University of New Mexico
UNM Digital Repository

Special Education ETDs   Education ETDs

2-14-2014

Preparing Teachers Who Can Effectively Assess Students with Disabilities

Karen M. Potter

Follow this and additional works at: https://digitalrepository.unm.edu/educ_spcd_etds

Karen M. Potter
_____
Candidate

Educational Specialties
_____
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Elizabeth Keefe, Chairperson
_____

Dr. Susan Copeland
_____

Dr. Tryphenia Peele-Eady
_____

Dr. Cathy Qi

_____

**PREPARING TEACHERS WHO CAN EFFECTIVELY ASSESS
STUDENTS WITH DISABILITIES**

by

KAREN M. POTTER

B.A. English, New Mexico State University, 1998
B.S. Secondary Education, New Mexico State University, 1998
M.S. Special Education, 2006

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Special Education

The University of New Mexico
Albuquerque, New Mexico

December, 2013

DEDICATION

This manuscript is dedicated to Joel Shirley, in grateful acknowledgement of good advice given in 1992.

ACKNOWLEDGEMENTS

I would like to acknowledge the invaluable support of the members of the dissertation committee, both across preparatory coursework and throughout the dissertation process.

To Liz Keefe, thank you for your continued guidance in the field of special education, beginning with the public school system and culminating with this manuscript. I am forever grateful for your professional knowledge and personal support throughout this process.

To Susan Copeland, thank you for your continual example of professional grace and for sharing from your vast knowledge of reading and behavioral support methods. It is always such a pleasure to work with you.

To Tryphenia Peele-Eady, thank you for being willing to apply your qualitative research knowledge to the special education field and for your dedication to the research process.

To Cathy Qi, thank you for the pleasant way in which you presented your careful considerations and thoughtful questions throughout this process.

I would also like to acknowledge Laurel Semsch for her valuable support as a peer going through this process.

Finally, I would like to acknowledge the personal support of Katherine Roberson and George and Isaac Potter. Their friendship and comforting presence was instrumental in the final completion of this manuscript.

Preparing Teachers Who Can Effectively Assess Students with Disabilities.

By

Karen M. Potter

B.A. English, New Mexico State University, 1998
B.S. Secondary Education, New Mexico State University, 1998
M.S. Special Education, University of New Mexico, 2006
Ph.D. Special Education, University of New Mexico, 2013

ABSTRACT

This qualitative interview study examined the classroom assessment knowledge and beliefs of five recent graduates of the University of New Mexico Special Education Dual License Program (SEDLP). Research questions were designed to gain understanding in three areas. First, in what ways did recent SEDLP graduates characterize their level of competence (theoretical understanding and practical application) in assessing the progress of students with disabilities in the classroom? Second, in what ways do SEDLP graduates report that they use classroom assessment to inform classroom instruction? Third, what features of the SEDLP teacher preparation program do graduates identify as having positively or negatively impacted their ability to effectively use classroom assessments? To answer these questions, each participant was interviewed twice using a semi-structured question format and constant-comparative methodology. The results showed participant knowledge and specific practices in addressing student affective needs, broadening the application of assessments by individualizing and differentiating, meeting district requirements for assessments, using measurable assessment criteria, and using frequent informal assessment. Classroom instruction was most impacted as participants determined next steps to address

knowledge gaps or intervene behaviorally. The SEDLP was characterized as positively

impacting classroom assessment in the areas of providing assessment models, multiple

examples of types, multiple informal assessment practice opportunities, and

organizational and resource availability. It was characterized as lacking instruction in the

areas of writing formal assessments and providing opportunities to learn and practice

assessments in math. These results lead to implications for future practice and research

that are discussed.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

Introduction

Public education is a worldwide concern in which the United States participates. The Universal Declaration of Human Rights (United Nations, 1948) describes this value as one in which the world community, including the U.S., is obligated:

Everyone has the right to an education. Education shall be free, at least at the elementary and fundamental stages. Elementary education shall be compulsory. Technical and professional education shall be made generally available and higher education shall be equally accessible to all on the basis of merit. (Art. 26)

Access to education has also traditionally been interpreted as a property right within the United States and applied to all citizens by the Fourteenth Amendment to the U.S. Constitution: Nor shall any state deprive any person of life, liberty, or property, without due process of law. As the United States has pursued this value, beginning with the Constitution and continuing past the Universal Declaration of Human Rights, many approaches to respecting the right to an education by improving the public education system have been attempted, abandoned, and sometimes revived.

One recurring approach to protecting education as a property right has been to publish an attention-gathering report delineating the failures of the public education system, predicting increasing negative results unless changes are made, and suggesting possibilities for improvement. A strong example of this approach is the 1983 report by The National Commission on Excellence in Education, titled *A Nation at Risk: The Imperative for Education Reform*. This report focused on the "rising tide of mediocrity" (para. 1) afflicting American schools and on the effect of low performance on the United

States' economic future. Another example of this type is *America 2000: An Education Strategy* (U.S. Department of Education, 1991). This report, created in response to *A Nation at Risk*, lists six goals to be accomplished by the year 2000 and emphasizes the need for data collection as a method of maintaining system accountability. In 2001, The No Child Left Behind Act (NCLB) was passed. This act adds action to the alarms and suggestions of the prior reports by setting forth a system of accountability in an effort to preserve the educational property rights of all students.

The purpose of NCLB is to ensure that *all* students make "adequate yearly progress" (2001). This is an important difference from prior approaches because it acknowledges the discrepancies in success rates on standardized tests among particular groups (including students with disabilities) when compared with the overall tested student population. Such an acknowledgement holds schools accountable for the success of all students and requires that research based interventions be used to increase the success of *any* delineated group that is not making adequate progress. NCLB interprets "research based" to mean those practices that can be quantitatively shown to create an increase in measures of students' grade-level skills proficiency on standardized tests.

The next and most recent large reform program is the Obama administration's Race to the Top (RTTT) grant program. As explained by McGuinn (2012), "RTTT may best be understood as an attempt by the Obama administration to respond to the failures of NCLB and adapt federal policy to a role more commensurate with limited federal power and administrative resources" (p.153). In a 2009 article in *The Washington Post*, Secretary of Education Arne Duncan announced the availability of program guidelines and described four program goals: to "reverse the pervasive dumbing-down of academic

standards" by working toward "common, internationally benchmarked K-12 standards" (para. 6), to provide appropriate data about "student achievement and identify effective instructional practices" (para. 7), to increase "quality of teachers and principals, especially in high-poverty schools and hard-to-staff subjects" (para. 8), and "to turn around the lowest-performing schools, states, and districts" (para. 9) by a variety of reform plans. RTTT is a new approach in that it awards funding to states on a competitive basis according to what state plans are judged to be most likely to address the goals described above.

Within the guidelines explained by the NCLB (2001) and in the current context of RTTT, students who receive special education services are expected to demonstrate continuing growth in skill proficiency as measured by standardized tests. Assessment also continues to be a major component in the field of special education as a whole; special education is driven by formal assessment, including initial assessment by certified diagnosticians and implementation of norm-referenced and criterion referenced assessments by special education teachers who follow the assessment protocols unique to each test. The skills of individuals are assessed to determine initial access to special education services, to develop learning goals, to determine progress, and to end services. The Individuals with Disabilities Education Act (IDEA, 2004) outlines the manner in which these assessments will be conducted, always with the goal of preserving individuals' property rights and the broader human right to an education (Luckasson, 2006).

Both the NCLB (2001) and IDEA (2004) mandate effective high stakes and procedural assessment. At the classroom level, teachers fulfill these long-term mandates

by using informal, classroom based assessments to measure progress toward larger goals and to adjust instruction to cause greater learning. This requires that teachers be skilled daily assessors, and according to NCLB (2001) and by extension IDEA (2004), use evidence based practices to more efficiently assess and instruct. Teacher education programs are tasked with providing enough graduates to fulfill the need for teachers who can assess proficiently, in a manner that provides an accurate representation of student learning growth and needs, and particularly struggle with this in the area of special education. Newly graduated teachers are expected to be able to fully participate in assessment activities and are held accountable for their proficiency even though they may not be prepared to enter the teaching profession as skilled assessors.

The use of evidence based practices, although mandated, continues to be an area of contention in the assessment field. Although the importance of using evidence based practices is obvious, what constitutes "evidence-based" is determined by a broad range of philosophical views and pragmatic understandings about quality research. As Cook, Tankersley, and Landrum (2009) explained in a review of methodological quality, "Despite the considerable interest in basing instructional practices on research evidence, special educators have not yet established definitively which practices are or are not evidence-based or settled on a systematic process for determining evidence-based practices" (p.366).

Even without consistency of understanding across the special education field, preservice teachers require some understanding of this issue in order to understand potential best practices, including assessment practices. Emmons et al. (2009), in the context of studying research and information literacy for preservice teachers, discussed

the critical need for preservice teachers to differentiate between academic research that contains well-supported recommendations for evidence based practices and research recommendations that are less substantiated. They concluded that directly teaching such information literacy skills to preservice teachers may lead to increased competence, confidence, and professional commitment in preservice teachers. They also addressed the difficulties presented above by Cook et al. (2009) and explained this concern particularly when considering the need to conform to the evidence based practice requirements of the U.S. Department of Education, which tend to emphasize quantitative measures that may not encompass the broad range of students who receive special education services.

In addition to debate about what practices can be determined to be evidence based, there is concern about providing preservice teachers with assessment understandings that are appropriate for the context of their eventual teaching assignments. The current research base is minimal in this area. As explained by Delano, Keefe, and Perner (2010), requirements for special education licensure vary widely by state and the quantity and quality of preservice teacher curriculum and training varies with these requirements. Further, alternative licensure programs and field experiences across grade levels may or may not match eventual teaching assignments but also affect the instruction provided to preservice teachers. Within this possibly limited instruction, preservice teachers may receive minimal instruction in methods to support specific disability categories and students categorized as needing greater supports. Little research in this area has been published but based on this study, it is logical to question the extent of assessment instruction available to preservice teachers, particularly in the area of students described as having more extensive need for support.

Although the larger frames of accountability based on high stakes testing and eligibility for services based on intensive individual evaluation are always relevant in the field of assessment, what is particularly relevant to educational practice is how teachers are assessing the day to day learning of content in classes, including both planning for assessment and actual practice. Campbell and Evans (2000) address planning for assessment in their study of assessment quality, which demonstrated minimal competence of preservice teachers in linking learning goals with assessment in lesson plans. This was based on examination of 309 lesson plans from 65 preservice general education teachers who had completed, among other courses, a course in assessing student work. Studies such as this clarify the need for further understanding of how much preservice teachers know about assessing classroom learning and how this knowledge impacts practice.

Little research is available about how new teachers implement assessment knowledge in actual practice. One example of this implementation from a practice perspective comes from Kohler, Henning, and Usma-Wilches (2008). This study discusses the perhaps limited pre-service understanding of the variety of assessment methods available. Among student teachers, the two predominant forms of formative assessment used were "listening to student talk" and "observing student behavior" (p. 2113). Although other methods of assessment of learning were provided, these methods were what preservice teachers actually used. It then becomes necessary to understand why student teachers choose particular methods (course work examples, ease of use, cooperating teacher expectations, gaps in assessment understanding, prioritizing non-assessment issues, etc.) and if these methods fulfill the purposes of effective assessment.

The Maclellan (2004) content analysis of pre-service teacher essays about assessment indicated consistent understanding of the varied purposes of assessment and the particular importance of formative assessment such as the types that would be used in classrooms during every day teaching and learning situations. These essays were examined in the context of the University of Strathclyde in Glasgow and commented on the need for self-assessment, reliability, and validity, although it was not clear from what was written whether or not the preservice teachers understood how to determine what assessments are reliable or valid in particular situations. The study discussed a lack of pre-service teacher knowledge of assessment methods, including knowledge of authentic assessment, performance based assessment, and the links between teaching methods and meaningful assessment.

Understanding how new teachers assess learning in the context of teaching using evidence based practices in a broad system of accountability is essential. Part of this understanding includes the knowledge students acquire from university courses. This is knowledge that can at least partially be understood from direct discussion with preservice teachers.

Research Context

The University of New Mexico (UNM) has a dual license education program, from which students graduate with a Bachelor of Science in Education degree in Special Education and Elementary Education. This includes coursework making students eligible for both a general education license (K-8) and a special education license (preK-12). The purpose of the Special Education Dual License Program (SEDLP) is to prepare preservice teachers to enter the teaching profession ready to provide an appropriate education for

students with a broad range of skill levels and learning needs. Approaching special education from the perspective of this broader base of knowledge allows greater opportunity of success in supporting students eligible for special education services within the context of the general education curriculum. It allows for generalizing knowledge, such as assessment knowledge, across a range of experiences.

Rationale

Assessment, and particularly assessment in the context of special education, is a broad topic with a complex history. Teaching preservice teachers to skillfully assess the classroom learning of students with disabilities is a challenge. Skillfully accomplishing this task is important to the future of preservice teachers and to all of the students with whom these educators will interact. The purpose of this study is to explore what SEDLP teachers understand about classroom assessment and how they use this assessment knowledge in practice, as well as to explore the strengths and continuing challenges of the SEDLP in preparing students for classroom assessment. Within this exploration it is important to understand how SEDLP students describe effective assessment for classroom purposes, characterize their level of competence in assessing the growth of skills for students with disabilities, and use or plan to use assessment to adjust instruction.

Conclusion

The remainder of this qualitative interview study is described in the following chapters. Chapter two provides a review of research of current classroom assessment topics, including assessment history, definition, trends and issues, and treatment in teacher preparation programs. Chapter three outlines the research methods for this study.

Chapter four reports study results, and Chapter five provides a discussion of the research

results and implications for future research.

CHAPTER 2

Literature Review

The purpose of this review of literature is to examine the research base within which teacher education programs provide instruction in classroom assessment. This context includes assessment definitions, laws, trends and issues, and methods. It also includes explanation of current teacher education programs' approaches to assessment, including preferences for particular methods and the effectiveness of these approaches. The assessment context also includes current understandings of the relationship between embedded assessments and IEP goals. In addition, because of the nature of this study, this review of literature also includes an examination of interviewing as a research strategy.

Method

Each main division of assessment inquiry required different search parameters. I began the literature search by searching the Zimmerman Library computer (LIBROS) using the words "special education", "assessment", "IDEA", and "embedding" to attempt a broad overview of assessment research. I located the general section for each area of interest and conducted a shelf search for books matching these topics, selecting 33 that met inclusion criteria based on the overall topic of assessment in education, with a preference for assessment in special education. Next, I used Academic Search Premier with the key words "assessment" and "special education" and added to these keywords, in separate searches, "definition" (37 entries), "legal" (29 entries), "IDEA" (52 entries) "embedding" (1 entry), "teacher preparation" (19 entries), and "history" (35 entries). Inclusion criteria included peer reviewed literature, English language format, and direct discussion of at least one keyword used in the search. I did not limit searches based on

year of publication or country of origin due to concern about limited information for

some topics. Exclusion criteria included a primary purpose other than those described by

the search keywords, language format other than English, and duplication. I then

expanded the search by locating articles listed in the references of those that met

inclusion criteria. Next, I examined my personal library of educational methods books for

relevant information about qualitative research methods. I conducted a separate search for

one specific recommended article. Finally, I used Academic Search Complete with the

search terms "assessment," "special education," and "teacher preparation" to determine if

additional relevant studies had been completed since the beginning of this literature

review.

Results

The searches yielded 33 books and 240 articles, of which 104 met inclusion

criteria and are included in the discussion below. In total, 14 books (some with more than

one cited section) and 73 peer reviewed articles were used. One government document

and one conference presentation was also included.

Fundamentals of Modern Educational Assessment

Definition and description. Educational assessment is a broad area. It includes

high-stakes standardized achievement tests, classroom assessments used regularly by

teachers, and the variety of assessments used in the special education system to determine

eligibility for special education services and to create and monitor Individualized

Education Plans (IEPs). Ysseldyke and Algozzine (2006) define assessment as "a process

of collecting data for the purpose of making decisions about students" (p.74). This

definition is simply stated and although perhaps not specifically explained in other

definitions, it sets forth the basic assumptions embodied in other definitions and descriptions. Assessment at its most basic always requires data collection, and its purpose is always to make some sort of decision regarding the education of a student or students. The authors elaborate on this definition as they further describe the purposes of assessment:

> Assessment information is used to make the following kinds of decisions: screening, provision of special help, referral to an intervention assistance team (IAT), provision of intervention assistance, referral to a child-study team, exceptionality, presence of special learning needs, eligibility or entitlement for special education, instructional planning, progress evaluation, program evaluation, and accountability decisions. (p.75)

Ysseldyke and Algozzine continue to explain that because the focus of any assessment is improvement of the educational experience, assessment should occur with the appropriate frequency and in the appropriate environment to produce reliable data and therefore appropriate improvements. What constitutes "appropriate" frequency and environment would be determined by the needs of the student and the type of information sought.

Other definitions of assessment not only include the assumptions of data and purpose explained by Ysseldyke and Algozzine (2006), but also expand the definition with details about methods or further explanations of purpose. Cummings (2000) cites Salvia and Ysseldyke (1981) to explain the functions of assessment in the special education context of schools: screening for further testing, determining eligibility, showing academic strengths and weaknesses, and "determining the impact of educational intervention on a class or group of students" (p.22). Schmitt (2000) cites Taylor (1984) to

explain an additional purpose of assessment: to assist in decision making about the least restrictive environment (LRE) for a particular student.

Cummings (2000) also adds to an understanding of assessment by explaining the variety of forms it may take. These forms are explained as dichotomies, such as norm-referenced vs. criterion-referenced assessments, individual vs. group administered assessments, and informal vs. formal assessments. Cummings (2000) cites Sedlak, Sedlak, and Steppe Jones (1982) to explain informal assessment in more detail: it includes "error analysis, behavioral observation, and the learner's relations to various instructional strategies" (p.24). These are all components of assessment that are used to make appropriate decisions about the purposefulness and systematicity that are described further in this paper in an explanation by the National Research Council (2008).

Hexom, Menoher, Plummer, and Stone (2008) explain the benefits of assessment when applied to both special education and general education environments. They state, "Many researchers have found a correlation to student success when principals, teachers, and the district use assessment to drive instruction" (p.389). They cite Kame'enui et al. (2000) to explain the strong and necessary link between assessment and curriculum, and identify the need for assessment tools that match the particular goals of each assessment situation. The link between assessment and instruction is even more strongly stated by Campbell and Collins (2007, p.9): "When assessment and instruction work in tandem, and assessment is implemented effectively, improvement in student achievement is likely to occur." Improving student achievement is a strong justification for applying appropriate assessment.

Hexom et al. (2008) also considered the effect of curriculum-embedded assessments and the scheduled review of the assessments to determine student growth. The authors explained that students with IEPs also received additional assessments beyond these embedded assessments, and that research in 33 California schools-in-need-of-improvement was conducted to understand how assessments reflected and informed progress. The result was that "refocusing on the achievement of all students does change school performance" and that "providing teachers with professional development and collaboration time does improve the quality of teaching and learning" (p.401). Both of these conclusions require the use of assessment and justify increased focus in this area.

The National Research Council (NRC) of the National Academies was asked by the U.S. government in 2006 to study "developmental outcomes and appropriate assessment of young children" (p.2) to find "important outcomes for children from birth to age 5 and the quality and purposes of different techniques and instruments for developmental assessments" (p.2). The NRC came to two conclusions (2008): first, "the purpose of an assessment should guide assessment decisions," and second, "assessment activity should be conducted within a coherent system of medical, educational, and family support services that promote optimal development for all children" (p.2). These findings are summarized in this publication as "purposefulness" and "systematicity." The concepts of both purposefulness and systematicity, though explored in the context of young children, are easily generalized to assessment of a broad group of students and to the field of assessment as a whole.

