was found between the observed and expected (i.e., average of the data) values for *questioning* items. For *discussion* items, there were significant differences between observed and expected values for Weeks 1 ($\chi^2(2) = 37.50$, $p < .0001$), 2 ($\chi^2(2) = 21.35$, $p < .01$), and 3 ($\chi^2(2) = 16.45$, $p < .05$) only. A significant difference was found between the observed and expected values for *writing process feedback* items for Weeks 1 ($\chi^2(4) = 269.39$, $p < .0001$), 4 ($\chi^2(4) = 59.71$, $p < .0001$), and 6 ($\chi^2(4) = 16.00$, $p < .05$)

only. For *written feedback* items, Weeks 3 ($\chi^2(3) = 76.64$, $p < .0001$) and 5 ($\chi^2(3) = 25.12$, $p <$ .01) showed significant differences between the observed and expected values. For *rubric feedback* items, a significant difference was found between the observed and expected values for Week 4 only ($\chi^2(1) = 37.00$, $p < .0001$), and rubrics were not reported as used at all during Week 6. The sole function for which there were no significant differences between observed and expected values for all six weeks were items involving *conferencing feedback*. These differences by week make sense, as they correspond to the variances in instructional practices and content foci that the teachers described in the interviews and had planned in their syllabi as part of the course curriculum.

Table 9

Weekly Item Use Frequency Analysis: Descriptives and Chi-square Tests

Week	Item #s	Function	*M*	*SD*	Range (min, max)	Chi-square (*df*)	*p* value
1	1-9	Questioning	69.33	49.00	142 (9, 151)	277.07 (8)	3.10E-55****
	10-12	Discussion	36.33	26.10	50 (7, 57)	37.50 (2)	9.29E-06****
	13-17	Process FB	28.20	43.58	102 (4, 106)	269.39 (4)	1.33E-53****
	18-19	Rubric FB	2.00	0.00	0 (2, 2)	0.00 (1)	1.0000
	20-23	Written FB	5.50	1.00	2 (4, 6)	0.55 (3)	0.9998
	24-25	Conferencing FB	8.00	2.83	4 (6, 10)	1.00 (1)	0.9982
2	1-9	Questioning	27.56	19.38	50 (3, 53)	109.02 (8)	6.04E-20****
	10-12	Discussion	13.33	11.93	22 (5, 27)	21.35 (2)	.0063**
	13-17	Process FB	1.80	1.64	4 (0, 4)	6.00 (4)	.6472
	18-19	Rubric FB	2.50	0.71	1 (2, 3)	0.20 (1)	1.0000
	20-23	Written FB	2.50	0.58	1 (2, 3)	0.40 (3)	.9999
	24-25	Conferencing FB	2.50	0.71	1 (2, 3)	0.20 (1)	1.0000
3	1-9	Questioning	26.78	13.12	40 (10, 50)	51.44 (8)	2.15E-08****
	10-12	Discussion	17.67	12.06	24 (5, 29)	16.45 (2)	.0363*
	13-17	Process FB	2.40	2.51	5 (0, 5)	10.50 (4)	.2317
	18-19	Rubric FB	3.50	2.12	3 (2, 5)	1.29 (1)	.9957
	20-23	Written FB	30.75	28.03	50 (5, 55)	76.64 (3)	2.31E-13****
	24-25	Conferencing FB	3.00	2.83	4 (1, 5)	2.67 (1)	.9535
4	1-9	Questioning	40.33	21.49	58 (15, 73)	91.64 (8)	2.16E-16****
	10-12	Discussion	13.00	6.93	12 (9, 21)	7.38 (2)	.4958
	13-17	Process FB	34.00	22.53	55 (16, 71)	59.71 (4)	5.32E-10****
	18-19	Rubric FB	18.50	26.16	37 (0, 37)	37.00 (1)	1.15E-05****
	20-23	Written FB	74.25	12.01	29 (58, 87)	5.83 (3)	.6665
	24-25	Conferencing FB	12.50	3.54	5 (10, 15)	1.00 (1)	.9982
5	1-9	Questioning	19.33	11.69	32 (5, 37)	56.59 (8)	2.17E-09****
	10-12	Discussion	2.00	2.00	4 (0, 4)	4.00 (2)	.8571
	13-17	Process FB	0.60	0.89	2 (0, 2)	5.33 (4)	.7214
	18-19	Rubric FB	0.50	0.71	1 (0, 1)	1.00 (1)	.9982
	20-23	Written FB	4.25	5.97	13 (0, 13)	25.12 (3)	.0015**
	24-25	Conferencing FB	3.50	2.12	3 (2, 5)	1.29 (1)	.9957
6	1-9	Questioning	25.00	11.75	35 (5, 40)	44.16 (8)	5.31E-07****
	10-12	Discussion	1.00	1.00	2 (0, 2)	2.00 (2)	.9810
	13-17	Process FB	0.80	1.79	4 (0, 4)	16.00 (4)	.0424*
	18-19	Rubric FB	0.00	n/a	n/a	n/a	n/a
	20-23	Written FB	1.50	1.91	4 (0, 4)	7.33 (3)	.5011
	24-25	Conferencing FB	1.50	2.12	3 (0, 3)	3.00 (1)	.9344

Notes. df = degrees of freedom. FB = feedback. †*p* < .1. *p < .05. **p < .01. ***p < .001. ****p < .0001. n/a = not applicable.

Participants' Selection of Assessment Practices

The second research question for this study, *How do secondary ELA teachers determine what assessment practices to use and when to use them?*, was addressed through qualitative analysis of interview data. When teacher participants discussed how they chose the assessment practices that they used, the following major themes emerged: (a) the purpose of assessment, (b) the intertwining roles of assessment and instruction, (c) their own knowledge and experience, and (d) the classroom context as factors that influenced their decision making. Other major emergent themes from these interviews included differences between ELA and other subjects and how those play into classroom assessment practices, and how the teacher participants' selection of assessment practices occurs for courses other than SPRUCE ENG 135.

Purpose of assessment. The first major theme, purpose of assessment, emerged as the interviewed teachers reported having various purposes for using particular assessment practices. For example, Mia talked about the amount of written feedback she provided to her students, saying, "It just depends on the time [that the student and I have to be able to conference with each other], and what we're doing at that point" (interview dated July 17, 2015). Several teachers commented on the need to give feedback quickly to students about their writing, such as through comments on think papers or drafts, so that they could incorporate those ideas into more refined versions of their work in a timely fashion. Helen, Krystal, and Charlotte spoke of writing questions and comments on their students' papers as a form of interaction with them that would be likely to prompt further discussion in a one-on-one meeting.

Questioning and discussion practices were also reported to have been used with a purpose by several of the teachers, especially with regard to checking for students' understanding of a topic. Although the term "formative assessment" was not used in the materials for this study that

were shared with teachers, Mia had learned of the phrase through her own education and

research. In our interview, she stated that the most common FA practice she used with her ENG

135 students was discussion. When I asked her to define what "formative assessment" meant to

her, she replied,

> Formative assessment is the assessment that happens before the end […] you kind
>
> of discuss something with [the students], or you have a paper that's not an end-of-
>
> the-unit assessment, and you can assess at that point what students are
>
> understanding, what they're not understanding, and being able to then use that
>
> information to change your teaching, and to change the classroom environment,
>
> or change what needs to be changed in order for kids to really understand what's
>
> going on. (interview dated July 17, 2015)

Although the other teacher participants did not explicitly discuss the term "formative

assessment," Mia's definition seemed to capture elements that were raised by the others. Similar

findings that emerged are presented below, and are organized by the following minor emergent

themes: (a) meeting learners' needs, (b) meeting teaching goals, and (c) fair evaluation of

students' work.

Meeting learners' needs. In their interviews, several teachers discussed how they used

assessment to help their students improve their learning, and how student needs impacted their

selection of assessment practices. Along with checking for students' understanding or

determining how well they were able to apply new knowledge, the teachers also talked about

how they use assessment to close gaps in student learning. During our conversation, Rachel

defined assessment as "a tool […] to gauge understanding […] a guide for the teacher to see

maybe where there are some areas of weakness or areas of strength." She also explained that

assessment also was "an opportunity […] to also dialogue with colleagues, too," and that

teachers at her school had conversations about how to improve learning for their students based

on assessment results (interview dated June 25, 2015). Both Helen and Allison also reported that

they frequently shared ideas with colleagues, especially other SPRUCE teachers, in order to help

address their students' learning needs.

     Meeting teaching goals. For some of the teacher participants, one of the purposes of

selecting an assessment practice was to help students meet the goals that they had in mind for

them. These particular goals were different from course or assignment objectives—which will be

addressed in a later section—in that they were often fluid and open to the teacher's interpretation

about whether or not they were met, and tended to be short-term. Charlotte, Rachel, and Krystal

all talked about goals that they personally planned for their students to achieve. For instance,

Rachel described having expectations of how she thought students should respond in class

discussion, and how students not meeting those goals guided her in making appropriate changes

in her questioning techniques (interview dated June 25, 2015). Krystal shared that she tended to

have students write in a sequence that built up to the final product. Her goal was to have

scaffolding built in to students' writing, and she noted that even if the students did not

understand the process at the time they were going through it, "I'd like them to understand it

eventually" (interview dated July 9, 2015).

     Fair evaluation of students' work. One other purpose of assessment that emerged as a

factor in selection of assessment practices was teachers' interest in being fair in evaluating their

students' work. Many of the teachers discussed their willingness to re-grade students' writing if

they were willing to put in the effort for revision. For Helen, working with students on revisions

led to more student ownership of their work and awareness of the effort they needed to invest to

earn a better grade. She said of her students, "I respect the work that they do, and I will evaluate it fairly […] I'm happy to work with you, and let's see how good we can make this." She continued to explain about her students' revisions: "they should — especially in English! — They should be able to influence the work. They should be able to get a second chance at writing" (interview dated June 26, 2015).

Assessment and instruction. Another major emergent theme underlying participants' selection of assessment practices was the relationship between assessment and instruction. Teachers reported using the information that they learned from their assessments to guide what they did with their students. There were distinctions made between in-class assessments, such as checking for student understanding through listening to students during discussions or reviewing their shorter papers, and higher-stakes assessments, such as those administered by the state. Rachel shared the following thoughts on accountability efforts and their impact on instruction:

I think in this whole data-driven society, it's more — it's sort of — it's just driven that, we're looking at the numbers, we're looking at the numbers, but we kind of need to step away from that, and just say, 'Ok, now — what's relevant, and how — how may I use this to assist students? How may I use this to help me as an instructor?' (interview dated June 25, 2015)

She continued by explaining that she felt that state assessments led to a number as a goal for students; however, they do not allow for teachers to be able to reflect on those scores and adjust their teaching in a timeframe that would help those particular learners. Rachel remarked that she believed SPRUCE courses allowed their instructors more time to reflect on assessment data and apply changes in a timely fashion. She expressed her thinking about this as follows:

I had to step away from 'This is the information I'm getting [from assessment]'
to, um, it's not just about that; it's about 'How does the information I'm getting,'
um, 'translate to what maybe — you know, what I need to say again, what I need
to do again, what, you know, evidence, do I talk to my colleagues about?'
(interview dated June 25, 2015)

Krystal also talked about assessment's role in her teaching, commenting that assessment
"helps me understand if I'm teaching something clearly. Like, if my pedagogy, my methods are
working, in terms of helping kids get at information" (interview dated July 9, 2015). She
explained that she used assessment mainly to determine the progress of individual students, but
that she also used it, in her words, "holistically" to adjust her instructional practices. She
explained:

If I have a lot of students in a given class who aren't getting a particular concept,
well then, that's telling me that perhaps that's a concept I need to revisit, that
something went wrong in my delivery, um, or in my setup, or in my scaffolding,
you know, whatever it is, and that's something I need to revisit. (interview dated
July 9, 2015)

Other participants also discussed how their choice of assessment practices influenced their
teaching, particularly through their use of what they learned and how they applied this
knowledge. Their use of assessment data included making decisions to clarify instruction;
deciding to either repeat, go back, or continue instruction based on students' reaction to a lesson;
and determining whether or not instruction needed to be better differentiated to meet students'
needs.

Teacher knowledge. The knowledge that teachers have gained through their experiences as educators was a third major emergent theme that influenced how they selected which assessment practices to use and when to use them. For the ENG 135 Instructors who participated in this study, this knowledge base comprises the following minor themes that emerged from analysis of the interview data: (a) teachers' intuition, (b) teachers' reflection on their practice, and (c) an awareness of assessment and its role in teaching.

Teachers' intuition. At the time that this study was conducted, all of the teacher participants had been educators for at least 10 years, and had been SPRUCE ELA Program Instructors for at least three years. Of these individuals, Helen had the most teaching experience, having taught for 27 years overall, including 15 years as a SPRUCE Instructor. In our interview, she discussed the kind of thinking about her students that has emerged from her experiences:

> I think all teachers probably, even before the first paper comes in you, you — you're figuring, you know, based on classroom performance and discussion and all that, you're kind of getting an idea of who's going to, you know — who's, who's going to be the better writers, who's going to have more to say, and all of that sort of thing. Um, I think…I think you just do a lot of, um, evaluating — I think you're evaluating all of the time. (interview dated June 26, 2015)

Other teachers expressed similar sentiments, especially when discussing student performance levels in their ENG 135 classes. All of the teachers were able to identify who of their students were lower-, average-, and higher-performing. Their identifications of their students and their reasoning for these choices are further explained in the "Performance levels of SPRUCE ENG 135 students" section later in this chapter.

Teachers' reflection on their practice. The majority of the interviewed teachers explicitly discussed their reflection on their own assessment use, including how these reflective practices developed over the years that they were teaching, and the forms that these reflective practices take. For example, Krystal simply stated, "I'm more reflective on my own practice now, than I was before." She attributed this shift to working with student teachers, and the frequent thinking aloud strategy she uses to explain her actions to them. In her words, "in demonstrating the 'how' and the 'why I do what I do,' it really has made me more cognizant of my practice overall" (interview dated July 9, 2015). Charlotte also explained that her thinking about and reflection on assessment came from her experiences:

> There are some things in the moment; after teaching long enough, you can usually figure out, 'If that's not working, what else can I do right now,' um, 'to try to achieve that?' And again, sometimes it's assessing how well my goal is met versus how well they are understanding or grasping something. There's a — I think a, a lot of multitasking happening with assessment: it's both, um, self-assessment, but also assessment of students. (interview dated June 29, 2015)

As another example, Rachel described using a particular reflection strategy that she had learned from her SPRUCE program training to think about her lessons and how effective they were with her students. She stated that she always kept a copy of her ENG 135 syllabus on hand for taking notes to guide decision making for the next time that she taught the course. Rachel added that her training and experiences in SPRUCE "[have] me very reflective in terms of, you know, how I do things" (interview dated June 25, 2015).

Teachers' awareness of assessment. Several of the participants remarked in their interviews that completing the weekly ELATE checklists influenced their thinking about their

assessment practices due to increased awareness about what tools and strategies they were using with their students. Rachel remarked that as she was completing the checklists each week, she thought to herself, "OK, now, I'm cognizant of this, so I'm now sort of watching […] how many times [I'm] doing this, and what does this mean" (interview dated June 25, 2015). Charlotte shared a similar sentiment, and commented, "your questions [on the ELATE checklist] made me very reflective," and "made me realize too, if I — if I can recognize that these things are working, why aren't I doing more of them?" (interview dated June 29, 2015).

Teachers also thought about assessment as part of their overall pedagogy and even their persona as an educator. In my interview with Helen, she emphasized how important she felt assessment is to teaching. She stated, "I mean, teachers take their assessment extr— [*sic*] very seriously, you know. It's an extension of who you are and your teaching practice, and your philosophy of education" (interview dated June 26, 2015). When talking about her understanding of "formative assessment," Mia expressed how integral the study of and reflection on assessment was to her identity as a teacher, saying, "it's important to me […] I think it's just the nature of what I do" (interview dated July 17, 2015).

Classroom context. During our interview, Helen said the following about the complex environment of education: "assessment never stands…you know, it's not like a scientific principle. It, it never stands outside of context" (interview dated June 26, 2015). The context of a classroom and its impact on how study participants chose which assessment practices to use was a fourth major emergent theme in addressing the second research question, and encompasses the following minor themes: (a) how teachers established expectations for their students, (b) the rapport and collegiality that existed between individuals, (c) the interactions that occurred

between students and their peers, and (d) the different types of student learners that were present in the classroom.

Established expectations. In describing their self-reported assessment practices, the SPRUCE ENG 135 Instructors in this study shared how they established expectations for their students for assignments and appropriate learning behaviors. These expectations, including learning goals, were explicitly stated and shared with students to help guide them in their writing and discussion practices. Students were aware of what their teachers wanted them to do, and were therefore better prepared to achieve the appropriate outcomes. Examples of how teachers established these expectations included using rubrics or grading criteria; leading guided activities, such as revisiting texts to improve understanding or having students find examples within the texts; the distribution of guides, such as potential questions for peer discussion or peer review of writing; and the use of modeling of expected practices by the teacher so that students could observe what they should do in their own work. With clear expectations established, teachers could tailor their assessment practices to help their students best meet these goals.

Rapport and collegiality. Several of the teachers stressed the importance of rapport and collegiality in their ENG 135 classrooms. Some of the teachers stated that they had the same students in multiple courses and/or years. This allowed the teachers and the students to be more familiar with each other, so that the teachers were able to address learners' needs more quickly and appropriately, and students knew of the approaches that the teachers might take in their assessment and instruction.

Rachel was one of the teachers who talked about the collegial respect that her ENG 135 students demonstrated for each other. She called it "such a safe environment," and noted that her students were "extremely supportive of one another" through their sharing of ideas and having

clarifying discussions on their own time (interview dated June 25, 2015). Having that kind of

support allowed for her use of more assessment practices that engaged students with each other

in the course, such as in-class discussion and the ongoing exchange of ideas via Internet-based

discussion forums.

Discussion played a large part in ENG 135, according to teachers' reports of their

practices. The interviewees noted that helping the students feel comfortable in voicing their

opinions was very important, and that they often used practices that allowed for students to be

more at ease with participating in class activities. For example, Krystal shared that

I didn't like being put on the spot as a kid, and it's not something I really

particularly enjoy doing to kids, either. […] I don't call them out in full-class

discussion. So I'll arrange these smaller group [discussions], so they're more

comfortable. (interview dated July 9, 2015)

Krystal continued to explain that she tended to "try to get a good balance between how much I'm

using discussion to gauge understanding and measure it, and then how much I'm using their

writing" to address students' strengths and comfort levels in the different ways in which they

demonstrated their knowledge to her (interview dated July 9, 2015).

Peer interactions. The teacher participants discussed how their students supported each

other in their learning as another contributing factor to classroom context that impacted their

selection of assessment practices. Several of the teachers mentioned that their students used

online discussion forums, including the learning management system for the course and social

media platforms such as Facebook, to post and respond to questions. The fact that these

exchanges took place and the content that was discussed in them would often be raised by the

students in class discussions. For instance, Rachel shared that her students had a Facebook chat

for the course, where they could ask her questions; she learned through class discussions that her students created their own chat to continue to talk to each other about ideas (interview dated June 25, 2015). Along with these forums, the teachers reported providing students opportunities in class to review and provide feedback on each other's work, such as a critique of drafts or commenting on and providing interpretations about quotes that other students extracted from a text.

    ***Addressing student differences*.** The different ways in which students learn is another factor that affected how the teachers in this study chose which assessment practices to use and when to use them. A few teachers specifically spoke about how and why they differentiated assessments along with instruction for their students. Mia shared in our interview:

> Well, you differentiate instruction everywhere. I mean, you say, OK, um, some of these kids are more successful in certain assessments, and some are more successful in other assessments. And you want to make them all successful at all of these types of assessments. […] So you vary them, you know, you don't do all of one type of assessment, because it just doesn't — you know, it's not beneficial to anybody. (interview dated July 17, 2015)

Krystal also commented on how she adjusted her assessment practices to meet her individual students' needs. She remarked that for some of her learners, "I know they get it, and I know they can demonstrate things better orally" than in writing; at the same time, she had other students who "don't like to give me things in writing" and who might be better served by "at least check-in in terms of discussion." She explained that she learned the best methods for assessing students' learning progress based on her experiences with each person, saying, "I usually kind of start with the same approach at assessment, and then as I go throughout the courses, I get to

know them better, and then I can kind of make adjustments as I need to" (interview dated July 9, 2015).

Differences between ELA and other content areas. Although the interviewed teachers were not asked specifically about the differences between ELA and other content areas and their potential impact on the selection of assessment practices, this major theme emerged when Helen brought up the topic during our conversation. At the time, she was talking about providing written feedback to students on their writing, and her willingness to re-read and re-grade students' papers if they put forth the effort to rewrite them. She told me,

> You know, I think — I think writing is, [*intakes breath*] you know — it's not math [*sighs*]. It's, uh — it evolves, and it has so much to do with maturity and worldview and language dexterity, and all of that stuff. And, and it just, um, I just say — I think you do a disservice if, if there's a hard and cold, fast — I mean, it's a hard and fast rule about 'You have to do [*knocks hand on table to emphasize the following words*] this and this and this. And if you don't get it first — the first time around, HAH! Gotcha!' You know? I hate that. And, so, if the — if the student has the desire to get better — and I think the only way to get better [at] writing is to write. The only way to get better at reading is to read. And, um, and so, I will always, you know, consider another draft. (interview dated June 26, 2015)

Helen added that individual conferencing with students helped her to meet their needs and helped the students to take the appropriate next steps, whether it be revising their work or creating a different thesis statement entirely. She stated that she was always willing to take these steps as an

ELA teacher in all of her classes, not just SPRUCE ENG 135, "when there's something worthwhile, something fundamental [for the students] to learn" (interview dated June 26, 2015).

Selection of assessment practices for other courses. The final major emergent theme from the participants' discussion of how they determined what assessment practices to use and when to use them was teachers' descriptions of how assessments varied between their SPRUCE ENG 135 students and their pupils in the other courses that they taught. As mentioned in the "Representativeness of practices" section earlier in this chapter, the questioning, discussion, and feedback assessment practices that these teachers used with their ENG 135 students were often used in other courses. However, the teachers noted incongruity between their SPRUCE and other ELA classes, which were often more driven by external expectations, such as meeting state standards and testing requirements. These differences included less flexibility in the course schedule for adding extra instructional time for a topic; focusing more on mechanics in students' writing; students in other courses lacking the same foundational skills as the ENG 135 students, and therefore needing more guidance, structure, and scaffolding of concepts; more use of multiple-choice assessments; differences in writing assignments, such as more papers in the course, but less of a focus on writing in general; and more of a focus on students stating or regurgitating what they had learned, rather than applying the knowledge to different texts, concepts, and assignments. All of these constraints impacted how teachers were able to conduct classroom assessments in their non-SPRUCE courses.

Participants' Perceptions of the Effectiveness of Assessment Practices

The third research question for this study, *What are secondary ELA teachers' perceptions of the effectiveness of the assessment practices that they use?*, was addressed primarily through statistical analyses of data collected from teacher participants' ELATE checklists. Supporting data to address this research question were drawn from interviews and email exchanges with the teachers. As explained in Chapter 3, SPRUCE ENG 135 Instructors participating in this study only indicated their perceived effectiveness of the assessment practices that they reported using. Graphs summarizing all teachers' perceptions of the effectiveness of the assessment practices that they reported using by ELATE checklist item are included as Appendix G. In addition, graphs comparing use and perceptions of effectiveness for each item are presented as Appendix H.

Performance levels of SPRUCE ENG 135 students. As mentioned in the "Student characteristics" section of this chapter, the teacher participants' perceptions of their SPRUCE ENG 135 students' performance levels varied. On their first ELATE checklists, both Mia and Helen indicated their perceptions of the effectiveness of the practices that they used for higher-performing students only. I asked Mia via email if this meant that she thought of all of her ENG 135 students were higher-performing, and she replied,

> Technically, all of my [ENG 135] kids are higher performing since my school [has] a pre-requisite for the class. Whether or not the students work to that level all of the time is questionable, but it most cases of research, these kids would be considered high performing. (email dated May 13, 2015)

Helen was asked the same question via email, and her response was quite similar:

> Technically, all our [ENG 135] students are 'higher performing' because they are
> only in the class by virtue of their past performance in English. To be in the class
> they need an A average in English for the first three years of high school. (email
> dated May 7, 2015)

In our interview, as we were discussing students peer reviewing each other's work, Mia did comment that "there are some kids that are just the top level, that are way ahead of everybody else" in their writing skills, and noted that in her SPRUCE ELA Program classes, there are "maybe two or three [of these students] in a class" (interview dated July 17, 2015).

The other five teacher participants all shared that there were differences in performance levels among their ENG 135 students, but the reasons offered to explain these differences varied. Allison described her lower-performing students as those who "were much less likely to be coming to class prepared," which impacted their ability to complete the assignments. Of her 17 ENG 135 students, she identified four as lower-performing, including "three students who were in danger of not graduating because of performance," four higher-performing students "that I would really put in the top tier," and "the rest [10 students] were in the middle." When asked what differences she saw between her average- and higher-performing students, Allison further clarified:

> the average students had read the texts and — very conscientiously and
> thoroughly, but didn't necessarily understand what they read, but they tried. Um,
> the lower-performing kids didn't even try, and then the high-performing kids were
> — had read it, got it, were thinking about it, and were sort of synthesizing,
> making connections between different texts. (interview dated July 1, 2015)

Charlotte said that out of her 17 ENG 135 students, she had five that were "high-achieving," and "four low students." Of those lower-performing students, she described one as always turning in work late, another as a pupil who performed well in other college-level courses, but who "really just couldn't get out of the C range" in ENG 135, and "two of them [who] may have had previous special ed[ucation] classifications." She said that her remaining eight students were "various levels in the middle, […] you know, high-middle, middle-middle," and said of all of her ENG 135 students, "If I had a name to them, I would definitely [be able to] rank them from top to bottom" (interview dated June 29, 2015). Later in our discussion, Charlotte explained that she made these performance level determinations primarily based on students' understanding of the material, their preparedness for class, and their effort in the course.