According to the NRC (2008), assessments need to be high-quality in three areas: "psychometric properties of the instruments used in the assessment system; the evidence

supporting the appropriateness of the assessment instruments for different ethnic, racial, language, functional status, and age group populations; and the domains that serve as the focus of the assessment" (p.3). The NRC gives three guidelines for assessments (within the context of purposefulness and systematicity) that help to clarify their purposes: the purpose of the assessment should be clear ("explicit and public," p.5), the assessment strategy should be determined by the purpose of the assessment, and each assessment strategy should be closely examined to determine the appropriateness (validity and reliability) for each student assessed.

An additional, though much less emphasized purpose of assessment, is to "accentuate student positives" (Tomlinson, 2009, p.17). This is particularly important when considering students who receive special education services for learning problems that are the result of a severe disability. By knowing what a student can do, educators can plan for how to build on strengths and prior knowledge to increase learning. It is too easy to focus on weaknesses for such students, and this decreases the likelihood that they will be seen as the capable learners that they truly are. In this manner, Tomlinson addresses the social construction of disability, and her work implies a construction of all students as learners.

When considering less formal, teacher generated assessments, the concepts described above remain important. Cole, Ryan, Kick, and Mathies (2000) add additional consideration of purpose for classroom assessment: "Accurate assessment requires that teachers construct avenues of valuing, tracking, and recording individual factors, such as growth, improvement, effort, reflection, risk taking, change, and so on" (p.4). This focus on growth over time is valuable across any content area and appropriate when

considering an individual student's learning, both for students with and without IEP

goals.

Authentic assessment is also an important concept in assessment definitions. Cole

et al. (2000) state that "a fundamental authentic assessment principle holds that students

should demonstrate, rather than be required to tell or be questioned about, what they

know and can do. Hence authentic assessment usually is classified as performance based"

(p.5). This type of assessment is an expansion of the purpose of assessment that is

particularly valuable to educators. Cole et al. include in their explanation a list of benefits

of authentic assessment:

> assists in learning, encourages good instruction, relates to curriculum outcomes,
>
> fosters higher order learning, follows developmental perspective, uses testing
>
> sparingly, supports time efficiency, reports meaningful information, promotes
>
> partnering of parents, educators, and students, fosters student metacognition and
>
> reflection, and is individualized. (p.7)

This explanation could also describe effective teaching, and therefore justifies authentic

assessment as a tool for educators.

Among specific content areas, Lipson and Wixson (2003) explain that "the

purpose of assessment, then, is to find patterns of interactions that allow us to make

relatively good decisions about instruction" (p.55). This is consistent with the definitions

described above and is here applied specifically to literacy instruction. As the authors

explain:

> The long-range goal of the entire assessment and instruction process is to produce
>
> strategic, motivated, reflective readers and writers and to develop mature readers

and writers who can and will apply their skills and strategies independently and in

a flexible manner- not to identify causes or provide labels. (p.56)

This additional point is important to consider in the context of special education, as well.

The purpose of special education services and IEP goals is not to identify causes or to

provide labels, but to produce skilled adults who can apply their skills in a flexible

manner.

An additional important point made by Lipson and Wixson (2003) is that "It is

important for classroom teachers to develop a repertoire of assessment techniques that

can be incorporated easily into their daily instruction" (p.56). This repertoire would

provide appropriate attention and increasing options for teachers to attend to the

purposefulness and systematicity described by the NRC (2008). This attention is essential

in making full use of assessment opportunities.

Legal requirements and guidelines for assessment. The legal, universal U.S.

requirement of assessment for students with disabilities began with The Education of all

Handicapped Children Act (PL 94-142) in 1975 (Kaufman, 2008). Levine and Wexler

(1981) explain that PL 94-142 created the requirement of an Individualized Education

Program (IEP) for all students receiving special education services. This is important

because as these authors explain, the IEP "sets out in writing what special education and

related services will actually be provided to the child" (pp.106-107). The creation of the

IEP therefore provided the framework about which assessment would be required- the

IEP goals- and provided the requirement of annual review of these objectives. Levine and

Wexler explain guidelines for assessment within PL 94-142, and particularly the

requirement that "tests and evaluations had to be non-discriminatory" (p.108). This

guideline appears to be intended more for the assessments that determined admittance to special education services than for assessment of goals. It is important to note that PL 94-142 guaranteed access to a free appropriate public education (FAPE), and assessment of student needs and success was to be conducted in this context. Funding for implementing PL 94-142, including its assessment provisions, required periodic reauthorization (Levine & Wexler), and discussion about PL 94-142 therefore continued.

Complementary legislation, according to Levine and Wexler (1981), included Section 504 of PL 93-112, which stated that discrimination against a person solely because of a disability would not be permitted for any program that received funding from the federal government. The authors explain that although PL 93-112 was passed in 1973, federal regulations for implementing this act were not written until 1975, when regulations for both PL 93-112 and PL 94-142 were written. According to the authors, regulations for both acts were complementary and informed each other, but access to the rights contained in PL 93-112 was not available until the regulations that accompanied it were written and went into effect in 1977. The regulations for PL 94-142 went into effect shortly after.

Levine and Wexler (1981) describe the due process components of PL 94-142 (including the requirement of non-biased assessment) as being particularly difficult to produce in a manner that provided satisfaction to the variety of stakeholders that participated in the writing of regulations. Regarding assessment, the authors explain that the final version of evaluation procedures and results must be available in the primary language of the family concerned and that assessments must be conducted to determine "specific areas of educational need and not be only those aimed at assessing a special

intelligence quota" (p.119). Additionally, only certified professionals may conduct

assessments. This refers to professionals certified as diagnosticians rather than teachers.

At this point, the assessment focus appeared to be on determining who would be eligible

to receive services. Additions continued to this act. Weishaar (2008) explains that among

the 1997 additions (now called IDEA), "all IEPs had to include measurable goals and

objectives or benchmarks" (p.68). These goals continued to be required, although the

exact requirements for the goals changed over time. They were a step toward the current

version of IDEA assessment.

Kaufman (2008) and Weishaar (2008) describe how The 2001 Elementary and

Secondary Schools Act, better known as the No Child Left Behind Act (NCLB),

influenced the development of the 2004 IDEA. Regarding assessment, the NCLB act

explains that improvement will "be accomplished by (1) ensuring that high-quality

academic assessments…are aligned with challenging State academic standards." The

NCLB act also explains that routine monitoring and intervention is needed to ensure that

student achievement is maintained. Campbell and Collins (2007) explain these sections of

the NCLB act and remark that, "It is likely that opportunities for intervention will go

undetected if student progress is (a) examined infrequently, (b) conducted poorly, or (c)

not monitored at all" (p.9). All of these remain assessment issues that are relevant for

both general education and special education.

In 2002, The President's Commission on Excellence in Special Education

presented *A New Era: Revitalizing Special Education for Children and Their Families*

(Kaufman, 2008). This report "called on schools to focus greater attention on delivering

opportunities and evidence-based practices promoting academic achievement and social

participation for students with disabilities" (Cushing, Carter, Clark, Wallis, & Kennedy,

2009, p.195). This is relevant to assessment in that it suggests increased opportunities for

achievement, such as those embodied in appropriate assessment. It continues the

expectation that appropriate teaching and learning will be available to all students.

The Individuals with Disabilities in Education Improvement Act of 2004 further

mandated assessment change (Kaufman, 2008). One change was described by Kaufman

as having been created to ease the burden of paperwork in the IEP system: "The 2004

IDEIA changes the requirement that IEPs include benchmarks on short-term objectives,

except in the instance of children who take alternate assessments based on alternate

achievement standards" (p.49). Gartin and Murdick (2008) also describe this change in

assessment procedure.

Weishaar (2008) explains that a shift was made in IDEA in order to focus on

educational outcomes and that RTI (Response to Intervention) was planned to deal with

learning disabilities in the 2004 IDEA rather than continuing the discrepancy model

previously used. So, learning disabilities are now considered to exist when a student's

response to educational interventions (to address skill deficits) is inadequate. Weishaar

explains that this type of assessment is addressed in depth in IDEA, and describes seven

principles of IDEA assessment, including "nondiscriminatory assessment" (p.73).

Hawkins and Riley (2008) further discuss assessment in relation to RTI intervention and

describe the need for precise and frequent assessments and data analysis to determine

exactly what skills the student is missing and how students have or have not responded to

the interventions provided. They raise a significant concern that is not adequately

addressed in IDEA:

The notion that these assessments will measure student performance in the so-called core curriculum is, as we have seen earlier, the first problem. The second is the extent to which assessments exist- even in schools and districts that have a coherent curriculum- at the quality and quantity required for them to provide the information that educators will need to make good decisions about individual student needs. (p.411)

According to this concern, core curriculums are not defined to the extent that they can be useful in measuring student performance in a manner that is reliable and valid. Race to the Top (RTTT) grant programs have attempted to address this area by providing "common core" curricular standards that all successful state grant applicants have adopted (Duncan, 2009).

Deisinger (2007) explains the overall impact of IDEA 2004 on academic assessment in a citation of Bateman and Herr that explains that "IDEA 2004 places increased emphasis on the need for students in special education to achieve academic progress. It also demands that such progress should be measured objectively and accurately…" (p.115). This refers to progress specifically across IEP goals, which are formed based on curricular standards and benchmarks. This is an attempt to create real change in expectations of success and in educator accountability for the learning of students with disabilities. It makes use of the understanding of purposefulness and systematicity (as described by the NRC, 2008) in assessment.

Additionally, citing Bateman and Herr, "documentation of measurable progress is consistent with the mandate that each child must receive a free and appropriate public education" (Deisinger, 2007, p.115). This link between assessment and appropriate public

education can be expressed using grades (citing Burkhardt, Hendrickson, Gable, &

Manning, Mertler, Sarkees- Wircenski & Scott), although (citing Munk and Bursuk)

letter grades seem to "place students with disabilities at a disadvantage relative to non-

disabled peers" (p115) and although (citing Mertler) "letter grades are inadequate for

communicating a student's academic strengths, and… do not convey information

concerning specific areas of deficit where improvement is needed" (p.115). Deisinger

(2007) attempts a reconciliation of grades and assessment, which includes an explanation

of an appropriate grading system. Deisinger's explanation is complex and, for practicing

teachers, seems unrealistic in that it attempts to use a limited system to express the much

broader system of assessment. IDEA does not specifically link grades to goal assessment,

although in practice grades may be used as part of the required assessments.

In addition to the legislated guidelines and regulations discussed above, court

decisions have affected assessment. The courts are particularly concerned with the

appropriateness of IEP goals. Assessment is a clear part of demonstrating whether or not

goals are appropriate. Yell (2006) explains several court cases that provide precedents for

future understandings of IEP goal requirements. One area in which IEP goals can be

inappropriate is if goals are not written for each area of need (Board of Education of the

St. Louis Central School District 1993, *Burlington School District 1994*, & *New Haven

Board of Education 1993*). Another area of concern explained by Yell is that:

> IEP teams must now ensure that a student's goals are measurable, say how they
>
> will measure the goals, and then actually measure them. Moreover, if a student is
>
> not progressing on a pace to meet his or her goals, the teacher must make
>
> instructional changes to the student's program and continue to monitor progress.

(p.295)

When these issues are ignored, judicial action is possible, such as that cited by Yell in

Board of Education of the Casadoga Valley Central School District, 1994; Cris D. v.

Montgomery County Board of Education, 1990; Lewis v. School Board of Loudoun

County, 1992; and Susquenita School District v. Raelee S., 1996.

Assessment trends and issues from 1975 to present. In addition to the laws and

guidelines established, several trends and issues in educational assessment have

developed and remain relevant. One issue has its roots in the ontological understanding

that informs assessment. For example, individuals with more positivistic life-views are

more likely to perceive assessment as something that genuinely provides the explanation

of the reality of the condition of understanding within the person assessed. Individuals

who view the reality of the world as a social construction are more likely to view

assessment as a more subjective process that reveals the current condition created by

interaction of the individual with external variables. The constructivist view is a more

hopeful view. It implies that changing environmental variables can create changes in the

quality being assessed, whereas the positivist view implies that the quality being assessed

is fixed and static in the individual.

Bourke and Mentis (2007) discuss the ontological perspectives described above.

Their description of applied positivism (embodied in a "psychometric model") is

particularly relevant and clearly stated:

[in the] *psychometric model* where often a deficit orientation is taken,…the

assumption is that difficulties lie within the learner. Emphasis is placed on the

diagnosis, prognosis, and etiology of the problem. There is less accountability on

the part of the teacher and minimal assessment of the curriculum, classroom

environment or context, because the innate qualities or deficits of the learner are

central to this approach. (p.321)

This discussion reflects both past and current assessment thought. Although there has

been some move to more authentic assessments that are grounded in constructivist

ideology, the world of special education remains focused on diagnosis and remediation of

deficit. Insofar as this provides concrete measures from which learning can be planned,

this is appropriate; however, the risk of labeling a disability as some innate part of a

person is high when using these methods. Assessment approaches that are diligently

focused on what has been learned but are more fully grounded in the educational

environment are generally more respectful of the learners being assessed in that they

provide some hope that the student can learn if environmental aspects are changed to

provide what the individual needs.

Weishaar (2008) briefly charts changing assessment trends from 1976 to 1995 to

2007. For 1976, the chart explains the presence of "little accountability other than teacher

feeling that progress had been made and parents being informed of progress" (p.66). By

1995, this changes to "assessments administered annually to determine if IEP goals were

met" (p.66), and by 2007, there exists "ongoing assessment to determine response to

interventions" (p.66). This path follows the legislative assessment path discussed in the

previous section. This path also allows, as Weishaar explains, for assessment of skills of

students in the general education program to indicate a need for intervention that can then

be provided without the student receiving special education services. Using this strategy,

every student is treated as an individual with a property right to education and a plan is

made to meet the need of any student who needs additional intervention to experience

success. This shift indicates a priority change- rather than a rush to identify students who

are having learning difficulties and assessing them for participation in special education

services, assessment is being used to attempt to improve student achievement, and

referral to special education occurs only if attempts at intervention are unsuccessful. I am

hopeful that this is indicative of a larger trend toward using assessment to provide what

each learner needs.

The National Research Council (2008) provides further information about

assessment trends and issues. The council explains the current push to create

"horizontally coherent,…vertically coherent…and developmentally coherent" (pp. 9-10)

assessments. Horizontally coherent assessments demonstrate alignment of curriculum and

assessment. This sounds simple, but has been difficult for educators to put into practice.

In many cases, coherent curriculums do not exist. In the cases where they do exist,

assessments must be selected that measure curricular understanding. They must validly

assess the curriculum and reliably assess across time and students. Vertically coherent

assessment as explained by the NRC involves a consensus among educators about what

skills are expected to be mastered at what educational levels. This is also more difficult

than it appears, because of the extreme variety of expectations even in seemingly

coherent skill sets. When more complicated material is introduced, vertical coherence is

stretched even further. Developmentally coherent assessments are, as the NRC

describes, developmentally appropriate across all contexts. This potentially addresses a

fairness issue in assessment. Assessments must be developmentally appropriate to

achieve reliable data. In other words, participants in assessment must be able to

understand the tasks they are given and be able to complete them, in order for the

assessment to be valid.

Another fairness issue in assessment includes the issue of who is conducting the

assessment. There are situations in special education practice in which the person

conducting the assessment is unfamiliar with the context and environmental norms in

which a student usually learns. Lipson and Wixson (2003) discuss this challenge and

label the experience as difficult. It is possible for this situation to occur when educational

diagnosticians conduct their assessment work and can lead to a view of a student as less

capable than he or she actually is. It is possible also for special education teachers

responsible for particular students' IEPs to experience this same problem in evaluating

progress toward IEP goals. This is an issue when the staff member responsible for

monitoring IEP goal success does not have direct daily contact with the students for

whom the staff member is responsible.

Another major issue is that of teacher education and preparation for educational

assessment. Campbell and Collins (2007) provide a list of research that explains the

status of teacher knowledge about assessment: "(Brookhart, 2001; Campbell & Evans,

2000; Carter, 1984; Fleming & Chambers, 1983; Mertler, 1999; Mertler & Campbell,

2006; Stiggins & Bridgeford, 1985)…(Mayo, 1976; Noll, 1955; Schafer & Lissitz,

1987)…Green and Mantz (2002)…Stiggins (1991)…Ediger (2000)" (pp.9-10).

According to Campbell and Collins, this body of research primarily describes missing

components of teacher pre-service education, and considers necessary assessment

knowledge from both general education and special education perspectives. These

authors reviewed current topics in introductory assessment course textbooks in an attempt

to understand what is considered essential assessment understanding for educators. Results of their study included thirteen categories of assessment instruction: "Decisions, Law, Technical Adequacy, Plan Assessment, Assessment Target, Assessment Type, Assessment Method, Interpret Assessment, Communicate Assessment Results, Assessment Population, and Computer-Assisted Assessment" (p. 11). Within these categories were 73 different topics, which were rated by their commonality in both special education and general education and by the degree of emphasis in the assessment textbook. Although Campbell and Collins provide a thorough explanation of assessment instruction before 2000 and some information for post 2000, pre-service assessment instruction for educators is an area in which more current information may be useful.

Assessment from the perspective of IEP goal quality and educator assessment knowledge is an issue that began with PL 94-142 and continues to the present. Schenck (1980) found that "the long-term goals and short-term instructional objectives of the IEP have limited foundation in the psychoeducational assessment…the extent to which current IEPs are addressing the unique needs of learner must be seriously questioned" (p.341). Tymitz (1981) also concluded that general education teachers were inadequately informed about IEPs and their goal assessments and that the special education teachers need more training in "how to use assessment information to establish goals and objectives" (p.260). Schenck's and Tymitz's concerns are also expressed by Catone and Brady (2005). These authors examine the change in IEP goals from elementary to high school in the areas of basic reading skills. They particularly consider decoding skills and assert that by the time students reach high school, IEP goals no longer address these types of reading needs. The authors studied student IEP goals over time, and determined that

not only did goals fail to specifically address reading needs, their specificity decreased over time as goals became more connected to learning the content for particular courses; therefore, reading skills became less assessed, even for students with particular difficulties in this area.

Walsh (2001) describes the anxiety special education teachers may feel when attempting to use their current knowledge for assessment of IEP goals. Walsh recommends using a variety of tools specifically designed to link curriculum content and IEP goals in both writing and assessing progress toward goals. This is important in that it embodies the assumption of general education curriculum in the IEP goals as well valid assessment, a view that is consistent with FAPE and the principles of the least restrictive learning environment. The link between needs, goals, and assessment is a continuing area of investigation in special education. As students progress in age through educational environments, it becomes an increasingly complex issue.

An additional assessment issue, also related to the assessment knowledge of educators, is selection of appropriate assessments for varied situations. Ysseldyke and Algozzine (2006) list some possible assessment options: "unit tests, portfolios, observations, and feelings and impressions" (p.15). The list continues with "Curriculum-based assessment, Curriculum-based measurement, Instructional diagnosis, Academic time analysis, Assessment of instructional environments, Outcomes-based accountability, and Performance assessment" (p.19). Authentic assessment (also embodied in parts of the list above) is also a concern. Whatever assessment method is used, efforts must be made to understand the appropriateness of a particular assessment for a particular purpose.