In talking about her ENG 135 students, Rachel emphasized that any difference in their performance levels "wasn't about ability. They're all extremely bright students." Of her 16 ENG 135 students, she identified "three who were on the lower end and struggling, […] about six were average," and seven "on the higher end, like, the high achievers." She attributed variations in students' performance to pupils that were "not as dedicated to putting in the time" or had a lack of interest in the material. In addition, Rachel commented that for "my struggling students, […] I would say confidence is part of it. You know, like having the confidence to say 'I think this is what I mean,' um, 'but I'm not quite sure.'" She added that

maybe exposure might have something to do with [students making connections]

as well, because I think some of my students who, you know, before have had

access to more, um, you know — the technology, um, the resources at an early

age think that, you know, they — they're a little more advanced in those areas.

(interview dated June 25, 2015)

Rachel later described differences in performance as related to students' academic experiences, using the example of a student that joined ENG 135 without having taken ENG 121, the SPRUCE writing course, in the fall with the other students. She said of him,

that was his first experience with, um, the ELA portion of [SPRUCE]. So, although he was quite able, it was really difficult for him to sort of come in halfway through the year, after missing [ENG 121] and some of the concepts that we learned there, and…pick up, as the other students had. (interview dated June 25, 2015)

She concluded about variations in students' performance levels, "I guess experiences sort of impact it too? I don't know" (interview dated June 25, 2015).

When describing her ENG 135 students' performance levels, Krystal noted differences between her two sections that she was teaching that semester. In her first section, she reported having "three on the lower-performing end" and "only two or three […] in the middle […] but the remaining [six or seven] really were pretty high flyers." Krystal shared that in her other section, she had two students who were "not very comfortable speaking in class, but [did] an excellent job with their writing." She added that she had "only three or four lower-performing, but […] the majority in that class were more kind of right in the middle" (interview dated July 9, 2015). When asked how she differentiated between the different performance levels, Krystal responded:

My lower-performing kids most likely struggled a bit with the content and with the writing, instead of one or the other. Because I think you could struggle with

one or the other, and still land yourself solidly in the middle. They would struggle

a little bit — and they all worked at it. I mean, I have to give them credit for that

— but would have more trouble either with expression of ideas, or having it come

together, or having a more nuanced understanding of the theoretical texts. […]

My…higher-performing tended to be…they, they came to me pretty good writers

already, […] they came to me with a really solid foundation. They were good

at…not only understanding the texts, but being able to kind of push back, being

able to ask questions, really grappling with the ideas, and what was so enjoyable

about so many of them is they really wanted to understand it. It wasn't just a,

'Well, I have to get this because I need a good grade.' No, like, they really

engaged with it, and that was fantastic. (interview dated July 9, 2015)

Finally, after he returned his first ELATE checklist, I asked Tony via email about what

differences he noticed between his students at the different performance levels. He responded

that he "may only have two or three students in my class I would label as high-performing," and

he attributed this small number to students' lack of work ethic. He went on to state that "many of

[the students] have fallen into the middle area," and that he put more effort to engage students

that semester than in his previous years as a SPRUCE ELA Program Instructor (email dated May

29, 2015). I was unable to learn more about Tony's thinking about his students' performance

levels, as he did not respond to multiple requests to participate in an interview after having

completed three ELATE checklists.

As demonstrated above, the seven teacher participants' thinking about their students'

performance levels varied greatly. Even though all ENG 135 students were expected to be

higher-performing due to their enrollment in the SPRUCE program and the potential for them to

earn college credit through successful completion of this course, the actual students in the classrooms were different from school to school (and in Krystal's case, even from section to section). Further discussion about these variations, as well as their implications for the study of assessment practices, will be presented in Chapter 5.

Having explored how teachers differentiated among their students—if at all—due to their performance levels, the following sections of this chapter will share the findings of participants' reports of their perceptions of the effectiveness of the assessment practices on the ELATE checklist that they used with their students. These results will be presented in the aggregate first, and then reviewed through the lenses of lower-, average-, and higher-performing students. Statistical analyses of these findings, which examine the perceived effectiveness of these practices along with reported ratings by student performance level, are also included.

Perceptions of effectiveness of practices. In addition to reporting their weekly use of assessment practices through the ELATE checklists, teachers also rated their perceived effectiveness of each practice used to improve student learning for their lower-, average-, and higher-performing students in their SPRUCE ENG 135 courses. Effectiveness was rated using the following six-point Likert-type scale: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; and 5 = *extremely effective*. Table 10 presents descriptive statistics for teachers' perceived effectiveness of the ELATE checklist items across all reports.

Table 10

Means, Standard Deviations, and Numbers of Reports of Teachers' Perceptions of the Effectiveness of Practices for Their Lower-, Average-, and Higher-performing Students

	Overall			Lower			Average			Higher		
Item	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
1	3.70	.85	53	3.19	.91	16	3.69	.60	16	4.10	.77	21
2	3.79	1.03	39	3.33	1.23	12	3.92	.79	12	4.07	.96	15
3	3.70	1.13	56	2.94	1.48	17	3.76	.83	17	4.23	.61	22
4	3.61	1.09	18	3.33	.52	6	3.80	.84	5	3.71	1.60	7
5	3.82	1.21	44	3.23	1.64	13	4.00	.68	14	4.12	1.05	17
6	3.73	1.12	37	3.55	1.29	11	3.55	1.13	11	4.00	1.00	15
7	3.47	1.01	30	3.11	1.27	9	3.67	.50	9	3.58	1.08	12
8	3.64	1.04	54	3.00	1.32	16	3.76	.66	17	4.02	.84	21
9	4.20	.93	49	3.86	1.23	14	4.20	.77	15	4.45	.76	20
10	4.31	.82	36	3.92	1.16	12	4.42	.51	12	4.58	.51	12
11	4.30	.67	27	4.00	.93	8	4.38	.52	8	4.45	.52	11
12	3.90	.89	50	3.53	1.13	15	3.93	.70	15	4.15	.75	20
13	4.39	.60	36	4.27	.65	11	4.45	.52	11	4.43	.65	14
14	4.50	.71	10	4.67	.58	3	4.67	.58	3	4.25	.96	4
15	4.67	.82	6	4.00	1.41	2	5.00	.00	2	5.00	.00	2
16	4.33	.62	15	4.00	.82	4	4.25	.50	4	4.57	.53	7
17	4.29	.82	31	4.20	.92	10	4.30	.82	10	4.36	.81	11
18	4.19	.40	16	4.20	.45	5	4.20	.45	5	4.17	.41	6
19	3.89	.75	36	3.82	.75	11	3.82	.75	11	4.00	.78	14
20	4.46	.77	37	4.36	.81	11	4.45	.82	11	4.53	.74	15
21	4.05	.85	37	3.91	.83	11	4.18	.87	11	4.07	.88	15
22	4.30	.73	33	4.20	.63	10	4.40	.70	10	4.31	.85	13
23	3.68	.91	37	3.45	1.04	11	3.82	.87	11	3.73	.88	15
24	3.84	.99	25	3.57	.53	7	3.88	.64	8	4.00	1.41	10
25	4.53	.71	34	4.40	.97	10	4.42	.67	12	4.75	.45	12

Note. Rating scale for teachers' perceptions of effectiveness of each practice to improve student learning is as follows: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Most effective practices overall. The overall perceptions of effectiveness (i.e., means averaged across all performance levels) of the ELATE checklist practices teachers used to improve student learning are displayed in Table 11, and are sorted from most to least effective. Teachers' perceptions of the effectiveness of all assessment practices included as ELATE checklist items were rated as at least *moderately effective*, with a lowest rated value of 3.47. More than half of the items (13 total) had an average rating above 4.0, or *very effective*.

Interestingly, in contrast to the teachers' reported use showing items related to questioning and discussion as the most frequently used practices, the six practices that were rated as most effective overall were related to feedback. This finding makes sense, as assessment practices involving feedback tend to be more time- and labor-intensive than practices such as asking students questions, and are therefore done less frequently, even if they are perceived to be just as or more effective when used to improve student learning.

Item 15, "Reviewed and gave feedback (oral or written) on students' portfolios of writing," was rated as being most effective overall ($M = 4.67$). However, due to teachers only rating the practices that they used, the efficacy of this item was only rated six times by all teachers across the six weeks. The third most highly-rated overall practice was item 14, "Reviewed and gave feedback (oral or written) on students' dialogic journals" ($M = 4.50$), but this also was not rated often (only 10 times).

A practice that was more frequently used—and therefore had a larger number of reports of ranking of its effectiveness—was item 25, "Conferenced with individual students about their writing to give them feedback and to review and discuss their work" ($M = 4.53$), which had been rated 34 times across students' performance levels. The total number of times this practice was reported as used (37; see Table 7) was relatively low compared to other practices, but the

majority of the teachers that reported using it rated it as *very effective* or *extremely effective*. For example, although Rachel did not report using individual conferences often, she did rate them highly in terms of effectiveness. When asked what impacted how often she used this practice, she responded that typically these conferences occurred after students received her written comments on an essay draft, and students would schedule a meeting with her. She shared that the focus of these meetings varied, based on students' needs:

> they would schedule an appointment and come to see me and talk about, you know, what they learned and what they thought, and sometimes they would scrap a whole complete idea and start all over, um, once they realized, like, 'I don't know if I'm exactly on the right track.' And sometimes they, you know — it was just a, a moment for clarifying. (interview dated June 25, 2015)

Allison also talked about individual conferences as a way of working with students to meet their specific needs, and said that "students like conferences, so they often seek them out." When asked how she used the information she learned from having these meetings with students, she replied:

> sometimes I'll realize that, um, they're at a different place in […] their individual projects than I thought they were, or where I think they should be, so I might have to restructure, even my calendar [for the lesson]. Um, and sometimes I'll realize when I've, I've met with […] three or four students, and […] they're all having the same problem; then you just stop and [say], 'OK, […] here's an impromptu, extemporaneous mini lesson on how to do this, because everyone's having this problem.' (interview dated July 1, 2015)

Table 11

Perceptions of Overall Effectiveness of Practices for Students, from Most to Least

Item	In teaching ENG 135 this week (last Friday through Thursday), I…	Item Type	*M*	*n*
15	Reviewed and gave feedback (oral or written) on students' portfolios of their writing.	F	4.67	6
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	F	4.53	34
14	Reviewed and gave feedback (oral or written) on students' dialogic journals.	F	4.50	10
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	F	4.46	37
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	F	4.39	36
16	Reviewed and gave feedback (oral or written) on students' *draft* versions of a writing assignment.	F	4.33	15
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	4.31	36
11	Used large group (5 or more students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	4.30	27
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	F	4.30	33
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	F	4.29	31
9	Asked questions intended to stimulate a general discussion.	Q & D	4.20	49
18	Used a rubric or rubrics for *informal* feedback to students on their writing.	F	4.19	16
21	Gave students *general or holistic* written feedback about their writing, such as comments at the end of an assignment about how they did overall.	F	4.05	37
12	Used whole-class discussion to determine how well students understood a concept, idea, or strategy.	Q & D	3.90	50
19	Used a rubric or rubrics for *formal* feedback to students on their writing.	F	3.89	36
24	Gave oral feedback to multiple students at once about their writing (e.g., discussing strategies with groups or the entire class).	F	3.84	25
5	Asked questions of students I thought would be more likely to respond well.	Q & D	3.82	44
2	Asked questions of individual students by name.	Q & D	3.79	39
6	Asked questions of reticent students to help improve their participation.	Q & D	3.73	37
1	Asked students questions to determine how well they understood a concept, idea, or strategy.	Q & D	3.70	53
3	Asked questions of the class as a whole.	Q & D	3.70	56
23	Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing.	F	3.68	37
8	Asked questions requiring more elaborated responses (e.g., a few sentences).	Q & D	3.64	54
4	Asked questions specifically of students I felt were not paying attention.	Q & D	3.61	18
7	Asked questions requiring brief responses (e.g., a word or phrase).	Q & D	3.47	30

Notes. Q & D = Questioning and Discussion. F = Feedback. Rating scale for teachers' perceptions of effectiveness of each practice to improve student learning is as follows: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Time limitations on when teachers were able to hold individual conferences with students also impacted how often they were able to conduct them. To allow for teacher-student conferencing during the time allotted for the class period, Krystal described meeting with students in-class when others were working in peer groups, and outside of class during the students' study hall periods (interview dated July 9, 2015). Mia also reported adjusting her schedule to accommodate time to meet with students, and said that she often met with them for conferences before the school day started (interview dated July 17, 2015).

Least effective practices overall. Of the six practices that SPRUCE ENG 135 Instructors in this study rated as least effective overall, five were related to questioning and discussion and one was related to feedback, as shown in Table 11. However, it is important to remember that the item rated as least effective, item 7, "Asked questions requiring brief responses (e.g., a word or phrase)," had an average rating of 3.47 (rated 30 times), which is still between *moderately effective* and *very effective*. The lowest-rated item related to feedback was item 23, "Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing" ($M = 3.68$, rated 37 times). In the interviews, teachers discussed correcting students' mechanics in their assignments, but stated that they preferred to focus more on content when reviewing the writing of their ENG 135 students. The other lower-rated questioning and discussion items included item 4, "Asked questions specifically of students I felt were not paying attention" ($M = 3.61$, rated 18 times); item 8, "Asked questions requiring more elaborated responses (e.g., a few sentences)" ($M = 3.64$, rated 54 times); item 3, "Asked questions of the class as a whole" ($M = 3.70$, rated 56 times); and item 1, "Asked students questions to determine how well they understood a concept, idea, or strategy" ($M = 3.70$, rated 53 times).

In our interview, Mia and I discussed her asking questions of her students in a whole-class environment. She expressed the following thoughts about what she learned from using this practice:

I do learn about their learning progress, but I think they [*chuckles*] learn about it, as well. […] I think they start to discover things — more of what they know, and — and start to ask questions of each other, and that gives me more idea of their understanding. Once they're able to ask questions of each other, and redevelop their own ideas, then I understand that they can understand what has been said and what has been going on, you know. I can ask them questions all they want, and they can answer them, but until they can start to develop on their own, um, they don't have a thorough understanding of the theory [I am teaching them].

(interview dated July 17, 2015)

This preference for having students take ownership of their learning through questioning and development of their ideas may explain the seemingly lower ratings overall for questioning and discussion as opposed to feedback practices.

To better understand how the effectiveness of assessment practices in the ELATE checklist were perceived by SPRUCE ENG 135 teachers, the efficacy ratings were separated by student performance level (i.e., lower, average, and higher). These findings are presented below.

Most effective practices for lower-performing students. Teacher participants' perceptions of effectiveness of the ELATE checklist practices that they used for improving student learning for their lower-performing students are presented in Table 12, and are sorted from most to least effective. While the perceptions of effectiveness ratings across the combined, or overall, groups ranged from 3.47 to 4.67 (see Table 11), the range of perceived efficacy

ratings of practices for lower-performing students was from 2.94 to 4.67, the lowest of the three student performance levels.

The top seven practices that were rated as most effective for improving student learning for lower-performing students were all related to teachers giving students feedback. Of those, the most highly rated practice was item 14, "Reviewed and gave feedback (oral or written) on students' dialogic journals" ($M = 4.67$). However, this practice was only rated three times for this student level. Item 25, "Conferenced with individual students about their writing to give them feedback and to review and discuss their work" ($M = 4.40$) was second-most effective for these students (rated 10 times), and item 20, "Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment" ($M = 4.36$) was third (rated 11 times). The other four items addressed feedback given from teachers to their students on their think/response papers (item 13, $M = 4.27$, rated 11 times) and their final versions of a writing assignment (item 17, $M = 4.20$, rated 10 times), feedback through rubrics used for informal purposes (item 18, $M = 4.20$, rated 5 times), and writing questions when giving feedback to students to prompt their thinking (item 22, $M = 4.20$, rated 10 times).

The value of written feedback and individual conferencing with students who needed extra guidance was discussed by Helen in our interview, who said that

> for a lot of kids who need that extra, you know, 'I just don't know how to get it out onto the paper,' um, I think that's — that's where, you know, we spend the most time […] when they do a second draft, […] I read it over again, and I write again, and then I meet with them again, and we go through it together. (interview dated June 26, 2015)

Charlotte also talked about the written feedback she provides to students to help them improve their writing, and commented, "if I don't write it down, then I feel like a student's not going to see either what I see, be it what's getting in my way, or […] I don't want to only give, um, things to correct" (interview dated June 29, 2015).

     Least effective practices for lower-performing students. The seven lowest-rated ELATE checklist items for improving the learning of lower-performing students were all related to questioning students (see Table 12). The lowest-rated practice was item 3, "Asked questions of the class as a whole" ($M = 2.94$, rated 17 times). The other items included asking questions that required more elaborate (item 8, $M = 3.00$, rated 16 times) and brief responses (item 7, $M = 3.11$, rated 9 times) and asking questions to check for understanding (item 1, $M = 3.19$, rated 16 times). Items involving singling out individual students also ranked as less effective for lower-performing students, including questioning students that would be more likely to respond well (item 5, $M = 3.23$, rated 13 times) or were not paying attention (item 4, $M = 3.33$, rated 6 times), and asking questions of individual students by name (item 2, $M = 3.33$, rated 12 times).

Table 12

Perceptions of Effectiveness of Practices for Lower-Performing Students, from Most to Least

Item	In teaching ENG 135 this week (last Friday through Thursday), I…	Item Type	*M*	*n*
14	Reviewed and gave feedback (oral or written) on students' dialogic journals.	F	4.67	3
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	F	4.40	10
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	F	4.36	11
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	F	4.27	11
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	F	4.20	10
18	Used a rubric or rubrics for *informal* feedback to students on their writing.	F	4.20	5
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	F	4.20	10
11	Used large group (5 or more students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	4.00	8
15	Reviewed and gave feedback (oral or written) on students' portfolios of their writing.	F	4.00	2
16	Reviewed and gave feedback (oral or written) on students' *draft* versions of a writing assignment.	F	4.00	4
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	3.92	12
21	Gave students *general or holistic* written feedback about their writing, such as comments at the end of an assignment about how they did overall.	F	3.91	11
9	Asked questions intended to stimulate a general discussion.	Q & D	3.86	14
19	Used a rubric or rubrics for *formal* feedback to students on their writing.	F	3.82	11
24	Gave oral feedback to multiple students at once about their writing (e.g., discussing strategies with groups or the entire class).	F	3.57	7
6	Asked questions of reticent students to help improve their participation.	Q & D	3.55	11
12	Used whole-class discussion to determine how well students understood a concept, idea, or strategy.	Q & D	3.53	15
23	Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing.	F	3.45	11
2	Asked questions of individual students by name.	Q & D	3.33	12
4	Asked questions specifically of students I felt were not paying attention.	Q & D	3.33	6
5	Asked questions of students I thought would be more likely to respond well.	Q & D	3.23	13
1	Asked students questions to determine how well they understood a concept, idea, or strategy.	Q & D	3.19	16
7	Asked questions requiring brief responses (e.g., a word or phrase).	Q & D	3.11	9
8	Asked questions requiring more elaborated responses (e.g., a few sentences).	Q & D	3.00	16
3	Asked questions of the class as a whole.	Q & D	2.94	17

Notes. Q & D = Questioning and Discussion. F = Feedback. Rating scale for teachers' perceptions of effectiveness of each practice to improve student learning is as follows: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Perhaps these students not responding well to questioning was due to their not completing their assigned work. Allison pointed out that in her ENG 135 classes,

> the issue with the questions [for the lower-performing students] is if they hadn't done the reading, or hadn't done it carefully, you could ask them anything, it's [*chuckles*]… [*says each word in next sentence slowly for emphasis*] It just doesn't work! [*speaks at normal pace*] I mean, […] what it mainly told me — reminded me is that they weren't coming to class prepared. (interview dated July 1, 2015)

On one of her ELATE checklists, Charlotte rated the effectiveness of "Asked questions that required brief responses (e.g., a word or phrase)" (item 7) with a value of 0, or *extremely ineffective*, for her lower-performing students. Like Allison, she discussed students not completing the readings prior to class, which impacted their ability to participate (interview dated June 29, 2015).

Most effective practices for average-performing students. Table 13 shows teacher participants' perceptions of effectiveness for ELATE checklist practices that they used with their average-performing students, sorted from most to least effective. These perceptions of effectiveness ratings ranged from 3.55 to 5.00, and therefore were all between *moderately effective* to *extremely effective* on the 0-5 scale provided.

The top four practices that were rated as most effective for improving student learning for average-performing students were all related to teachers giving feedback. Of those, the highest rated practices were item 15, "Reviewed and gave feedback (oral or written) on students' portfolios of their writing" ($M = 5.00$) and item 14, "Reviewed and gave feedback (oral or written) on students' dialogic journals" ($M = 4.67$). However, similar to item 14's high rating and low frequency of reported ratings of effectiveness for lower-performing students, these practices

were also seldom rated, with two and three ratings of efficacy, respectively. The other practices that were rated highly effective for average-performing students both had a mean rating of 4.45. These were item 13, "Reviewed and gave feedback (oral or written) on students' think/response papers," and item 20, "Gave students **targeted** written feedback about their writing, such as comments on certain parts of their assignment," which were both rated 11 times.

In our interview, Charlotte and I discussed her use of targeted written feedback with her students, and what that might look like in her teaching. She described her use of questions for students, such as prompts to consider texts and resources for the course to help guide their thinking. She quipped about the exchange with students that she was trying to promote, "I interact like mad, be it on something like an outline, or, um, an annotated bibliography, or, uh, their papers." She went on to explain how she used her written feedback with her ENG 135 students:

> I'd like my feedback to be enough where they could revise without me. […] I like having enough feedback on there where they could start revising without meeting with me, and then when they come to the meeting, the appointment, they get to choose how we spend the time. 'Do you want to review any of the things I've written? Or, do you want me to look at some of the changes you've made, based on my comments?' It gives them, I think, a lot more…a lot more that they can do before coming to see me. If they choose not to do anything before seeing me and want to go over everything, that's fine; I'll do that with them too. But at least I feel like there's some reminders there for me, as well, for the meeting. (interview dated June 29, 2015)

Table 13

Perceptions of Effectiveness of Practices for Average-Performing Students, from Most to Least

Item	In teaching ENG 135 this week (last Friday through Thursday), I…	Item Type	*M*	*n*
15	Reviewed and gave feedback (oral or written) on students' portfolios of their writing.	F	5.00	2
14	Reviewed and gave feedback (oral or written) on students' dialogic journals.	F	4.67	3
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	F	4.45	11
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	F	4.45	11
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	4.42	12
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	F	4.42	12
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	F	4.40	10
11	Used large group (5 or more students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	4.38	8
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	F	4.30	10
16	Reviewed and gave feedback (oral or written) on students' *draft* versions of a writing assignment.	F	4.25	4
9	Asked questions intended to stimulate a general discussion.	Q & D	4.20	15
18	Used a rubric or rubrics for *informal* feedback to students on their writing.	F	4.20	5
21	Gave students *general or holistic* written feedback about their writing, such as comments at the end of an assignment about how they did overall.	F	4.18	11
5	Asked questions of students I thought would be more likely to respond well.	Q & D	4.00	14
12	Used whole-class discussion to determine how well students understood a concept, idea, or strategy.	Q & D	3.93	15
2	Asked questions of individual students by name.	Q & D	3.92	12
24	Gave oral feedback to multiple students at once about their writing (e.g., discussing strategies with groups or the entire class).	F	3.88	8
19	Used a rubric or rubrics for *formal* feedback to students on their writing.	F	3.82	11
23	Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing.	F	3.82	11
4	Asked questions specifically of students I felt were not paying attention.	Q & D	3.80	5
3	Asked questions of the class as a whole.	Q & D	3.76	17
8	Asked questions requiring more elaborated responses (e.g., a few sentences).	Q & D	3.76	17
1	Asked students questions to determine how well they understood a concept, idea, or strategy.	Q & D	3.69	16
7	Asked questions requiring brief responses (e.g., a word or phrase).	Q & D	3.67	9
6	Asked questions of reticent students to help improve their participation.	Q & D	3.55	11

Notes. Q & D = Questioning and Discussion. F = Feedback. Rating scale for teachers' perceptions of effectiveness of each practice to improve student learning is as follows: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

***Least effective practices for average-performing students*.** Similar to teachers'
perceptions of effectiveness of ELATE checklist items for lower-performing students, the six
lowest-rated items for improving the learning of average-performing students were all related to
questioning students (see Table 13). The lowest-rated practice was item 6, "Asked questions of
reticent students to improve their participation" ($M = 3.55$, rated 11 times). The other less
effective items included asking questions: requiring brief responses (item 7, $M = 3.67$, rated 9
times); to determine students' understanding (item 1, $M = 3.69$, rated 16 times); requiring more
elaborated responses (item 8, $M = 3.76$, rated 17 times); of the class as a whole (item 3, $M = 3.76$, rated 17 times); and of students that seemed to not be paying attention (item 4, $M = 3.80$,
rated 5 times).

As previously stated, several of the ENG 135 Instructors who participated in this study
indicated that their average-performing students likely put forth the effort to learn, but struggled
more with the material than their higher-performing peers. Because of this, it is possible that
comments Krystal made during our interview can help to shed some light on why questioning
practices were not as effective with average-performing students as other assessment practices.
She stated that in one of her ENG 135 sections, she had

> a couple who I would say were perfectly average students, who were also
> interested in trying, and they just didn't necessarily grasp [the material] as well,
> but they, they contributed [to discussion]. But I had more in that class who were
> also just kind of willing to coast, and let others do the talking. And they do OK
> with their papers, but they weren't as invested as I would've liked to have seen
> them. (interview dated July 9, 2015)

This willingness of students to let others talk instead of contributing themselves may align with observations by other teachers. For example, Rachel had commented about her "struggling students" that they were more impacted by a lack of confidence in their thinking (interview dated June 25, 2015). Feeling unsure about one's own ideas may have affected average-performing students' participation in class discussions, including responding to questions from the teacher.

Most effective practices for higher-performing students. Participants' perceptions of the effectiveness of the ELATE checklist practices that they used for their higher-performing students are presented in Table 14, and are sorted from most to least effective. The range of perceived efficacy ratings of these practices for higher-performing students was from 3.58 to 5.00 (as compared to the overall ratings range from 3.47 to 4.67, as shown in Table 11), the highest of the three student performance levels. The number of times that each practice was rated was also the highest of the three performance levels; however, as previously discussed, this can be attributed to the fact that Mia and Helen only rated the effectiveness of practices for higher-performing students, based on the students they had in their ENG 135 classes.

As with the efficacy ratings for average-performing students, the most highly rated practice for higher-performing students was item 15, "Reviewed and gave feedback (oral or written) on students' portfolios of their writing" ($M = 5.00$). Again, however, the number of ratings that were given for this item were very low—only two, as were given for the same item for the other two student performance groups. The second-highest item, which was rated 12 times, was item 25, "Conferenced with individual students about their writing to give them feedback and to review and discuss their work" ($M = 4.75$).

During my conversation with Helen, she discussed the importance of getting to know individual students and their different needs, even with her higher-performing ENG 135 pupils. After describing her one-on-one work with a particular student, she said:

> So I just think it's, um — you know, it's still laboratory-like, […] you're still trying to figure out what prescription can I give you that will be different from you, and — and at the same time, trying to keep everybody's interest […] I try to keep all of the conferencing and all of that stuff between, you know, private — between me and whoever needs to talk to me. (interview dated June 26, 2015)

Helen later remarked, "I think assessment is — it's never one-size-fits-all" (interview dated June 26, 2015).

Unlike the predominance of items relating to feedback in the practices that were most effective for improving student learning for lower- and average-performing students, the third highest rated assessment practice for higher-performing students was related to questioning and discussion. Item 10, "Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy" was rated 12 times. Of note, this practice's rating increased as student performance level increased from lower to average to higher ($M =$ 3.92, $M = 4.42$, and $M = 4.58$, respectively). In our interview, Mia had commented,

> I think that there is a lot of value to that individual [conferencing with the teacher]/small group [discussion], because now I can really understand. […] When you do break them up into smaller groups, you can see who, um…where very specific people's strengths are, and where people's weaknesses are. (interview dated July 17, 2015)

Table 14

Perceptions of Effectiveness of Practices for Higher-Performing Students, from Most to Least

Item	In teaching ENG 135 this week (last Friday through Thursday), I…	Item Type	*M*	*n*
15	Reviewed and gave feedback (oral or written) on students' portfolios of their writing.	F	5.00	2
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	F	4.75	12
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	4.58	12
16	Reviewed and gave feedback (oral or written) on students' **draft** versions of a writing assignment.	F	4.57	7
20	Gave students **targeted** written feedback about their writing, such as comments on certain parts of their assignment.	F	4.53	15
9	Asked questions intended to stimulate a general discussion.	Q & D	4.45	20
11	Used large group (5 or more students) discussion to determine how well students understood a concept, idea, or strategy.	Q & D	4.45	11
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	F	4.43	14
17	Reviewed and gave feedback (oral or written) on students' **final** versions of a writing assignment.	F	4.36	11
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	F	4.31	13
14	Reviewed and gave feedback (oral or written) on students' dialogic journals.	F	4.25	4
3	Asked questions of the class as a whole.	Q & D	4.23	22
18	Used a rubric or rubrics for **informal** feedback to students on their writing.	F	4.17	6
12	Used whole-class discussion to determine how well students understood a concept, idea, or strategy.	Q & D	4.15	20
5	Asked questions of students I thought would be more likely to respond well.	Q & D	4.12	17
1	Asked students questions to determine how well they understood a concept, idea, or strategy.	Q & D	4.10	21
2	Asked questions of individual students by name.	Q & D	4.07	15
21	Gave students **general or holistic** written feedback about their writing, such as comments at the end of an assignment about how they did overall.	F	4.07	15
8	Asked questions requiring more elaborated responses (e.g., a few sentences).	Q & D	4.02	21
6	Asked questions of reticent students to help improve their participation.	Q & D	4.00	15
19	Used a rubric or rubrics for **formal** feedback to students on their writing.	F	4.00	14
24	Gave oral feedback to multiple students at once about their writing (e.g., discussing strategies with groups or the entire class).	F	4.00	10
23	Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing.	F	3.73	15
4	Asked questions specifically of students I felt were not paying attention.	Q & D	3.71	7
7	Asked questions requiring brief responses (e.g., a word or phrase).	Q & D	3.58	12

Notes. Q & D = Questioning and Discussion. F = Feedback. Rating scale for teachers' perceptions of effectiveness of each practice to improve student learning is as follows: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

***Least effective practices for higher-performing students*.** As with the most effective

practices for this student group, the least effective ELATE checklist practices for higher-

performing students were a mix of questioning and discussion and feedback practices (see Table

14). The lowest-rated practices were item 7, "Asked questions requiring brief responses (e.g., a

word or phrase)" ($M = 3.58$, rated 12 times) and item 4, "Asked questions specifically of students

I felt were not paying attention" ($M = 3.71$, rated 7 times). The next lowest practice was item 23,

"Corrected students' spelling, grammar, and other mechanical errors when giving them feedback

on their writing" ($M = 3.73$, rated 15 times). The following three items all rated the same ($M =$

4.00, or *very effective*), and dealt with giving oral feedback to multiple students about their

writing (item 24, rated 10 times), using rubrics for formal feedback on students' writing (item 19,

rated 14 times), and asking questions of reticent students to improve their participation (item 6,

rated 15 times).

The comments of Mia, one of the two ENG 135 Instructors who reported on her students

only as higher-performing, may provide some insight on why these practices, which tended to

coax reluctant students into discussion, may not have been as helpful with this group of students.

She remarked that

> early in the semester, you do do [*sic*] more talking […] you're kind of modeling
>
> the kind of inquiry that you're asking the students to do. And, by the end of the
>
> semester, I sit there, and I can not talk for, you know, 40 minutes. (interview
>
> dated July 17, 2015)

This may explain why, when this study was conducted at the end of the school year, students did

not need as much guidance in their discussions.

Helen, the other ENG 135 Instructor that thought of all of her students as higher-performing, had spoken about assessing the spelling and grammar of students' writing during our interview. In discussing students drafting, revising, and refining their writing, she referred to these kinds of errors as "tiny mechanics" several times. Like Mia, she noted that changes have happened over the course of the semester, and said that "all of that kind of stuff, by this time, has taken care of itself. And so, we're now really more evaluating ideas and expression of ideas and, you know, that sort of thing" (interview dated June 26, 2015).

Further discussion about how teacher participants perceived the effectiveness of the ELATE checklist practices for their lower-, average-, and higher-performing students is presented in Chapter 5. What follows in the next sections are statistical analyses of whether there are significant differences between perceptions of these practices' effectiveness for different student performance levels, and an examination of potential correlations between the reported use and perceived efficacy of these assessment practices.

Repeated measures ANOVA and post hoc analysis of perceptions of effectiveness. A one-way repeated measures ANOVA was conducted to determine whether there were statistically significant differences between teachers' average perceptions of effectiveness for the assessment practices they reported using with their lower-, average-, and higher-performing students. The following assumptions were tested: (a) independence of observations, (b) normality, and (c) sphericity. Independence of observations and normality were met. Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 13.23$, $p = .001$. Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .70$).

The results indicated that teachers did perceive the effectiveness of their reported assessment practices differently for these three levels of student performance. These differences were statistically significant, $F (1.39, 33.39) = 32.77$, $p < .001$, partial $\eta^2 = .58$. According to Richardson (2011), the partial η^2 value can be benchmarked against Cohen's (1969) criteria of small, medium, and large effects (with benchmarks of η^2 at 0.01, 0.06, and 0.14, respectively); in this case, the value suggested high practical significance.

The means and standard deviations for the effectiveness ratings by student performance level are presented in Table 15. Examination of these means suggests that teacher respondents rated their reported assessment practices as more effective for higher-performing students than for students at lower or average levels. Polynomial contrasts indicated, in support of this, that there was a statistically significant linear trend, $F (1, 24) = 37.10$, $p < .001$, partial $\eta^2 = .61$.

Table 15

Means and Standard Deviations of the Three Student Performance Levels

Variable	*M*	*SD*
Lower-performing	3.76	.48
Average-performing	4.12	.36
Higher-performing	4.23	.33

Post hoc analysis with a Bonferroni adjustment revealed that teachers' perceptions of effectiveness of the reported assessment practices was statistically significantly increased from lower- to average-performing students ($M = .36$, 95% CI [.21, .50], $p < .001$), from average- to higher-performing students ($M = .11$, 95% CI [.00, .22], $p = 0.46$), and from lower- to higher-performing students ($M = .46$, 95% CI [.27, .66], $p < .001$).

Analysis of potential correlations between use and effectiveness of practices. Pearson correlations were computed to examine potential relationships between teachers' self-reported frequency of use of assessment practices and their perceptions of the effectiveness of these

practices. Since teachers only rated the practices that they used, there were fewer than 25 comparison pairs per item for all of the data. To compensate for these smaller sample sizes (Howell, 2002), adjusted correlation coefficients (r_{adj}) values were calculated using the formula:

$$r_{adj} = \pm\sqrt{1 - \frac{(1 - r^2)(N - 1)}{N - 2}}$$

As shown in Table 16, the only statistically significant relationship at the $p < .05$ level was between frequency and overall effectiveness for item 10, "Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy," $r_{adj}(12) = .620$, $p = .031$. According to Cohen (1988), this would be a strong positive correlation. This finding means that having students work in pairs or small groups was a practice used by teachers that was also perceived as very effective.

At the $p < .1$ level, there was a statistically significant positive relationship for this same practice between frequency and effectiveness for both lower-performing, $r_{adj}(12) = .534$, $p = .074$, and average-performing students, $r_{adj}(12) = .571$, $p = .053$, both with a strong correlation, according to Cohen (1988). Of note, a significant relationship was not found with higher-performing students for this item. Another positive correlation was found at this level between frequency and perception of effectiveness for higher-performing students for item 5, "Asked questions of students I thought would be more likely to respond well," $r_{adj}(17) = .438$, $p = .079$.

A strong negative correlation, according to Cohen (1988), was found at the $p < .1$ level between frequency and effectiveness for average-performing students for item 24, "Gave oral feedback to multiple students at once about their writing (e.g., discussing strategies with groups or the entire class)," $r_{adj}(8) = -.693$, $p = .056$. In other words, group oral feedback was not used as much with students about their writing, but was perceived to be effective with average-performing students when it was used.

Table 16

Adjusted Pearson Correlations between Teachers' Reported Frequency of Use of Practices and Perceptions of Their Effectiveness for Lower-, Average-, and Higher-performing Students

		Effectiveness							
	Frequency (Times Used)	Overall		Lower		Average		Higher	
Item		r_{adj}	N	r_{adj}	N	r_{adj}	N	r_{adj}	N
1	211	.170	53	.137	16	.267	16	.203	21
2	291	.238	39	.308	12	.067	12	.216	15
3	285	.201	56	−.253	17	−.092	17	.042	22
4	48	.199	18	.270	6	−.447	5	.089	7
5	114	.136	44	.229	13	.282	14	.438[†]	17
6	81	.210	37	.327	11	.203	11	.186	15
7	226	.088	30	.234	9	.308	9	.327	12
8	291	−.222	54	.241	16	−.105	17	.224	21
9	328	.142	49	.163	14	.142	15	.144	20
10	28	.620*	36	.534[†]	12	.571[†]	12	.452	12
11	93	.270	27	.207	8	.272	8	.319	11
12	129	.336	50	−.274	15	−.264	15	.345	20
13	152	.234	36	.204	11	.316	11	.278	14
14	28	.685	10	−.538	3	−.538	3	.685	4
15	20	n/c	6	n/c	2	n/c	2	n/c	2
16	84	.253	15	.555	4	.724	4	.189	7
17	55	.467	31	.492	10	.448	10	.363	11
18	7	−.447	16	−.500	5	−.500	5	−.447	6
19	47	.331	36	.202	11	.202	11	.419	14
20	95	.112	37	.108	11	.161	11	.136	15
21	118	.386	37	.060	11	.231	11	.351	15
22	121	.420	33	.439	10	.105	10	.290	13
23	141	.192	37	−.312	11	.330	11	.232	15
24	25	.243	25	−.198	7	−.693[†]	8	.332	10
25	37	.276	34	−.267	10	−.288	12	.163	12

Notes. †$p < .1$. *$p < .05$. **$p < .01$. ***$p < .001$. ****$p < .0001$. n/c = correlation could not be calculated due to too few paired observations.

Participants' Determination of the Effectiveness of Assessment Practices

The fourth and final research question for this study, *How do secondary ELA teachers determine the effectiveness of their assessment practices?*, was addressed through qualitative analysis of interview data. Overall, the following major themes emerged from the interviews as factors that influenced how the SPRUCE ENG 135 teachers determined the efficacy of assessment practices: (a) teachers' knowledge of and familiarity with their students, (b) students' understanding of the content, (c) students' performance levels, and (d) the benefits to students of the use of these practices.

Of all of the teacher participants who completed ELATE checklists, Helen was the only one who rated the effectiveness of all of the practices she reported used as 5, or *extremely effective*. During our interview, I asked if she felt, in retrospect, that all of these practices truly were that effective. She replied, "I found that the ones I used, I used because they are the most effective." She added that the assessment practices that she uses as a whole is "the system that works for me," as long as the students are "getting something out of it," which she felt was reflected in the improvement she sees in their writing (interview dated June 26, 2015). This focus on what is appropriate for the particular students in class, augmented by teacher knowledge of and familiarity with their students and how well they respond to these practices, was a similar motif in the responses of the other interviewed teachers.

Teachers' knowledge of students. The first major theme that emerged from discussions of how participants determined the effectiveness of the assessment practices that they used was teachers' knowledge of their students. All of the interviewed teachers discussed their knowledge of their SPRUCE ENG 135 students and their backgrounds; this knowledge tended to be even deeper if the teacher had previously taught a particular student. For instance, Charlotte had many

of her ENG 135 students in her ELA courses the year before. She said of this, "I came in [to

ENG 135], you know, with pre-existing relationships and already knowing something about a lot

of them, not all of them, but a lot of them and their writing." She added that in one of her ENG

135 sections, her students were not as likely as the other group to be as diligent in their work nor

to pay careful attention, but noted that this "was not shocking to me, […] most of them I had the

year before, too" (interview dated July 9, 2015). Another example is Krystal, who when asked

how she makes decisions about which assessment practices to use for her students, replied, "I

kind of get a sense [of what each student's needs are] — and I'm, again, I'm reasonably lucky in

[SPRUCE] that I've had many of them before" (interview dated July 9, 2015). Her familiarity

with these students also permitted her to tailor her assessment practices to best suit them as

learners.

Students' understanding of content. A second major theme that emerged about how

teachers determined the effectiveness of their assessment practices was students' understanding

of the course content. Participants reported using evidence of both individual and class

understanding, including reactions and responses to practices, and individual learners' ability to

demonstrate that they had learned a topic as ways of evaluating effectiveness. These minor

emergent themes are further explored below under the headings of (a) individual understanding,

(b) class understanding, and (c) demonstration of student knowledge.

Individual understanding. Individual students' reactions to assessment practices were a

way in which teachers evaluated the effectiveness of these practices. Along with students

improving their writing as a result of her assessment practices, Helen said that students'

responsiveness to the practices that she used was another way she determined their efficacy

(interview dated June 26, 2015). Charlotte also used students' reactions to her practices as a way

to determine their effectiveness. She recalled how she observed particular individuals when using an assessment activity where her students passed around and commented on text excerpts that their peers had selected:

> it's a lot of, of visual monitoring. […] When I know the students didn't fully read the reading, and I'm doing the pass-around, […] I would see [my two lower-level students] finding the quote for the activity, but going back and marking it. So even though their page was completely empty, they found the quote, and they wrote it down, or noted [it] or something, as a possibility for themselves. […] They didn't just do it to get through the activity, they saw, 'Oh, I may be able to use this.' […] I know if those two students were doing it, then it was working. You know, the [other] students who already had it marked, well, it worked already anyway. (interview dated June 29, 2015)

Class understanding. Several of the teacher participants shared that they frequently determined the effectiveness of an assessment practice through in-class discussions and observations. For instance, in talking about adjusting class time to help her ENG 135 students understand a topic, Rachel recounted that she "spent more time going back […] when it seemed like students didn't, you know, understand [the more challenging theories and ideas]." To determine the understanding of the class as a whole, such as when using questioning practices, Rachel reported that she observed students' body language and evaluated their responses:

> when I have a certain expectation of what I think their answers might be, and they're not exactly hitting the target, um, then that's sort of what guides me. Other than body language and facial features that shows you that 'I'm confused!'

[*makes face and moves head to side slowly; I chuckle and do the same*]. (interview dated June 25, 2016)

Krystal described similar experiences in her ENG 135 classes when her students were participating in full-class discussion, and gave an example of how she might evaluate the effectiveness of this practice. She said that using discussion was helpful in revealing the students who understood the material well, as "they're wanting to answer questions, or refute somebody else's point: I think that's pretty effective for them; they're very engaged in it." She continued to explain that "For kids who…maybe don't understand the text as well—for some, that's still a helpful practice. Listening to that conversation and hearing other people talk about it helps them understand." However, Krystal said that using discussion to determine the learning progress of her students that do not participate was not effective, "which is why it's necessary for me, then, to have those [written] thinking pieces, so that I can see […] what they know" (interview dated July 9, 2015).

Demonstration of student knowledge. The interviewed teachers also described students' demonstration of their knowledge as a means of checking for student understanding of content and evidence of the use of effective assessment practices. As an example, Mia noted that her students would tell her if a particular assessment practice was not working for them, such as peer assessment. She said of practices that were not effective with certain students, "that I can change," but emphasized that the main way she determines the effectiveness of an assessment is

> basically in the product. So you can see as […] you walk around [the classroom].
> You listen, you look, you see what they're writing. As you're doing that, you
> know if they're — if what you're doing is positive or not. (interview dated July
> 17, 2015)

Allison also responded that her ultimate determination of the efficacy of a certain practice in SPRUCE ENG 135 was "often in the […] final product, the paper." She explained that for larger projects, students were required to produce drafts "where they're working through the ideas," and on which they received a lot of feedback for appropriate revisions. Ultimately, though, she stated that "[what] it comes down to in that class is a portfolio analysis, and […] it depends on how well they end up doing on the [final] paper. Did they get it or not?" (interview dated July 1, 2015).

A third example of a teacher who spoke of the importance of the attainment of end goals for ENG 135 as a way of determining whether assessment practices were effective or not was Krystal. When asked how she made decisions about the efficacy of certain practices, she replied:

I'm looking for two things. I want to see if students are able to demonstrate their understanding of content, so if whatever it is I'm asking them to do, I can then look at if, one, they get it, […] and then, [two,] is it something that can serve as scaffolding? […] Does it serve a purpose, ultimately, in getting us somewhere? […] If I'm going to have them put in time and effort, I want to make sure that it's worthwhile. So, [an effective assessment practice] needs to serve both that content and that working towards the ultimate goal of that unit. (interview dated July 9, 2015)

Student performance levels. The performance level of students (i.e., lower, average, and higher) in SPRUCE ENG 135 was a third major emergent theme that was tied to how teacher participants determined the effectiveness of their assessment practices. Several teachers reported through their ELATE checklists that certain practices were perceived as less effective for their lower-performing students. These differences tended to occur for similar reasons among the

teachers. For example, Allison noted that issues of preparedness and attendance impacted student performance in her ENG 135 class. She explained, "the kids who had read and were engaging with the ideas from the get-go; I think they just benefited more from assessment because they were farther along" (interview dated July 1, 2015). Likewise, Rachel observed more difficulties for her students who did not invest as much time in the course (interview dated June 25, 2015), and Krystal reported more issues for her students who were not as engaged in class discussions (interview dated July 9, 2015). Both Rachel and Krystal shared that they had struggling students who put in the effort but did not grasp the material as well, and that these were also students for whom their assessment practices were not perceived to be as effective.

Benefit to students. The final major theme that emerged from the interview data about how SPRUCE ENG 135 teachers determined the efficacy of their assessment practices was the potential benefit that the use of the practices would have for students. During the interviews, teachers reported that effective practices were those that would be empowering and aid in the development of student autonomy. Interviewees also spoke of the feasible, long-term benefits that would positively impact students based on their current use of effective assessment practices. More about these minor emergent themes are presented below under the headings (a) empowerment and autonomy and (b) longer-term effects.

Empowerment and autonomy. During the interviews, teacher participants spoke of how they regarded assessment practices to be effective when they empowered students. For instance, Helen recounted how students were able to take ownership of their ideas—and feel confident in generating new ones—through class discussions, where they shared their opinions with their peers (interview dated June 26, 2015). Mia permitted her ENG 135 students to choose their peer editors, saying, "if they're going into college, they need to be able to do these types of things"

(interview dated July 17, 2015), and Krystal had her students choose their own groups for discussion and the completion of projects (interview dated July 9, 2015).

Participants also discussed how they developed and encouraged student autonomy through SPRUCE ENG 135. For instance, Mia said of her students' writing over the semester, "as they progress in being able to write [in] the kind of style that a think paper is looking for […] they begin to develop, and they begin to be more successful." She also remarked on the evolution of students' participation in class discussion over the course, commenting that "by the end of the semester, they can do it — they run it themselves, and I, I simply am there to facilitate" (interview dated July 17, 2015).

As another example, Helen shared that she always permitted her students to revise their writing, because "I always want them to feel the ball is in their court." She later described how she expected her ENG 135 students to develop independence in their writing and thinking after their having also taken the SPRUCE writing course, ENG 121, the previous semester:

> in the fall, we're really pushing the fundamentals, and kind of, um, understanding what makes something good, and how you yourself can look at what you've done and decide: is this good? Did I do what I was supposed to do? […] So then, in the spring, we're looking for you to be able to do all of that by yourself, and now try to, to reach these more sophisticated levels of expression and connection, and do it, you know, with a sophisticated grasp of language. That, um, no one has to be in the classroom discussion to understand what you're doing; it should be that your stuff — your work — stands on its own, and it stands — you know, makes sense, without having it to be just a part of whole class discussion on something. (interview dated June 26, 2015)

Longer-term effects. Besides sharing how students may benefit from being in the SPRUCE ENG 135 course, teachers also discussed how effective assessment practices may positively impact these learners in the future. Overall, this included guiding students into being reflective on their work, such as revisiting and revising drafts and creating and reviewing portfolios; supporting students in feeling more comfortable in asking for help; and having ongoing discussions, both in small and large groups, to develop students' confidence in sharing and refining their ideas.

Summary of Findings

A summary of the quantitative and qualitative findings that were described in this chapter is presented below. Further discussion of these findings, their relationship to related research, and the implications of this study are presented in Chapter 5.

The teachers who participated in this study, while teaching similarly-structured versions of SPRUCE ENG 135 with 12th grade students in Spring 2015, were all teaching in different contexts that varied in terms of the districts and schools in which the classes took place, the students in the course, and the teachers' own backgrounds. Across the 28 ELATE checklists completed by the seven teachers, questioning and discussion practices were reported as most frequently used. The overall assessment practices teachers reported using were statistically different from week to week, but this aligned with their planned curriculum for the course.

Through interviews, teacher participants reported that their selections of assessment practices were due primarily to the following factors: (a) the purpose of the assessment to be used, including meeting learners' needs and teaching goals, and fair evaluation of students' work; (b) how the assessment connected to their instruction; (c) their own knowledge as teachers, including intuition, reflection on their practice, and their awareness of assessment as a

whole; and (d) the classroom context, including expectations, rapport, peer interactions, and student differences. Other major factors impacting teachers' selection of assessment practices that emerged were differences teachers noted between ELA and other content areas and the students in the teachers' classrooms, as their assessment practices were adjusted to suit the needs of students in their non-ENG 135 courses.

As reported through the ELATE checklists, the most reliable highly-rated effective practice across all student performance levels was item 25, "Conferenced with individual students about their writing to give them feedback and to review and discuss their work," with a mean rating of 4.53 on a 0-5 scale. As a whole, the practices with the highest perceptions of effectiveness for all students were related to feedback practices. Although the assessment practices on the ELATE checklist were rated with the highest values for higher-performing students by teacher participants, almost all of the included practices were rated at 3.0 (*moderately effective*) or higher for all student performance levels. (The sole exception was item 3, "Asked questions of the class as a whole," with a perceived effectiveness value of 2.94 for lower-performing students.)