Curriculum based measurement (CBM) is an area that has experienced much development. Alonzo, Ketterlin-Geller, and Tindal (2007) provide explanations of curriculum based measurement (as this term has changed from 1979 to present). The current definition that the authors provide explains that in CBM, "time series data are generated to reflect improvement using frequent comparable measures" (p.308). The authors also explain that CBM is used to preview and review curriculum and that it uses norms, criteria, and progress in comparison to self to provide an assessment. Codding, Skowron, and Pace (2005) demonstrate that CBM can be used to create more effective IEP goals and that teachers can be trained to more effectively use CBM. In many cases, CBM answers the requirements of purposefulness and systematicity.

An additional relevant trend continues to be the use of assessment primarily for accountability purposes rather than as a guide for instruction. The discussion about this topic centers on summative vs. formative assessment. Bourke and Mentis (2007) summarize the benefits of both types, and emphasize that the assessment used should be selected by the purpose of the assessment, to provide the information needed. The authors make an argument for formative assessment to follow a path that is more constructivist, humanistic, and ecological. Within this context they explain the use of student self-assessment as a method for guiding instruction. Meta-cognitive and behavioral models of assessment are still possible within the framework of these approaches, and the learner retains his or her identity and value as a person of capability.

The types of assessments used have also been linked to the types of disability categories used to describe individual students. Although the validity of an assessment is not based on disability category, the type of assessment selected is likely to be different

based on category. For example, students who have IEPs to address behavioral concerns

are more likely to have functional behavior assessments (FBAs) as a regular part of the

ongoing IEP goal assessment routine (Johns, Crowley, & Guetzloe, 2002). Students with

autism spectrum disorder are also more likely to have FBAs included in their

assessments. The use of FBAs involves structured observational tools to describe

behavior in classroom environments.

     Whatever assessment approach is used, the tool should match the purpose. As

educators attempt to improve understanding of assessment, the focus should remain on

that which allows improved instruction and the highest possibility of learning for all

students.

Assessment in Teacher Preparation Programs

     The National Council for Accreditation of Teacher Education (NCATE) is

responsible for setting the standards by which pre-service education programs are

measured (Conderman, Katsiyannis, & Franks, 2001). As Conderman et al. explain, this

includes several categories of knowledge, including assessment skills. The authors

conducted a survey of 58 special education programs across the United States to

determine the frequency of program internal review and the methods used to assess pre-

service teachers' knowledge and skills. Among their findings, summative paper-and-

pencil testing remains the most common form of assessing pre-service teacher

coursework success, and although institutions evaluate student-teacher success, they tend

not to thoroughly review the quality of the evaluation that faculty supervisors provide for

student teachers. This study suggests the need for examination of how pre-service

teachers are assessed for content knowledge, including knowledge of assessment.

Conderman, Morin, and Stephens (2005) further examined the issue of assessing special education student teachers in a survey of 100 undergraduate educational institutions. They discovered that 92% of respondents required student-teachers to use informal assessment, 76% required student-teachers to use formal assessment, and 58% required student-teachers to write an assessment report. Further, they confirmed that student-teacher assessment instruction continues to be formed of primarily traditional assignments, rather than more authentic assessments of assessment. Pre-service and in-service teacher assessment skill is an essential competency that warrants continuing examination.

Begeny and Martens (2006) conducted a survey of pre-service teachers to determine their professional knowledge. Their survey participants included student teachers and recently graduated students from masters-level programs through K-12 instructional levels and both general education and special education programs. One-hundred-ten participants from six northeastern U.S. educational institutions were included in the results. One finding was that areas of minimal instruction (as identified by participants) were "behavioral instruction concepts, strategies, programs, and assessment practices" (p.279). A second finding was that pre-service teachers received less instruction time in assessment than in general instructional strategies, with the exception of special education teachers, who reported more coursework and class time addressing assessment topics. The authors state, "The general lack of training in academic assessment strategies raises concern as to how teachers are assessing students in their classrooms" (p.280). Although the increased assessment instruction in special education programs reported in this study is encouraging, the question remains as to whether or not

and to what extent preservice special education teachers generalize this knowledge to their eventual teaching assignments.

McCombes-Tolis and Feinn (2008) surveyed in-service general education and special education teachers licensed to teach K-3 students. Among the many concerns raised in their survey was that 30% of general and special educators did not express agreement with the statement, "I know how to use the results of assessment to improve instruction for a given child or group of children" (p.257). Additionally, as described by the authors, "more than 50% of elementary teachers and nearly 37% of special education teachers responded that they were not required to complete as part of their professional preparation a course dedicated to assessing K-3 students' reading profiles" (p.263). This finding continues the trend described by Conderman et al. (2005) and Begeny and Martens (2006) in which limited assessment coursework is required of preservice teachers.

An indirect method of addressing limited coursework in assessment was demonstrated in Jenkins, Pateman, and Black's (2002) examination of a Hawaiian preservice teacher education program. This program used dual-license preparation in special education and either elementary or secondary education. One benefit of this particular dual-licensure training was situating mandatory assessment coursework (usually only required for special education programs) among the mandatory knowledge base available for preservice teachers in general education. Even in this situation, in which there was an assessment requirement, a collection of recommendations for program improvement (from mentor teachers and school administrators) requested that the program "provide cohort students with more experience in unit and long-term

planning and assessment" (p.368). This type of assessment fits the profile of formative assessment and suggests either inadequate training in such assessment or lack of application from coursework to classroom practice.

Mertler (2005, p.79) lists American Federation of Teachers (AFT) principles as a minimum standard that pre-service and in-service teachers should be skilled in: (a) "choosing assessment methods appropriate for instructional decisions," (b) "developing assessment methods appropriate for instructional decisions," (c) "administering, scoring and interpreting the results of both externally produced and teacher-produced assessment methods," (d) "using assessment results when making decisions about individual students," (e) "developing valid pupil grading procedures that use pupil assessments," (f) "communicating assessment results to students, parents, other lay audiences, and other educators," and (g) "recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information." These principles seem to cover the variety of assessment situations teachers face, with the exception of the need to assess progress toward IEP goals; however, if teachers are skilled in (b), which describes using assessment for instructional decisions, they must include assessments for their students who have IEP goals. Mertler characterized his results as limited due to instrument reliability, but found a general lack of pre-service and in-service knowledge in (e), using appropriate grading based on assessments. Mertler also noted that in-service teachers scored significantly better than pre-service teachers in five areas, and explains that this suggests a lack of valuable assessment training for preservice teachers.

Campbell and Evans (2000) studied pre-service teachers' use of assessment after participating in assessment coursework. They cite several studies to demonstrate 1980's

and early 1990's research about in-service teachers' use of assessment (Brookhart 1993,1994; Frary et al., 1993; Gullickson & Ellwein, 1985; Stiggins et al., 1989); for this more recent study, they examined 309 lesson plans that were developed by 65 preservice teachers for use during their student teaching experiences. The lessons were required to fit the particular framework specified by their student teaching program, including an assessment section for each lesson. The authors report that out of the 309 lesson plans, only 250 included an assessment component and only 82 of the 250 provided a copy of the assessment tool or procedure to be used. Additionally, 32 of the plans without assessments did not include instructional goals. Of the 113 lessons that required a rubric, a rubric was included in only 13 lessons, and a complete rubric was included in only 8. The result was that "the overwhelming majority (approximately 88%) of assessments within the lesson units did not demonstrate preservice teachers' appropriate use of measurement principles when assessing pupil achievement during student teaching" (Campbell & Evans, p.353). The authors concluded that in the instance of this study, pre-service teacher assessment training was not being put into practice during student teaching.

        In more recent pre-service teacher education research, Delano, Keefe, and Perner (2010) explain that requirements for special education licensure vary widely by state and the quantity and quality of preservice teacher curriculum varies with these requirements. Further, alternative licensure programs and field experiences across grade levels that may or may not match eventual teaching assignments also affect the instruction provided to preservice teachers. Preservice teachers may then receive minimal instruction in methods

to support specific disability categories and students categorized as needing greater supports.

In-service teacher assessment does not always fare better. Smith and Gorard (2005) examined in-service assessment practice in the context of Welsh general education. They designed a study to determine the effect of providing grades as feedback versus providing no grades but detailed comments. Data were gathered using school provided standardized assessment information, a demographic information survey, and unstructured group interviews. The researchers found that, across the subjects English, math, and Welsh, the participant group showed less progress than the control group, and "the treatment has been ineffective (or worse) as a method for improving student learning" (p.30). They conclude that likely reasons for this result were formative assessment results that were not used to directly inform teaching, comments that were not specific or understandable enough to inform student learning, and a strong desire among student participants to know their grades. Additionally, they suggest that the theory of formative assessment and general progress monitoring did not translate to in-service teacher practice in this instance.

In-service teacher assessment practice was also examined by Mertler (1999), who used a survey to examine classroom assessment practices in Ohio. He discovered that teachers newer to the profession used a greater variety of assessments than those teaching for 30 or more years, that teachers do not regularly use statistical analysis to study their assessment data, that middle and high school teachers use fewer informal assessments than elementary school teachers, and that, "with respect to insuring that assessments are both valid and reliable, teachers believe that they are doing a better job than most other

teachers" (p.295). Mertler's study did not include assessment specific to special education services, but did generate information about common assessment practices among in-service teachers.

In particular areas of special education, assessment training and coursework may be more lacking than in other areas. Dodd and Scheetz (2003) describe the lack of assessment coursework available in deaf-education programs for students with both disabilities and deafness or hard-of-hearing concerns. They cite weakness in this area to be a challenge in appropriately meeting IDEA requirements within deaf education programs. There is a need for assessment coursework across a variety of disability categories to ensure teacher ability to fully implement the letter and the spirit of IDEA regulations.

McNicholas (2002) was the only study found that specifically addressed assessment for students with more severe disabilities. McNicholas used a survey, several semi-structured teacher interviews, and several observations in the context of four schools in the British education system. He found that assessments of students with severe disabilities were generally summative and generally not connected to curricular planning. Additionally, almost 25% of teachers conducting assessments did not hold special education certification, and "more than 70% of respondents…stated that they required further training in assessment" (p.152). It would be instructive to see a broader picture of assessment for students with moderate to severe disabilities across both British and U.S. environments.

Assessment Methods

Relating curriculum to assessment. The connection of instructional goals to

assessments is a fundamental assessment component necessary for both pre-service and

in-service teachers. Campbell and Evans (2000) cite Mehrens & Lehmann (1991) to

explain that a table of specifications is necessary in assessment to ensure that there is

alignment between the instructional goals and assessments. This is relevant for both

formative and summative assessments. The authors describe the table as having a

horizontal axis that represents the instructional content and a vertical access that has the

six levels of Bloom's taxonomy. Instructional goals determine which level of the

taxonomy will be addressed in which lesson areas, and assessments are planned to match

the content at that level. As the authors explain, having the table of specifications ensures

that assessments match instructional purposes and that higher order thinking is addressed.

A table of specifications appears to be useful for unit planning because it requires

detailed matching of goals and assessment. Some systematic method of determining and

documenting what content will be assessed is needed before teachers decide what

formative and summative classroom assessment methods to use. A table of specifications

would additionally meet the best practice requirements of backward design in planning

curriculum and assessment. This is complementary with Parrish and Stodden's (2009)

description of using backward design to ensure that students with "significant"

disabilities are accessing the general education curriculum within the context of their IEP

goals and learning needs.

Stecker, Lembke, and Foegen (2008) created a list of progress monitoring

strategies recommended for use and recommended for avoidance in improving student

learning outcomes, based on a summary of research. They recommend using consistency,

fidelity to method, and actual *use* of the data to make instructional decisions. They

particularly recommend avoiding using uncontrolled mixing and matching of norms and

cut-off scores from one program with another program. Keilty, LaRocco, and Casell

(2009), from the perspective of early childhood educators, add the use of authentic

assessment environments to the list of appropriate progress monitoring behaviors. They

suggest home involvement in assessment and explain the need for increased professional

development for early childhood special educators and interventionists to facilitate more

effective interaction with families and therefore obtain more authentic assessment data.

Knowledge of progress monitoring methods is essential for appropriate

assessment. Fuchs and Fuchs (1986) discuss appropriate long-term and short-term

monitoring of progress toward instructional goals. In this context, they describe

classroom based assessment (CBA) as the situation in which:

> measurement procedures are designed to match students' program
>
> objectives…student progress data are evaluated regularly with reference to the
>
> performance criteria specified in objectives; and individualized programs are
>
> tested formatively and modified over time as required to ensure effective
>
> instructional programs and attainment of objectives. (p.69)

The authors explain the importance of ensuring that short term CBA goal monitoring is

conducted in the context of long term objectives, and does not become a series of

unrelated skill assessments that fail to lead to an overall objective.

Jones (2001) further describes CBA (quoting Mercer, 1997) as "any approach that

uses direct observation and recording of a student's performance in the school curriculum

as a basis for obtaining information to make instructional decisions" (p.36) and as

(quoting Idol, Nevin, & Paolucci-Whitcomb, 1996) "a criterion-referenced test that is

teacher constructed and designed to reflect curriculum content" (p.36). Shapiro and Ager

(1992) add to the CBA discussion by describing five models of CBA in a mini-review of

research. They explain that one model of CBA "focuses on identification of subsequent

modification of instructional content" (p.284) with the intent of controlling curriculum

difficulty levels (citing Gickling & Havertape, 1981; Gickling, Shane, & Croskery, 1989;

Gickling & Thompson, 1985; Hargis, 1987). A second model "involves teaching specific

skills, and using mastery criterion levels derived from the assessment" (p.284; citing

Blankenship, 1985; Idol, Nevin, & Paolucci-Whitcomb, 1986). A third model is "based

on task analysis and error analysis of skill subcomponents" (p.284; citing Howell &

Morehead, 1987). A fourth model is CBM, and "uses repeated, brief skill measures taken

from instructional curriculum as a means of evaluating the effectiveness of instructional

changes" (p.284, citing Deno, 1985, 1986; Shinn, 1989). The fifth model is "a model of

direct assessment of academic skills that incorporates components of an early version of

CBA…along with components from behavioral assessment paradigms used in assessing

nonacademic skills" (p.284, citing Shapiro & Lentz, 1985, 1986; Shapiro, 1987, 1989;

Deno & Mirkin, 1977). Gansle and Noell (2008) describe CBA in the context of the

consultative role of school psychologists as "the collection of data that will allow

consultants, together with consultees, to develop plausible hypotheses for why academic

problems have occurred for a student" (p.203) and explain that this curricularly-

referenced data explores the relationship among progress "relative to the teacher's

expectations, and relative to other students" (p.204). Although these descriptions are not

thorough enough to truly understand the differences in these models, they do present a

picture of CBA understandings. Shapiro and Ager conclude that using the CBA type that is appropriate to the particular situation and using combinations of types might yield the most useful results for students.

Brookhart and Bronowicz (2003) conducted a multiple-case descriptive study of CBA from the perspective of the students being assessed. They examined "student perceptions of the assessment task, including its interest and importance…their perceived self-efficacy to do that specific assessment; and…goal orientations or reasons for their work" (p.222). Participants were selected from four schools and included third graders, fifth graders, high school World Cultures students, and high school Anatomy students. Students were not described by whether or not they received special education services, but the authors did claim as full of a range of student abilities as might be present in any class. Brookhart and Bronowicz found that the assessment environment was not as relevant to testing effort and success as student individual concerns about testing. As they explain, "in high school, students who found an assessment interesting almost always also thought that it was important in some way. But the converse was not true; students who did not think an assessment was interesting often thought that it was important" (p.239). Elementary school students tended to express interest in the assessment based on their perceived self-efficacy. This study is particularly relevant in that it raises the issue of self-efficacy and classroom based assessments. Investigation into the self-efficacy beliefs of students who receive special education services could demonstrate self-efficacy levels in which students would report lack of interest or effort on assessments. This needs to be considered individually with students, and adjustments to assessment are necessary to ensure students can succeed and are aware of their capacity for success.

The role of student self-assessment in CBA is also worth examining. Brookhart, Andolina, Zuza, and Furman (2004), in a qualitative action research project, examined this topic with third grade mathematics participants, including students with and without IEPs. Participants predicted success levels in weekly Minute Math multiplication activities, selected and used a strategy to practice multiplication, graphed the results after testing, reflected, and began the process again with prediction. Predictions became more accurate over time. All students progressed, but students were placed into groups for additional teacher intervention based on their rate of progress. An additional benefit cited by teachers was that students learned how to use bar graphs, a mathematically appropriate activity. According to the authors, a major benefit of student self-assessment, as demonstrated by students and reported by teachers, was the increase of metacognition in student learning, even with an activity requiring memorization. The student self-monitoring increased student control over the activity and possibly gave students a reason to care about what they were learning. Brookhart, Moss, and Long (2008) further describe the positive learning results of increasing student self-direction in formative assessment.

Allen, Ort, and Schmidt (2009) describe approaches to ensuring that the theory of formative assessment makes it into in-service teacher practice. This addresses a need identified in the survey studies previously described. Allen et al. include in their description a list of areas of attention needed for this to occur: goals, criteria, support, process and product, and feedback to students and teachers. Anderson, Zuiker, Taasoobshirazi, and Hickey (2007), in the context of science instruction, conducted a case study to examine the role of planned student conversation in formative assessment.

They conclude: "the role and nature of group discourse and teacher intervention affect the quality of students' movement along a discursive trajectory from formative feedback conversations to more formal assessment activities" (p.1742). These studies collectively provide support to the variety of possible methods of formative assessment.

Roehrig, Duggar, Moats, Glover, and Mincey (2008) use a grounded theory approach to explain the use of progress monitoring for literacy instruction. They studied CBM as used in literacy instruction for schools receiving Reading First grants and found that participants (ten K-1 teachers, 4 reading coaches) described three ways they used progress monitoring data: monitoring strength and weakness areas, adjusting groups to individualize instruction, and choosing appropriate instructional methods. Roehrig et al. created a model titled "Theoretical Model of Using Progress Monitoring Data to Inform Literacy Instruction" (p.374), which demonstrates how teacher training, reading coaches, and environmental barriers precede reading instruction (including the interactions among the three ways progress monitoring data was used). The reading instruction then has an outcome of either student progress or increased concern and focus on particular student needs. The model clearly demonstrates one way that progress monitoring (CBM with coaching, particularly) affects instructional practice and outcomes.

In-service special education teacher use of CBA has also been studied. Roehrig et al. (2008) cite Capizzi and Fuchs (2005) to state that "Special education teachers who received CBM data and diagnostic feedback (general instructional focus recommendations) created lesson plans in reading that were more tailored to the needs of targeted students than did teachers who received CBM data only" (p. 366). This tailoring was also in evidence in the Reading Recovery program. O'Connor and Yasik (2007)

explain that targeted instruction is possible in both monitoring and creating goals in this program. As they state, "Reading Recovery teachers, because of their individual instruction, daily observation, and assessment of a child, are in a unique position to help create IEP goals that are specifically customized to a child's distinctive needs" (p.147). In this program, progress assessments are conducted often and inform instruction. From an assessment research standpoint, the frequency of this monitoring may be one aspect of Reading Recovery success.

Allsopp et al. (2008) describe "mathematics dynamic assessment" (MDA) as a CBA method to assist students with low success in math. As the authors explain, MDA uses four processes to assist struggling learners: (a) "assessment of student interests and experiences," (b) "concrete-representational-abstract assessment within authentic contexts," (c) "error pattern analysis," and (d) "flexible interviews" (p.6). These strategies are described as research based, although the research is not cited. Allsopp et al. explain that this process is a type of informal formative assessment that helps identify with specificity student misconceptions about particular mathematical constructs. In addition, the authors explain that other benefits of MDA include lower student anxiety and greater access to student prior knowledge. After using MDA and instructional planning to address the content errors identified, Allsopp et al. suggest that other formative methods, such as CBM, can be used to monitor progress toward learning. Dynamic assessment is also described as a strategy useful for students classified as English Language Learners (ELLs) (Spinelli, 2008). Spinelli describes dynamic assessment as being "a sensitive predictor of progress" (p.110), meaning that it is able to detect differences in growth that some other measures may miss.

Gansle, VanDerHeyden, Noell, Resetar, and Williams (2006) explain the concern that although CBA may be part of in-service teacher practice, "the technical adequacy of the secondary measures beyond those derived from CBM is not established" (p.437). They particularly describe this problem regarding the Six Trait writing programs used in many schools, stating that "empirical studies of the technical characteristics of the Six Trait system per se are absent from the peer-reviewed literature and are not provided by the publisher" (p.438). The authors examined writing samples of 538, 1st through 5th grade students, with 6.5% receiving special education services. They examined interrater reliability of both Six Trait assessment and CBM assessment and concluded that "the strongest Six Trait relationships are with themselves. Despite the publisher's assertions that the Six Trait model assesses distinct dimensions of writing that can be used to guide instructional planning, no differentiation arose in this study" (p.444), leading to the conclusion that assessment was measuring a holistic writing effort rather than specific traits. Therefore, the researchers describe Six Trait measurement as technically inadequate as a CBA. They conclude that more research is needed to implement CBM procedures to measure the characteristics intended to be measured by Six Trait. This does not imply that Six Trait writing programs are not instructionally sound, just that they may not be valid or reliable assessment methods.

CBM research and potential for teacher practice. CBM is a type of CBA that has been more fully researched than other types. Fuchs and Fuchs (1991) define CBM as "a standardized measurement system for indexing student proficiency in the basic skills, including reading, spelling, math, and written expression" (para. 6). They explain that

this indexing is possible because "every CBM test administered within an academic year represents the entire year's curriculum" (Fuchs & Fuchs, para. 7).

Madelaine and Wheldall (2004) review CBM research in reading. They particularly discuss the differences between ORF (oral reading fluency) and maze measures. ORF takes a measure of how many words students read correctly in one minute. Maze procedures measure accuracy in selecting from appropriate words to fill in gaps in passages in which every nth word is absent. Madelaine and Wheldall explain advantages of ORF; ORF may be more sensitive for students with less developed reading skills (citing Faykus & McCurdy, 1998), ORF "has the most theoretical and empirical support of any form of CBM" (p. 59; citing Kranzler, Brownell, & Miller, 1998), and ORF showed steeper growth rates (citing Shin, Deno, & Espin, 2000) than maze methods. It is important to note that although ORF was cited as superior for these reasons, the strength of maze was also acknowledged, and a primary advantage of using a maze strategy was noted; its ease of implementation while still preserving the benefits of progress monitoring.

Madelaine and Wheldall (2004) cite Wheldall and Madelaine (1997) and Deno, Fuchs, Marston, and Shin (2001) to explain the urgency of setting ambitious goals for students with disabilities. It is possible to then use the most situationally appropriate form of CBM to help students achieve these goals. Concerns about appropriate forms are varied. Challenges to CBM implementation described by Madelaine and Wheldall include the perception that it is time consuming (citing Eckert, Shapiro, & Lutz, 1995; Faykus & McCurdy, 1998; Foegen, Espin, Allinder, & Markell, 2001; Fuchs, 1998; Hasbrouk, Woldbeck, Ihnot, & Parker, 1999; Marston, Diment, Allen, & Allen, 1992;

and Pemberton, Dyck, Horton, & Kaff, 2002), teacher "resistance to change" (p.69, citing

Swain & Allinder, 1997), limited teacher implementation skills (citing Foegen et al.,

2001), and "lack of face validity" (p.69, citing Foegen et al., 2001; Deno, 2003; and

Faykus & McCurdy, 1998). As Madelaine and Wheldall explain, "lack of face validity"

refers to the perception that ORF cannot measure reading comprehension.

An interesting finding (Madelaine & Wheldall, 2004, citing Swain & Allinder,

1997) was that teachers who actually used CBM for reading did not find it to be

inappropriately time consuming. Additionally, concerning the lack of face validity, pre-

service teachers shared the in-service teacher belief that CBM (ORF) did not measure

progress in reading comprehension (Foegen, Espin, Allinder, & Markell, 2001). Foegen

et al. found that the preservice teachers were convinced of the usefulness of CBM, but

not of its validity as an assessment tool, which led to the description of lack of face

validity. De Ramirez and Shapiro (2006) studied the validity of using CBM measures for

students classified as ELLs. They found that "Spanish-speaking ELLs were making

progress, albeit at a slower rate than general education students" (p.365), and concluded

that CBM was a valid progress monitoring measure for ELL students. Despite perceived

disadvantages, as a progress monitoring strategy, CBM (and ORF, specifically) appears

to be thoroughly supported by a variety of research.

Hosp and L. S. Fuchs (2005) continued CBM research in an exploration of

whether differences in elementary grade level affected the use of CBM to measure

reading skills. They found that "the relation between CBM and decoding was generally

higher in Grades 2 and 3 than Grades 1 and 4" (p.23), that "the relation between CBM

and word reading was higher at Grades 1, 2, and 3 compared to Grade 4" (p.23), and that

"comparisons for CBM and comprehension revealed no difference among the grades" (p.23). Based on this, the authors suggest that CBM may be measuring different skills across early elementary grade levels, based on student progress through stages of learning to read. This is an area that requires more research to form conclusive results.

Wayman, Wallace, Wiley, Tichá, and Espin (2007) also conducted a literature review on CBM for reading. They explain significant duplication in understandings gained from prior literature reviews, including confirmation of the idea that the reading aloud (ORF) measure is a valid indicator of comprehension ability. They express some clearer limitations than were expressed in the earlier research reviews, including the understanding that "reading aloud may not be the best choice for very young and older students" (p.109) due to a floor effect for younger readers and limited reflection of growth for older readers. The authors suggest a word identification CBM method for younger students and the maze strategy for older students, with some overlap and lack of clarity of best methods at the middle school level. They also explain current limited research at the high school level and concern about overestimation of performance levels for African American students and Hispanic students whose first language is Spanish. More research would be valuable in these areas.

Additional findings of Wayman et al. (2007) relate to the effects of curriculum on CBM reading progress-monitoring. The authors found three themes across research; that "level of performance differs significantly with curriculum source" (p.110; Tindal, Marston, Deno, 7 Germann, 1982; Tindal, Flick, & Cole, 1992; Hintze, Shapiro, Conte, & Basile, 1997; Bradley-Klug, Shapiro, Lutz, & DuPaul, 1998; Powell-Smith & Bradley-Klug, 2001; Brown-Chidsey, Johnson, & Fernstrom, 2005), that "although technical

adequacy does not vary with curriculum source, rates of growth may" (p.110; Fuchs & Deno, 1992; Hintze et al., 1997; Hartman & Fuller, 1997; Brown-Chidsey et al., 2005; Hintze, Shapiro, & Lutz, 1994; Hintze & Shapiro, 1997; Bradley-Klug et al., 1998), and that "it is not necessary to match instructional and progress monitoring material" (p.110; Hintze et al., 1994; Hintze & Shapiro, 1997; Tindal et al., 1992; Powell-Smith & Bradley-Klug, 2001; Riley-Heller, Kelly-Vance, & Shriver, 2005). These findings are useful in that they indicate a need to consider what materials will produce the most usable progress monitoring results for individual students while at the same time assuring educators that there are many types of curricular materials that can provide valid results. An additional caution of Wayman et al. is that CBM be used for its intended purpose-monitoring progress. If high-stakes decisions are being made using CBM, different and lengthy procedures are needed to ensure technical adequacy. Current research, the authors explain, does not support validity and reliability for purposes other than progress monitoring.

Jenkins, Graff, and Miglioretti (2009) studied the needed frequency and quantity of CBM for reading progress monitoring. They asked, "Is validity of growth estimates degraded by measuring progress less often, by minimizing the number of scores per measurement occasion, or by minimizing the number of scores used to assess baseline" (pp.151-152)? Jenkins et al. discovered that monitoring could be used less often without decreasing the reliability of CBM (at a minimum of every three weeks), but that the number of scores at each monitoring session and during baseline could not be reduced without negatively affecting reliability. The exception to negative reliability impact with fewer scores in each session was when progress was monitored every week. Although the

authors cite study limitations as a reason for further research before generalization, this is an important possible finding because of the time constraints that may affect in-service teacher practice in implementing CBM and other assessment measures. The authors also caution that there may be other reasons for weekly assessment, such as being able to quickly determine the effect of instructional intervention.

Deno et al. (2009) examined the use of CBM in informing school wide progress monitoring of reading. They used a single-level maze procedure across all grades and reading levels so that results could be examined at a systems level. Maze is described in detail in author citations of Espin, Deno, Maruyama, and Cohen (1989), Fuchs and Fuchs (1992), and Jenkins and Jewell (1993), who explain that "The MAZE measure…requires students to silently read text passages. Every seventh word is deleted and replaced with three word choices. Students select the correct word" (p.46). The authors describe main components of CBM reading measures, and particularly maze, as being desirable to schools because of "its efficiency, effectiveness (validity and utility), and the clarity of the data…generated" (p.54).

Although the majority of CBM studies have been in the area of reading, McMaster and Espin (2007) conducted a literature review of CBM for writing. Interesting findings for secondary education include the contradictory findings that "validity …does not appear to depend on the type of writing prompt…or sample duration (3 min vs. 5 min.) for middle school students" (p.80, citing Espin et al., 2000) and that "longer samples yielded stronger validity than did 50-word samples" (p.80, citing Espin, De La Paz, Scierka, & Roelofs, 2005). McMaster and Espin also discuss the decreasing validity of more simple scoring procedures as grade level increases (citing Jewell &

Malecki, 2005). A further concern described by McMaster and Espin for both elementary and secondary writing is that "procedures developed thus far have yielded more modest criterion validity coefficients than have those obtained in other areas of CBM research" (p.82). These are issues that may be addressed in future research.

Hessler and Konrad (2008) describe possible applications of CBM for monitoring student writing progress and the connection between this monitoring and the creation of IEP goals. They explain progress monitoring problems using holistic writing assessment methods, including ceiling effects and limited sensitivity to growth, and cite Binder (2003) to explain that "CBM is a better alternative than the percent correct measures often used in IEP development" (p.30). They explain CBM as requiring three-minute writing probes based on a prompt that can be analyzed based on individual student needs (using total words written, words spelled correctly, total writing sequences, correct writing sequences minus incorrect writing sequences, and correct punctuation marks). The authors cite Espin, Scierka, Skare, and Halverson (1999) and Espin et al. (2000) to explain that of these, "total words written, words spelled correctly, and correct writing sequences" (p.31) have been found to be reliable measures.

Math has also been an area of CBM research. Foegen, Jiban, and Deno (2007) conducted an extensive review of research regarding using CBM as a method of progress monitoring for math. They concluded progress monitoring in mathematics is a relatively new research field, and that much of this research has only been present since 2000 and has been conducted by a fairly small group of researchers (particularly Fuchs and colleagues). Additionally, Foegen et al. concluded that CBM for math tends to be less sensitive to change than it does for reading, and that "for students who experience severe

difficulties in mathematics, this limited degree of sensitivity to growth may not meet

teachers' needs" (p.136) as a progress monitoring tool. This is a concern for students with

disabilities that affect their progress in math. Foegen (2008) conducted an additional

study using general education middle school participants, and recommended further study

for students with disabilities, but did not predict the generalizability of her research

without further study.

　　　　Stecker and L. S. Fuchs (2000) studied in-service teacher use of math CBM to

make instructional changes for students with learning disabilities, and found that although

CBM information is useful in individualizing student instruction, using "a steering-group

strategy, in which they monitor the progress of one student and generalize their

instructional decisions to larger instructional groups" (p.133) is not effective in

improving outcomes for all students. This implies the need for individualizing instruction

for students, including those who receive special education services across a full

continuum of educational environments.

　　　　Several problems have been discussed with implementation of CBM across varied

content areas. In a description of CBM issues, Cannon (2006) cites Gersten, Morvant,

and Brengelman (1995) to explain that, "in many cases, teachers do not even look at

curriculum based measurement data unless these data serve as a focus for a discussion

with a consultant" (p.10). This, if true, would seem to indicate that teachers do not feel an

immediacy about the value of CBM data. In the context of a study on how to improve

teacher instruction of CBM, Codding, Skowron, and Pace (2005) describe the situation in

which CBM is not connected to IEP goals, and cite Schenck (1980) and Smith (1990) to

explain that there are "documented inconsistencies between assessment information and

annual goals" (p.166). If unaddressed, this reduces the utility of CBM for students who receive special education services.

In Allinder, Bolling, Oats, and Gagnon's (2000) study of adding teacher self-monitoring to CBM implementation, the authors found that "the self-monitoring group was significantly different from the CBM-alone group and the control group; there were no significant differences between the CBM-alone and the control groups" (p.223). The authors conclude that the reason for this is that "the addition of the self-monitoring component enhances teacher's abilities to reflect meaningfully on how students were responding to the instruction they were providing" (p.224). This study replicated parts of prior research (the authors cite Allinder & BeckBest, 1995; Belfiore & Browder, 1992), but added the component of teacher monitoring of "qualitative information about student performance" (p.225) to improve instruction. As demonstrated by prior research (the authors cite Browder, Liberty, Heller, & D'Huyvetters, 1986; Deno, 1985; Forness, Kavale, Blum, & Lloyd, 1997, Kaplan & Carter, 1995), self-monitoring of formative assessment can be an effective strategy. This study suggests that when math teachers who provide special education services self-monitor formative assessments to adjust instructional planning, student learning outcomes improve.

Variations of CBM in combination with other methods have also been suggested. Garcia (2007) describes how to combine CBM with miscue analysis (MA). She explains that the CBM oral passage reading procedure is first used to establish baseline skill levels for individual students, and then MA is used to examine the error types present in the CBM reading. Garcia cites Goodman (1995) to explain that at least 25 miscues are needed to effectively use MA to determine areas for further instruction. This is an

interesting strategy because it has the potential to move teachers from progress monitoring to instructional changes that can then be examined using further progress monitoring.

Wallace, Espin, McMaster, Deno, & Foegen (2007) explain a vision of future CBM research: "a comprehensive system of measurement designed to assess individual progress within a standards-based educational system" (p.66). Such a system would be sensitive to the impact of instructional changes and allow for improved outcomes for students across general and special education, across a variety of grade levels, and across a variety of strength and challenge areas. I would add to this vision the component of educators who are knowledgeable and skilled at implementing CBM and other CBA methods and who use this knowledge to improve educational outcomes for all students.

Assessment Based on Embedding IEP Goals in Classroom Instruction

Description of embedding. It is necessary to address IEP goals with methods that are effective for individual students with disabilities. "Curriculum and instructional adaptations, embedded instruction, parallel instruction, circles of friends, peer tutoring…direct instruction by paraprofessional staff" (McDonnell et al., 2003, p.231), and differentiated instruction are some options to address these goals. To promote generalization of goals, instruction in environments where this is most likely to effectively occur is needed. As Bricker (2000) explains, "youngsters with disabilities did not generalize the responses to home and classroom that they had acquired in a pull-out session" (p.15), and therefore methods that promoted greater generalization were needed. One way to create better generalization possibilities is to embed learning goals in classroom instruction and assessment. Bricker (2001) continues to explain, "Embedding

children's goals and objectives in routine and play activities should be a generalized approach adopted by all…personnel" (p.24).

Horn, Lieber, Li, Sandall, and Schwartz (2000) investigated how IEP goals could be embedded in inclusive settings for early childhood learners. These authors cite Bricker, Pretti-Frontczak, and McComas (1998) for a definition of embedding:

> *Embedding* is defined as "a procedure in which children are given opportunities to practice individual goals and objectives that are included within an activity or event in a manner that expands, modifies or adapts the activity/event while remaining meaningful and interesting to children." (p.209)

VanDerHeyden, Snyder, Smith, Sevin, and Longwell (2005) cite the Bricker et al. (1998) definition, as well as describing embedded instruction as hierarchical. The authors attribute a description of this to Kaczmarek (1999) and to Sandall and Schwartz (2002) when they state that:

> Hierarchical embedded instruction approaches promote "leveling" of intervention, beginning with a developmentally appropriate environment as the foundation, followed by implementing environmental modifications and adaptations, creating distributed trial opportunities for children to practice targeted objectives during routine or planned activities, and using explicit discrete trial naturalistic instructional techniques when necessary to provide children with sufficient instructional trials. (p.81)

When authors are discussing embedded instruction, they appear to be discussing the hierarchical embedding as described above, using the degree of intervention that is deemed necessary for the student and the skills taught.

VanDerHeyden et al. (2005) further discuss the value of embedded instructional goals using the idea of discrete trials, which they cite McBride and Schwartz (2003) as defining as "an instructional cue, a child response, and a teacher-delivered consequence" (p.82). They consider this ABC (antecedent, behavior, consequence) relationship as being important when investigating the efficacy of embedding IEP goal instruction, and cite Binder (1996) and Skinner, Fletcher, and Henington (1996) to explain that providing sufficient practice units arranged in this manner is essential for the success of hierarchical embedding of goals. The authors explain that incomplete ABC units limit the positive effects of embedded learning.

The definition of embedded instruction is further attempted by Sigafoos et al.(2006): "*Embedded instruction* is characterized by the use of existing routines as the context for instruction, with the teacher inserting learning opportunities into the existing flow of the activity" (p.196). Unlike the description provided by VanDerHeyden et al. (2005), the Sigafoos et al. description does not place discrete trial training in the realm of embedded learning. Sigafoos et al. instead describe discrete trial training as a component of self-contained instruction. Using the Sigafoos et al. description, both embedded instruction and discrete trial training can occur in self-contained educational environments; however, this does not appear to fit the "naturalistic instruction" components described especially by VanDerHeyden et al. Placing discrete trial training outside the realm of embedded learning clearly conflicts with the other definitions and descriptions above.

Jameson, McDonnell, Johnson, Riesen, and Polychronis (2007) add to the definition of embedding instruction by further describing distributed trials: "The primary

difference from traditional teaching formats is that the instructional trials are distributed

within and across class activities rather than being presented rapidly one after the other in

a massed practice format" (p.24). In the Jameson et al. study, distributed instructional

trials in inclusive settings were contrasted with mass practice trials in self-contained

settings.

McBride and Schwartz (2003) emphasize the role of child preference in

embedded instruction; "Embedded instruction typically involves incorporating explicit

instruction on preselected target objectives into play activities, focusing on child-initiated

activities within common classroom routines" (p.5). The authors created a study to

examine teacher training needs in the process of embedding goals into ABI (activity

based instruction) in which goals were imbedded in the format of discrete trials in play

activities of preschool children with disabilities. McBride and Schwartz conclude that

"understanding instructional trials, attaining competence in conducting discrete trials, and

knowing how to organize instruction in the context of discrete trials are important

components of naturalistic, embedded instruction such as ABI" (p.14).

Horn et al. (2000) explain that embedded learning in preschool classes occurs in

general education environments with specifically planned adaptations for students with

disabilities. The authors cite Goodman and Bond (1993) to explain that often goals do not

fit well into the general education preschool environment and they cite Notari-Syverson

and Schuster (1995) to explain that "Implementation of the IEP goals should neither

supplant the classroom curriculum nor restrict the child's participation in classroom

activities. Implementation should occur within the context of the existing classroom

activities and routines" (p.208). This identifies the process of embedding IEP goals in preschool environments.