When examining perceptions of the efficacy of practices from lower- to higher-performing students, there appears to be a shift of practices that are valued from those that provide students more guidance in the writing process to practices that allow for more student independence. A one-way repeated-measures ANOVA of the perceptions of effectiveness found statistically significant differences between what teachers perceived as effective for their lower-, average-, and higher-performing students. A post hoc analysis indicated that there were statistically significant increases in teachers' perceptions of effectiveness from their lower- to average- to higher-performing students, as well. Pearson correlations were conducted to explore

potential relationships between frequency of the use of items and teachers' perceptions of their efficacy, but few significant relationships were found.

Finally, teachers' determinations of the effectiveness of the self-reported assessment practices that they used were discussed through interviews. The following major themes emerged as factors that impacted whether a teacher viewed a practice as effective or not: (a) teachers' knowledge of the students in their classroom; (b) students' understanding of the content, including the individual and class levels, as well as demonstration of that knowledge; (c) student performance levels; and (d) the potential benefit of the use of the practice to the students, including empowerment and autonomy, as well as longer-term benefits.

Overall, these findings indicate that ENG 135 Instructors in the SPRUCE ELA Program were able to report and discuss their assessment practices through the use of the ELATE checklists and the semi-structured individual interviews, which provided practical insight into the FA practices of secondary ELA teachers. Participants were able to share what they did in their ENG 135 classrooms and reflect on these practices and the decisions that led to their use, as well as make their implicit knowledge explicit about the perceived efficacy of the assessment practices that they use. As mentioned above, further discussion of these findings, including connections to prior research, implications, and conclusions, are presented in the next chapter.

Chapter 5: Discussion and Conclusions

This research examined the formative assessment practices of secondary ELA teachers through exploration of the assessment practices that they used, how they selected these practices, how effective they believed these practices to be, and how they determined the effectiveness of these practices. For the purposes of this study, there was a focus on teachers' questioning, discussion, and feedback practices. The study was conducted using a mixed methods approach with a sample of secondary ELA teachers who were all teaching the same course within a concurrent enrollment program in the northeastern United States. Participating teachers completed weekly checklists to report their use of assessment practices with their students in this course. Follow-up interviews were then conducted to further explain the results from the checklists.

A discussion of the findings of this study, which were presented in Chapter 4, is presented in this chapter. The discussion, like the findings, is organized by the following research questions:

1. What assessment practices do secondary ELA teachers use?

2. How do secondary ELA teachers determine what assessment practices to use and when to use them?

3. What are secondary ELA teachers' perceptions of the effectiveness of the assessment practices that they use?

4. How do secondary ELA teachers determine the effectiveness of their assessment practices?

Following a discussion of the research findings are a review of the study's limitations and strengths, the implications of this study to relevant fields, suggestions for future research, and final conclusions.

Variations among Participants and Their Contexts

Although all seven of the teachers who participated in this study were SPRUCE ENG 135 Instructors at the time that this study was conducted, there were variations among the participants. These teachers ranged in teaching experience from 10 to 27 years, and in experience as a SPRUCE Instructor from 3 to 15 years. All of them had participated in the relevant SPRUCE training for their position and taught a similarly-structured version of the ENG 135 course. Through the ELATE checklists, each of the Instructors reported varying usage of questioning, discussion, and feedback assessment practices with their students, as well as differing perceptions of the effectiveness of the practices that they used with their students (see Appendix I for snapshots of each teacher).

There were also differences in the contexts in which each Instructor taught. The districts and schools in which these teachers worked were geographically located across the state in suburban and rural locales. Additionally, differences in student populations included variations in ethnic and economic representation at the school level, as well as variations in the students who were taking ENG 135 at each school. These class-level differences included students' reasons for taking the course; their performance, motivation, and ability levels; and the prerequisites that students had met prior to enrollment in the course. Further exploration into the impact of context on teachers' assessment practices, including the teaching of ELA in secondary schools, is presented later in this chapter.

Use of Assessment Practices

The first research question for this study, *What assessment practices do secondary ELA teachers use?*, was addressed through the quantitative analysis of teacher participants' completed ELATE checklists. Additional support came from the qualitative analysis of interviews and emails with the teachers.

  Questioning and discussion. The six practices that were reported by all teachers as most frequently used were related to questioning and discussion. ELATE checklist item 9, "Asked questions intended to stimulate a general discussion," was used the most, with a total of 328 times reported by the teachers over the six weeks the checklists were used. Participants described these practices as being a means for quickly determining where students were in their learning progress, as well as an indicator for them about whether or not they should shift their instructional practices. Questioning and discussion practices that allowed the students to share their ideas and hear the thinking of others helped them to develop strategies for dealing with challenging course texts and grappling with the application of the theories that they learned in the course. As this study took place at the end of the semester and students were completing multimedia projects, the interviewed teachers noted that questioning and discussion occurred with more frequency than if students were focusing on individual writing assignments.

The use of discussion-based practices to both develop and assess student understanding has previously been studied in secondary ELA education. For example, Nystrand (1997, cited in Applebee, Langer, Nystrand, and Gamoran, 2003) found that discussion-based approaches were significantly related to middle and high school ELA students' literacy performance, and were effective for students at all achievement levels. The features of effective discussion that led to larger improvements in student performance included "more use of authentic questions, which

were used to explore differing understandings rather than to 'test' what students might already know;" more time for open discussion, or whole-class discourse; and "more 'uptake,' in which a teacher's question 'took up' and built on a student's previous comment, creating continuity in the discourse" (Applebee et al., 2003, p. 690). All of these features were present in the descriptions that teacher participants provided of how classroom discussion might take place in ENG 135. Teachers described encouraging students to explore their thinking aloud with each other, frequent use of whole-class discussion so students could learn from their peers, and their own role as facilitators in class discussions.

The use of questioning and discussion in the secondary ELA classroom can also be viewed as a use of instructional dialogues to assess students' learning. Ruiz-Primo (2011) described such practices as part of informal FA, where "the activities that teachers and students do in their classrooms can be described as potential assessment opportunities for collecting evidence of students' understanding" (p. 15). In order for instructional dialogues to be effective assessment conversations as well as effective informal FA strategies, they must meet several conditions. These conversations need to be learning goal-guided, dialogic and interactive in nature, applied as instructional scaffolding tools, support social participation and social cognition, and enculturate students into the language of the discipline (Ruiz-Primo, 2011). Based on the descriptions of the SPRUCE ENG 135 Instructors who participated in this study, these conditions were all met through their own use of questioning and discussion assessment practices with their students.

Feedback on students' writing. Although they were not used as frequently as questioning and discussion assessment practices, teachers did report providing feedback on students' writing. The most frequent use of this practice was ELATE checklist item 13,

"Reviewed and gave feedback (oral or written) on students' think/response papers." During the interviews, teachers often discussed how they used students' think papers to develop their writing skills over the course of the semester, and how these skills were an extension of what students learned in the SPRUCE writing course, ENG 121, in the previous fall semester. Participants noted that they were able to provide feedback to students on these think papers, so that students could use the feedback to revise and refine their work as they prepared final drafts of longer papers. The other most frequently used feedback practices, involving correcting mechanics in students' papers, writing questions to prompt thinking, and providing both general and targeted feedback to students, were all also related to helping students in the revision of their work and improvement of their writing skills.

Other research on using feedback to help students revise their writing is related to these findings. For example, Fitzgerald (1987) stated that generally, high school students are more skilled writers than younger pupils, and their products improve in quality when revised. She added that what she termed "naturalistic classroom support," or "intervention procedures such as questioning, conferencing, having dialogues, and providing lots of opportunity to write and revise," has produced much more positive results in developing writers than simply providing directions (p. 494).

More recently, Parr and Timperley (2010) studied the use of feedback on writing within an assessment for learning (a synonym for FA) framework, and found that guiding feedback that identified where students' performance was, where it needed to be, and what was needed to be done to close that gap significantly improved students' writing achievement. They did note, however, that considerable teacher pedagogical content knowledge (such as what the teachers in this dissertation study exhibited) was necessary to provide appropriate formative feedback. Other

researchers also concluded that feedback that provided students with strategies when developing complex competencies was important to writing development, and that other effective feedback practices included teachers' applying their knowledge of students' abilities to their feedback, providing feedback in a timely fashion, and using feedback systematically in their teaching (Bruno & Santos, 2010). As with their use of questioning and discussion assessment practices, SPRUCE ENG 135 teacher participants' reported use of feedback to their students on their writing seems to apply practices that are known to be effective in the improvement of student learning.

Variations in use of practices. Teachers' observed use of assessment practices on the ELATE checklist varied significantly from their expected (i.e., average of the data) use during the time that this study was conducted, as shown through the results of the chi-square analysis of use (Table 8). These results demonstrate how much teachers' use of these items varied overall. Although the use of rubrics by teachers can help to improve student writing (e.g., Andrade, 2000; Panadero & Jonsson, 2013; Spandel, 2006), they were rarely used by study participants. However, this low reported use of rubrics for assessing students' writing makes sense, as ENG 135 students were primarily working on multimedia projects and not written assignments at the point in the semester when this research was conducted.

Further analysis of the items, with the practices sorted by their function and with analysis conducted by week of report (see Table 9), found that there were significant differences between observed and expected use of almost all of the practices reported as used through the ELATE checklists across the six weeks. In other words, teachers' actual use of assessment practices changed from practice to practice and from week to week, as they changed their instructional practices based on students' responses to various circumstances. This included addressing the

content being taught and the instructional strategies being used as part of the curriculum, using varied assessment practices to suit students' learning needs, addressing assignments that were being completed by their students, and revising planned practices to accommodate schedule changes at the end of the school year. These variations in practice may serve as an example of the many different processes that can occur within an ELA classroom as teachers assess their students' understanding with an eye to improving their learning, especially when the teachers are reflective about their practice.

Selection of Assessment Practices

The second research question for this study, *How do secondary ELA teachers determine what assessment practices to use and when to use them?*, was addressed through qualitative analysis of the interviews that were conducted with teacher participants. The factors that teachers described as impacting their selection of assessment practices included the interaction between assessment and instruction, relying on their own knowledge and experience, and the specific contexts of their ENG 135 classrooms.

The synergy of assessment and instruction. Interviewees in this study frequently talked about how they used the information they learned from their questioning, discussion, and feedback assessment practices to guide their instruction. These teachers discussed how they were able to use classroom assessment practices to impact their instruction, and that they could do so in a much more timely fashion than they were able to apply the results of more formal assessments, such as state-administered examinations. In addition, they shared how the purpose of using certain assessment practices was often to determine where a student was in her/his learning progress, and what next steps they needed to take in their instruction.

In the FA literature, the use of assessment data to drive instructional decision making is critical (e.g., Birenbaum et al., 2009; Black & Wiliam, 1998a; Griffin, 2009; Phelan et al., 2011). In fact, Shermis and Di Vesta (2011) stated, "Assessment is an integral part, not a separate appendage, of instruction" (p. 83). The findings of this study suggest that secondary ELA teachers are just as likely to incorporate assessment into their instruction as instructors in other content areas and grade levels. Due to the nature of ELA teaching, this use of assessment in instruction is likely to be highly individualized to suit students' needs, a topic which is further discussed later in this chapter.

Teachers' knowledge and intuition. The knowledge and intuition of the teachers in this study played a large part in how they chose which assessment practices to use. Since this population of teachers was selected for study because of their role as expert teachers (Berliner, 2001, 2004), they were probably able to better articulate their experiences and explain their decision making than novice teachers. For example, Krystal's comment that her work with student teachers over the years and her frequent explanations of how and why she uses certain practices "really has made me more cognizant of my practice overall" (interview dated July 9, 2015) showed that she was often thinking about and reflecting on her teaching.

By encouraging teachers to share their expertise and make implicit knowledge explicit, teachers can become more aware of their own practices (Brevik, 2014) and contribute to research on teaching in real-world conditions (Fang, 1996). In addition, teachers' sharing of their practical knowledge and how they reflect on their practice can be used to benefit the training of new teachers and guide them toward becoming reflective practitioners (Griffiths, 2000; Verloop, Van Driel, & Meijer, 2001).

The impact of context on assessment practices. Perhaps the factor that had the biggest impact on teacher participants' selection of assessment practices was the context in which these practices would be used. As briefly discussed at the beginning of this chapter, the schools and districts in which the teachers were working in Spring 2015 varied, and there were differences in student ability and performance levels and the prerequisites that students needed to meet to enroll in the SPRUCE ENG 135 course. When discussing how they chose which assessment practice to use, the interviewed teachers spoke of what would be best suited to help their students achieve their learning goals. All of these practices took place within the context of teaching and assessing ELA at the secondary level; others' research about these contextual factors is presented below.

Factors in the study of teaching and learning. Several factors that contribute to context and have the potential to impact teaching and learning have been identified in the literature. White (1985) proposed three major features of context in educational research: (a) physical conditions, which involve place and time; (b) people features, which include the age of the learners, the diversity of the participants, and the pattern of their interpersonal relationships; and (c) social features, which encompass the values of the participants, their institution, and their society. When examining context variables that affect measures of teacher efficacy, Guskey (1987) also proposed three variables: (a) the nature of the student performance outcome, either positive or negative; (b) the ability of the students involved, either high or low; and (c) the scope of influence, or whether the teacher was working with an individual or a group of students. Of these six features of context, the latter five are most important to how the contexts of each teacher's ENG 135 classes in this study can be viewed. The diversity of the students' ability and performance levels and their relationships between them; students' performance outcomes; the values placed on these SPRUCE ELA Program courses; and the number of students with whom a

teacher was working were all described by this study's participants as factors that influenced how they determined which assessment practices to use.

Tessmer and Richey (1997) highlighted several assumptions about the role of context from an instructional design standpoint. Of these, two of the assumptions relate to the information that the teachers shared about how assessment informed their instructional decisions. One assumption, "The impact of context varies with the nature of the learner, the content, and the intensity of the contextual elements" (p. 88), can be associated with the differences that teacher participants observed between their SPRUCE ENG 135 students and their students in other courses, and even differences between their ENG 135 sections, if applicable. A second assumption, "Successful instructional designs must be, to some extent, situation-specific" (p. 88), ties to teachers' description of adjusting their assessment and instruction to meet their students' needs. Participants in this study used the information they gathered from their assessment practices to inform their instruction at the individual, small group, and whole-class levels in order to improve student learning. A major aspect of the contexts in this study was the students themselves, especially with regards to their performance levels. This "nature of the learner" and their understanding of the content, whether individually or with others, impacted the assessment practices used by the teachers in this study, and often was the primary influence in the "situation-specific" learning that was taking place.

Differences between ELA and other subjects. During our interview, Helen remarked on the differences between ELA and math, stating that it would be a disservice for teachers in the former subject to hold students to "a hard and fast rule" and rigid expectations, rather than allowing them to revise their work (interview dated June 26, 2015). There are indeed differences across disciplines in schools, including how they are taught, the content needed for

understanding by learners, their subjectivity or objectivity, and the ways in which educators

assess their students in them (Grossman & Stodolsky, 1995; Marshall, 2007; McMillan, 2010;

Siskin, 1991; Stodolsky & Grossman, 1995).

As previously discussed in Chapter 2, ELA instructors conduct much of their assessment

in the moment, and without using preplanned feedback and responses (McMillan, 2010). FA-

related practices, including questioning, discussion, and feedback, will occur as they do in other

subject areas, but are more of a "meandering, organic affair" in the ELA classroom (Marshall,

2007, p. 137). The findings of this dissertation study substantiate the predictions of prior

research: the SPRUCE ENG 135 Instructor participants reported that they were responsive to

their students and their needs, were able to provide immediate feedback to students orally and

timely feedback in written form, and were likely to restructure their assessment practices to suit

the dynamic ELA context in which they were teaching. In addition, the participants in this study

were examples of highly thoughtful and reflective practitioners that were able to make in-the-

moment judgments based on their knowledge and experience in order to best suit the needs of

their students at all performance levels.

Perceptions of the Effectiveness of Assessment Practices

The third research question for this study, *What are secondary ELA teachers' perceptions*

of the effectiveness of the assessment practices that they use?, was addressed through the

quantitative analysis of teacher participants' completed ELATE checklists. Additional support

came from the qualitative analysis of interviews and emails with the teachers. As discussed in

earlier chapters of this dissertation, SPRUCE ENG 135 Instructors only rated the effectiveness of

the checklist items that they reported using with their students. Effectiveness was rated using the

following six-point Likert-type scale: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; and 5 = *extremely effective*.

Overall effectiveness. Teacher participants' overall perceptions of the effectiveness of the ELATE checklist items that they reported using indicated that items related to feedback were most effective with their students. The practice that was highly rated and reported as used often was item 25, "Conferenced with individual students about their writing to give them feedback and to review and discuss their work" ($M = 4.53$, rated 37 times). The use of and the highly favorable rating attributed to this assessment practice indicate not only its value to ENG 135 Instructors, but also its potential importance to the assessment of secondary ELA students in general. Although the focus of this practice is the use of feedback, it is likely that questioning and discussion between the teacher and the student also took place during these conferences. The teacher participants were able to individualize their assessment practices to each student, their performance and ability level, and their needs to aid them in improving their writing.

These findings align with prior research on the use of teacher-student conferences in writing and how they can be used to promote student self-efficacy (Pajares, Johnson, & Usher, 2007) and actively involve students in the evaluation and revision of their work through structured modeling, discussion about, and use of these behaviors (Sperling, 1990; Walker & Elias, 1987). However, as effective as this practice may be perceived to be, the amount of time needed for individual conferencing with students is a challenge to its implementation (e.g., Dunn, 2011; Moss & Bordelon, 2007). For the participants in this study, this limitation was mitigated by teachers holding individual conferences in class while other students were working with their peers, or holding conferences outside of class time, including during students' study hall time and before or after the school day.

There were two assessment practices that were rated as highly effective, but that were seldom used: item 15, "Reviewed and gave feedback (oral or written) on students' portfolios of their writing" ($M = 4.67$, rated 6 times), and item 14, "Reviewed and gave feedback (oral or written) on students' dialogic journals" ($M = 4.50$, rated 10 times). These practices were also not reported as being used often on the ELATE checklists (28 and 20 times, respectively, as shown in Table 6). Although these practices were perceived as very effective by the teachers that reported using them, they were likely not used frequently during the timeframe of this study due to the ENG 135 course coming to a close and a focus on students creating multimedia projects, rather than introducing new material (which the students would tackle using their dialogic journals) or concentrating on writing assignments, as was done earlier in the semester.

Effectiveness for lower-performing students. The top seven assessment practices that were rated as most effective for lower-performing students were all related to teachers providing feedback to students. As noted in Chapter 4, the highly-rated and frequently used practices for this student group involved individual conferencing with students ($M = 4.40$); targeted written feedback to students on their writing ($M = 4.36$); feedback on students' think/response papers ($M = 4.27$); feedback on students' final versions of writing assignments ($M = 4.20$); the use of rubrics for informal feedback on students' writing ($M = 4.20$); and teachers writing questions to students to prompt their thinking and develop their writing skills ($M = 4.20$). Taken as a whole, these practices seem to focus on providing more structured guidance through feedback than the other ELATE checklist items. These students likely needed more assistance in the development of their ideas, more prompts to encourage their thinking, and more support in the cultivation of their writing skills.

Effectiveness for average-performing students. The top six assessment practices that were rated as most effective for average-performing students were mostly related to teachers providing feedback to students, with one practice tied to questioning and discussion. The highly-rated and frequently used practices for this student group involved feedback on students' think/response papers ($M = 4.45$); targeted written feedback to students on their writing ($M = 4.45$); the use of paired or small group discussion with students ($M = 4.42$); and individual conferencing ($M = 4.42$). Most of these practices are similar to what was rated as effective for the lower-performing learners, but the order has changed and the efficacy means have increased. In addition, discussions involving students working in pairs or small groups was more highly rated. It seems that the practices that are most effective for this particular student group still provide structure, but permit more independent expression than the top-rated practices for the lower-performing students.

Effectiveness for higher-performing students. The top seven assessment practices that were rated as most effective for higher-performing students were a combination of practices related to teachers providing feedback to students and practices associated with questioning and discussion. The highly-rated and frequently used practices for this student group involved individual conferencing with students ($M = 4.75$); the use of paired or small group discussion with students ($M = 4.58$); feedback on students' draft versions of writing assignments ($M = 4.57$); targeted written feedback to students on their writing ($M = 4.53$); asking questions to stimulate a general discussion ($M = 4.45$); and the use of large group discussion with students ($M = 4.45$). For this group of students, it seems as though the most effective practices are ones that allow for increased student ownership of learning. Rather than a practice focusing on feedback on final writing assignments, as with the lower-performing students, a practice focusing on

feedback on students' draft writing was highly rated. Also, as these students are more likely to have stronger academic autonomy, they are more likely to benefit from student-based discussions that permit them to share their ideas with their peers.

Taking into account student differences. The teacher participants in this study were all able to identify the performance levels of their SPRUCE ENG 135 students, and had these differences in mind when completing the ELATE checklists. Along with interpretations that can be made from the descriptive effectiveness data, as described above, statistical analyses were conducted to determine if there were significant differences between these teacher-identified student performance levels.

The results of a one-way repeated measures ANOVA indicated that there were statistically and practically significant differences between teachers' perceptions of the effectiveness of assessment practices for their lower-, average-, and higher-performing students. There was a statistically significant linear trend of the perceived effectiveness of practices, as efficacy increased from the lower- to higher-performing student groups. A post-hoc analysis with a Bonferroni adjustment also indicated statistically significant increases between the lower- to average-, average- to higher-, and lower- to higher-performing student groups. These statistical tests showed differences between the student performance levels; practically, the teachers were also able to indicate and explain how their perceived effectiveness of the assessment practices that they used varied among these groups.

These findings are all based on a relatively homogeneous student population; even still, teachers were able to anecdotally determine the performance level to which each of their ENG 135 students belonged. These teacher participants made distinctions between their learners based

on these levels, and through differentiation of their assessment practices, were likely able to better meet the needs of all of their students.

Determination of the Effectiveness of Assessment Practices

The fourth and final research question for this study, *How do secondary ELA teachers determine the effectiveness of their assessment practices?*, was addressed through qualitative analysis of the interviews that were conducted with teacher participants. The factors that SPRUCE ENG 135 teacher participants described as impacting their determination of the effectiveness of assessment practices included a focus on the students that they were assessing, students' performance levels, and the potential benefits of the assessment practices for the students.

Focus on students in assessment. The teachers who participated in this study described how their assessment practices were focused on addressing their students' learning needs. This was facilitated by teachers' knowledge of the students, which was increased if the instructor had taught the student in a previous course; student's knowledge of the content that was being taught, which teachers would evaluate at the individual and class levels; and students' demonstration of their knowledge.

Heritage (2013) detailed how classroom teachers develop and use a learning progression for their students to help them learn discipline-specific content and skills. She wrote that this progression

> is based on [the educator's] experience of teaching children. Their sources for
> developing the progression are curricula, their views of what is best taught when,
> and their knowledge of children's learning. In this context, validation involves
> working together, testing each other's hypotheses against their professional

knowledge, making refinements accordingly, trying out the progression to see if

their model actually predicts what happens in terms of student learning, and then

making further refinements from this experience. A by-product of teacher-

developed progressions is an associated deepening of teacher knowledge about

learning in a domain, which can have considerable payoff for evidence gathering

and use. (p. 189)

This form of hypothesizing and testing out approaches to improve student learning, while

concurrently contributing to teacher knowledge and experience, is congruent with the

descriptions of participants in this study about how they determined the effectiveness of the

practices that they used. Teachers made references to their experience, knowing what for them

was most effective, and getting a sense of their students when determining what practices to use

to best suit their learners.

Student performance. The performance level (i.e., lower, average, and higher) of the

students of the SPRUCE ENG 135 Instructors who participated in this study was another factor

that impacted participants' determination of the effectiveness of the assessment practices that

they used. Teachers reported evaluating the efficacy of practices, in part, through observations of

student effort, motivation, and engagement. These findings relate to the research of Brookhart

(1997), who described the classroom assessment environment as a context in which teachers set

performance criteria and standards, appraise performance and provide feedback to students on

their learning progress, and monitor learning outcomes.

The findings of this study also align with research on student performance and FA.

Andrade (2010) emphasized the importance of fostering students' self-regulated learning as a

means of closing gaps in understanding. As discussed in previous sections of this chapter, the

assessment practices used by ENG 135 teacher participants varied among student performance levels in how effective they were perceived to be. Practices that supported students' thinking and scaffolded their writing skills were perceived as most effective for lower-performing students; teachers perceived that average-performing students benefited from a combination of practices that were structured and allowed them some ownership in their learning (e.g., paired and small group peer discussions); and the practices that were perceived as most effective for higher-performing students allowed for the most student academic autonomy. Teachers' ongoing cultivation of ELA-related skills with their students were factors that influenced their determination of the efficacy of these assessment practices.

Short- and long-term benefits to students. A final factor that impacted how teacher participants determined the effectiveness of the assessment practices that they used was the potential benefit the practice would have to students, either in the short or long term. The SPRUCE ENG 135 Instructor participants discussed the importance of the role of the assessment practices that they used in helping to empower students and develop their academic autonomy, traits also found in the FA literature (e.g., Andrade, 2010; Brookhart et al., 2009; Hodgen & Marshall, 2005; Olson et al., 2010; Wiliam et al., 2004). This included the use of questioning, discussion, and feedback to develop students' self-confidence and comfort in sharing their ideas with the teacher and each other, and the selection of peers for editing papers and for working and discussion groups. The longer-term benefits included students taking the knowledge and skills that they learned through ENG 135 and applying them elsewhere, such as being reflective on their work, seeking support when warranted, and continuing to share and refine their ideas in collaboration with others.

Connections to Research on Formative Assessment

As stated in the introductory chapter of this dissertation, I had previously conducted a literature review of 25 studies involving professional development for teachers in grades K-12 on FA to attempt to operationalize the construct of "formative assessment" (Tolley, 2012). This literature spanned grade levels and content areas (including nine studies targeting science teachers and seven targeting teachers of mathematics), used varied research methods, and detailed studies that were conducted in the U.S., United Kingdom, New Zealand, Israel, and Portugal. Although definitions of "formative assessment" or "assessment for learning" varied from study to study, there were six common elements of FA that emerged: (a) teachers setting and making explicit clear learning goals; (b) the use of questioning and feedback from teachers and students; (c) data collection, analysis, and use; (d) teacher revision/modification of instruction; (e) the development of student academic autonomy; and (f) evaluation to determine if a gap between the learning goals and current student performance exists.

All six of these elements were present in the results of this dissertation research study. Below are brief summaries of how each of these elements are described in the literature, and how the findings of this study contribute to the research on each particular FA component.

Teachers setting and making explicit clear learning goals. One major element of FA is teachers establishing clear learning goals for their students to attain, and more importantly, making those goals explicit to their students (e.g., Aschbacher & Alonzo, 2006; Birenbaum et al., 2009; Blanchard, 2008; Borko, Mayfield, Marion, Flexer, & Cumbo, 1997; Brookhart et al., 2009; Dixon & Williams, 2003; Fontana & Fernandes, 1994; Sato et al., 2008; Sondergeld et al., 2010; Wiliam et al., 2004). Having clear, shared objectives helps teachers guide their instruction and assessment effectively, and helps students in their awareness of how their actions impact

their achievement, including how to make appropriate changes in their performance to meet these learning goals. In this study, teacher participants described how they established clear expectations for their students through the use of shared grading criteria (such as rubrics), providing guiding questions for peer discussions and peer editing of writing, and modeling expected practices so students knew what they should do in their own work.

The use of questioning and feedback from teachers and students. As discussed in earlier chapters of this dissertation, the use of questioning and feedback by both teachers and students is another major element of FA (e.g., Aschbacher & Alonzo, 2006; Ash & Levitt, 2003; Brookhart et al., 2009; Dixon & Williams, 2003; Olson et al., 2010; Sondergeld et al., 2010; Wiliam et al., 2004). Rather than lecture-based instruction, where questioning and feedback practices are conducted primarily by the teacher, in FA the students are just as involved and engaged in these processes as their instructor. Instead of interactions only occurring in one direction—from teacher to student—FA allows for teacher-student, student-teacher, and student-student (either peer- or self-reflective) exchanges.

The SPRUCE ENG 135 teachers who participated in this study reported frequent use of questioning and discussion practices with their students. At the time that this research was conducted, there was more discussion occurring than feedback on students' writing due to the end-of-the semester assignment, a multimedia project. Teachers shared that they often used questioning and discussion with their students to informally assess where students were in their learning progress, and that they used the information from these practices to inform their instructional decision making. As with other research on FA practices, teachers also reported having their students very involved in their own learning through questioning and discussion, in the form of class and smaller group peer discussions and interactions.

Data collection, analysis, and use. A third element of FA is the collection, analysis, and use of assessment data to inform the improvement of student learning (e.g., Bell & Cowie, 2001; Birenbaum et al., 2009; Borko et al., 1997; Herman et al., 2010; Lee & Wiliam, 2005; Nelson & Eddy, 2008; Phelan et al., 2011; Schneider & Meyer, 2012). Although formal assessments, such as standardized tests, and detailed analyses of their results may impact what a teacher does a few months after initial administration of the exams, FA practices serve a more immediate purpose, and are tailored to react to students and their specific learning needs at the classroom level.

As Rachel commented during our interview for this study, rather than focusing on the numbers generated by higher-stakes assessments, she thought teachers should say to themselves, "'what's relevant, and how — how may I use this to assist students? How may I use this to help me as an instructor?'" (interview dated June 25, 2015). The teacher participants reported using what they learned from their questioning, discussion, and feedback practices to inform what they needed to do next in their instruction. Additionally, they used this information to self-assess their teaching, and to determine what changes they should make both now and in the next time they taught a similar lesson in the future. Teachers' reflection on their practice, which combined with their knowledge and experience and the in situ data they gathered from interactions with their pupils and their learning progress, strengthened their ability to best meet their students' needs.

Teacher revision/modification of instruction. Another common element of FA, related to data collection, analysis, and use, is teachers' revision and/or modification of their instruction based on the results of their assessment practices (e.g., Aschbacher & Alonzo, 2006; Gearhart & Saxe, 2004; Herman et al., 2010; McGatha et al., 2009; Phelan et al., 2011; Sato et al., 2008). For teachers, this is likely to involve a combination of instructional and assessment practices that can be changed as necessary to help students meet appropriate learning goals.

In this study, teachers most often discussed revising or modifying their instruction as a reaction to student understanding, including students' responses to the assessment practices that they used. These actions were part of their use of assessment data, and included making decisions to clarify instruction; deciding to either repeat, go back, or continue instruction based on students' reaction to a lesson; and determining whether or not instruction and/or assessment needed to be better differentiated for the individual learners in their classroom.

Development of student academic autonomy. As teachers adjust their instruction and assessment practices to improve student learning, they ultimately are striving for a fifth common element of FA, the development of student academic autonomy (Blanchard, 2008; Brookhart et al., 2009; Fontana & Fernandes, 1994; Hollingworth, 2012; Olson et al., 2010; Sondergeld et al., 2010; Wiliam et al., 2004). When teachers use FA practices, they are involving students in these processes, and mentoring and guiding them in how to engage in and take ownership of their own learning.

The SPRUCE ENG 135 teachers interviewed in this study shared how they facilitated the development of rapport and collegiality among their students, and how students were enabled to gain self-confidence in their ideas through class discussion. Participants reported encouraging students to take ownership of their ideas and responsibility for their own work, including revising their writing based on feedback to improve both the product and their skills. The teachers also helped students in reaching learning goals through clearly established expectations and appropriate guidance, and noted the strides that their students had made in their thinking, discussion, and writing over the course of the semester.

Evaluation of potential gaps between learning goals and student performance. A final common element of FA is the evaluation of potential gaps between where students currently

are in their learning and where they need to be to meet the appropriate learning goals (e.g.,

Birenbaum et al., 2009; Leung, 2004; McGatha et al., 2009; Olson et al., 2010; Yin et al., 2008).

For secondary ELA teachers, this may involve a teacher knowing a great deal about each

individual student, her/his performance level and capabilities, and how to help each of them

make progress toward achieving her/his goals.

In this study, interviewed teacher participants described their knowledge of their students

and their learning needs, which frequently was augmented due to their having taught the students

before in a previous course. All of the teachers were able to identify the performance levels of

the students in their ENG 135 classes, as well as which assessment practices might be most

effective for each student. Several of the interviewed teachers described the purpose of

assessment as a way to meet students' needs, including closing gaps in their learning. The

SPRUCE Instructors in this study often addressed those gaps through data-informed

differentiation of their assessment and instructional practices to support the learning of all of

their students.

Limitations and Strengths

There were several limitations and strengths of this research due to factors such as the

nature of the study design, the researcher's familiarity with the content area and context, and the

use of the ELATE checklists and interviews. The most relevant limitations and strengths are

highlighted below.

Limitations of the study. The primary limitations of this study were the small sample

size; the use of self-reported data from teachers; the use of student performance level labels; the

timing of the study; the need for more appropriate anticipation of response patterns to the

ELATE checklists; and the reliability and validity of the ELATE checklist instrument. Each limitation is briefly summarized below.

Small sample size. The sample size for this study was small, with seven teachers who volunteered to participate (from a potential sample of 49), six of whom participated in interviews. The point might be raised that this limited number of participants undermines the generalizability of the findings. However, this study was focused on researching a particular population as a means of determining the feasibility of learning about secondary ELA teachers' assessment practices from expert instructors, rather than the generalizability of the results. Future work can involve the revision and use of the instruments created for this research with larger numbers of participants.

Self-reported data from teachers. The data from this study were self-reported by the participants, which raises concerns about the reliability and validity of the findings. The use of experience sampling methods alleviated some of these concerns, as this approach asks participants about something that just happened, and serves to capture events and actions that would be challenging to capture otherwise. Zirkel et al. (2015) noted that experience sampling method research "affords a tremendous reduction in retrospective bias inherent in all self-report data," as it has the potential to capture a more accurate picture of actual practice than asking participants to reflect over a longer period of time due to "asking people to report on their activities, affect, and actions in situ and on many small occasions" (p. 9). The recall period for this study was only seven days. In addition, teacher self-reports that are focused (e.g., specified subject matter, class group, and time frame for retrospective reporting) can be used to gather reliable data about their practices (Koziol & Burns, 1986). For this study, the ELATE checklist

focused specifically on participants' SPRUCE ENG 135 students in Spring 2015 and the

questioning, discussion, and feedback practices that they used with them.

***Use of student performance level labels*.** In hindsight, the use of tracked student

performance levels (e.g., lower, average, and higher) may not have been the best approach for

accounting for student differences on the ELATE checklist. As the students of the teacher

participants in this study were relatively homogeneous (i.e., all were likely higher-performing in

comparison to students in other ELA courses), the inclusion of these labels cued teacher

participants to use them to differentiate between their ENG 135 students in potentially artificial

ways. In future research, one way in which variations among students may be more appropriately

studied would be to ask teachers to describe how they perceive differences among their learners,

and then use each teacher's self-described labels when examining the reported use and perceived

effectiveness of their assessment practices.

***Timing of the study*.** This study was conducted at the end of the 2014-2015 academic

year, in particular from May through June for the ELATE checklists, and June and July for the

interviews. If the study had been conducted at the beginning or middle of the Spring 2015

semester, there may have been a greater number of participants, and end-of-the-year events,

including AP examinations, state testing, and senior-year celebrations for the students would not

have affected when teachers were able to use and report appropriate assessment practices.

Interviewed teachers did remark that there was more of a focus on discussion rather than writing

at the end of the course, so the findings may also have been impacted by the study timeframe.

***Anticipation of response patterns to the ELATE checklists*.** The design of the ELATE

checklist for this study did not appropriately anticipate and/or facilitate the capturing of teacher

participants' responses when using the instrument. For example, Mia and Helen only rated the

effectiveness of the practices that they used for higher-performing students, as that level of learners is who they believed they had in their ENG 135 courses. Overall, the means of effectiveness ratings for all student performance levels, whether viewed individually or in the aggregate, were skewed positively. The lowest mean was 2.94 for the efficacy of "Asked questions of the class as a whole" for lower-performing students, and the means for the effectiveness of ELATE checklist practices for average- and higher-performing and overall students were greater than 3.0, with 3 equal to *moderately effective* on the rating scale.

The number of times that ratings were reported for practices that were used also varied. Besides being dependent on whether or not a teacher used a particular practice that week, that practice may not have been used with all students. For example, teachers may have only "Asked questions specifically of students I felt were not paying attention" of their lower-performing students.

Finally, the trends that emerged from these data are relevant only to the particular population of teachers who participated in the study and their impressions of their students in their SPRUCE ENG 135 courses. Although the teachers relayed differing explanations of how they determined the performance levels of these students, as a whole, these pupils were still high school seniors enrolled in a course for which they could earn college credit, and were therefore likely to be more higher-performing than their peers.

Reliability and validity of the ELATE checklist. Due to the design of this study, and especially the ELATE checklist, reliability (e.g., Cronbach's alpha or test-retest reliability) and validity (e.g., content validity measures such as factor analysis) statistics could not be computed. Since respondents were asked to only rate the effectiveness of the assessment practices they used on the ELATE checklist, practices varied from week to week even for the same teacher, and two

of the seven Instructors rated the effectiveness of assessment practices for higher-performing

students only. As a result, data were missing by design, which led to SPSS errors in attempts to

analyze the data that were reported. These limits of the data persisted, even when multiple

imputation of missing data values was attempted. The number of participants also impacted the

ability to conduct factor analysis on the 25-item ELATE checklist data. According to research on

sample size in factor analysis (e.g., Bentler and Chou, 1987; Brown, 2006; Gorsuch, 1983;

MacCallum, Widaman, Zhang, & Hong, 1999), 10 respondents per item would be appropriate to

conduct this kind of analysis, and a sample of 125-250 respondents would be very conservative.

As this study was exploratory in nature, however, the trends observed from the

descriptive statistics and analyses from this relatively small sample can help to inform the

administration of similar methods with a larger sample and/or a different population. In addition,

the use of interviews after the ELATE checklists were completed did help to validate teachers'

responses. It is anticipated that a larger number of respondents, especially when combined with

stronger validity measures, such as classroom observations, will aid in the development of the

ELATE checklist as a valid and reliable tool that can be used to better understand secondary

ELA teachers' assessment practices.

Strengths of the study. The primary strengths of this study were the use of a mixed

methods design; my own familiarity with both ELA teaching and the SPRUCE ELA Program;

and the study's focus on underlying components of "formative assessment," rather than on the

term itself. Each strength is briefly summarized below.

Use of a mixed methods design. The use of a sequential explanatory mixed methods

design for this study (Creswell & Plano Clark, 2011; Ivankova et al., 2006) allowed for a more

in-depth exploration of secondary ELA teachers' assessment practices than if only quantitative or

qualitative methods had been used. The interview questions for this study were developed only after initial quantitative analyses had been conducted, which helped to explain the findings. In addition, the data from both phases of the study, taken together, improved the validity of the findings and informed the research as a whole.

Researcher's familiarity with ELA teaching. A second strength of this research was my own familiarity with ELA teaching. As a former high school English teacher, I understood the strategies that participants were discussing with me, and I could use my own pedagogical content knowledge (Shulman, 1986) to ask appropriate questions of them. It also seemed as though teachers were more willing to participate in a study that was being conducted by a researcher that was both familiar with and had experience in teaching in their subject.

Researcher's familiarity with the SPRUCE ELA Program. A further strength of this study was my familiarity with the SPRUCE ELA Program. As discussed in the first chapter of this dissertation, I conducted preliminary fieldwork with three administrators and five teachers from the program in 2014. The purpose of that initial qualitative study, conducted through semi-structured interviews and a focus group, was to understand how SPRUCE ELA Program Instructors talked about their practices to improve student learning. Knowing the terms that those participants used to describe their assessment practices, as well having an understanding of the course that they taught or oversaw in the ELA Program, permitted me to create instruments for this dissertation study that used appropriate language and that were easily understood by the participants in the current research.

Focus on the underlying components of "formative assessment." Since a major finding of the preliminary fieldwork, conducted prior to this dissertation study, was that all of those eight participants defined "formative assessment" differently, I was aware to not use that particular

term in the instrumentation for this research. As a result, I was able to focus specifically on the components of FA (i.e., questioning, discussion, and feedback) that were most important to this study, and the phrasing I did use had the same meaning for me and the teacher participants.

Implications of This Study

The findings of this study have potential implications for ongoing research into other related topics in the fields of education and evaluation. The examples presented here address teachers' reflection on their own practice, appropriate and relevant teacher professional development, the educational accountability movement, educational evaluation, and ongoing research into formative assessment.

Teachers reflecting on their practice. The teachers in this study shared how they reflected on their teaching, including their assessment practices, and how their thinking about what they used in their classroom came from the knowledge that they constructed from their experiences. These teachers also talked about how they discussed their practices with others, whether they were other ELA teachers or other SPRUCE teachers in their schools, and their awareness of assessment in general, and how this informed their teaching. These participants already had established professional networks; in order to foster teacher reflection on their practice, support systems are needed in the schools to encourage this kind of professional growth for all teachers (Lieberman, 1995; Lieberman & McLaughlin, 1992; Nelson & Eddy, 2008). Ongoing reflective practice that is supported organizationally, such as through professional learning communities, can help to increase teachers' perceptions of their efficacy and mastery (Kennedy & Smith, 2013).

Appropriate and relevant professional development. As discussed in Chapter 3, all SPRUCE Instructors take part in extensive professional development, including summer

workshops, on-site training, and ongoing check-ins with SPRUCE staff. This professional development and continuous support helped the participants to learn about and share effective assessment practices specifically for the uniform assignments that they used in ENG 135. Having structured training and a shared language to discuss their teaching contributed to participants' effective use of questioning, discussion, and feedback practices. In addition, these resources facilitated the ongoing reflection of the participants in this study on their teaching. As research has shown that relevant professional development has a positive impact on teachers' use of FA practices (e.g., Birenbaum, Kimron, & Shilton, 2011; Heitink, Van der Kleij, Veldkamp, Schildkamp, & Kippers, 2016; Smith, 2011; Yin, Olson, Olson, Slovin, & Brandon, 2015), programs such as SPRUCE could be further studied and even used as models of best practices for effective professional development.

Relevance to accountability movements. The findings of this study, although not explicitly about high-stakes assessment nor teacher evaluation, are still relevant to the current culture of educational accountability in the U.S. Teacher evaluations that are based, often in large part, on assessments of students' performance, can also be tied to classroom assessments (Popham, 2013) and teachers' efficacy and agency. The latter two factors are further explored below.

Teacher efficacy. The experienced teachers in this study exhibited a strong sense of efficacy, as evidenced by their ability to justify how and why they made the decisions they did about their use of assessment practices. Teachers with higher senses of self-efficacy are likely to be more open to instructional innovations (Guskey, 1988; Tschannen-Moran & Hoy, 2001) and, when working together, can positively impact an organization's entire culture (Kennedy & Smith, 2013). Interestingly, meta-analyses of classroom-level factors that have the greatest

influence on teaching effectiveness include several components that emerged from this study, including structuring and scaffolding lessons, teaching modeling, and application of knowledge (Kyriakides, Christoforou, & Charalambous, 2013), as well as teacher expectations, classroom climate, the questioning level used (i.e., higher-order or lower-order), and feedback (O'Neill, 1988). Further study into teachers' use of practices that are related to FA can simultaneously inform research into factors that contribute to teacher efficacy.

Teacher agency. In an age of accountability and corresponding school reform, teachers' professional identity is often evaluated according to external definitions of progress and achievement (Day, 2002; Lasky, 2005). Within these environments, teachers and their professional agency, or "the power to act, to affect matters, to make decisions and choices, and take stances" (Vähäsantanen, 2015, p. 1) can be greatly impacted. The teacher participants in this study noted that they had the freedom to select and implement appropriate practices with their students in their SPRUCE ELA courses. A continual fostering of teachers' agency in their work, such as what is supported by SPRUCE—and by extension, the schools that partner with the program—with their Instructors, can be used as an example of a way in which teachers' motivation, efficacy, and commitment (Day, 2002) can be cultivated by other schools and programs.

Implications for educational evaluation. Since FA involves teachers using evaluative skills and strategies to guide and improve student learning through timely feedback, it is akin to program evaluators offering recommendations for improvement (Ayala & Brandon, 2008). Continued study of FA practices, including questioning, discussion, and feedback, and how they inform teachers' instructional decision making can ultimately help to inform educational program evaluation, especially for formative evaluation purposes. Studies such as this one can

add to empirical research about evaluation processes in practice, which, in turn, can contribute to the ongoing development and application of empirically grounded theory in evaluation (Christie, 2012). Additionally, the creation, development, and use of appropriate instruments as means of measuring what occurs within a classroom will assist evaluators in their study of educational programs and their outcomes through the use of context-specific tools.

Implications for the study of formative assessment. This study contributed empirical research to the field of formative assessment in a grade level and content area that is subjective and not frequently studied. Besides learning more about teachers' use of questioning, discussion, and feedback practices with their students, the ELATE checklists also captured teachers' perceived effectiveness of these practices with students at three different performance levels. Further information was learned through interviews of the teachers to better understand their decision making and professional knowledge in their use of and thinking about these practices.

These findings address McMillan's (2003) call for more exploration of how teachers address individual differences among students in their assessment, and Brookhart's (2003) suggestion for more attempts at understanding how assessment takes place within the context of a classroom. This study not only explored teachers' in-class use of assessment practices, but also examined their thinking behind making decisions about assessment. Participants were able to identify and differentiate between their ENG 135 students at different performance levels, and report and justify their perceptions of the effectiveness of the assessment practices that they used with those students. In addition, they were able to provide examples of how these practices are used with their students in the specific contexts of their respective classrooms.

Of particular note, the findings of this study did align with prior research on FA in terms of the major components that make up this concept. An implication for the results of this study is

that prior research on FA, regardless of content area, can be utilized to better understand FA in less-understood subject areas, such as ELA at the secondary level.

Future Research

Although the results of this research are specific to the population under study, it is expected that the findings will contribute to ongoing classroom assessment and FA research. For example, the mixed methods approach and the instruments used in this study may be adapted for use with other populations of teachers, given their relatively successful use with a sample of SPRUCE ENG 135 Instructors. Work that stems from this study will continue to add to our collective understanding of both explicit and implicit teacher knowledge about the effective use of assessment in the classroom.

Future research that builds upon this dissertation study will likely begin with revision of the ELATE checklist. For example, rather than asking for the total number of times a practice was used in one week, I would separate the "Times Used" category into three parts for the lower-, average-, and higher-performing students in the class under study, and have teachers report the frequency for each group. That would facilitate the interpretation of assessment practice use when the teachers then rate their perceptions of the effectiveness of that practice across the three levels.

Other work that is needed is further exploration into the reliability and validity of the ELATE checklist. This would involve recruiting and working with a much larger number of teacher participants, so missing data will not have as great an impact on the calculation of reliability statistics. The addition of a classroom observation component, such as an observation checklist, to teachers' self-report through the ELATE checklists would help to improve validity. Use of the instruments at different points in the academic year, such as at different times in the

semester, or even following teachers from September through June to determine if their assessment practices change over time, can provide additional insight into their use of and thinking about their assessment practices.

Along with refinement of the instruments, another research area to pursue is the study of different populations. For example, how might an ELA teacher responsible for instructing 9th graders of varying performance and ability levels in a general-level course respond to the same questions that were used here? Revised instrumentation could also be used with other populations of teachers with varying levels of experience, rather than focusing on expert teachers, as was done in this study. The data collection tools might be further expanded to examine the FA practices of teachers in other subjective content areas besides ELA.

Longer-term research ideas stemming from this study may include an exploration into how teachers use the data they learn from their assessment practices to inform their instruction. Teachers' perceptions of the effectiveness of assessment practices could also be compared with students' perceptions of what is most efficacious for them, using data from sources such as student focus groups or surveys.

There are also potential applications for the findings of this study. For example, a stronger understanding of experienced teachers' perceptions of FA practices as a form of classroom assessment and how these practices inform instruction may be applied to training pre-service teachers in ELA and other subjective content areas. The findings may also inform future work in evaluating teacher preparation programs and teachers' pre- and in-service training in classroom assessment and instructional practices, as well as evaluating the support systems in place that enable context-appropriate implementation of these practices.

Conclusions

In an age of educational accountability for teachers and students in K-12 education, careful attention must be paid to the assessment and instruction practices that are actually occurring in classrooms. Although formative assessment has emerged as a viable approach to improving student learning and achievement, there is currently a dearth of empirical knowledge on how the practices that fall under the aegis of this concept are used in more subjective content areas, such as secondary ELA.

This study examined the questioning, discussion, and feedback assessment practices of secondary ELA teachers, using a sample of SPRUCE ENG 135 Instructors, in Spring 2015. Using a mixed methods sequential explanatory approach that involved teachers completing a series of weekly self-report checklists about their in-class practices and semi-structured individual interviews, this study explored (a) what assessment practices secondary ELA teachers use, (b) how these teachers determined which assessment practices to use and when to use them, (c) what these teachers' perceptions of the effectiveness of the assessment practices they used were, and (d) how teachers determined the effectiveness of these assessment practices. Through this research with expert teachers, a greater understanding was developed about the use of FA-related practices by secondary ELA teachers, specifically with regard to questioning, discussion, and feedback.

The findings of this dissertation study align with prior research on the operationalization and use of FA in other contexts. Previously studied FA components, which involved teachers setting and making explicit clear learning goals, using questioning and feedback practices with students, collecting and using assessment data in their instruction, revising and modifying their

instruction, developing student academic autonomy, and evaluating potential gaps between current and expected student learning progress, all emerged through this research.

In addition, more was learned about teachers' decision making for the selection and determination of the effectiveness of the assessment practices that they used with their students. The results of this study suggest that secondary ELA teachers' use of assessment practices is closely tied to the context of the classroom, especially in terms of students' ability and performance levels. Along with more insight into expert teachers' thinking about assessment, a major contribution of this study to the field is the finding that there were statistically significant differences between teacher participants' perceptions of effectiveness of the assessment practices that they reported using with their lower-, average-, and higher-performing students.

In summary, the following are the five major "lessons learned" from this study:

1. Teachers can be very self-reflective about assessment as they teach.

2. Teachers vary their assessment practices depending on their perceptions of student performance (e.g., lower, average, and higher) and ability levels, as well as their perceptions of student learning progress.

3. Teachers reason about what assessment practices to use, when to use them, and to what effect they can be used.

4. Teachers evaluate the utility of the assessment practices that they use with their students and change their instruction based on them.

5. Teachers' use of assessment practices is an iterative, emergent process conducted thoughtfully and with full attention to changing contextual factors; it is an organic process with the ultimate aim of improving student learning.

Appendix A: SPRUCE ELA Program ENG 135 Instructor Recruitment Email

Subject: Invitation to collaborate in assessment practices research study

Dear <<firstname>>:

When I was a secondary ELA teacher in an urban district outside of Philadelphia, I found that assessing student progress was one of my most challenging tasks. Teaching ELA to high schoolers requires such careful attention to individual pupil differences, subtle indicators of learning and growth, and thoughtful judgments about how to help each student. I know that as a SPRUCE ELA Program Instructor currently teaching ENG 135, you have to deal with these issues on a daily basis. How do we best assess student learning in our daily teaching?