Benefits of embedding, as described by Horn et al. (2000), include augmenting/adapting rather than attempting to replace curriculum in general education environments, the ability to implement goal instruction without additional staff, and the idea that "learning opportunities for objectives should be possible in nearly all classroom activities and areas and could be used to help children develop many different types of functional, meaningful skills" (p.210). The authors use "embedded learning opportunities" or "ELO" (p.210) to describe this process.

Zeece, Graul, and Hayes (2004) discuss embedding for early learners in the context of literature use in inclusive classrooms. These authors provide several examples of specific goals and how they can be embedded in the class literature study. They also provide a philosophy statement of embedding goals in literacy instruction:

> Effective literature-based activities provide opportunities for all children to thrive and learn. As such, children with special needs are treated first and foremost as children. Strategies for embedding IEP goals and objectives in literature-based activities are best driven by this philosophy. (p.256)

In keeping with this philosophy, the authors recommend that instruction always occurs with high interest, exciting literature experiences. The role of child preference (and motivation) as described by Horn et al. (2000) and Zeece et al. (2004) is an important component in understanding embedded instruction.

Although many sources discuss embedding IEP goals in instruction, few discuss imbedding goals in assessment. Assessment of IEP goals can be conducted using

curriculum based assessment, portfolio assessment, and other authentic assessments. Embedding assessment of IEP goals into the classroom experiences of students is not defined in the "embedding" literature.

Embedding practices and methods. Ganley (2000), speaking from a speech and language perspective, places IEP goals (and therefore the embedding of IEP goals) in context:

> Although the IEP must be educationally relevant, it should not duplicate the
>
> general education curriculum in every content area and list everything the student
>
> is expected to learn in a year. Instead, it should describe the supports needed to
>
> progress toward achieving learning standards established for all students and
>
> address a student's other disability needs. (p.2)

This highlights both the purpose of IEP goals and the attention required to preserve access to the general education curriculum.

Kanne, Randolph, and Farmer (2008) propose the strategy of using a "Bridge Document" (p.372) to connect diagnostic work with the planning of IEP goals into instruction for students who have Autism Spectrum Disorder (ASD). The authors do not specifically consider embedding goals into practice, but their work lends itself easily to this process. They explain that there is limited research about the connection between diagnostic work (from which students are admitted to special education services and have the label ASD applied to them) and the construction, use, and assessment of IEP goals. They further explain that in practice this gap may be due to a lack of method for diagnosticians to report results in a way that educators find usable. Schenck (1980) also describes this gap for students with learning disabilities, emotional disturbance, and

cognitive disabilities. To remedy this problem, Kanne et al. suggest that diagnosticians

work with a teacher strongly connected to the student to use a template for what they call

a "Bridge Document." They state:

> The special educator can more narrowly define the recommendations into specific
>
> suggestions based on education interventions considered best practice. This may
>
> be of great benefit, since many recommendations made by health professionals
>
> may not be deemed "best practice" in an educational setting. (p.375)

The Bridge Document is composed of a cover letter, summary page, and full evaluation

document. It addresses the areas of environmental needs, instructional needs, and

behavioral needs, which the authors address based on the needs of students with ASD as

explained by the National Research Council (NRC) report of 2001. The summary page

provides recommendations that can be translated to needed goals.

Wolery, Brashers, and Neitzel (2002) use an "Ecological Congruence

Assessment" (p.131), (ECA), to determine the appropriate manner to embed IEP goals

into classroom activities. They cite Thurman and Widerstrom's (1990) work as the basis

for this assessment and describe the ECA process in three parts: collecting information

about child participation in classroom activities, summarizing this information, and

sharing this information with a team to accomplish the goal of instructional planning.

They explain that within the collecting information section the goal is to describe the

differences between a child and the other children, to explain the child's functional skill

levels within particular contexts, and to explain the "level of tolerance for difference"

(p.133) that is based in adult perceptions of the child and that is based in the child's

perceptions of the context. Using this process is expected to help teachers develop

appropriate functional IEP goals for young children. Although it is concerning that this appears to be a rather extreme model of deficit rather than a more balanced view of the child as a whole person with strengths and weaknesses, the model described does potentially increase communication among a child's teachers and base its recommendations on observed behavior. The authors explain that this process "will always be idiographic- dependent (more than other measures) on the skills and dispositions of the user" (p.140).

Wolery and Anthony (1997) conducted a study to examine the teacher training needed to enable educators to effectively embed instruction of IEP goals. They concluded that "a training package consisting of a written manual, individual training session of 30 to 45 min, and verbal feedback was sufficient to help teachers embed a response prompting procedure, constant time delay, into their ongoing classroom activities" (p.12). They further concluded that the accuracy and frequency of the use of constant time delay increased after the training, but the behavior was not consistently maintained across time by all participants. Although minimal wide ranging conclusions can be drawn from this study, perhaps one conclusion is that greater preservice and in-service education may be needed for teachers to effectively embed IEP goals, including the use of strategies such as constant time delay prompting.

Horn et al. (2000) also raise the concern about general education teacher involvement in goal setting. They cite McDonnell, Brownell, and Wolery (1997) to state that "early childhood educators were not consistently a part of Individualized Education program (IEP) development and implementation" (p.208). If this is this case, the

selection of appropriate goals, including those with appropriate embedding options, and how goals are assessed in the general education classroom also become concerns.

Horn et al. (2000) investigated embedding used a multiple-baseline design in the context of two case studies in which three students' IEP goals were embedded in general education preschool environments, and a single case (AB) design for one other student. The process began with ELO (embedded learning opportunity) planning, including videotaping to inform IEP goal selection and to analyze the embedding opportunities within the general environment. Following planning, baseline data was collected and ELO training was provided to teachers. Then, "dressing up the IEP objective" (p.212, citing as adapted from Giangrego, Dennis, Edelman, & Cloninger, 1994) and creating an "ELO-at-a-glance" (p.211, cited as an adaptation from McCormick & Feeney, 1995) were used to create a comprehensive embedding plan. Horn et al. found in Case 1 that the teacher cues (to the student) and teacher response to the goal behavior increased, while student attainment of IEP goals reached 75% on probes. In Case 2, teacher implementation behavior was inconsistent and student goal attainment varied from 38% to 100%. In Case 3, teacher provided opportunities for practicing IEP goals increased, and average student goal attainment increased to 75%. The study culminated with teacher interviews. Although study limitations were acknowledged, the authors concluded that the ELO process did increase student success at achieving IEP goals. The authors further explained that results were linked to teacher perceptions of their roles in inclusive environments, with greater success when teachers viewed individualization as a goal for all students in the class and less success as teachers felt cognitively overwhelmed.

In considering the relationship between IEP goals and effective instruction to reach these goals, Jameson et al. (2007) examined differences in learning possible when using distributed instructional trials in inclusive classrooms and massed instructional trials in self-contained classrooms. They found that the distributed "response prompting strategy did not disrupt the educational program of students without disabilities in the general education classes" (p.38) and that "both the special education teacher and paraprofessional were able to implement the embedded instructional strategy with a high degree of fidelity" (p.39). The authors also found that although both strategies were effective, the efficiency of distributed and massed trials differed by student. They suggest that one possibility to describe this difference was in the difficulty of the learning goals; "Sarah had no difficulty learning the target stimuli that had distinctive discrimination characteristics, it was only on the discriminations were [*sic*] there was minimal difference that the massed trails [*sic*] seemed to have been a more effective instructional strategy" (p.40). It makes sense that strategy effectiveness might vary by student needs and by goal difficulty.

Wolery, Anthony, Caldwell, Snyder, and Morgante (2002) explain the need for generalization probes to determine if IEP goal performance was achieved using embedded instruction. These authors studied the issue of how to embed goal instruction in an elementary summer day camp during circle time and transitions by distributing trials throughout these times, as opposed to massing trials into one segment of instruction. The authors note that "A problem…of embedding and distributing trials into ongoing activities is delivering enough trials to ensure learning" (p.14). In addition to needing to ensure that students receive enough opportunities to learn the goal material,

the authors explained the need to see if generalization occurred outside of these times. The authors conclude that teachers could be taught to accurately embed and distribute the trials, all child participants learned the goal material, and probes were successfully used to measure generalization across teachers and materials.

Coyne, McCoach, and Kapp (2007) also examine the need for frequent trial opportunities in their discussion of the effects of embedding vocabulary instruction for kindergarten students, including those with learning delays and at-risk conditions that may lead to the eventual description of students as having learning disabilities. They found that when fewer contacts with the target vocabulary words were created, embedded instruction was not as effective as extended vocabulary instruction. The authors do not examine particular words in the context of IEP goals, but their work is important because it highlights the effect of frequency of learning opportunity.

Riesen, McDonnell, Johnson, Polychronis, and Jameson (2003) discuss the most effective methods of embedding vocabulary instruction in inclusive environments. They examined whether, within the context of embedded instruction, constant time delay or simultaneous prompting would be most effective in teaching new vocabulary (presumably indicated in IEP goals) to students with moderate to severe disabilities. Constant time delay is described as "a near errorless procedure that requires the instructor to systematically fade the initial or controlling prompt to the natural discriminative stimulus" (p.243) and simultaneous prompting as a similar procedure but one in which "no attempt…is made to fade assistance provided by the teacher" (p.243). Riesen et al. found that the effectiveness of each strategy varied by student. The authors explain that this differs from research by others (citing Schuster, Griffen, & Wolery, 1992; Tekin &

Kircaali-Iftar, 2002) that found simultaneous prompting to be more effective. Both constant time delay and simultaneous prompting methods were found to be effective by Riesen et al., although different students learned more efficiently from each method.

VanDerHeyden et al. (2005) conducted a study to measure the effect of using discrete trials (ABC units) in embedded instruction to increase preschool students' engagement in learning activities. They found that for the preschool students with disabilities who participated in this study, engagement in learning increased with increased discrete trials presented using embedded instruction. The increase was such that participation increased to the level of peers without disabilities in the same environment, even in a lower quality preschool program. As the authors state, "Under controlled experimental conditions, naturalistic instruction is powerful, producing robust results by using methods that are often preferred by parents and teachers" (p.93). The authors are suggesting that embedding goals in inclusive environments often meets the qualification of parent-preferred educational environment.

Daugherty, Grisham-Brown, and Hemmeter (2001) investigated how embedded skill instruction could be used to teach both target skills (such as IEP goals) and non-target skills (incidental learning). The authors embedded counting skills in preschool activities in the context of general education environments that included students with developmental delays. They found that, using constant time delay in an embedded format, the three students with disabilities did learn the target goals (counting). They further found that one participant also learned the non-target goals. This is an interesting model because it allows for addressing the goals while still differentiating in a manner that does not limit other learning.

Wertz, Wolery, Venn, Demblowski, and Doren (1996) also explain the usefulness of presenting non-target learning goals along with embedded IEP goal opportunities. They conducted a study in general education kindergarten classes that included students with moderate to severe disabilities. They examined the use of choral response and instructive feedback (similar to instruction of the non-target goals discussed above) to understand if participants would benefit from these embedded learning methods. Learning opportunities were embedded in transition times. The authors concluded that these methods were effective with 5 of 6 children without disabilities, but not effective for any of the students with disabilities until the modification of individual embedding with regular review trials were used. The authors also discovered that for one child with "oppositional behavior" (p.85), embedded learning with individual modification was not effective. This indicates that care must be taken to match methods with individual needs.

Johnson, McDonnell, Holzwarth, and Hunter (2004) address the gap in understanding how to embed instruction of IEP goals for post-early-childhood students in inclusive environments. They examined the goal embedding process for three elementary school students with moderate to severe disabilities, learning in inclusive environments. The embedding methods they used "included constant time delay, error correction, and social reinforcement" (p.219). Most of the embedded learning opportunities were conducted during activity transition times. The authors describe teacher perceptions of embedding as a component of student learning success in this study: "These professionals perceived embedded instruction as an effective approach in meeting the educational needs of students and compatible with the organizational structures of typical general education classes" (p.224). A particularly interesting finding in this study was that:

Anecdotal evidence also suggests that embedded instruction had secondary

benefits for each of the students participating in the study. For example, the

researchers observed that Wendy began to raise her hand in response to the probe

questions presented by the teacher to the entire class following embedded

instruction in each science unit. (p.224)

Additionally, "Chuck was the first student to correctly identify a word when it was

presented to the entire class" (p.225), which led to spontaneous communication acts with

his peers, and Brenda "began to spontaneously use her communication device to request

"help" and "break" throughout the day" (p.225), which was a demonstration of the

generalization possible when goals are embedded. It is likely that other environments in

which IEP goals are embedded could have similar unforeseen positive results.

Johnson et al. (2004) also commented on the teacher ability to provide multiple

opportunities for embedded goal practice throughout the instructional day, and cited this

as a likely reason for successful learning outcomes. The researchers also issued a caution

about embedded instruction- it may not be appropriately effective for IEP goals that

include "complex chains of behavior" (p.226); nevertheless, in this instance, the

embedded discrete skill learning provided the opportunity for child initiation of practice

of more complex social behaviors.

Polychronis, McDonnell, Johnson, Riesen, and Jameson (2004) examined the

efficiency of prompting schedules for embedding instruction of IEP goals within general

education environments in elementary schools. The authors concluded that although

"embedded instruction allowed students to participate more fully in the general education

curriculum under both trial distribution schedules" (p.147), embedding within a 30-

minute schedule was more effective than embedding within a 120 minute schedule. The authors also noted that "the teachers did not perceive the embedded instruction as a disruption to the educational programs of students without disabilities in the class" (p.147) and that teacher perceptions of the embedding process were generally positive.

Embedding IEP goals also includes consideration of goals in the social domain. Smoot (2004) discusses using nominal sociometry to assess for social goal needs and attainment. She used this process to assess whether students who had only recently begun to be included in general education classes were becoming friends with the other students in these classes. This is an interesting concept because of the difficulty in measuring social relationships. Sociograms are one method to measure relationships that could be described as embedded assessment. Although the author does not specifically make this link, she does comment on its applications in class and briefly review the literature on sociometric strategies. Macy and Bricker (2007) also address embedding social skills expressed in IEP goals. They conclude that embedded opportunities for social skills practice can make inclusive general education environments more successful options for students with disabilities.

In a publication for parents, Hammond, Casteneda, and Ortega (2006) suggest a process for including a student's IEP goals in home activities; although not specifically described as such, it is a process of embedding goals into the student's home environment in a natural way. Parents are encouraged to chart differing home environments, identify during which times goals would naturally be addressed, and assess progress using the chart. This is similar to other at-a-glance systems, with the addition of an assessment component.

Cook, Rule, and Mariger (2003) also suggest that for young children, embedding IEP goals in home settings can be effective. They explain this idea with statements such as, "any time children and adults are together, there are opportunities for learning" (p.21), rather than by use of the term embedding. They examined parent learning by having parents evaluate a web site to teach the strategies of this process- the Strategies for Preschool Intervention in Everyday Settings (SPIES) web site. The authors conducted this study by using an electronic demographic survey and a questionnaire for each of the six sections of the study. Evaluation of the data collected led the authors to conclude that parents were able to use the site to gain information to more effectively embed IEP goal learning in naturally occurring home activities, although some changes to the site could improve parent learning.

Trent, Kaiser, and Wolery (2005) examined the use of embedding goals in home play environments in cooperation with siblings of students with moderate to severe disabilities. The participants were elementary school sibling dyads, with the older participant being described as typically developing and the younger participant described as having Down syndrome. Older siblings were taught to use nonverbal mirroring and verbal responses to equalize and invite communication opportunities for younger siblings. Effects of this procedure were positive; the students with Down syndrome learned to initiate verbal interactions during play activities. One conclusion reached was that "teaching typically developing siblings to use responsive interaction strategies may be one way to support children with disabilities without placing the former in a teaching role" (p.117).

Embedded assessment of IEP goals has received much less attention than embedded instruction. Keenan-Takagi (2000) discusses embedded assessment from a general education perspective in a chorus class, and does not mention IEP goals; however, her comments are potentially useful in understanding how IEP goal assessment may be embedded. Keenan explains that embedded assessment must be quick, clean, simple, and based on what was practiced. She suggests using photo evidence as one method. She also suggests that teachers make an assessment bank of their own assessment methods from which to choose to address assessment needs throughout the year.

Bricker and Gumerlock (1988) describe three levels of assessment appropriate for IEP goal development and success. First, norm-referenced assessment can be used to understand the overall picture of student skills and point to direction of concerns. This is of course assuming that assessments are normed appropriately for the students being assessed. Next, criterion- referenced or curriculum-based assessments are used because such a test "provides the content for developing individual IEPs for participating children, and it provides quarterly feedback to staff on children's progress toward the acquisition of the IEP goals and objectives" (p.80). Finally, weekly monitoring of goals is needed. The authors discuss this in the context of short term objectives (these were still included in standard IEPs at this time), but even without short term objectives written into IEPs, the need for assessment in the short-term is still clearly needed, as is explained when the authors state "the collection of such information permits the systematic evaluation of children's progress and a mechanism for introducing modification in programming efforts in a timely manner" (p.80).

Teacher embedding practices. Although there are several studies that explain effective ways to embed IEP goals into inclusive environments, there are few with detailed examples on how to actually do this outside of the context of a study. Physical education instructors seem to have a better grasp on this than many other content areas.

Yun, Shapiro, and Kennedy (2000) provide a brief review of literature that discusses how to embed IEP goals in physical education classes. The authors explain that modern physical education curriculum often make use of multi-activity models which are short units of study that "teach sport-specific skills across the physical, cognitive, and affective learning domains" (p.34). The authors give three recommendations:

> To achieve IEP goals in general physical education classes that use such a model,
> teachers must (1) consider different ways of organizing the general physical
> education curriculum, (2) develop appropriate annual goals and short-term
> objectives in light of this curriculum, and (3) communicate with the IEP team.
> (p.34)

They also cite Graham, Holt-Hale, and Parker (1998) to explain that motor skill growth of general education students as well as for students with IEPs is often limited in these traditional games-based physical education curricula. Switching to a skill-based program that has a circular teaching pattern (my words) could be a way to embed IEP goals in a manner that provides benefit to students with particular physical goals as well as for students who do not receive special education services.

Kowalski, Lieberman, Pucci, and Mulawka (2005) discuss curricular individualization for all students: "Whether instructing the class using a partner/group activity, station work, relay race, obstacle course, lead-up game, or drill, good physical

educators are constantly modifying and changing activities to meet student goals and objectives" (para. 13). These authors also refer to embedding goals as "infusing the student's objectives" (para. 26) and provide a table of relevant advice for how to effectively accomplish this:

1. Do not exclude the student with a disability from the class activity just to work on goals and objectives. Having an IEP does not mean that the student's objectives are the only skills he or she can or should work on.

2. Objectives do not need to be covered all at once. If you can fit only one or two objectives into each unit, that is fine. Objectives do not need to be covered in every class period. Incorporate where appropriate.

3. If you must pull the student out to work on a specific objective, make sure you also have same-age peers with him or her so as not to exclude the student.

4. Utilize teacher assistants and paraprofessionals when working on goals and objectives in general physical education.

5. The goal of the instructor is to improve on objectives and to try to reach the STOs of the student. Once this is done, the STOs can be increased, or the student can work on something else.

6. Do not be afraid to ask the adapted physical education specialist, classroom teacher, or special education teacher if you have questions about the student's IEP.

7. If the student has a physical or occupational therapist, this person is a great resource and can provide ideas for implementing goals and objectives into games and activities.