As a doctoral candidate in the Instructional Design, Development and Evaluation (IDD&E) program at SU, I am studying how we can improve our assessment procedures, but first I need to better understand how ELA teachers—how you—assess your students. What works and what doesn't? How do you decide?

Might you be willing to collaborate with me in this study for research purposes by sharing your practices and thoughts?

If you are willing to work with me, I would like to talk with you about your assessment practices. This chat would take about 45-60 minutes, and would be conducted via Skype; to help maintain your privacy, I will be in an office behind closed doors. In order to prepare for our conversation, I think it would be useful for us to have some specific information about your recent assessment practices, including which ones you are using and why. If you were to record what you are currently doing on a very short checklist, which I will email to you and you will return to me via email once a week for four weeks, we would then have some concrete information to consider in our conversation.

I believe that our collaboration will provide insights into effective assessment practices to improve student learning. After our ongoing discussion about your individual practices, I will share with you a personalized summary of what we discussed as a thank you for your time and for working with me in this process.

If you might be interested in participating in this research, or would like to discuss the study with me before making a decision, please contact me at lmtolley@syr.edu, or call my cell phone at [number]. I am happy to answer any questions that you may have.

Thank you in advance, and I look forward to hearing from you soon!

With best regards,

-Leigh

- - - - -
Leigh M. Tolley, M.S.Ed., M.S.
Doctoral Candidate
Instructional Design, Development and Evaluation
School of Education, Syracuse University

Appendix B: Instructions for Participation in the Study

Dear <<firstname>>:

Hello, and thank you for your interest in participating in my study! Enclosed with this message are several files:

1. The informed consent document for this study.
Please read the informed consent document (besides the attached file, the text is pasted below this message), which provides an overview of the study, its benefits and risks, and contact information, should you have any questions about this research. **Please let me know via email (**lmtolley@syr.edu**) if you are willing to participate in the study, and if so, if I have your permission to audio record our interview when we speak**.

After you have decided to participate in the study, please complete the other two files. Both are PDF forms created in Adobe, so that you can type and save your responses and return them to me as email attachments at lmtolley@syr.edu. If you prefer to write your responses on paper, you can print the forms, hand-write your responses, and then scan the forms and email the images to me. I will contact you to let you know when I have received your completed forms.

Please note that there are no "right" or "wrong" answers. This study is addressing matters of professional judgment, and you are sharing your professional experiences as an ENG 135 Instructor. Your personally identifying information will be kept *confidential* and will not be shared with others, including my dissertation committee and SPRUCE staff and administrators.

2. A background questionnaire.
The background questionnaire will help me to learn about you and your teaching experiences, which may be related to your assessment practices. This questionnaire will take about 5-10 minutes to complete. It asks about you, your background as a teacher, and the courses that you are teaching this year.

3. A blank copy of the English/Language Arts Teachers' Experiences (ELATE) checklist.
The ELATE checklist is a list of assessment practices, related to questioning, discussion, and feedback, that you may have used with your ENG 135 students over the course of the past week (last Friday through Thursday). It will take about 15-20 minutes to complete. There are instructions on the first page of the checklist about how to indicate which practices you used and how effective you felt those practices were. Please return your completed form to me by the end of the week. I will let you know when I have received your completed checklist, and will send you a blank copy for the next week.

If you have any questions as you are completing the forms, or about this study in general, either now or in the future, please do not hesitate to contact me via email (lmtolley@syr.edu) or phone ([number]).

Thank you, and I look forward to working together!

Best regards,

-Leigh

Appendix C: Teachers' Background Questionnaire

Background Questionnaire

Teachers' Background Questionnaire

The purpose of this study is to learn more about the assessment practices that you use when teaching English/language arts to high school students. In particular, I am interested in learning about the questioning, feedback and discussion assessment practices you are using on a weekly basis with your Student Preparation and Readiness through University Concurrent Enrollment (SPRUCE) English/Language Arts (ELA) Program students in the ENG 135 (Critical Reading: Issues, Theory, and Applications) course.

This questionnaire will ask about your background as a teacher. By completing this questionnaire, you assure that you are at least 18 years of age and that you consent to participate in this study. Your participation in this study is voluntary. The information collected for this study will be confidential, and only de-identified data will be reported.

Background Questions

1. What is your gender? _____

2. How many years have you been teaching? *(Please round to the nearest year.)* _____

3. How many years have you been teaching in the SPRUCE ELA Program? *(Please round to the nearest year.)* _____

4. How many sections of SPRUCE ELA Program courses are you teaching this academic year? _____

5. On average, how many students do you have in each of your SPRUCE course section(s)? _____

6. How many students do you currently have in your SPRUCE ENG 135 section(s)? _____

7. What is the total number of students that you are teaching this academic year across all of your courses? _____

8. What other courses are you teaching this academic year? Please list below the course name(s), the number of sections you are teaching, and the grade level(s) of the students taking those courses. *(If not applicable, please write N/A in the first blank space.)*

Course Name	*Number of Sections*	*Grade Level(s)*
a.		
b.		
c.		
d.		
e.		

Thank you for completing this questionnaire!

Appendix D: English/Language Arts Teachers' Experiences (ELATE) Checklist

ELATE Checklist

English/Language Arts Teachers' Experiences (ELATE) Checklist

Think about your teaching with your SPRUCE ENG 135 students this past week (last Friday through Thursday). The checklist on the following pages includes a list of assessment practices that you may have used. These selected practices are related to questioning, discussion, and feedback.

First, on the lines next to **Dates Being Reported** at the top right of each page of the checklist, please indicate the dates for which the checklist is being completed.

Next, read each practice that is listed in the checklist. In the **How Many Times Did You Use This Practice?** box, please indicate approximately how many times you used that particular practice in teaching ENG 135 this week by writing in a whole number. If you did not use that practice, please write **0** for zero.

Then, for each practice you indicated *using* this week, please rate how effective you feel it was for the purposes of improving student learning for your lower-performing, average-performing, and higher-performing ENG 135 students. In the appropriate boxes in the **Perception of Effectiveness of This Practice to Improve Student Learning** section, write in a whole number from 0-5, using the following scale:

$$0 = \text{extremely ineffective}$$
$$1 = \text{very ineffective}$$
$$2 = \text{moderately ineffective}$$
$$3 = \text{moderately effective}$$
$$4 = \text{very effective}$$
$$5 = \text{extremely effective}$$

If you used any additional questioning, discussion, or feedback assessment practice(s) with your ENG 135 students that is/are not included on this list, please write in the name of the practice(s) used in the **Other Practices** section at the end of the checklist. Please then indicate *how many times you used each additional practice* and rate *how effective you feel it was for the purposes of improving student learning*.

By completing this checklist, you assure that you are at least 18 years of age and that you consent to participate in this study. Your participation in this study is voluntary. The information collected for this study will be confidential, and only de-identified data will be reported.

ELATE Checklist

Dates Being Reported: _____ **to** _____

In teaching ENG 135 this week (last Friday through Thursday), I…	How Many Times Did You Use This Practice?	Perception of Effectiveness of This Practice to Improve Student Learning		
		0 = extremely ineffective 3 = moderately effective 1 = very ineffective 2 = moderately ineffective 4 = very effective 5 = extremely effective		
		Lower-Performing Students	Average-Performing Students	Higher-Performing Students
Questioning and Discussion				
1. Asked students questions to determine how well they understood a concept, idea, or strategy.				
2. Asked questions of individual students by name.				
3. Asked questions of the class as a whole.				
4. Asked questions specifically of students I felt were not paying attention.				
5. Asked questions of students I thought would be more likely to respond well.				
6. Asked questions of reticent students to help improve their participation.				
7. Asked questions requiring brief responses (e.g., a word or phrase).				
8. Asked questions requiring more elaborated responses (e.g., a few sentences).				
9. Asked questions intended to stimulate a general discussion.				
10. Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.				
11. Used large group (5 or more students) discussion to determine how well students understood a concept, idea, or strategy.				
12. Used whole-class discussion to determine how well students understood a concept, idea, or strategy.				

ELATE Checklist

Dates Being Reported: _____ **to** _____

In teaching ENG 135 this week (last Friday through Thursday), I…	How Many Times Did You Use This Practice?	Perception of Effectiveness of This Practice to Improve Student Learning		
		0 = extremely ineffective 3 = moderately effective 1 = very ineffective 2 = moderately ineffective 4 = very effective 5 = extremely effective		
		Lower-Performing Students	Average-Performing Students	Higher-Performing Students
Feedback				
13. Reviewed and gave feedback (oral or written) on students' think/response papers.				
14. Reviewed and gave feedback (oral or written) on students' dialogic journals.				
15. Reviewed and gave feedback (oral or written) on students' portfolios of their writing.				
16. Reviewed and gave feedback (oral or written) on students' *draft* versions of a writing assignment.				
17. Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.				
18. Used a rubric or rubrics for *informal* feedback to students on their writing.				
19. Used a rubric or rubrics for *formal* feedback to students on their writing.				
20. Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.				
21. Gave students *general or holistic* written feedback about their writing, such as comments at the end of an assignment about how they did overall.				
22. Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.				
23. Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing.				
24. Gave oral feedback to multiple students at once about their writing (e.g., discussing strategies with groups or the entire class).				
25. Conferenced with individual students about their writing to give them feedback and to review and discuss their work.				

ELATE Checklist

Dates Being Reported: _____ **to** _____

In teaching ENG 135 this week (last Friday through Thursday), I…	How Many Times Did You Use This Practice?	Perception of Effectiveness of This Practice to Improve Student Learning 0 = extremely ineffective 1 = very ineffective 2 = moderately ineffective 3 = moderately effective 4 = very effective 5 = extremely effective		
		Lower-Performing Students	Average-Performing Students	Higher-Performing Students
Other Practices				
26. Other:				
27. Other:				
28. Other:				
29. Other:				
30. Other:				

Thank you for completing this checklist!

Appendix E: General Semi-Structured Interview Protocol

Hello _____! Thank you so much for speaking with me- it's great to finally have the chance to talk with you! As I've mentioned before, I am very grateful for the opportunity to learn from you about your teaching, and I really appreciate your working with me.

To begin our chat today, I'd like to give you a brief overview of what I'm anticipating for our discussion. As you know, I'm researching the assessment practices of secondary ELA teachers, specifically SPRUCE ENG 135 Instructors. I have several questions I'd like to ask you about the assessment practices that you reported using to me over the past few weeks. I would also like to talk about how you choose which practices to use and how you make decisions about which practices are effective or not. Of course, I'm happy to discuss other related topics that come up. I greatly appreciate you sharing your expertise with me!

Before we get started, do you have any questions for me?

Thank you for letting me know via email that I have permission to audio record our conversation. Once I start recording, I will re-confirm that I have your permission, and record your consent on tape. May I start recording now?

I have started audio recording, and I am speaking with _____.

So that I have confirmation on the audio recording, are you willing to participate in this interview? (Yes/No)

To confirm your consent, do I have your permission to audio record this interview? (Yes/No)

_____, thank you again for helping me learn about you and your teaching.

1. To get started, tell me about the students that you taught in ENG 135 this semester.

    1a. What are their backgrounds (such as demographics, socio-economic status, or prior education)?

    1b. What similarities or differences do you see between your ENG 135 students and the other students in your school?

Next, let's talk about your assessment practices.

2. You reported that you used [*practice*]³ many times in teaching ENG 135 during these past few weeks. Tell me what that looks like in your classroom—can you give me or describe some actual examples of your use of that practice?

    2a. What kinds of information does [*practice*]⁴ give you about your ENG 135 students?

    2b. How do you use that information?

3. [*individualized question for teacher re: practices*]⁵ When I reviewed the practices that you reported using with your ENG 135 students, …

4. In thinking about the assessment practices you reported using across all of the checklists, how representative are they of what typically takes place over an ENG 135 semester?

---

³ Example from interview with Mia: "You reported that you reviewed and gave students feedback on their think/response papers many times in teaching ENG 135 during these past few weeks."

⁴ Example from interview with Mia: "What kinds of information does reviewing students' response papers (and giving them feedback) give you about your ENG 135 students?"

⁵ Example from interview with Allison: "When I reviewed the practices that you reported using with your ENG 135 students, I noticed you used individual and group conferencing with your students fairly often. How does conferencing help you to understand students' learning progress? How do you use the information from these conferences in your teaching?"

5. What are the similarities and differences between the assessment practices you use with your ENG 135 students and the assessment practices you might use with your other ELA classes?

Now I'd like to talk about the effectiveness of assessment practices.

6. [*individualized question for teacher re: efficacy*][6] When I reviewed your perceptions of the efficacy of the assessment practices you used, …

7. How do you typically determine if an assessment practice is effective or not?

Next, I'd like to discuss assessment in general.

8. How do you define "assessment"?

    8a. How do you understand assessment's role in the way that you teach? (i.e., How does assessment impact your teaching?)

    8b. How do you decide which assessment practices to use for students at different ability levels?

    8c. How have your thinking about and practicing of assessment changed over the years that you've been teaching?

Thank you very much—that's it for my questions. Is there anything else that you would like to share with me about the assessment practices that you used in teaching ENG 135 these past few weeks?

In case I have any questions as I'm reviewing my notes, what would be the best way to contact you over the next few months (same email, phone, etc.)?

---

[6] Example from interview with Charlotte: "When I reviewed your perceptions of the efficacy of the assessment practices you used, I noticed that you rated giving students feedback on their think/response papers as 'very effective' for all of your students, but you only reported using this practice once. Why do you view this practice as "very effective"?"

Once I've reviewed and analyzed everything, I would like to share with you a summary of what we discussed through the ELATE checklists, email, and this conversation. To what mailing address should I send this information (school, other)?

Thank you again for your time, _____! I hope that you enjoy your summer, and I'll let you know if I have any additional questions.

Appendix F: All Teachers' Reported Use of Practices by Item

All Teachers—Times Each Practice Used Per Week, as Reported on ELATE Checklists

W1 = Week 1, 05/01/15-05/07/15
W2 = Week 2, 05/08/15-05/14/15
W3 = Week 3, 05/15/15-05/21/15
W4 = Week 4, 05/22/15-05/28/15
W5 = Week 5, 05/29/15-06/04/15
W6 = Week 6, 06/05/15-06/11/15
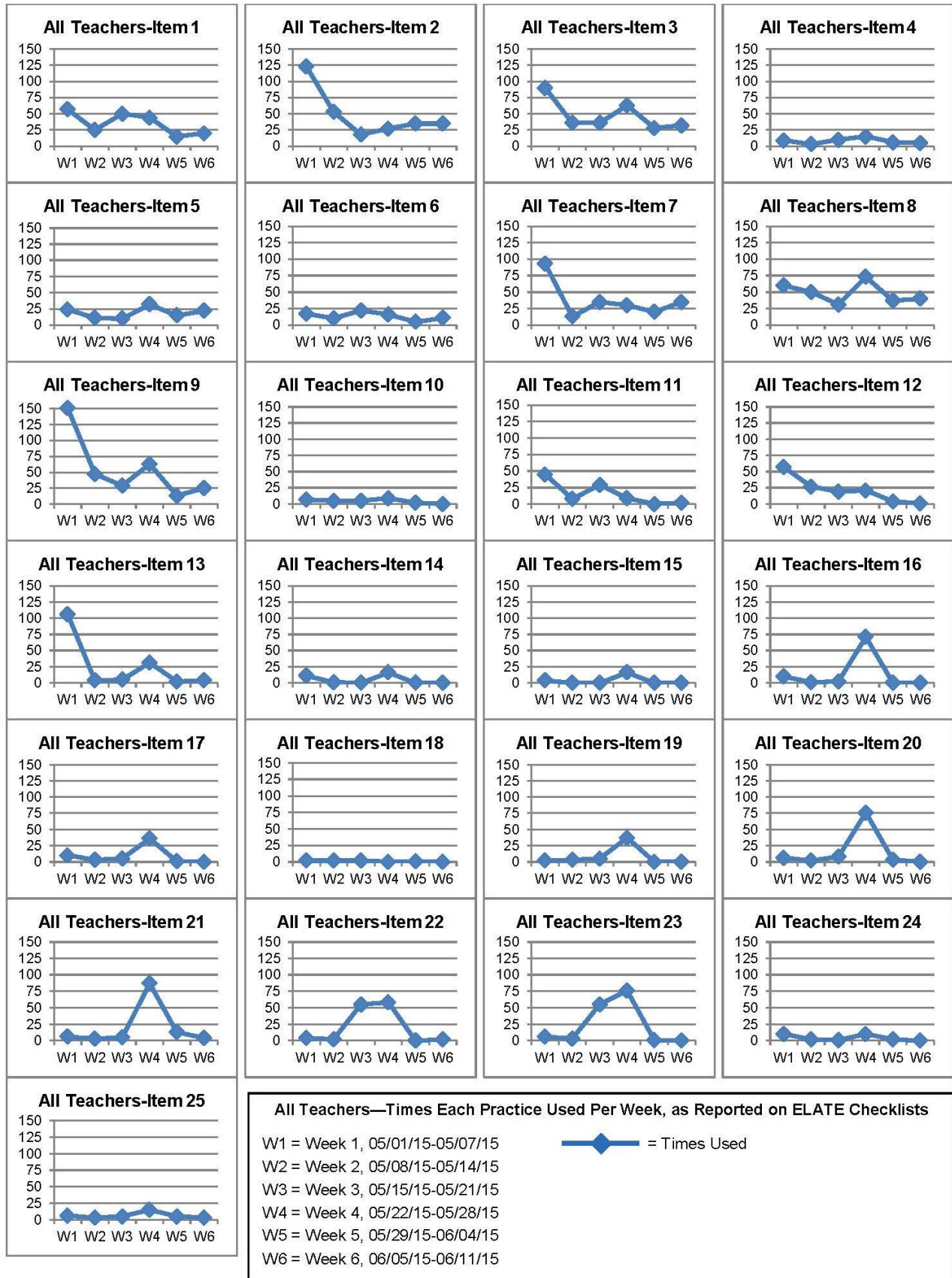
= Times Used

Appendix G: All Teachers' Perceptions of Effectiveness of Practices Used by Item

All Teachers—Perceptions of Each Practice's Effectiveness to Improve Student Learning, as Reported on ELATE Checklists

W1 = Week 1, 05/01/15-05/07/15   ■ = Perceived Effectiveness for Lower-Performing Students
W2 = Week 2, 05/08/15-05/14/15   ▲ = Perceived Effectiveness for Average-Performing Students
W3 = Week 3, 05/15/15-05/21/15   ● = Perceived Effectiveness for Higher-Performing Students
W4 = Week 4, 05/22/15-05/28/15
W5 = Week 5, 05/29/15-06/04/15
W6 = Week 6, 06/05/15-06/11/15

Effectiveness Scale: 0 = extremely ineffective; 1 = very ineffective;
2 = moderately ineffective; 3 = moderately effective;
4 = very effective; 5 = extremely effective

Appendix H: All Teachers' Reported Use of Practices
and Perceptions of Effectiveness of Practices Used by Item

All Teachers—Times Each Practice Used Per Week and Perceptions of Each Practice's Effectiveness to Improve Student Learning, as Reported on ELATE Checklists

W1 = Week 1, 05/01/15-05/07/15   ◆ = Times Used
W2 = Week 2, 05/08/15-05/14/15   ■ = Perceived Effectiveness for Lower-Performing Students
W3 = Week 3, 05/15/15-05/21/15   ▲ = Perceived Effectiveness for Average-Performing Students
W4 = Week 4, 05/22/15-05/28/15   ● = Perceived Effectiveness for Higher-Performing Students
W5 = Week 5, 05/29/15-06/04/15
W6 = Week 6, 06/05/15-06/11/15

Effectiveness Scale: 0 = extremely ineffective; 1 = very ineffective; 2 = moderately ineffective; 3 = moderately effective; 4 = very effective; 5 = extremely effective

All Teachers-Item 16

All Teachers-Item 17

All Teachers-Item 18

All Teachers-Item 19

All Teachers-Item 20

All Teachers-Item 16

All Teachers-Item 17

All Teachers-Item 18

All Teachers-Item 19

All Teachers-Item 20

All Teachers-Item 21

All Teachers-Item 22

All Teachers-Item 23

All Teachers-Item 24

All Teachers-Item 25

All Teachers-Item 21

All Teachers-Item 22

All Teachers-Item 23

All Teachers-Item 24

All Teachers-Item 25

All Teachers— Times Each Practice Used Per Week and Perceptions of Each Practice's Effectiveness to Improve Student Learning, as Reported on ELATE Checklists

W1 = Week 1, 05/01/15-05/07/15   ◆ = Times Used

W2 = Week 2, 05/08/15-05/14/15   ■ = Perceived Effectiveness for Lower-Performing Students

W3 = Week 3, 05/15/15-05/21/15   ▲ = Perceived Effectiveness for Average-Performing Students

W4 = Week 4, 05/22/15-05/28/15   ● = Perceived Effectiveness for Higher-Performing Students

W5 = Week 5, 05/29/15-06/04/15

W6 = Week 6, 06/05/15-06/11/15   *Effectiveness Scale: 0 = extremely ineffective; 1 = very ineffective;*
2 = moderately ineffective; 3 = moderately effective;
4 = very effective; 5 = extremely effective

Appendix I: Teacher Snapshots

SPRUCE ENG 135 Teacher Snapshot: Spring 2015	
Teacher Name	Mia
Total Number of Years Teaching	17
Total Number of Years Teaching in the SPRUCE ELA Program	7
School Locale	Suburb: Large
Total Number of SPRUCE ENG 135 Students	50
Total Number of Students in All Courses	75
Weeks in which ELATE Checklists Were Completed	May 1-7, 8-14, 15-21, and 22-28, 2015
Date of Interview	July 17, 2015

What were the performance levels of this teacher's SPRUCE ENG 135 students?

"Technically, all of my [ENG 135] kids are higher performing since my school [has] a pre-requisite for the class. Whether or not the students work to that level all of the time is questionable, but it most cases of research, these kids would be considered high performing" (email dated May 13, 2015).

Most Frequently Used Assessment Practices, as Reported on the ELATE Checklists

Item #	Assessment Practice	Total Times Reported as Used
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	52
3	Asked questions of the class as a whole.	48
12	Used whole-class discussion to determine how well students understood a concept, idea, or strategy.	48
9	Asked questions intended to stimulate a general discussion.	35
11	Used large group (5 or more students) discussion to determine how well students understood a concept, idea, or strategy.	33

Why did this teacher choose to use a certain assessment practice (or practices) often?

Mia said the following about reviewing her students' think papers: "they're different, obviously, than the unit papers, and it's more of a, a — a quicker assessment, I write all my feedback on those, on those papers for them. So, um, any time they have one submitted, it's an extensive commentary on what they're writing that is given back to them, and then if they have questions about what I've written, then we'll have a conference based on that. […] I would have individual conferences with them to work on their current writing, and that way they can take those practices, and then move — and then develop their final piece from there" (interview dated July 17, 2015).

Assessment Practices Perceived to be Most Effective for Lower-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
N/A	[*Not Applicable (N/A)—only perceptions of effectiveness of practices for higher-performing students were rated*]	N/A

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Average-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
N/A	[*Not Applicable (N/A)—only perceptions of effectiveness of practices for higher-performing students were rated*]	N/A

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Higher-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
9	Asked questions intended to stimulate a general discussion.	5.00
16	Reviewed and gave feedback (oral or written) on students' **draft** versions of a writing assignment.	5.00
1	Asked students questions to determine how well they understood a concept, idea, or strategy.	4.67
20	Gave students **targeted** written feedback about their writing, such as comments on certain parts of their assignment.	4.67
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	4.50

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

How does this teacher determine the effectiveness of an assessment practice?

"You can kind of see it in either, you know, in their writing, or their discussion, or they'll tell you. Like, some kids, especially the, the high-level kids, they're like, 'I can't stand the peer editing. I'm so….' And I'm 'OK, I'm sorry!' [*laughs*] You know? And then they'll tell you! And so, there, there are times where you're doing something specific, and the kids are like, 'This isn't working for me.' And, you know, that I can change. But it's not — it's, it's basically in the product. So you can see as — you know, and you, you walk around. You listen, you look, you see what they're writing. As you're doing that, you know if they're — if what you're doing is positive or not" (interview dated July 17, 2015).

What is this teacher's definition of "assessment"?

"Assessment is the discovery of what the students have learned throughout a specific unit, or a specific time period, and, um, it can come in many different forms. But that's generally what assessment is: assessment of what is being learned" (interview dated July 17, 2015).

What does an example of an assessment practice look like in this teacher's ENG 135 class?
Mia described how whole-class discussion occurred in her ENG 135 class:

"I like whole-group discussion with [SPRUCE], because there are so few kids in the classroom. If I had 30 kids, or 35 kids, it wouldn't work — it doesn't work like that, and that's when small group becomes more successful for me, as, as a formative assessment. But I think because of the small groups [SPRUCE] lends itself to…I […] can do these whole group, and, and talk to every single kid. Whereas, u— [*sic*] usually when you're doing whole-group discussion, you don't get the opportunity to do that" (interview dated July 17, 2015).

When asked what her role would be during whole-class discussion, Mia responded:

"You're a facilitator. Uh, you start the questioning — in the — early in the semester, you do do more talking, um, because […] you're kind of modeling the kind of inquiry that you're asking the students to do. And, by the end of the semester, I sit there, and I can not talk for, you know, 40 minutes. [*says while smiling*] I, I mean, I do — I'll redirect, if I need to, or, um, I'll say, 'Hey, you know what, that's really interesting. Let's go back to that, and talk about that.' But, by the end of the semester, they can do it — they run it themselves, and I, I simply am there to facilitate" (interview dated July 17, 2015).