8. Brag! When you are successful in implementing a students objectives into class activities, be sure to share it with the interdisciplinary team members. (Table 1)

This is a detailed list, and this detail demonstrates how much more clearly physical education teachers understand and can explain embedded instruction of IEP goals. An additional relevant point made by Yun et al. (2000) is that it is likely that physical education teachers are not regularly part of a student's IEP team and goals may be set without consultation. Kowalski et al. (2005) and Kowalski et al. (2009) discuss this problem in the context of the expectations of direct service providers. All of the authors emphasize that for success, PE teachers must take an active role in IEP goal-setting. Yun et al. also clearly explain the positive results of such involvement: "When physical educators carefully develop activity-based short-term objectives in accordance with the general curriculum, they provide a clear means of integrating students with disabilities into general physical education classes and helping them achieve IEP goals" (p.36).

Aside from the physical education context, I was unable to find specifics about how teachers do or do not embed instruction and assessment of IEP goals. This is concerning. One possibility for this is that embedding is discussed using different terminology across different content areas. Another possibility is that different content areas do not have a collective understanding of how to embed; possibly, pockets of embedding occur without generalization to content areas as a whole. One tool that may be used to embed instruction is a program-at-a-glance. A program-at-a-glance lists a student's IEP goals and the content areas / classes in which these goals will be addressed. Programs-at-a-glance are good tools to begin embedding, and certainly increase goal

understanding among special education teachers and general education teachers, but they do not generally specify the manner in which goals will be embedded and do not often explain the manner in which goals will be assessed. Perhaps an instructional-methods-and-assessments-at-a-glance could be used to facilitate the embedding process.

Factors that affect embedding IEP goals in instruction and assessment. Many factors affect how IEP goals are embedded in instruction and assessment. A list and description of some factors indicates some gaps in knowledge about what causes high-quality embedding of instruction and assessment to occur.

Location of services. In general, the location of services in the literature I examined is in inclusive general education environments (Coyne et al., 2007; Daugherty, et al., 2001; Horn et al., 2000; Jameson et. al., 2007; Johnson et al., 2004; Macy & Bricker, 2007; McBride & Schwartz, 2003; McDonnell et al., 2003; Polychronis et al., 2004; Riesen et al., 2003; Smoot, 2004; VanDerHeyden et al., 2005; Wertz et al., 1996; Wolery & Anthony, 1997; Wolery, Anthony, et al., 2002). Only one study, Sigafoos et al., 2006, considered embedding instruction in the context of a self-contained classroom. Two studies addressed services embedded at home by parents or siblings (Cook et al., 2003; Trent et al., 2005).

School schedules. VanDerHeyden et al. (2005) discuss the concern that although the literature supports using "embedded instruction approaches for targeted skill acquisition, generalization, and maintenance" (p.81), it is debatable whether or not there are enough opportunities for learning using this system and whether such learning is systematically applied. The issues of school schedules that allow enough opportunities and the systematicity needed must be addressed in order to ensure the robustness and

efficiency of embedding IEP objective instruction in the general education environment. If the teachers of a student communicate effectively about how to embed his or her IEP goals across environments, it is likely that embedding will occur; however, it takes a certain amount of educator stamina to overcome the entropy inherent in educational systems and to provide the consistency needed for well-embedded and appropriately assessed goals.

    Disability category. There were several disability category descriptions, some of which overlap. Several studies included participants who were assigned to different disability categories:

(a) Developmental delays (Daugherty et al, 2001; Horn et al., 2000; Macy & Bricker 2007; McDonnell et al., 2003; Polychronis et al., 2004; VanDerHeyden et al., 2005; Wertz et al., 1996; Wolery, Anthony, et al., 2002)

(b) Down syndrome (Jameson et al., 2007; McBride & Schwartz, 2003; Polychronis et al., 2004; Trent et al., 2005; Wertz et al., 1996; Wolery & Anthony, 1997)

(c) Intellectual disabilities (Jameson et al., 2007; Johnson et al., 2004, McDonnell et al., 2003; Riesen et al., 2003; Smoot, 2004; Wolery & Anthony, 1997)

(d) Learning Disabilities (Coyne et al., 2007)

(e) Behavior Disorder: (Wolery, Anthony, et al., 2002)

(f) ADHD: (Wolery, Anthony, et al., 2002)

(g) Cerebral Palsy (Horn et al., 2000)

(h) Autism (Johnson et al., 2004; McBride & Schwartz, 2003; McDonnell et al., 2003; Polychronis et al., 2004; Riesen et al., 2003; Sigafoos et al., 2006;

Wertz et al., 1996)

(i) Multiple disabilities (McDonnell et al., 2003; Polychronis et al., 2004; Riesen
et al., 2003).

This data demonstrates an extreme lack of study in embedding instruction for
students with learning disabilities, behavior disorders, ADHD, and cerebral palsy, and
minimal research in embedding instruction for students with multiple disabilities.

Student education level. There is a pattern of decreased study in embedding IEP
goals as students move upward in their education: (a) preschool/early childhood
(Daugherty et al., 2001; Horn et al., 2000; Macy & Bricker, 2007; McBride & Schwartz,
2003; Smoot, 2004; VanDerHeyden et al., 2005), (b) elementary school (Coyne et al.,
2007; Johnson et al., 2004; McDonnell et al., 2003; Polychronis et al., 2004; Trent et al.,
2005; Wertz et al., 1996; Wolery & Anthony, 1997), and (c) middle school (Jameson et
al., 2007; Riesen et al., 2003; Sigafoos et al., 2006). I found no data for high school or
post-secondary education. One possible reason for this is decreased focus on inclusive
experiences as students age.

Teacher attitudes/dispositions. Studies were conducted using teacher participants
who had positive attitudes about the possibility of embedding instruction and who
demonstrated positive dispositions at the study conclusion. Although some studies used
interviews to describe the social validity of findings, this was not a focus of the literature
reviewed, and there was not enough information to explain this in the studies I reviewed.

Teacher preparation. Particularly in early childhood education, teacher
preparation was minimal and quite varied. Even experienced teachers lacked specific
preparation in how to embed instruction. Studies noted teacher experience but were not

specific about what teachers knew about inclusive practices and embedded instruction. More information is needed in this area.

  Caseload management models. I was unable to locate specific information about caseload management models. It was not discussed in the literature I reviewed, but it is possible more information could be found by expanding the search to broader special education topics.

<div align="center">

Conclusions and Future Research

</div>

  There is extensive past and current research about the theory and practice of assessment. This research includes examination of assessment definitions, descriptions, and methods and answers questions about best assessment practices, particularly highlighting the need for purposefulness and systematicity in assessment (NRC, 2008). Based on these concepts, legislation and court decisions delineate the assessment rights of students who receive special education services (Cushing et al., 2009; Deisinger, 2007; Gartin & Murdick, 2008; Kaufman, 2008; Levine & Wexler, 1981; Weishaar, 2008; Yell, 2006). School districts can be held liable for failing to uphold these rights by fulfilling assessment requirements. To preserve all students' property rights to education, teachers must therefore be aware of assessment requirements and skillfully use appropriate assessment practices.

  Although the need for this understanding is clear, the links among scholarly knowledge of assessment best practices, teacher knowledge of assessment best practices, and actual teacher practice have not been adequately explained in assessment research. Research only partially considers the manner by which teacher education programs promote assessment knowledge in students. Current research suggests that teacher

preparation programs do not sufficiently teach preservice teachers how to effectively use assessment practices (Begeny & Martens, 2006; Campbell & Collins, 2007; Campbell & Evans, 2000; Conderman et al., 2001; Conderman et al., 2005; Dodd & Scheetz, 2003; Jenkins et al., 2002; McCombes-Tolis & Feinn, 2008; McNicholas, 2002; Mertler, 1999; Mertler, 2005; Smith & Gorard, 2005) and does not adequately answer the questions of whether preservice teachers understand assessment theory or why this understanding may or may not be translated into assessment practice.

Many methods of assessment are relevant to students who receive special education services. Of particular importance is embedding assessment of IEP goals into instruction. This area of continuing research is essential for teachers who provide special education services because it unites an appropriate method with the legal and natural rights contained in IEPs. There are gaps in this research that appear to be based on many factors: location of special education services, school-day schedules, type of disability, student grade level, teacher attitudes, and teacher preparedness. It is important for future research to address incomplete areas, such as embedding assessment of IEP goals for students with multiple disabilities or with specific learning disabilities.

It is necessary to gain a more thorough understanding of how teachers understand and implement assessment practices. Using appropriate qualitative interviewing methods can provide the environment in which new teachers can comfortably discuss their current understandings, how they came to these understandings, and how these understandings influence their practice.

CHAPTER 3

Methods

Conceptual Framework and Positionality

I am assuming a relativistic ontology in this research. The data collected will therefore be viewed as potentially explaining one aspect of a complex reality that remains relative and situational. Analyzing data will yield a picture of assessment preparation for recent graduates of the Special Education Dual License Program (SEDLP).

Epistemologically, I am using social-constructivist assumptions. These assumptions are made both about the nature of disability and the research methods used. Regarding disability, I assume that medical models and deficit models of assessment are not sensitive enough to measure growth and are dehumanizing, particularly since they have been used historically to dehumanize students with special education needs. I assume that the assessments conducted by recent SEDLP graduates are capturing one aspect of a student's skill at a particular time and in particular circumstances, and that altering environmental components will construct a different result. I further assume that self-assessment of recent SEDLP graduates' assessment skills requires similar assumptions about these graduates.

An additional epistemological concern comes from critical theoretical perspectives, including that of disability studies. I assume that individuals with disabilities are allowed limited power in a variety of situations, and that this affects day-to-day existence for such individuals. Assessment is one method in which power may be exercised against individuals with disabilities, so care must be taken that assessment is

conducted in a manner that will benefit those assessed. I assume that carefully

constructed, respectful assessment is also part of the mandate of the SEDLP program.

These assumptions recommend a research design composed of semi-structured

interviews, in which the rich context of recent SEDLP graduates' learning can be

explored for multiple individuals from their perspectives and such individual

understandings can be combined to form a larger snapshot of assessment perspectives

within a teacher preparation program. The expected result would be an understanding of

the situation in which SEDLP graduates gain or do not gain adequate understanding of

assessment knowledge and of their ability to implement this understanding in assessment

practice.

Interviewing as a Data Collection Technique

Interviewing techniques are particularly relevant to qualitative research designs.

Creswell (2007) explains this idea when he explains that interviews are considered

important for many types of qualitative research, including narrative study,

phenomenology, ethnography, grounded theory, and case study. He provides several

suggestions for interviews: unstructured interviews, semistructured interviews, focus-

group interviews, e-mail interviews, and telephone interviews. Recording methods may

include writing notes and audiotaping and transcribing notes. Glesne (2006) suggests

audiotaping and (citing Merriam, 1998) creating an interview log in order to decrease the

time requirements of full transcription and focus on relevant data. McMillan and

Schumacher (2006) also describe several interview approaches: informal conversation

interview, interview guide approach, standardized open-ended interview, key informant

interviews, career and life history interviews, elite interviews, and phenomenological

interviews. Interview methods must be chosen to match the interview purpose.

A definition of interviewing comes from Glesne (2006), where she describes

interviewing as the situation in which:

> researchers ask questions in the context of purposes generally important primarily
>
> to themselves. Respondents answer questions in the context of dispositions
>
> (motives, values, concerns, needs) that researchers need to unravel in order to
>
> make sense out of the words that their questions generate. (p.79).

Despite the self-serving nature presented in this construct, it is possible within this

description to include situations in which interviewees have a strong interest in sharing

the information requested in the interview. The process ideally is mutually beneficial.

Creswell (2007) identifies several procedures for interviewing: "identify

interviewees," "determine the most effective type of interview," "use adequate recording

procedures," "design and use an interview protocol form," "refine interview questions

and procedures through pilot testing," "determine the place for conducting the interview,"

"obtain consent from the interviewee to participate in the study" (pp.132-134), and follow

appropriate interview etiquette such as completing the interview in the scheduled time,

being respectful, and listening rather than speaking whenever possible.

Creswell (2007) places particular emphasis on creating an interview protocol.

This creates a well organized data collection device. Creswell recommends particular

strategies for using an interview protocol, such as memorizing the questions and their

order so that eye contact with the interviewee can be appropriately maintained. Another

Creswell suggestion is to "provide appropriate verbal transitions from one question to the

next" (p.135), which can be planned in advance. Additionally, as Creswell explains, prewritten closing comments will ensure that appropriate closure is reached, and of course, the researcher should thank the interviewee.

Glesne (2006) particularly describes techniques to develop appropriate questions. She cites Patton (2002) to describe question types as "experience/behavior questions, opinion/values questions, feeling questions, knowledge questions, sensory questions, and background/ demographic questions" (p.82). Additionally (citing Patton, 2002), Glesne explains that questions provide more in-depth data when asked about the past and present than when asked about the future, and that pre-supposition questions (as opposed to leading questions) can generate valuable data. Further questioning techniques she discusses (citing Clark, 1999) include grand-tour questions (big picture questions), and scene setting/mood provoking questions. Maxwell (2005) adds to the understanding of interview questions with a discussion of how to anticipate whether or not planned interview questions will yield needed information.

Glesne (2006) also describes some interview difficulties. As she explains, technical (recording) difficulties are usually the easiest to solve. A more difficult problem is when interviewees do not answer the question asked. Glesne suggests that off-topic talking that is a result of question-avoidance can be prevented by explaining at the beginning of the interview that the interviewer purpose is to gain information about the specific topic rather than to make judgments about participants. Glesne also explains that interviewees may talk too little or too much for the purposes of the researcher. The researcher's job is to then elicit more information in a way that is acceptable to the interviewee. Creswell (2007) more generally explains interview challenges:

Many inexperienced researchers express surprise at the difficulty of conducting

interviews and the lengthy process involved in transcribing audiotapes from

interviews. In addition, in phenomenological interviews, asking appropriate

questions and relying on participants to discuss the meaning of their experiences

require patience and skill on the part of the researcher. (p.140).

Additionally, Creswell cites several researchers (Kvale, 2006; Nunkoosing, 2005; Weis

& Fine, 2000) to explain that "recent discussions about qualitative interviewing highlight

the importance of reflecting about the relationship that exists between the interviewer and

interviewee" (p.140). He concludes that it is important to be sensitive to what is occurring

during the interview process. Glesne (2006) describes this sensitivity as building and

maintaining rapport.

In addition to this sensitivity, skilled interviewers anticipate possible needs and

methods, put aside personal assumptions to understand only what the interviewee is

communicating, listen analytically to determine areas in which more questions should be

asked, indicate listening without directing opinion, and use patience throughout the

interview process (Glesne, 2006). Other qualities Glesne uses to describe skilled

interviewers include "nonthreatening" (p.98), "aware of power and hierarchy" (p.98), and

"caring and grateful" (p.99). Both Glesne and Bogdan and Biklen (2005) explain the need

for the researcher to approach the participant as an expert on the interview topic, and to

understand that the researcher is in the position of a learner. Corbin and Strauss (2008)

add that skillful interviewing involves not just asking questions, but using facial

expressions and silent pauses effectively. Interviewing, as they explain, requires skill

built by practice.

Corbin and Strauss (2008), citing Corbin and Morse (2003), explain that unstructured interviews are "the most data dense" (p.27) and therefore provide the most information. Glesne (2006) contradicts Corbin and Strauss somewhat by explaining that rich data is secured by interview questions that are structured, open, and depth probing. This contradiction can be explained by Glesne's focus on prepared probing, which Corbin and Strauss discuss in less detail. Bogdan and Biklen (2007) add to this discussion with the idea that interview schedules (more structured approaches) are useful, but adhering to them strictly is a waste of qualitative resources, particularly if appropriate probe questions are not used to help participants elaborate. These authors suggest the interview schedule as an appropriate beginning, but that researchers need to keep focused on thoughtful listening and using this focus to ask probe questions and therefore to generate as much topically appropriate information as possible.

The time required for an effective interview varies based on the research purpose and the comfort level of participants (Glesne, 2006), but "an hour of steady talk is generally an appropriate length before diminishing returns set in for both parties" (p.88). Glesne continues to explain that although the number of interviews varies by the nature of the research, at least two interviews are usually required "to obtain trustworthy results" (p.89).

Corbin and Strauss (2008) also explain interview ethics. Interviewers must follow Institutional Review Board guidelines, treat interviewees respectfully, respect confidentiality, and refrain from judgment. The authors explain that "Sometimes a researcher feels uncomfortable or awkward with interview material or something that is observed. However, participants are not. In fact, they want their stories out there" (p.29).

A respectful interviewer neither censors the speech of those who want others to know their stories nor shares the stories of those who do not want them shared.

Some limitations in using interviews are expressed by Corbin and Strauss (2008). One difficulty that contributes to the limitations of interviewing as a technique is that "persons may not be consciously aware of, or be able to articulate, the subtleties of what goes on in interactions between themselves and others" (p.30). For this reason, the authors recommend using additional methods with interviewing, such as observation, and using these other sources for data triangulation. Triangulating data is an appropriate strategy for most types of research.

For dissertation purposes, I am using a series of two interview sessions with each participant. Using two interviews allows the possibility of elaboration in areas discovered throughout the interview process and across participants. It allows analysis and coding between interviews in order to better understand what assessment issues are of importance to teachers and therefore allows asking about issues I may not have considered. It also allows for necessary time to build some level of rapport with participants. For this reason, I also am using forty-five minute to hour long interviews. I believe this will be enough time to build the interviewer/interviewee relationship to the extent that useful information about interviewee experiences will be shared, while still remaining respectful of participants' time.

I am using a semi-structured approach. Although some researchers advocate for a less structured format, I do not believe I can, as a novice interviewer, adequately conduct interviews without some sort of concrete guide. I am using an interview protocol, such as that described by Creswell (2007), to ensure that I appropriately introduce and close

interview sessions and ask a certain set of questions. As interviewees respond, I plan to listen thoughtfully and use probe questions for elaboration. To assist in this process, I am (with interviewee consent) audio recording interviews for transcription and later to code content. Although I recognize the advantages of creating interview logs such as those described by Glesne (2006) (these logs extract relevant information while reducing transcription time between two interviews in a series), I am fully transcribing interviews because of my position as a novice researcher.

It is also important to consider the effects of interview transcription choices on the manner in which meaning is constructed. If transcription choices are not explained, it can appear that the researcher is assuming a positivistic perspective in which what is written is presented as truth without acknowledgement of choices that have constructed the truth presented (Lapadat & Lindsay, 1998). Ochs (1979) explained the direct impact of these choices in the context of describing children's language. She emphasized the importance of selecting methods that provide the desired information, calling this "selectivity" (p.44). Selectivity is important because it allows focus on particular aspects of the data. Selectivity for the adults who will be participating in this study includes the choice to only transcribe the words and sounds that are spoken and to ignore environmental aspects.

Lapadat and Lindsay (1998), in a review of transcription methodology, continue the idea of matching transcription to purpose in the explanation that "the impetus has shifted away from establishing one standardized set of transcription conventions" (p.6). Davidson (2009) reviews transcription literature from 1979 to 2009. She references the work of Bucholtz (2000) and Oliver et al. (2005) to explain a continuum of naturalized

and denaturalized transcription. For this study I am using a more denaturalized

transcription, which "preserves the features of oral language" (Davidson, 2009, p.39). I

am transcribing everything that is spoken, including "ums" and other similar utterances,

but ignoring pauses and involuntary noises such as coughing. An advantage of using this

method is that a record is preserved of the entire spoken content, allowing greater

exploration of meanings without cutting out material that may have differing

interpretation possibilities. Additionally, some participants are likely to use a larger

variety of non-word speech than others and may use this type of speech as a way to

process or refine their thoughts and it could be interesting to see this process. A possible

disadvantage of this approach is that the transcripts may appear more cluttered and may

present the image of participants as being less verbally skilled or knowledgeable than

they actually are based on apparent disfluency. This disadvantage can be countered by

preserving awareness of the purposes in utterances and viewing them as participant

attempts to clarify their thoughts and tell their complete views.