SPRUCE ENG 135 Teacher Snapshot: Spring 2015	
Teacher Name	Helen
Total Number of Years Teaching	27
Total Number of Years Teaching in the SPRUCE ELA Program	15
School Locale	Suburb: Large
Total Number of SPRUCE ENG 135 Students	51
Total Number of Students in All Courses	102
Weeks in which ELATE Checklists Were Completed	May 1-7, 22-28, and May 29-June 4, 2015
Date of Interview	June 26, 2015

What were the performance levels of this teacher's SPRUCE ENG 135 students?

"Technically, all our [ENG 135] students are 'higher performing' because they are only in the class by virtue of their past performance in English. To be in the class they need an A average in English for the first three years of high school" (email dated May 7, 2015).

Most Frequently Used Assessment Practices, as Reported on the ELATE Checklists

Item #	Assessment Practice	Total Times Reported as Used
21	Gave students *general or holistic* written feedback about their writing, such as comments at the end of an assignment about how they did overall.	63
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	51
16	Reviewed and gave feedback (oral or written) on students' *draft* versions of a writing assignment.	51
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	51
23	Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing.	51

Why did this teacher choose to use a certain assessment practice (or practices) often?

Helen described how she gave students feedback on their writing to help develop their writing skills. She said her ENG 135 students were at a point where "it's more about, um, a kind of more sophisticated way of looking at the writing. […] The writing course in the fall really sets them up with the expectation of, you know, clear thesis, developing, et cetera, et cetera. So, by the time we are doing this, um, now it's — they have to sort of blend what they truly think about a topic […] it becomes more personal, in a way, but not casual."

She added, "in the fall, we're really pushing the fundamentals, and kind of, um, understanding what makes something good, and how you yourself can look at what you've done and decide: is this good? Did I do what I was supposed to do? And then — so then, in the spring,

we're looking for you to be able to do all of that by yourself, and now try to, to reach these more sophisticated levels of expression and connection, and do it, you know, with a sophisticated grasp of language. That, um, no one has to be in the classroom discussion to understand what you're doing; it should be that your stuff — your work — stands on its own, and it stands — you know, makes sense, without having it to be just a part of a whole class discussion on something" (interview dated June 26, 2015).

Assessment Practices Perceived to be Most Effective for Lower-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
N/A	[*Not Applicable (N/A)—only perceptions of effectiveness of practices for higher-performing students were rated*]	N/A

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Average-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
N/A	[*Not Applicable (N/A)—only perceptions of effectiveness of practices for higher-performing students were rated*]	N/A

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Higher-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
1-9, 11-13, 16, 20-25	[***Note***: *Helen rated her perception of the effectiveness of all of the assessment practices that she used with her higher-performing students as 5.0. She did not report using items #10 (pair/small group discussion), 14 (feedback on dialogic journals), 15 (feedback on portfolios), 17 (feedback on final versions of writing) and 18-19 (use of rubrics for informal/formal feedback).*]	5.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

How does this teacher determine the effectiveness of an assessment practice?

Helen said of her assessment practices, "I found that the ones I used, I used because they are the most effective. And, you know, and it's probably, um, a combination of my personality and my way of doing things. I personally, my system is the system that works for me. And I just feel that, as long as it works for me, and they're getting something out of it, and — and I see by the improvement in their writing that they are getting something out of it, um, then it's — then it's good, you know, and — and it's what I'm happy doing" (interview dated June 26, 2015).

What is this teacher's definition of "assessment"?

 "What it means to me is an honest attempt to evaluate how much learning has gone on. And, you know, you have to start from September, so you start at zero, but…I think it's very important to — for each student, to remember where the student started, and where the student is now. And, um — and I think assessment should in— [*sic*] involve, um, an appreciation of effort, and, um, and — and the amount of progress that a person makes, you know. Um, some people make a lot of progress, because they work really hard and they've — they progressed, some people are content to just kind of skim the surface, you know. OK. But, then, you know, I — I'm thinking you're missing an opportunity here to get better, and that sort of thing, and so I'll try to nudge 'em along, and, um, you know. But, I also think that assessment is, uh, I think it's — it should never be hard and fast, it should never be, um…you know, definitive for the sake of being definitive."

 Helen later added, "assessment, I think in the beginning, is — as much your, your gut instinct as to who you're talking to. You know, who your learners are, and that. And then as you get to know each other, assessment becomes, 'Oh, come on! You're still doing the same thing I've been telling you not to do,' or, or 'Wow! Look at you! You started here and now you're here,' you know. So I think assessment is — it's never one-size-fits-all" (interview dated June 26, 2015).

What does an example of an assessment practice look like in this teacher's ENG 135 class?

 Helen described how she used class discussion to address challenging topics and encourage students to share their ideas in her ENG 135 class:

 "What I've learned is, that a lot of people say, 'I never even thought of that until I heard somebody else say it.' And, you know — and with [ENG 135], you're talking about culture, ideology, um — things that a typical high school senior is not even aware [*laughs a little*] that he or she is a part of a culture, or subscribes to an ideology, you know. And, um, and so, bringing up those things, and challenging them with, you know, whether we use, you know, YouTube videos or, uh, you know, popular television — you know, like, taking a popular television show, and kind of taking it apart and, you know — 'Well, so what is this?', you know. And, and — and the kids will go, 'Oh yeah! That guy always has this point of view' kind of thing, you know. And, um — and then, you know, 'OK, so in your life, where do you see yourself?' So the classroom discussion is — I think for a lot of students, it generates ideas that they have, um, that they didn't even perhaps know that they possess, but when they hear it from another student, or when somebody says, 'But what about this? Is that…?' And you — and then the student can say, 'Oh yeah, that is,' you know.

 "So it's empowering, and it — it opens minds, and, um — and I think too, it's that — it's just that, like, thing that that we all do, that when you think of — so many times when you think of your reaction to something, or your idea about something, you think you're the only one, or, you know, 'I don't know what other people think, but…', you know. And then you find out, 'Oh, you're not alone,' you know. 'Other people do think this way'" (interview dated June 26, 2015).

SPRUCE ENG 135 Teacher Snapshot: Spring 2015	
Teacher Name	Allison
Total Number of Years Teaching	25
Total Number of Years Teaching in the SPRUCE ELA Program	10
School Locale	Rural: Fringe
Total Number of SPRUCE ENG 135 Students	17
Total Number of Students in All Courses	95
Weeks in which ELATE Checklists Were Completed	May 1-7, 8-14, 15-21, 22-28, and May 29-June 4, 2015
Date of Interview	July 1, 2015

What were the performance levels of this teacher's SPRUCE ENG 135 students?

Allison described her lower-performing students as those who "were much less likely to be coming to class prepared," which impacted their ability to complete the assignments. Of her 17 ENG 135 students, she identified four as lower-performing, including "three students who were in danger of not graduating because of performance," four higher-performing students "that I would really put in the top tier," and "the rest [10 students] were in the middle" (interview dated July 1, 2015).

Most Frequently Used Assessment Practices, as Reported on the ELATE Checklists

Item #	Assessment Practice	Total Times Reported as Used
3	Asked questions of the class as a whole.	29
8	Asked questions requiring more elaborated responses (e.g., a few sentences).	16
2	Asked questions of individual students by name.	15
9	Asked questions intended to stimulate a general discussion.	15
1	Asked students questions to determine how well they understood a concept, idea, or strategy.	14

Why did this teacher choose to use a certain assessment practice (or practices) often?

Allison talked about how she used whole-class discussions to assess her students' learning progress: "very often, it gives me an i— [*sic*] a, a pretty good idea of whether or not they understood [*laughs a little*] basic concepts. […] it helps me figure out what gaps I need to fill in. And sometimes, though, I'll find that they're — they've really handled it well, and then we can move on to something else. But, sometimes, you know, on Bloom's taxonomy, we're pretty low if we're dealing with something really challenging" (interview dated July 1, 2015).

Assessment Practices Perceived to be Most Effective for Lower-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
6	Asked questions of reticent students to help improve their participation.	5.00
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	5.00
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	5.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Average-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	5.00
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	5.00
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	5.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Higher-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
2	Asked questions of individual students by name.	5.00
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	5.00
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	5.00
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	5.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

How does this teacher determine the effectiveness of an assessment practice?

"In [ENG 135], it's often in the product, in the final product, the paper. […] I think we've developed some interesting practices in that class. For example, for any big project we have students do, instead of like a traditional outline, they might do, um, a PowerPoint, where they're working through the ideas — or some kind of a slideshow, where they're working through the ideas, um, including the theory. We have certain things that they have to put into that slideshow very early on, and it's like a pre-rough draft. And so, they get a lot of feedback on that, so you can often help shape a good project early on, um, but really, [what] it comes down to in that class is a portfolio analysis, and, and, uh, it depends on how well they end up doing on the paper. Did they get it or not?" (interview dated July 1, 2015).

What is this teacher's definition of "assessment"?

"It's a way of, um, checking to see whether or not students have mastered something I want them to master" (interview dated July 1, 2015).

What does an example of an assessment practice look like in this teacher's ENG 135 class?

Allison described how feedback on students' writing and individual conferencing might occur with her ENG 135 students:

"I'm often looking at their, their work; I'm looking at — even if I'm just looking at how they crafted a claim, or, um, maybe we're talking through what their plan is for a, a writing project. Remember, like, when you came into, uh, into the picture, we were at the end of the course, where we've done a lot of the, um, theory work already, and we're just working on the writing and the projects. So, often I can see where a student needs to sort of maybe restructure a, a paper. Very often, you know, they're afraid they're not going to have enough to say, so they take on a huge project, and so maybe my role is to help them cut it down to a manageable size, or maybe to shift direction. Um, if I — if I see that they're not saying anything particularly original, but they have this, like, really nice little kernel over here, sometimes we'll talk through that, and then the project takes a whole different direction."

She added, "if it's an individual conference, it might be that I'm, I'm looking at their project, and I see that they really could be bringing in this theory, that theory, this text, that text. But they haven't considered [it], because they just weren't thinking in those terms. So, individually, I can really help them shape projects, too" (interview dated July 1, 2015).

SPRUCE ENG 135 Teacher Snapshot: Spring 2015	
Teacher Name	Charlotte
Total Number of Years Teaching	20
Total Number of Years Teaching in the SPRUCE ELA Program	6
School Locale	Suburb: Large
Total Number of SPRUCE ENG 135 Students	17
Total Number of Students in All Courses	101
Weeks in which ELATE Checklists Were Completed	May 1-7, 8-14, 15-21, 22-28, and May 29-June 4, 2015
Date of Interview	June 29, 2015

What were the performance levels of this teacher's SPRUCE ENG 135 students?

Charlotte said that out of her 17 ENG 135 students, she had five that were "high-achieving," and "four low students." Of those lower-performing students, she described one as always turning in work late, another as a pupil who performed well in other college-level courses, but who "really just couldn't get out of the C range" in ENG 135, and "two of them [who] may have had previous special ed[ucation] classifications." She said that her remaining eight students were "various levels in the middle, […] you know, high-middle, middle-middle" (interview dated June 29, 2015).

Most Frequently Used Assessment Practices, as Reported on the ELATE Checklists

Item #	Assessment Practice	Total Times Reported as Used
8	Asked questions requiring more elaborated responses (e.g., a few sentences).	32
3	Asked questions of the class as a whole.	24
2	Asked questions of individual students by name.	19
9	Asked questions intended to stimulate a general discussion.	19
12	Used whole-class discussion to determine how well students understood a concept, idea, or strategy.	13

Why did this teacher choose to use a certain assessment practice (or practices) often?

Charlotte explained that asking questions of her ENG 135 students "helps me see how much they are thinking about the work, that — many times, you know, almost a guaranteed question […] in every unit that we do this, is to say, 'All right, how does this fit into the unit?' You know, and try to get them to take that step back and look at the bigger picture […] it also serves as a springboard for 'What could you possibly write your paper on, if these are the big issues coming up?'" (interview dated June 29, 2015).

Assessment Practices Perceived to be Most Effective for Lower-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	4.00
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	4.00
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	4.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Average-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	4.00
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	4.00
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	4.00
24	Gave oral feedback to multiple students at once about their writing (e.g., discussing strategies with groups or the entire class).	4.00
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	4.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Higher-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	4.00
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	4.00
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	4.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

How does this teacher determine the effectiveness of an assessment practice?

"Sometimes it was that the, um, either the — you know, the eye contact and seeing if they were taking down what I said. Um, many times — again, it's a lot of, of visual monitoring. If we did, like, the — when I know that the students didn't fully read the reading, and I'm doing the pass-around — again, thinking of those two low-middle students… I would see them finding the quote for the activity, but going back and marking it. So even though their page was completely empty, they found the quote, and they wrote it down, or noted or something, as a possibility for themselves, so, um, they didn't just do it just to get through the activity, they saw, 'Oh, I may be able to use this.' So I would see it, how often — or it's not so much how often that it happened, but I know if those two students were doing it, then it was working. You know, the students who already had it marked, well, it worked already anyway.

"Um, but usually it's how I can tell when I start the discussion that's going nowhere, you know, I might ask a few different questions. I might try, you know, pulling out a quote to have a class discussion. And if those things don't work — and that's how I usually go to Plan B, the pass-around" (interview dated June 29, 2015).

What is this teacher's definition of "assessment"?

"I define it as…very generally, something that an instructor does to determine how successfully the students are, I guess, embracing, using, or understanding material. And that 'something' is really broad, because I think, like I said, that visual assessment of walking around the room and seeing what they're doing versus also, you know, collecting something that they've written that I give them feedback on. So, it's just something that lets the instructor know how well something is being accomplished, achieved, understood" (interview dated June 29, 2015).

What does an example of an assessment practice look like in this teacher's ENG 135 class?
Charlotte described how she gave feedback to her ENG 135 students on their writing:

"I…how do I put it? I read a lot! [*laughs a little*] I do, because I feel like, I want a student to either — well, know that (a) I read it, and (b) um, you know, some — some people might think, 'Oh my gosh,' you know, 'you're writing too much!' But at same time, I feel like if I — if I don't write it down, then I feel like a student's not going to see either what I see, be it what's getting in my way, or what's — you know, I don't want to only give, um, things to correct.

"I, I do like to compliment. I — and they, they laugh. Like, 'Ooh, what is this? Ooh, what does this facial expression mean? How would you say it, like…,' you know. So, I do interact a lot when I get their writing. If it's something like, 'Give me a proposal of what you think your paper's going to be,' you know, and they outline based on what their sources have provided, you know, I will, you know, put some questions like, 'Did you think about using [the textbook section on the topic] for this?' So some of them are big hints, and some of them are, you know, 'Are you sure this is gonna — going to take you in, in the right direction?', you know, 'How is this going to relate to…?' So, um, I, I do, I interact like mad, be it on something like an outline, or, um, an annotated bibliography, or, uh, their papers" (interview dated June 29, 2015).

SPRUCE ENG 135 Teacher Snapshot: Spring 2015	
Teacher Name	Rachel
Total Number of Years Teaching	15
Total Number of Years Teaching in the SPRUCE ELA Program	3
School Locale	Rural: Fringe
Total Number of SPRUCE ENG 135 Students	16
Total Number of Students in All Courses	76
Weeks in which ELATE Checklists Were Completed	May 1-7, 8-14, 15-21, and 22-28, 2015
Date of Interview	June 25, 2015

What were the performance levels of this teacher's SPRUCE ENG 135 students?

In talking about her ENG 135 students, Rachel emphasized that any difference in their performance levels "wasn't about ability. They're all extremely bright students." Of her 16 ENG 135 students, she identified "three who were on the lower end and struggling, […] about six were average," and seven "on the higher end, like, the high achievers" (interview dated June 25, 2015).

Most Frequently Used Assessment Practices, as Reported on the ELATE Checklists

Item #	Assessment Practice	Total Times Reported as Used
2	Asked questions of individual students by name.	167
9	Asked questions intended to stimulate a general discussion.	155
7	Asked questions requiring brief responses (e.g., a word or phrase).	120
1	Asked students questions to determine how well they understood a concept, idea, or strategy.	74
23	Corrected students' spelling, grammar, and other mechanical errors when giving them feedback on their writing.	74

Why did this teacher choose to use a certain assessment practice (or practices) often?

Rachel said of using class discussions and asking questions of her students, "I find […] those types of questions helpful, especially since asking them of students, it seemed to, um, sort of… I don't know, sort of understood what was going on a little bit more. When they heard from their peers, it was a lot easier sometimes, for them than coming from the instructor. To hear, you know, them say, 'Oh, OK, well, I can understand that context.' Because it's like teen language [*laughs, says next few words while laughing*] — they speak their own language, right?" (interview dated June 25, 2015).

Assessment Practices Perceived to be Most Effective for Lower-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
9	Asked questions intended to stimulate a general discussion.	5.00
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	4.75
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	4.75
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	4.75

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Average-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
9	Asked questions intended to stimulate a general discussion.	5.00
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	5.00
15	Reviewed and gave feedback (oral or written) on students' portfolios of their writing.	5.00
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	5.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Higher-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
9	Asked questions intended to stimulate a general discussion.	5.00
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	5.00
15	Reviewed and gave feedback (oral or written) on students' portfolios of their writing.	5.00
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	5.00
20	Gave students *targeted* written feedback about their writing, such as comments on certain parts of their assignment.	5.00
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	5.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

How does this teacher determine the effectiveness of an assessment practice?

Rachel said that the effectiveness of an assessment practice "would especially surface in — in discussion, you know. Um, if we're […] analyzing a particular article and, you know, it's clear that the concept […] — if it wasn't [*chuckles*], I don't know — absorbed the way I wanted it to be absorbed, then, um — I'd hear a little, sort of — comments that were off-task, for lack of a better word. […]

"But like I said, […] the students in this group, they were comfortable with one another, so it wasn't unusual for them to say, um, 'I don't understand,' 'How do we,' you know — 'Can we go back to this?' and 'What does this mean?' Um, so, in essence, you know, in listening, I could hear things that were being said that didn't sound right, um, and the students themselves would say, 'I don't understand'" (interview dated June 25, 2015).

What is this teacher's definition of "assessment"?

"To me, assessment is a tool, uh, to gauge understanding. Um, it's also sort of a…a guide for the teacher to see maybe where there are some areas of weakness or areas of strength. […] I guess I see it as an opportunity, sort of, to also dialogue, with colleagues, too, because, um, when we're seeing things, sort of, you know, at the same time, or, you know — those are times when we can have conversations about, 'OK, so maybe at this level, we go back to 7th grade, or 8th grade, or 9th grade, or, you know — what can we do differently? What can we…' you know — it — it's just a moment to open up conversation. So, I would say those three things" (interview dated June 25, 2015).

What does an example of an assessment practice look like in this teacher's ENG 135 class?

Rachel described how individual conferences with her ENG 135 students about their writing usually occur:

"Typically that had to do with essays, when they were writing, and it was after, um, they looked at my comments on the draft, and then they would schedule an appointment and come to see me and talk about, you know, what they learned and what they thought, and sometimes they would scrap a whole complete idea and start all over, um, once they realized, like, 'I don't know if I'm exactly on the right track.' And sometimes they, you know — it was just a, a moment for clarifying, and, sort of, you know, sort of — I like to let them think about what it is they're trying to do — um, you know, little activities like, well, first I tried to pull them away from sort of recall, recall, recall, um, and I have them do things like — strange things, like, 'Highlight the sections where you're just doing a bunch of summary,' and when they can look at it, they can see, like, 'Oh, wow, I have, like, [*laughs a little*] 80 percent —' […] So that's why it helps — just to sit down and just sort of, you know, act like you do" (interview dated June 25, 2015).

SPRUCE ENG 135 Teacher Snapshot: Spring 2015	
Teacher Name	Krystal
Total Number of Years Teaching	13
Total Number of Years Teaching in the SPRUCE ELA Program	9
School Locale	Suburb: Large
Total Number of SPRUCE ENG 135 Students	32
Total Number of Students in All Courses	110
Weeks in which ELATE Checklists Were Completed	May 8-14, 15-21, 22-28, and June 5-11, 2015
Date of Interview	July 9, 2015

What were the performance levels of this teacher's SPRUCE ENG 135 students?

Krystal noted differences between her two sections that she was teaching that semester. In her first section, she reported having "three on the lower-performing end" and "only two or three […] in the middle […] but the remaining [six or seven] really were pretty high flyers." Krystal shared that in her other section, she had two students who were "not very comfortable speaking in class, but [did] an excellent job with their writing." She added that she had "only three or four lower-performing, but […] the majority in that class were more kind of right in the middle" (interview dated July 9, 2015).

Most Frequently Used Assessment Practices, as Reported on the ELATE Checklists

Item #	Assessment Practice	Total Times Reported as Used
21	Gave students *general or holistic* written feedback about their writing, such as comments at the end of an assignment about how they did overall.	20
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	17
17	Reviewed and gave feedback (oral or written) on students' *final* versions of a writing assignment.	16
19	Used a rubric or rubrics for *formal* feedback to students on their writing.	16

Why did this teacher choose to use a certain assessment practice (or practices) often?

Krystal said of giving students feedback on their response papers, "it depends on the purpose a little bit, but if my purpose is to check for understanding, or especially if it's in regards to a seminal piece — so for example, if they're reading Foucault's *Panopticism*, and writing to me about it, as I go — look at that paper, chances are, like, I'm underlining the things that real— [*sic*] like, where I can tell, 'Yeah, you've got it. This is really good.' You know, like, if I'm marking it on the paper, and then I'm interacting with something, that's one of their things that I want to call attention to" (interview dated July 9, 2015).

She continued, "I might even just write, like, 'Yes,' 'Good,' you know, something affirmative in the comments. I'll usually also then write questions, like 'Have you considered…' or 'You might want to think about…', just to kind of push them a little further, if I think that they could take, you know, where their — their thoughts a little bit further.

"Usually with those response papers, I'm not…I'm not really using it to give feedback on grammar or mechanics or anything — I take it into account, generally, when I'm grading it, and if there's something really egregious, I'll mark it. But, that — you know, that's not the purpose of those papers. It really is just to see their thought process. So, I try to engage it on that level" (interview dated July 9, 2015).

Assessment Practices Perceived to be Most Effective for Lower-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
2	Asked questions of individual students by name.	5.00
6	Asked questions of reticent students to help improve their participation.	5.00
13	Reviewed and gave feedback (oral or written) on students' think/response papers.	4.67
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	4.67

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Average-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
2	Asked questions of individual students by name.	5.00
6	Asked questions of reticent students to help improve their participation.	5.00
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	5.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Higher-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
2	Asked questions of individual students by name.	5.00
6	Asked questions of reticent students to help improve their participation.	5.00
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	5.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

How does this teacher determine the effectiveness of an assessment practice?

"I'm looking for two things. I want to see if students are able to demonstrate their understanding of content, so if whatever it is I'm asking them to do, I can then look at if, one, they get it, which is helpful — like, understand what I'm asking of them, um, which means most of them all do it correctly or from what I asked, so my directions are clear. Which is not — not something I usually have an issue with in [SPRUCE], I, I've worked — I've been doing this a while. I've worked it out, [*says while laughing a little*] for the most part.

"Um, so, I want to be able to see their understanding the content, of the, whatever the thing is that they're writing about, and then…[two,] is it something that can serve as scaffolding? For me, that's really what's effective. Like, they don't necessarily have to understand why it's scaffolding, at the time. I'd like them to understand it eventually. I don't always let them in on my method to my madness, not officially. But, you know, yeah, does it serve a, a purpose, ultimately, in getting us somewhere?" (interview dated July 9, 2015).

What is this teacher's definition of "assessment"?

"Assessment would just be any tool that helps me gauge student understanding" (interview dated July 9, 2015).

What does an example of an assessment practice look like in this teacher's ENG 135 class?
Krystal described how peer feedback and individual conferencing with her ENG 135 students about their writing usually took place as they progressed through a unit of the course:

"I try to set up some time in class, when they're doing peer conferencing, then any sort of, um…we usually do — I'd say, like, within a typical unit, the first two, maybe three weeks are spent reading the seminal texts, reading the theoretical texts, doing some work with them, and then pulling in whatever […] other cultural pieces […] I want to pull in, given that particular unit. We always do at least one round robin discussion of their ideas for their paper, where they have to come with, like — it doesn't have to be a fully developed claim, but a, 'Here's my topic; here's what I'm thinking about; here's where I think I want to go; here's what I'm not sure about.' And, so that way, they get feedback from me, but also from other people in the class, so that's a guarantee.

"Usually after that, they start drafting, depending on how much time we have. Sometimes we'll come back, and do another round robin, when they've had more time to really solidify, like, 'This is my claim; I think this is where I'm going to go.' And then we always have at least one day for peer editing, with — sometimes not a fully completed draft, but a pretty close to completed draft, so that they have time to give each other feedback. During that block, they can conference with me. Like, they know they can pull out of any peer editing pair to come talk to me, and usually they do. I mean, usually — I'd say, I usually see [*inhales while thinking*] at least four or five kids. Some of them just want to talk through their ideas, and just kind of want confirmation, like, 'This is what I'm thinking, and then this, this, and this. Does that make sense?' And I can offer them a little bit of feedback, or — and some of them just really need assurance, like, 'Yes, that makes sense,' and then they're fine. Some of them, you know, will write a sample paragraph, have me read it, see if it makes sense, give them some ideas, so, it depends" (interview dated July 9, 2015).