Interview techniques that are ethical are essential. I am preserving interviewee

confidentiality by using pseudonyms, removing individually identifying information

during transcriptions, and sharing audiotapes only as required by specific data collection

questions directed to my advisor.

Research Questions

This is an interview study that uses qualitative methodology and includes two

semi-structured interviews for each participant and the option of a follow-up e-mail, in-

person communication, or focus group participation. This design is appropriate because

of the research goal of understanding the case in which multiple participants have gained

some level of assessment understanding and proficiency.

This study addresses three research questions:

1. In what ways do recent SEDLP graduates characterize their level of

   competence (theoretical understanding and practical application) in assessing

   the progress of students with disabilities in the classroom?

2. In what ways do SEDLP graduates report they use classroom assessment to

   inform classroom instruction?

3. What features of the SEDLP teacher preparation program do graduates identify

   as having positively or negatively impacted their ability to effectively use

   classroom assessments?

To answer these questions, two semi-structured interviews of approximately

forty-five minutes were used with each participant. Interviews were audio-recorded with

the consent of participants and fully transcribed. Participants were offered an e-mail, in-

person, or focus group summation and discussion of the data they provided.

Terms

Assessment can be defined using the Ysseldyke and Algozzine (2006) definition:

"Assessment is a process of collecting data for the purpose of making decisions about

students" (p.74). For the purposes of this study, "assessment" refers to data collected by

classroom teachers for the purpose of making instructional decisions. This data may have

been collected using a variety of methods. For this study, assessment does not consider

data collected by specialists for the purposes of determining eligibility for special

education services or data collected *solely* for the purposes of monitoring progress toward

Individualized Education Plan (IEP) goals.

Setting

Research was exclusively in the context of the University of New Mexico Special

Education Dual License Program (SEDLP), which is an undergraduate program in the

Department of Educational Specialties in the College of Education. Approximately 30 –

40 students are enrolled in this program each year. In this program, students earn a

Bachelor's of Science in Education degree in elementary education and in special

education. Students are then eligible for preK-12 Special Education licensure and K-8

General Education licensure. Students in the SEDLP follow a four semester proscribed

sequence of classes (see Appendix A) which includes a full year of student teaching that

encompasses both special education and general education environments. The program

goal is to produce graduates who are able to "be effective across the full continuum of

educational settings for all students" (UNM College of Education, 2013). My relationship

to the SEDLP has included serving as a cooperating teacher for five preservice teachers

enrolled in the program. As a cooperating teacher, I gained familiarity with the SEDLP

program as it related to my work at the middle school level in a large southwestern public

school district. I was also part of the community of educators providing joint support to

the cohort of SEDLP preservice teachers who were experiencing student teaching at our

school.

IRB

This study was approved on September 12, 2012, using the University of New

Mexico's Institutional Review Board expedited review process. The research protocol

explained the research context, recruitment methods, methods for obtaining informed

consent, data collection procedures (two interviews and a follow-up e-mail, in-person

conversation, or focus group), and data protection procedures. The protocol was followed

throughout the study and is described below.

Participants

Participants in this study were drawn from SEDLP graduates who graduated in

the past one to two years. Participants were recruited by advertising on college of

education notice boards, and by email contact of recent graduates using the graduate

listserve. Ultimately, these methods resulted in seven e-mail contacts with potential

participants and yielded five participants. It is unclear why so few participants responded

to the advertising, but it is possible that this was partially a product of the intense

workload of new teachers. Additionally, it is expected that teachers continue their

educations to advance across salary levels, so new teachers may often be taking

additional coursework while also trying to provide high quality instruction in an

environment that is new to them while at the same time maintaining some sort of balance

with their personal family lives.

Before beginning interviews, I explained informed consent in-person using the

consent documents approved by the IRB. Participants signed consent forms before each

of two interviews. In one case, the participant had moved to a different state between the

first and second interview and provided informed consent for the second interview by

signing the form, scanning it and saving it as a PDF, and e-mailing it back to me.

All participants completed a brief demographic survey before the first interview

(Appendix B) and selected a pseudonym. All participants are current public school

teachers hired to provide special education services and employed in one of two different

school districts. All participants list this as the beginning of their teaching career with no

prior career in other fields. In Table 1, in the Grade/Content section, Elementary Gifted

refers to students who are identified as twice-exceptional, with both areas of giftedness

and disability. Inclusive service setting (for the purposes of this study) refers to settings

in which instruction is provided 80% or more of the day with general education peers.

Self-contained refers to settings in which instruction with general education peers is

provided less than 80% of the day. Although there are many levels of inclusion, the

instructional setting information provided by participants for this study clearly showed

only these two arrangements. Additionally, service setting refers to the "program" in

which the teacher was hired to provide services, and not individual student IEPs.

Table 1

Participant Demographic Survey Information

Participant	Ethnicity	Age	Graduated	Grade/Content	Service Setting
Carmen	Hispanic	18-25	2012	Middle Science	Self-Contained
Cathy	Caucasian	26-36	2012	Middle Science	Inclusive
Dawn	Caucasian	26-36	2011	Elementary 1-3	Self-Contained
Lynn	Caucasian	18-25	2012	Preschool	Self-Contained
Shelly	Caucasian	18-25	2011	Elementary Gifted	Self-Contained

Carmen identified herself as a Hispanic female between the ages of 18 and 25 who had graduated from the SEDLP in 2012. She is currently teaching middle school science in a large southwestern school district to students receiving special education services in a segregated environment. According to her interview, her school has experienced a great deal of staff turn-over and prior to her hire students had had a series of long term substitutes who were not licensed to teach science. Carmen's school is located in an area of town identified with a high degree of economic need and she reports that her students are primarily Hispanic and include many who are classified as English Language Learners. Carmen reports that although she is not a fluent Spanish speaker she is able to understand most of what her students say when they use Spanish. Carmen also stated that she is currently enrolled in coursework with the goal of eventually pursuing a degree in law.

Cathy described herself as Caucasian and between 26 and 36 years old. At the beginning of this study she identified herself as teaching inclusive science at a middle school in an area of economic need in a large southwestern school district. She had also provided support in one section of inclusive social studies. Part way through the study Cathy moved with her family to a northern state and was seeking employment teaching there.

Dawn described herself as Caucasian between 26 and 36 years old and having graduated in 2011. At the time of the study she was employed in a moderate sized southwestern school district in a segregated 1st to 3rd grade special education classroom in an area that she reports to be middle class. She stated that she was anticipating temporarily leaving the teaching profession to pursue family interests.

Lynn identified herself as Caucasian between 18 and 25 years old and graduated in 2012. She is teaching in a segregated preschool setting in a large southwestern school district in a suburban area that she reports to be middle class.

Shelly also described herself as Caucasian between 18 and 25 years old and graduated in 2011. She is teaching in a large southwestern school district in a middle class area of town and teaches a self-contained class of elementary students identified as twice exceptional, meaning that they have both identified areas of giftedness and disability. Shelly stated that because she does not have gifted licensure she is taking coursework in this area with the goal of becoming licensed.

Participant confidentiality was maintained by using the pseudonyms provided by participants and removing any identifying information during transcription. It will be further maintained by destroying audio recordings after two years.

Data Collection and Analysis

Prior to beginning the first interview, and after IRB approval, each participant provided informed consent as described above and completed the demographic survey. First and second interviews were conducted between October 30, 2012 and August 28, 2013. All interviews were recorded using a digital voice recorder and a back-up micro-cassette recorder. Interviews were conducted in-person in three types of locations based on participant preference: the homes of participants, the work-places of participants during hours of the day in which they were not engaged with students, and my home. Two second round interviews were conducted over the telephone to accommodate one participant who had moved out of state and one participant who was otherwise unable to schedule a second round interview. After each interview, transcription was conducted by

transferring the audio-recording from the digital voice recorder to the computer using the USB port and then typing word for word using repeated listening and checking for accuracy by reading the entire transcript as the recording played and correcting errors. During this process I removed any potentially individually identifying information, particularly that which named a specific school or school district. After transcription was completed, I destroyed the micro-cassette recordings and deleted the digital-voice recordings. I will retain the audio-recordings as protected computer files for two years.

First round interview questions included a Grand Tour question (Spradley, 1979) to gain an overview of participants' understanding of assessment without limiting responses based on researcher expectations. The remaining questions are intended to seek more data but also seek to gain this data without limiting responses in any predetermined way other than for basic relevance. First round interview questions included:

1. Please tell me how you define assessment.

2. In what ways have you been able to use assessment in your classroom?

3. Please describe what you do to determine if students with disabilities are learning.

4. Please describe any aspects of classroom assessment that you believe you do well or that you believe you do not do well.

5. Please share any elements of the teacher preparation program or other experiences that have influenced you in the area of classroom assessment.

6. What do you wish you'd learned more about in the area of classroom assessment?

I did not begin to analyze data until all first round interviews and transcriptions were completed (Seidman, 2006). I began first round data analysis by printing transcripts on a different color of paper for each participant (Appendix C). This color coding linked responses with participants and was intended to aid in analysis by making it easier to determine which eventual emergent themes were consistent across participants and which were only addressed by some participants. I repeatedly read transcripts and marked them with hash marks to break participant responses into units of thought or "*units* of information" (Lincoln & Guba, 1985, p.344). Lincoln and Guba explain the difficulty in creating an operational definition of these units, but are clear that these units should be "heuristic, that is, aimed at some understanding or some action that the inquirer needs to have or to take," and "the smallest piece of information that can stand by itself" (p.345). After I had marked the transcripts, I cut out each unit of thought and glued it to an eight by five index card, with one card for each unit of thought.

Once separation into units of thought was complete, I began coding the data for the content presented in each unit of thought. Then, I physically sorted the index cards into categories that addressed similar content. I created an initial categorization key and labeled the back of each card for later use (Appendix D). Throughout this analysis, I used a constant comparative process described by Lincoln and Guba (1985) as an extension of work by Glaser and Strauss (1967). As described by these authors, the constant comparative process is flexible and varied based on the data and includes looking for repeated words and concepts and continually examining differences, similarities, and other relationships among potential categories. I used this process to identify areas in the data which needed further exploration to reach conceptual saturation (Corbin & Strauss,

2008), and formed second round interview questions from this process (Seidman, 2006).

Second round interview questions included:

1. Please tell me about challenges you have had when creating your own informal

   assessments.

   -How do you know in informal assessment whether or not learning has

    occurred?

   -What are your next steps after oral (verbal) assessment?

2. Please discuss any assessments you do of social or behavioral issues.

   -What are your next steps after these assessments?

3. Can you give some advice to new teachers about how to develop or implement

   assessments that will meet the needs of all learners?

Before engaging in second round interviews, I wrote an analytic memo to consider the implementation of interview best practices. This process is explained by Maxwell (2005): "Memos can be written on methodological issues, ethics, personal reactions, or anything else" (p.12) and can be used as a method to aid deeper thinking about issues relevant to the research. In this memo I reconsidered some of the material from Seidman (2006), particularly regarding how to respond to participants (don't interrupt, limit continual positive acknowledgement, allow think time, etc.). Next, as a trustworthiness measure, I had a critical friend (listed in the IRB) sort the data cards into categories. I wrote an analytic memo describing this process and our reconciliation of categories. This critical friend, identified later as Critical Friend 1, is a doctoral student from the University of New Mexico Department of Educational Specialties. Concurrent with this study she successfully completed her doctoral comprehensive exams and is at

this time preparing to begin the first steps in her dissertation. As documented in the IRB, she has completed CITI training. She has also worked with the SEDLP program in the past in the context of working collaboratively with myself and others to support the success of students in this program during their student teaching at the middle school level.

Second round interviews followed the same procedure as the first round, including consent forms, interviews, transcriptions, breaking content into units of thought, and using a constant comparative process to analyze data. The goal was to create and continue categories that could be used to develop themes. Again, I met with Critical Friend 1 who sorted a random sample of 1/3 of the new cards to determine if new categories would emerge. We compared her categorization to mine, noted differences and similarities, and discussed the results until we developed a categorical consensus. Next, I met with a second critical friend (listed in the IRB). We discussed the categories and by consensus united categories into broader constructs from which to develop themes. I provided a sample of 100 randomly selected data cards and the critical friend sorted the cards into these categories. From this sort it was possible to obtain a percentage of agreement between my sorting and her sorting, as another measure of trustworthiness. The second critical friend (later identified as Critical Friend 2) is a faculty member in the University of New Mexico SEDLP program.

Finally, I conducted a member check in order to ensure that the data obtained and used in analysis was consistent with the intent of participants in expressing their beliefs about assessment. As explained by Lincoln and Guba (1985), "If the investigator is able to purport that his or her reconstructions are recognizable to audience members as

adequate representations of their own (and multiple) realities, it is essential that they be given the opportunity to react to them" (p.314). All participants were offered the opportunity to react in three ways: to receive an e-mail summary of their ideas and respond via e-mail, to have an in-person follow-up meeting, or to participate in a focus group. Although the data that would most likely contribute to an understanding of trustworthiness would have been most likely to be obtained from a focus group, all participants chose an e-mail summary, perhaps because of the complicated schedules that result from raising families, teaching, taking courses (two participants), and the difficulty of traveling from another state (one participant). I sent each participant an e-mail summarizing the main points they presented in their interviews and requested feedback about whether or not the summary matched what they recalled saying and/or what they intended. Results of the member check are reported in chapter four.

CHAPTER 4

Results

This chapter is a summary of data supplied from transcriptions of a total of ten interviews provided by five participants who were graduates of the SEDLP program, conducted between October 30, 2012 and August 28, 2013. As described in Chapter 3, transcripts were broken into units of thought and placed one unit per index card, using color coding as a link to individual participants. Cards were first coded for meaning, then sorted into categories which were later united in broader categories and then used to determine themes relevant to the research questions. Constant comparative methods and collaboration with two critical friends were used to identify and test categories for theme development.

Categorization

Initial sorting included four categories with many responses and one category that encompassed fewer responses, labeled "other":

1) purposes of assessment

2) challenges in assessing

3) types of assessment

4) assessment best practice beliefs

5) other

    a) examples of competence / resolved challenges

    b) student affect (nerves, expectations, readiness, motivation)

    c) teacher affect (personal connections and experiences)

    d) assessment coursework

e) high stakes testing / standardized testing

Reexamination of categories with critical friend participation showed a basic congruence in categories. Critical Friend 1 categorized all of the round one data cards, allowing a comparison to the original investigator categories (Table 2).

Table 2

Category Congruence

Critical Friend	Investigator
What is assessment	Purposes Types
Mandated assessments	Standardized testing
Data collection in the classroom Practice as part of the learning process Teacher practice (pedagogical behavior) Student learning Differences for ages/grade levels/content	Assessment best practice beliefs
ELL learners Reading problems Insecurity about assessment	Continuing challenges
	Resolved challenges
Preservice/training	Assessment coursework
Teaching life skills/social support Accountability as a student trait	Student affect about assessment
	Participant personal connections to assessment

This comparison showed that all data cards were categorically accounted for with congruent though not exact categorization, with the exception of two areas: resolved challenges and participant personal connections to assessment.

After the second round of interviews, data cards were sorted by the investigator. No new categories were developed but coding additions were made in the "Type" category to include movement and tasks and in the "Best Practice" category to include collaboration. Critical Friend 1 was given a random sample of 1/3 of the cards and she resorted them. She identified one new category: Individualizing assessment. Creating a new category would highlight the importance of this topic, although it could be included as an area of congruence with "Assessment Best Practice Beliefs."

Next, collaboration with Critical Friend 2 was conducted to reorganize data categories into a more consistent, united system. After discussion, the prior categories were consolidated into four broader categories:

1) Purposes and Types of Assessment

2) Implementation of Assessment

3) Challenges in Assessing

4) Preparation for Assessment

A random sample of 100 cards was selected and sorted independently by both the investigator and the critical friend. Of these cards, 61% were placed in the same categories by both the investigator and the critical friend, with the greatest consistency in category four- preparation for assessment. After close examination of the disparately sorted cards, it was clear that the sorting discrepancies were greatest between category two and category three. Categorization in these areas depends on whether or not the

sorter views the card as describing a challenge or describing a participant practice. For

example, one card representative of this problem states:

> I was gonna do it with my sixth graders but they're not used to projects yet so I
>
> want to start them off with a mini project and then sometime next semester start
>
> them off with like a science project. (Carmen)

This participant has a strategy- a practice- in which she evaluates readiness to determine

what assessment to use. She also has a challenge in that she has determined that her

students are not ready for her planned assessment and therefore she must find another

way to assess. Although it would be satisfying as a researcher to have a higher percentage

of agreement about categories, reasonable conclusions can be made from either scenario

described above, and although greater consistency would indicate greater trustworthiness,

it seems important not to over-identify this agreement percentage with trustworthiness.

Themes

The four categories demonstrate themes and sub-themes that are defined and

illustrated by participant responses.

Purposes and types of assessment. Each participant gave a similar basic

definition of assessment, although some participants added greater detail: "I would define

assessment as getting a reading of the student's knowledge of any given subject area"

(Cathy), "being able to determine whether or not something that is being taught is being

learned," (Shelly), "a way to understand what students or individuals understand about a

topic" (Lynn), and "to see where they're at" (Carmen). These definitions encompass

concepts of assessment purpose such as determining a baseline skill level and using

formative assessment, although only two participants used the term 'baseline' and none

used the term 'formative' in a manner consistent with that used by the education field.

Regarding baselines, "they start the school year off with like two days of assessment so

getting a basis. Yeah, their baseline. So then…you know where they are." (Dawn).

Participants described using this baseline for instructional grouping and for sequencing

instruction: "like in reading groups that would change depending on where they are in

their reading" (Dawn), and "it has helped tremendously you know as a teacher to

understand where all they've come and how far I can take them" (Shelly). Additional

examples of participant definitions of formative assessment include: "trying to figure out

where they are, where we needed to go in small groups" (Dawn) and "once you know the

target area what they need to work on that's probably what you would work on more in

the classroom" (Shelly). Summative assessment as a concept was also captured in

descriptions of testing: "They always have an end of the…chapter test" (Carmen).

Another purpose explained across participants was "to see how much progress the

students made" (Lynn) in comparison to earlier scores, peers, and IEP goals.

        These definitions are consistent with the more formally stated Ysseldyke and

Algozzine (2006) definition of assessment as "a process of collecting data for the purpose

of making decisions about students" (p.74). When participants talk about "getting a

reading" (Cathy) "being able to determine" (Shelly) and "a way to understand" (Lynn)

they are talking about having a process of collecting data. The processes used by the

participants in this study are further explained in the next theme section. The "making

decisions" component of the Ysseldyke and Algozzine definition is also captured by

participants: "in reading groups this would change depending on where they are in their

reading" (Dawn), "what I do after assessment is I try to see okay who are the ones that

keep getting the A's, who are the ones that keep getting the F's…and I try to look at that and figure…okay what are they doing?" (Carmen), and, as above, "once you know the target area what they need to work on that's probably what you would work on more in the classroom" (Shelly).

These participant responses also address The National Research Council (NRC) (2008) description of assessment to include purposefulness and systematicity. Purposefulness was summarized by Lynn above when she states that assessment is "a way to understand what students or individuals understand about a topic" and the value of systematicity is described by Shelly: "assessment is…really rigorous and it is time consuming but it's well worth it 'cause I think you can really get to know your students on a deeper level." The assessment systems used by participants are further described in the next theme section.