SPRUCE ENG 135 Teacher Snapshot: Spring 2015	
Teacher Name	Tony
Total Number of Years Teaching	10
Total Number of Years Teaching in the SPRUCE ELA Program	7
School Locale	Suburb: Large
Total Number of SPRUCE ENG 135 Students	14
Total Number of Students in All Courses	105
Weeks in which ELATE Checklists Were Completed	May 22-28, May 29-June 4, and June 5-11, 2015
Date of Interview	[*interview not conducted*]

What were the performance levels of this teacher's SPRUCE ENG 135 students?

Tony wrote via email that he "may only have two or three students in my class I would label as high-performing," and he attributed this small number to students' lack of work ethic. He went on to state that "many of [the students] have fallen into the middle area," and that he put more effort to engage students that semester than in his previous years as a SPRUCE ELA Program Instructor (email dated May 29, 2015).

Most Frequently Used Assessment Practices, as Reported on the ELATE Checklists

Item #	Assessment Practice	Total Times Reported as Used
8	Asked questions requiring more elaborated responses (e.g., a few sentences).	115
3	Asked questions of the class as a whole.	95
7	Asked questions requiring brief responses (e.g., a word or phrase).	85
2	Asked questions of individual students by name.	65
5	Asked questions of students I thought would be more likely to respond well.	60
9	Asked questions intended to stimulate a general discussion.	60

Why did this teacher choose to use a certain assessment practice (or practices) often?
[*unknown—interview not conducted*]

Assessment Practices Perceived to be Most Effective for Lower-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	4.00
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	4.00
2	Asked questions of individual students by name.	3.67
9	Asked questions intended to stimulate a general discussion.	3.67

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Average-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
5	Asked questions of students I thought would be more likely to respond well.	4.00
9	Asked questions intended to stimulate a general discussion.	4.00
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	4.00
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	4.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

Assessment Practices Perceived to be Most Effective for Higher-performing Students, as Reported on the ELATE Checklists

Item #	Assessment Practice	Mean Rating
5	Asked questions of students I thought would be more likely to respond well.	5.00
9	Asked questions intended to stimulate a general discussion.	4.33
3	Asked questions of the class as a whole.	4.00
10	Used paired or small group (2-4 students) discussion to determine how well students understood a concept, idea, or strategy.	4.00
22	Wrote questions when giving feedback to students on their writing to help prompt their thinking and develop their writing skills.	4.00
25	Conferenced with individual students about their writing to give them feedback and to review and discuss their work.	4.00

Rating scale for perceptions of effectiveness of each practice to improve student learning: 0 = *extremely ineffective*; 1 = *very ineffective*; 2 = *moderately ineffective*; 3 = *moderately effective*; 4 = *very effective*; 5 = *extremely effective*.

How does this teacher determine the effectiveness of an assessment practice?
[*unknown—interview not conducted*]

What is this teacher's definition of "assessment"?
[*unknown—interview not conducted*]

What does an example of an assessment practice look like in this teacher's ENG 135 class?
[*unknown—interview not conducted*]

References

American Federation of Teachers, National Council on Measurement in Education, & National Education Association (1990). *Standards for teacher competence in educational assessment of students*. Washington, DC: National Council on Measurement in Education. Retrieved from http://files.eric.ed.gov/fulltext/ED323186.pdf

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*(5),13-18.

Andrade, H. L. (2010). Students as the definitive source of formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 90-105). New York, NY: Routledge Taylor & Francis Group.

Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal, 40*(3), 685-730.

Aschbacher, P., & Alonzo, A. (2006). Examining the utility of elementary science notebooks for formative assessment purposes. *Educational Assessment, 11*(3&4), 179-203.

Ash, D., & Levitt, K. (2003). Working within the zone of proximal development: Formative assessment as professional development. *Journal of Science Teacher Education, 14*(1), 23-48.

Ayala, C. C., & Brandon, P. R. (2008). Building evaluation recommendations for improvement: Insights from student formative assessments. In N. L. Smith & P. R. Brandon (Eds.), *Fundamental issues in evaluation* (pp. 159-176). New York, NY: The Guilford Press.

Beavers, A. (2009). Teachers as learners: Implications of adult education for professional development. *Journal of College Teaching & Learning (TLC)*, *6*(7), 25-30.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, *85*(5), 536-553.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5-25.

Bentler, P. M. and Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16*(1), 78-117.

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research, 35*(5), 463-482.

Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science Technology Society, 24*(3), 200-212.

Birenbaum, M., Kimron, H., & Shilton, H. (2011). Nested contexts that shape assessment *for* learning: School-based professional learning community and classroom culture. *Studies in Educational Evaluation, 37*(1), 35-48.

Birenbaum, M., Kimron, H., Shilton, H., & Shahaf-Barzilay, R. (2009). Cycles of inquiry: Formative assessment in service of learning in classrooms and in school-based professional communities. *Studies in Educational Evaluation, 35*(4), 130-149.

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). Los Angeles, CA: SAGE.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Berkshire, England: Open University Press.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Black, P., & Wiliam, D. (2005). Lessons from around the world: How policies, politics and cultures constrain and afford assessment practices. *Curriculum Journal, 16*(2), 249-261.

Blanchard, J. (2008). Learning awareness: Constructing formative assessment in the classroom, in the school and across schools. *The Curriculum Journal, 19*(3), 137-150.

Bonner, S. M. (2013). Validity in classroom assessment: Purposes, properties, and principles. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 87-106). Los Angeles, CA: SAGE.

Borko, H., Mayfield, V., Marion, S., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. *Teaching & Teacher Education, 13*(3), 259-278.

Brevik, L. M. (2014). Making implicit practice explicit: How do upper secondary teachers describe their reading comprehension strategies instruction? *International Journal of Educational Research, 67*, 52-66.

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L., & Yin, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice, 31*(4), 13-17.

Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education, 10*(2), 161-180.

Brookhart, S. M. (2003). Developing measurement theory for classroom assessment purposes and uses. *Educational Measurement: Issues and Practice, 22*(4), 5-12.

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice, 30*(1), 3-12.

Brookhart, S. M., Moss, C. M., & Long, B. A. (2009). Promoting student ownership of learning through high-impact formative assessment practices. *Journal of MultiDisciplinary Evaluation, 6*(12), 52-67.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research: Methodology in the social sciences*. New York, NY: The Guilford Press.

Bruno, I., & Santos, L. (2010). Written comments as a form of feedback. *Studies in Educational Evaluation, 36*(3), 111-120.

Christie, C. A. (2012). Advancing empirical scholarship to further develop evaluation theory and practice. *The Canadian Journal of Program Evaluation, 26*(1), 1-18.

Clemente, M., & Ramírez, E. (2008). How teachers express their knowledge through narrative. *Teaching and Teacher Education, 24*(5), 1244-1258.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Pearson.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Crossouard, B. (2011). Using formative assessment to support complex learning in conditions of social adversity. *Assessment in Education: Principles, Policy & Practice, 18*(1), 59-72.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2014). *The Framework for Teaching evaluation instrument: 2013 edition*. Princeton, NJ: The Danielson Group, LLC. Retrieved from http://danielsongroup.org/download/?download=448

Danielson, C., Axtell, D., Bevan, P., Cleland, B., McKay, C., Phillips, E., & Wright, K. (2009). *Implementing the Framework for Teaching in enhancing professional practice: An ASCD action tool*. Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L. (1988). The futures of teaching. *Educational Leadership, 46*(3), 4-10.

Day, C. (2002). School reform and transitions in teacher professionalism and identity. *International Journal of Educational Research, 37*, 677-692.

Day, C., & Gu, Q. (2007). Variations in the conditions for teachers' professional learning and development: Sustaining commitment and effectiveness over a career. *Oxford Review of Education*, *33*(4), 423-443.

Dixon, H., & Williams, R. (2003). Formative assessment and the professional development of teachers: Are we focusing on what is important? *Set: Research Information for Teachers, 2*, 35-39.

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation, 14*(7). Available online: http://pareonline.net/getvn.asp?v=14&n=7

Dunn, M. W. (2011). Writing-skills instruction: Teachers' perspectives about effective practices. *Journal of Reading Education, 37*(1), 18-25.

Executive Office of the President. (2015, December). *Every Student Succeeds Act: A progress report on elementary and secondary education*. Retrieved from https://www.whitehouse.gov/sites/whitehouse.gov/files/documents/ESSA_Progress_Report.pdf

Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research, 38*(1), 47-65.

Filsecker, M., & Kerres, M. (2012). Repositioning formative assessment from an educational assessment perspective: A response to Dunn & Mulvenon (2009). *Practical Assessment, Research & Evaluation, 17*(16). Available online: http://pareonline.net/getvn.asp?v=17&n=16

Fisher, D., Lapp, D., Flood, J., & Moore, K. (2006). Linking literacy teaching with assessment: A continuing professional development initiative for secondary schools. *Literacy, 40*(2), 115-122.

Fitzgerald, J. (1987). Research on revision in writing. *Review of Educational Research, 57*(4), 481-506.

FitzPatrick, D. (2008). Constructing complexity: Using reading levels to differentiate reading comprehension activities. *English Journal, 98*(2), 57-63.

Fontana, D., & Fernandes, M. (1994). Improvements in mathematics performance as a

    consequence of self-assessment in Portuguese primary school pupils. *British Journal of*

    Educational Psychology, 64, 407-417.

Frey, N., & Fisher, D. (2013). A formative assessment system for writing improvement. *English*

    Journal, 103(1), 66-71.

Gearhart, M., & Saxe, G. B. (2004). When teachers know what students know: Integrating

    mathematics assessment. *Theory into Practice, 43*(4), 304-313.

Gorlewski, D. A. (2010). Overflowing but underused: Portfolios as a means of program

    evaluation and student self-assessment. *English Journal, 99*(4), 97-101.

Gorlewski, J. (2008). Formative assessment: Can you handle the truth? *English Journal, 98*(2),

    94-97.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Griffin, P. (2009). Teachers' use of assessment data. In C. Wyatt-Smith & J. J. Cumming (Eds.),

    Educational assessment in the 21st century: Connecting theory and practice (pp. 183-

    208). London, England: Springer.

Griffiths, V. (2000). The reflective dimension in teacher education. *International Journal of*

    Educational Research, 33(5), 539-555.

Grossman, P. L., & Stodolsky, S. S. (1995). Content as context: The role of school subjects in

    secondary school teaching. *Educational Researcher, 24*(8), 5-23.

Guskey, T. R. (1987). Context variables that affect measures of teacher efficacy. *Journal of*

    Educational Research, 81(1), 41-47.

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of

    instructional innovation. *Teaching & Teacher Education, 4*(1), 63-69.

Guskey, T. R. (2003). How classroom assessments improve learning. *Educational Leadership, 60*(5), 6-11.

Halcomb, E. J., & Davidson, P. M. (2006). Is verbatim transcription of interview data always necessary? *Applied Nursing Research, 19*(1), 38-42.

Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review, 17*, 50-62.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Thousand Oaks, CA: SAGE Publications.

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from http://www.edweek.org/media/formative_assessment_next_generation_heritage.pdf

Heritage, M. (2013). Gathering evidence of student understanding. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 179-195). Los Angeles, CA: SAGE.

Herman, J. L., Osmundson, E., & Silver, D. (2010). *Capturing quality in formative assessment practice: Measurement challenges* (CRESST Report 770). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Hill, J., Hawk, K., & Taylor, K. (2002). Professional development: What makes it work? *Proceedings of the NZARE Conference, Christchurch, New Zealand, December 2001*. Retrieved from http://www.educationgroup.co.nz/uploads/Publications/Professional%20Development%20-%20what%20makes%20it%20work.pdf

Hodgen, J., & Marshall, B. (2005). Assessment for learning in English and mathematics: A comparison. *The Curriculum Journal, 16*(2), 153-176.

Hollingworth, L. (2012). Why leadership matters: Empowering teachers to implement formative assessment. *Journal of Educational Administration, 50*(3), 365-379.

Howell, D. C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury.

Hoyle, E. (1982). The professionalization of teachers: A paradox. *British Journal of Educational Studies, 30*(2), 161-171.

Hunter, D., Mayenga, C., & Gambell, T. (2006). Classroom assessment tools and uses: Canadian English teachers' practices for writing. *Assessing Writing, 11*(1), 42-65.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods, 18*(3), 3-20.

Kelleher, J. (2003). A model for assessment-driven professional development. *Phi Delta Kappan*, *84*(10), 751-756.

Kennedy, S. Y., & Smith, J. B. (2013). The relationship between school collective reflective practice and teacher psychological efficacy sources. *Teaching and Teacher Education, 29*, 132-143.

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice, 30*(4), 28-37.

Koziol, S. M., Jr., & Burns, P. (1986). Teachers' accuracy in self-reporting about instructional practices using a focused self-report inventory. *The Journal of Educational Research, 79*(4), 205-209.

Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education, 36*, 143-152.

Labaree, D. F. (1992). Power, knowledge, and the rationalization of teaching: A genealogy of the movement to professionalize teaching. *Harvard Educational Review, 62*(2), 123-155.

Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teaching and Teacher Education, 21*(8), 899-916.

Lawrence, S. A., Rabinowitz, R., & Perna, H. (2009). Reading instruction in secondary English Language Arts classrooms. *Literacy Research and Instruction, 48*(1), 39-64.

Lee, C., & Wiliam, D. (2005). Studying changes in the practice of two teachers developing assessment for learning. *Teacher Development, 9*(2), 265-283.

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2011). *IBM SPSS for intermediate statistics: Use and interpretation* (4th ed.). New York, NY: Routledge.

Lendrum, A., & Humphrey, N. (2012). The importance of studying the implementation of interventions in school settings. *Oxford Review of Education, 38*(5), 635-652.

Leung, C. (2004). Developing formative teacher assessment: Knowledge, practice, and change. *Language Assessment Quarterly, 1*(1), 19-41.

Lieberman, A. (1995). Practices that support teacher development: Transforming conceptions of professional learning. *Phi Delta Kappan*, *76*(8), 591-596.

Lieberman, A., & McLaughlin, M. W. (1992). Networks for educational change: Powerful and problematic. *Phi Delta Kappan, 73*(9), 673–677.

MacCallum, R. C., Widaman, K. F., Zhang, S. B., & Hong, S. H. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84-99.

Marshall, B. (2007). Formative classroom assessment in English, the humanities, and social sciences. In J. H. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 136-152). New York, NY: Teachers College Press.

McGatha, M. B., Bush, W. S., & Rakes, C. (2009). The effects of professional development in formative assessment on mathematics teaching performance and student achievement. *Journal of MultiDisciplinary Evaluation, 6*(12), 32-43.

McManus, S. (2008). *Attributes of effective formative assessment*. Washington, D.C.: Council of Chief State School Officers. Retrieved from http://www.dpi.state.nc.us/docs/accountability/educators/fastattributes04081.pdf

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice, 20*(1), 20-32.

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. (2010). The practical implications of educational aims and contexts for formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 41-58). New York, NY: Routledge Taylor & Francis Group.

McMillan, J. H. (2013). Why we need research on classroom assessment. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 3-16). Los Angeles, CA: SAGE.

McMillan, J. H., Venable, J. C., & Varier, D. (2013). Studies of the effect of formative assessment on student achievement: So much more is needed. *Practical Assessment, Research & Evaluation, 18*(2). Available online: http://pareonline.net/pdf/v18n2.pdf

Morgan, G. A., Leech, N. L., Gloeckner, G. W., & Barrett, K. C. (2013). *IBM SPSS for introductory statistics: Use and interpretation* (5th ed.). New York, NY: Routledge.

Moss, B., & Bordelon, S. (2007). Preparing students for college-level reading and writing: Implementing a rhetoric and writing class in the senior year. *Reading Research and Instruction, 46*(3), 197-221.

Murphy, S. M. (1997). Who should taste the soup and when? Designing portfolio assessment programs to enhance learning. *The Clearing House: A Journal of Educational Strategies, Issues, and Ideas, 71*(2), 81-84.

Murphy, S. M., & Smith, M. A. (2013). Assessment challenges in the Common Core era. *English Journal, 103*(1), 104-110.

National Center for Education Statistics. (2015, May). Concentration of public school students eligible for free or reduced-price lunch. In *The Condition of Education*. Institute of Education Sciences, U. S. Department of Education. Retrieved from http://nces.ed.gov/programs/coe/indicator_clb.asp

National Center for Education Statistics. (2016). Common Core of Data (CCD). Institute of Education Sciences, U. S. Department of Education. Retrieved from http://nces.ed.gov/ccd/

Neira, M. (2015). What is TED? In *TED: Teacher evaluation and development*. Latham, NY: New York State United Teachers. Retrieved from http://www.nysut.org/resources/special-resources-sites/ted/what-is-ted

Nelson, M., & Eddy, R. M. (2008). Evaluative thinking and action in the classroom. In T. Berry & R. M. Eddy (Eds.), *Consequences of No Child Left Behind for educational evaluation. New Directions for Evaluation, 117*, 37-46.

New York State Education Department. (2015, February 11). *Guidance on New York State's annual professional performance review for teachers and principals to implement Education Law §3012-c and the Commissioner's regulations*. Albany, NY: EngageNY, NYSED. Retrieved from https://www.engageny.org/resource/guidance-on-new-york-s-annual-professional-performance-review-law-and-regulations

Noskin, D. P. (2013). Toward a clearer picture of assessment: One teacher's formative approach. *English Journal, 103*(1), 72-80.

Olson, J., Olson, M., Slovin, H., Gilbert, M., & Gilbert, B. (2010, January). *The design and delivery of a professional development program to implement formative assessment in a networked classroom*. Paper presented at the 8th Annual Hawaii International Conference on Education, Honolulu, Hawaii.

O'Neill, G. P. (1988). Teaching effectiveness: A review of the research. *Canadian Journal of Education, 13*(1), 162-185.

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376-407.

Pajares, F., Johnson, M. J., & Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Research in the Teaching of English, 42*(1), 104-120.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129-144.

Pappageorge, T. (2013). Checking in: Informal communication to assess learning in the English language arts classroom. *English Journal, 103*(1), 54-59.

Parr, J. M., & Timperley, H. S. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing Writing, 15*(2), 68-85.

Phelan, J., Choi, K., Vendlinski, T., Baker, E., & Herman, J. (2011). Differential improvement in student understanding of mathematical principles following formative assessment intervention. *The Journal of Educational Research, 104*(5), 330-339.

Popham, W. J. (2008). *Transformative assessment*. Portland, OR: Association of Supervision and Curriculum Development.

Popham, W. J. (2013). Can classroom assessments of student growth be credibly used to evaluate teachers? *English Journal, 103*(1), 34-39.

Randel, B., & Clark, T. (2013). Measuring classroom assessment practices. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 145-163). Los Angeles, CA: SAGE.

Rich, C. S., Harrington, H., & Kim, J. (2008). Automated essay scoring in state formative and summative writing assessment. Paper presented at the meeting of the National Council on Measurement in Education, Chicago, IL. Retrieved from http://www.ctb.com/ctb.com/control/openFileShowAction?mediaId=905.0

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review, 6*(2), 135-147.

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing data* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation, 37*(1), 15-24.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144.

Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board Certification. *American Educational Research Journal, 45*(3), 669-700.

Schneider, M. C., & Meyer, J. P. (2012). Investigating the efficacy of a professional development program in formative classroom assessment in middle school English language arts and mathematics. *Journal of MultiDisciplinary Evaluation, 8*(17), 1-24.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds)., *Perspectives of curriculum evaluation* (pp. 39-83). Chicago, IL: Rand McNally & Company.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Shermis, M. D., & Di Vesta, F. J. (2011). Formative assessment: Using assessment for improving instruction. In *Classroom assessment in action* (pp. 83-118). Lanham, MD: Rowman & Littlefield Publishers, Inc.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-23.

Siskin, L. S. (1991). Departments as different worlds: Subject subcultures in secondary schools. *Educational Administration Quarterly, 27*(2), 134-160.

Smith, K. (2011). Professional development of teachers—A prerequisite for AfL to be successfully implemented in the classroom. *Studies in Educational Evaluation, 37*(1), 55-61.

Sondergeld, T. A., Bell, C. A., & Leusner, D. M. (2010). Understanding how teachers engage in formative assessment. *Teaching & Learning, 24*(2), 72-86.

Spandel, V. (2006). Speaking my mind: In defense of rubrics. *English Journal, 96*(1), 19-22.

Sperling, M. (1990). I want to talk to each of you: Collaboration and the teacher-student writing conference. *Research in the Teaching of English, 24*(3), 279-321.

Sprinthall, R. C. (2007). *Basic statistical analysis* (8th ed.). Boston, MA: Allyn and Bacon.

Stake, R. E. (2010). *Qualitative research: Studying how things work*. New York, NY: The Guilford Press.

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, 18*(1), 23-27.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan, 83*(10), 758-765.

Stodolsky, S. S., & Grossman, P. L. (1995). The impact of subject matter on curricular activity: An analysis of five academic subjects. *American Educational Research Journal, 32*(2), 227-249.

Tessmer, M., & Richey, R. C. (1997). The role of context in learning and instructional design. *Educational Technology Research and Development, 45*(2), 85-115.

Tolley, L. M. (2011, November). *Everything matters: Understanding the impact of context on formative assessment*. Paper presented at the American Evaluation Association (AEA) Annual Conference, Anaheim, California.

Tolley, L. M. (2012, October). *But does it work? A review of teacher professional development on formative assessment*. Poster presented at the American Evaluation Association (AEA) Annual Conference, Minneapolis, Minnesota.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805.

U. S. Department of Education (2011a). *No Child Left Behind legislation and policies*. Retrieved from http://www2.ed.gov/policy/elsec/guid/states/index.html

U. S. Department of Education (2011b). *Race to the Top Assessment Program*. Retrieved from http://www2.ed.gov/programs/racetothetop-assessment/index.html

U. S. Department of Education (2016). *Every Student Succeeds Act (ESSA)*. Retrieved from http://www.ed.gov/essa

Vähäsantanen, K. (2015). Professional agency in the stream of change: Understanding educational change and teachers' professional identities. *Teaching and Teacher Education, 47*, 1-12.

VanDeWeghe, R. (2006). Learning from the arts. *English Journal, 95*(6), 71-75.

Verloop, N., Van Driel, J., & Meijer, P. (2001). Teacher knowledge and the knowledge base of teaching. *International Journal of Educational Research, 35*(5), 441-461.

Walker, C. P., & Elias, D. (1987). Writing conference talk: Factors associated with high- and low-rated writing conferences. *Research in the Teaching of English, 21*(3), 266-285.

Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research*, *79*(2), 702-739.

White, R. (1985). The importance of context in educational research. *Research in Science Education, 15*(1), 92-102.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3-14.

Wiliam, D., Lee, C., Harrison, C., & Black, P. J. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education, 11*(1), 49-65.

Yin, Y., Olson, J., Olson, M., Slovin, H., & Brandon, P. R. (2015). Comparing two versions of professional development for teachers using formative assessment in networked mathematics classrooms. *Journal of Research on Technology in Education, 47*(1), 41-70.

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., Tomita, M. K., & Young, D. B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education, 21*(4), 335-359.

Zirkel, S., Garcia, J. A., & Murphy, M. C. (2015). Experience-sampling research methods and their potential for education research. *Educational Researcher, 44*(1), 7-16.

Vita

NAME OF AUTHOR: Leigh M. Tolley

PLACE OF BIRTH: Pennsylvania, USA

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:
Syracuse University, Syracuse, New York
University of Pennsylvania, Philadelphia, Pennsylvania

DEGREES AWARDED:
2012   M.S. in Instructional Design, Development and Evaluation (Specialty Area: Evaluation), Syracuse University
2002   M.S.Ed. in Secondary Education, English, University of Pennsylvania
2000   B.A. in English (Concentration: Celtic Studies; Minor: Anthropology), University of Pennsylvania

AWARDS AND HONORS:
2014   Instructional Design, Development and Evaluation (IDD&E) Award for exceptional teaching and service to the IDE 201/301/401 teaching team from 2010-2014, School of Education, Syracuse University
2013   IDD&E Outstanding Leadership Award for work as a senior teaching assistant, School of Education, Syracuse University
2012   IDD&E Evaluation and Research Award, School of Education, Syracuse University
2012   Outstanding Teaching Assistant Award, The Graduate School, Syracuse University
       Certificate in University Teaching, The Graduate School, Syracuse University
2010   Phi Kappa Phi Honor Society inductee
       Co-recipient of the CNY BEST Learning and Performance Consultant Award
2007   Recognized in *Madison Who's Who*, Spring 2006-2007 edition
2006   Recognized by the National Honor Roll as an Outstanding American Teacher in 2006
2005   Recognized in *Who's Who Among America's Teachers*, 9th Edition
2004   Recognized in *Who's Who Among America's Teachers*, 8th Edition
1995   Girl Scout Gold Award

PROFESSIONAL EXPERIENCE:
Assessment Specialist, Office of the Associate Provost for Academic Programs, Syracuse University, 2015-present
Research Assistant, Hezel Associates, LLC, Summers 2009-2012 and March 2014-March 2015
Teaching Mentor, Teaching Assistant Program, The Graduate School, Syracuse University, August 2011, 2012, 2013 and January 2012, 2013, 2016
Graduate Teaching Assistant/Instructor of Record, IDD&E Department, School of Education, Syracuse University, Spring 2010-Spring 2014
Online Teacher, Educate Online, January-July 2009
Adjunct Instructor, English Department, Le Moyne College, Fall 2008-Spring 2009
English Teacher, Penn Wood High School, William Penn School District, 2002-2007
Clinical Research Coordinator, Center for Hereditary Retinal Degenerations, Scheie Eye Institute, 1995-2001 and August 2007-July 2008