One additional purpose of assessment was explained by Carmen: "for my own data just to see how I'm doing as a teacher." This was an interesting addition to the purposes of assessment because although it was made in the context of low-stakes, formative and summative classroom assessment, it matches the purpose of assessment as used by school systems in the larger climate of student high-stakes testing as a measure of teacher quality.

Participants each described an array of assessment types, including informal assessments, formal assessments, projects, verbal/oral, games, paper folds, check-ins, hands-on projects, observation, presentations, exit tickets, posters, journal entries, tests and quizzes, task completion, hidden purpose, and physical movements. In addition to these descriptions they specifically recommended using frequent informal assessment (4

participants), multiple choice tests (1), "special needs" versions of grade level content

tests (1), vocabulary quizzes (1), projects (2), task analysis (1), environmental assessment

(1), oral assessment (3), behavioral observation (3), adaptations of existing assessments

(3), checklists (1), and rubrics (2). One participant also recommended informally

embedding assessment in classroom assignments: when assessing, "I look largely at their

um assignments that they do in class because they always get different assignments that

are directly tied to you know whatever they are learning" (Cathy).

        Within this array of assessment types, there were some interesting areas of focus.

What was considered mandatory formal assessment varied by grade level. The two

middle school science teachers, Carmen and Cathy, both discussed the need to make

general education summative science assessments accessible to students with disabilities

and mentioned low student reading abilities as a barrier to accessing such assessments.

They seemed to consider this assessment format (written chapter or unit tests provided by

the textbook publisher) as mandatory. One interpretation of this is that it may

demonstrate their concern about the validity of other types of summative assessments;

they have the knowledge of other formal assessment types but believe that these types are

not real indicators of knowledge. A different interpretation is that this may be a reflection

of their understanding of the unequal power role of the special education teacher

compared to the general education teacher; the general education teacher provides a

particular test, so the special education teacher must provide access to the test rather than

proposing other types of assessments. A third interpretation of this is that it is a reflection

of a district perspective of assessment validity in which the participants of this study are

required to participate as part of their continuing employment.

At the preschool and elementary levels participants focused more on authentic assessments across the range of subjects taught as well as behavioral and social needs. Implementation of these types is discussed in the next section.

Implementation of assessment. Participants viewed a broad range of assessment strategies as effective or worth implementing as demonstrated by the variety of data in this area. It is noteworthy that lack of participant discussion of any particular strategy cannot be used to conclude that the strategy is viewed as inappropriate by the participant; rather, the participant did not comment on it.

Two strategies were addressed by every participant: provide for student affect needs and individualize and differentiate. Regarding affect, participants consistently mentioned student nerves, stress, or past negative experiences regarding assessment. These concerns were viewed as causing feelings not conducive to effective assessment, as demonstrated by Carmen and Shelly: "My answer is always you're not stupid, you need that extra time, just like I need" and "some kids need that processing time to reflect and think and um even if it's like you know a nerves thing I think they need the opportunity." Cathy explained the need to understand "what they actually know when they're not uncomfortable and…under the pressure of the test." These are all comments on the situation in which students who are receiving special education services experience negative feelings regarding testing. As students age, they end up with more and more negative testing experiences and greater need for grade level reading skills in order to do well on summative tests. As experienced test takers it appears that they know they are not very good at scoring well on these types of tests, which may create greater anxiety when testing. Although participants commented on the need to address this topic

for students, little was stated about how to resolve the issue of negative affect about testing beyond modeling positive talk to students. Carmen addressed it partially when explaining comments she makes to students about testing: "I tell them, you know we're trying to make you guys better….as teachers we're trying to do it because we want the best for them."

Participants explained a variety of ways in which they individualized assessments. Although individualizing assessment and differentiating assessment are different concepts, it appeared from interviews that participants were using the terms somewhat interchangeably. Greater probing would be needed to determine exact participant definitions, but for the purpose of this study, it is clear that all participants considered it important to broaden the scope of assessments in some way by either making individual accommodations and modifications or by differentiating. Cathy explained this as "I just used as many resources as possible from textbooks and…altered them based on what was right for my students." Lynn discussed individualized assessment based on individualized interventions: "Say it's a social story you put that intervention in and see, see if there's a certain behavior occurring…if that behavior decreases then that's the way you, you know, assess the situation." Shelly explains: "I think it's really crucial that as teachers it's not you know one assessment fits all. It really is each kid needs some type of tweaking. Even if it's in general ed. I think they need their own tweak." It was clear from this type of comment that this concept was important to participants. Participants' comments show the idea that every student needs to have the opportunity to demonstrate knowledge and it is the teacher's responsibility to provide a way for students to do this.

There were many examples of how participants individualized. As Lynn explained, "We have one student that loves snakes and so when we made the letters look like snakes that was the thing that made it so he would be able to tell us what letters he knows." She provides other examples of this same concept: "maybe pointing to letters in a book isn't very much fun so putting…sparkly letters or things like that is another way." Dawn also explained individualizing for a student: "we did a Velcro on one side and he was able to match them…saying the colors that was more challenging for him but he was able to match them." She continues this thought in explaining the manner in which assessment had to be conducted to get valid results for a particular child: "I knew with him just his ability to focus and pay attention and attend to a task…I had to be creative and quick on your feet."

Shelly explained individualizing by using an assessment method that matches the strengths of individual students. She stated, "it's based a lot upon how well, what areas my kids excel in. Um, if it's written then they'll do it written, if it's oral then I'll have them do it oral." This describes authentic assessment of concepts. Lynn explained this idea by stating that "I definitely believe in you know the making sure you find out what-some people are visual learners, some are auditory, some are kinesthetic." Lynn continued to explain this idea with the comment that "sometimes it's just putting them in a situation where they can show what they know." This situation appears to be different for elementary and middle school level teachers, as expressed by participants.

Cathy and Carmen discussed individualizing assessments in the middle school science context mostly by adapting summative tests. As Cathy explains, "the format was maybe not working like…maybe multiple choice isn't working for everybody…I would

change the format but maybe keep the basic structure." Carmen explains trying to adapt

tests using multiple methods: by "depending on the student I've had to not give them less

but shorten the question…and then they can go ahead and move on," by extensively

previewing test questions in review sessions, and by using the "special needs" version of

the test provided by the textbook publisher. She explained that although the "special

needs" test uses simpler language, it is often still too difficult for students without teacher

assistance in reading the questions. Carmen also explained that many of her students were

labeled as English Language Learners, and could in some cases provide greater

information about what they knew using Spanish, which she then translated and wrote for

them in English. Although both Cathy and Carmen spoke a lot about the need to

individualize assessments, they did not express any degree of hope for student success on

these types of tests. They described the situation in which they appeared to put a lot of

effort into making accommodations and modifications without discussing the possibility

that doing these things would cause students to succeed on the test.

Three other areas had data generated from four of the five participants: meeting

district requirements, having measurable criteria, and frequent informal assessment.

District requirements included standardized testing and formal assessments such as the

DRA and assessments of standards or common core standards. Although no questions

were asked about standardized assessments, participants commented about the

expectation to participate in them (explained further in the next category). Regarding

formal assessments, Dawn explained that "Reading came pretty easy with as far as like

up keeping up with the DRA's or the running records." These were required by her

district for her grade level. Other formal assessments viewed as mandated included

SPIRE reading assessments, Brigance assessments (for some participants), and unit tests for middle school science classes.

Having measurable criteria was most clearly explained by Shelly: "it has to be measurable if you're collecting data." She went on to explain that this includes both academic and behavioral concerns and both formal and informal assessment. Lynn gave the example of task analysis for learning to use the "potty," which has clear measurable criteria, such as washing hands: "I have some sort of criteria or rubric" and use data sheets to record this data. Cathy describes criteria as the degree to which they match the content goal: "some of them will maybe not perhaps get the complete correct answer but you can tell that they are getting the concepts." Dawn references common core standards as criteria: "Doing like a number sense assessment …I would have um the kids count rote count and see how high they could count." These examples provide evidence of participants' understanding of the "systematicity" component of the NRC's description of assessment (2008).

Frequent informal assessment was described in several ways by Carmen, Cathy, Lynn, and Shelly. Frequency is another component that addresses "systematicity." Carmen explained it as largely oral assessment in which all students were invited to participate and as the product of exit tickets. As she explained, when using exit tickets:

> If I can tell they don't know the answer to something or I know they don't know
> how to explain it the next morning is usually when we'll review it again and then
> I'll have them tell me about their thoughts on it.

Cathy described it as checking-in with specific questions at least ten times during a typical science class of 45 minutes. Lynn described it similarly: "I check for

understanding, understanding the way they're able to explain it- the way they're able to

explain their answers." Shelly added, "informally it's every day that I'm assessing them."

It is clear from these comments that regardless of how often participants actually assess,

which is something not measured in this study, participants consider it important to

conduct frequent informal assessment. Their understanding is that frequent assessment is

important in valid formative assessment.

A great deal of data was generated in the area of assessment implementation with

a range of responses across participants. Table 3 provides a summation of which

participant commented on each topic.

Table 3

Teacher Implementation of Assessment

Assessment Strategies	Carmen	Cathy	Dawn	Lynn	Shelly
During the Assessment					
Repetition	x	x			
Prompting/Help Navigate	x	x			
Preview Test Questions	x				
Quick Feedback			x		
Breaks			x		
Peer Modeling	x				
Processing Time					x
Reassurance / Praise	x		x		
Presented as Play			x		
Choices/ Options					x
For Individualizing					
Trial and Error	x				
For Low Reading Skills	x				
For Low Writing Ability	x				
For Focus / Attention Needs	x		x		x
For ELL Learners	x				
For Behavioral Needs		x	x		x
For Student Interests			x	x	x

Assessment Strategies	Carmen	Cathy	Dawn	Lynn	Shelly
For Individual Learning Goals	x		x	x	x
Matching Learning Style/Strengths			x	x	x
Decreasing Length	x				
Accessing Using Technology	x				
Changing Formats		x			
To Provide Challenge			x		x
General Practices					
Assess Standards / Common Core			x		
Multiple Methods per Concept	x	x			x
Set Up Assessment Routines	x				
Collaborate			x	x	x
Consult Families					x
Preserve Student Privacy					x
Have Measurable Criteria		x	x	x	x
Keep Current / Up to Date	x				
Meet Requirements of District	x		x	x	x
Give Useful Feedback to Students			x		x
General Practices					
Be Positive / Enjoy Assessing				x	x
Provide for Student Affect Needs	x	x	x	x	x
Make It Fun for Students	x		x	x	
Individualize and Differentiate	x	x	x	x	x
Grade Level Appropriate	x				
Create It If You Can't Find It			x	x	x
Use Detailed / Complete Assessments		x		x	x
Keep Detailed Records			x	x	x
Use Results to Plan Learning		x	x	x	
Use Valid Assessments		x			
Assess Informally Frequently	x	x		x	x

Challenges in assessing. This category parallels and extends the previous category at times in that many strategies described in the last section could be viewed as practices that are useful but perhaps difficult to implement. 'Challenge' ideas that are parallel to 'Strategy' ideas include addressing student affect, differentiating and individualizing for different skill levels, using valid assessments, and finding and/or creating appropriate assessments.

Challenges regarding standardized assessments were described by four participants, although semi-structured interview questions did not include this topic as it is outside the focus of this study. Participants discussed their concerns about student nervousness (affect), test validity for students with disabilities, and the consequences for poor results. Lynn also explained a lack of knowledge about "what they're trying to hit on. Is it like related to state standards completely or is it just national" and Dawn explained that "some of the areas that the SBA hit I don't think I covered very well." This was a clear area of interest to participants. The new teacher evaluation system being implemented in New Mexico this year bases 50% of a teacher's effectiveness rating on growth in standardized test scores. Teachers who do not score at least "effective" are required to participate in an improvement plan. Their employment may be in jeopardy and they may feel shame at being labeled ineffective or minimally effective. This has created a situation in which all teachers and special education teachers specifically may experience negative consequences for an assessment that cannot validly measure their students' levels of growth. Based on participant responses, standardized testing is a relevant, omnipresent assessment issue that needs to be addressed.

Differentiating and individualizing assessments also included concerns of four participants. Carmen particularly described the number of ways in which she was struggling to adapt grade level assessments to address significantly lower reading levels and the needs of English Language Learners: "I've had to sometimes make two sets of the same test just different wording" and "it's like one question. And to me for them that will take three hours." Cathy had a similar concern about low student reading levels in that "when they take the tests I spend a lot of time breaking it down and helping them to

figure out what it's really asking." Nowhere do Carmen and Cathy state that it is possible, with adaptations, for their students to be successful on these tests. It is interesting that Carmen particularly provided multiple examples of other methods to formally assess (posters, presentations, hands-on projects) but still assessed using adapted grade level chapter and unit tests rather than replacing them with another method.

Of additional concern in this area is providing the right level of challenge and an appropriate assessment for each student. Carmen explains that "my challenge is just finding…what's the balance…to meet both the high, middle, and low." Shelly explains the problem when using a district required reading assessment (the DRA). One of her students had both giftedness in conceptual understandings and a learning disability that affected reading. She explains:

> He was reading two grade levels below so you know content reading about a farm
> or a cat and it was just you know this kid had no interest in it…so he could listen
> to lectures and he could retain it all but when it came to the DRA…I kept saying
> he could do so much better

This is not a challenge that classroom teachers will necessarily be able to address. If the DRA is a required test, the teacher may have reservations about its value but must still give it and find another way to capture actual knowledge or skill level. Dawn explained the challenge of appropriate individual assessment for one of her students: "He would-wasn't able to sit long through assessments like the other kids were…I had to be creative" in order to find a way in which he could be assessed.

Participants also described assessment validity as a challenge. Shelly and Cathy make similar comments on this topic: when using an assessment, "a big problem or issue

would be if you know what I'm trying to assess for is really um what I need to be looking

for" (Shelly) and "that's actually kind of tricky to figure out how to ask them or what to

ask them to know if they are getting it" (Cathy). Recognizing this as a challenge requires

a degree of assessment knowledge beyond just a procedural knowledge of how to use

informal assessments.

> Cathy further explains:

> > Sometimes it's hard to make sure that like my informal assessments are
> >
> > actually really telling me what the kids know you know that it's actually
> >
> > really targeted like really specific and like gives me the information that
> >
> > I'm trying to get.

Other participants have expressed a procedural knowledge of informal assessment types

available but do not mention an awareness of challenges that are directly tied to the

purpose of assessment, as this challenge is.

> Additional validity concerns relate further to the type of assessment or format of

the assessment used. Lynn most clearly expresses this as "sometimes I feel like man, I

know they know this but how am I going to get it so that they can show me what they

know?" All five participants expressed a desire for greater understanding in this area.

Shelly was concerned about using this knowledge to better understand social and content

skill acquisition for the "couple of hours a week" in which her students worked with

general education peers. Carmen stated that she would "like to learn more about

assessments, though- different types," including "how can you like kind of trick them so I

am assessing them but not, them not know that's what's happening" and "project based

assessments." Tricking them appears to be related to reducing negative student affect to get a more reliable assessment.

Carmen also stated a desire to know "what other assessments to use for special ed. 'cause all you get the whole time is multiple choice. That's the easiest and best way to go for special ed., but there has to be another way…" These comments provide interesting contradictions. Carmen appears to be explaining the situation in which she knows that other types of formal assessments are available but she does not believe that she is allowed to use them and is stuck with multiple choice summative tests, which are the easiest for her to use within the confines of adapting Chapter tests. She is asking for knowledge of how to incorporate her actual assessment knowledge into the broader perceived requirements of her school or school district.

On a similar topic, Cathy made several comments summarized in the statement "I don't think that I write a formal assessment very well. I haven't had a lot of practice in it." She then explained the need to be able to create this type of test from time to time and being unclear about "whether or not they should be multiple choice or question and answer and like the format." She explained the desire to write a formal assessment that "would incorporate like a broad range of students and abilities" and that used appropriate language. As she explained, "I don't want to not expose them to that sort of language but I know that in the tests that they get we spend a lot of time breaking apart the language." Again, Cathy is explaining the requirement for students to be able to engage in this type of assessment but the challenges of providing access to these tests and the situation in which these tests are not authentic assessments of science knowledge.

Two other challenges were described. Dawn briefly described the difficulty in keeping up with math assessments when changing to a new collaborative team at her school: "Sometimes I wasn't sure as far as the team like what are we teaching this week like trying to keep up with that." This is an important comment because it situates assessment in the context of working with other teachers. Even when providing segregated special education services, as Dawn did, there is an expectation that teachers work as a part of a team and coordinate the content or common core standards that are being taught and assessed. The other challenge was described by Shelly in relation to writing Individualized Education Plans (IEPs). As she explained, "the IEP is a big assessment piece in itself you know and as much as they prepared you for what an IEP is…we don't actually do it" before beginning teaching. Shelly explains that although this issue has been resolved for her, she was unclear about how to use assessment to inform IEP goals: "I know that you know it has to meet their needs and be specific but they never show you how to write one." This concern highlights a possible lack of understanding, at least at the time Shelly began teaching, about how classroom assessment and IEP goals are related. Other participants did not discuss how regular classroom assessment is related to IEP goals. I did not ask interview questions about this topic but now wonder if participants understand the connection between these topics.

Preparation for assessment. 'Challenges' as a thematic category naturally flows into a discussion of teacher preparation for assessment. Preparation is one manner in which challenges may have been or could have been addressed. Participants commented exclusively on preparation in the context of what they learned when enrolled in SEDLP coursework and practicum experiences.

First, participants commented generally on their preparation coursework: "the dual license program was fantastic…There was several courses that explicitly talked about assessment" (Shelly), "we had several assessment classes or at least one um I'm not remembering exactly (Dawn), "especially my last two semesters probably influenced a lot of the way that I do my assessments" (Cathy), and "I've thought about retaking that class, the assessments class, and seeing or even just going to some training or something for it" (Carmen).

Next, participants commented on positive aspects of their preparation program. Lynn had particularly positive comments about her instructor/professor: "he was awesome he talked about all the different forms so he was very knowledgeable." She continues to describe the plentiful examples provided and the memorable class discussions he caused about how to cause students to show what they know, positives and negatives of formal and informal assessments, and thinking about the reality of testing in older grades. She characterized the professor as "kind of silly" in modeling "constant check in" and positively characterized her professors as a whole; "they make it sound fun and data collection is fun and so that's what puts the passion" in assessment preparation. It was surprising to hear a participant characterize assessment in such a positive manner, and clear that SEDLP coursework contributed to this view. Carmen also explained a positive trait of her professor: "She didn't do a lot of teaching it was more of us doing teaching so I did like that. That was nice because then you get comfortable with the subject and you can teach it." Other participants did not specifically describe the positive traits embodied by the faculty members teaching assessment coursework.

Three participants described their current use of assessment knowledge gained from the SEDLP program. Cathy described a more comprehensive understanding of what assessment is: "before…I wasn't really aware of the fact that that could be considered an assessment you know before that I thought that it had to be a test. So that definitely influenced me a lot." This is an interesting comment in the context of Cathy's currently described use of summative assessments. She states an awareness of formal assessment options but not the use of these options in her actual assessment practice. Cathy also explained learning to do informal assessments in the SEDLP: "I feel like I get a much better understanding of what students are learning." She does explain using this knowledge as part of her assessment practice, as discussed in the implementation section.

Shelly explained that "I saved all of the stuff from the class 'cause I refer back to it when you know I need some more ideas or um maybe a method's not working as well as I would have hoped it would." This demonstrates the use of SEDLP coursework as an introduction to methods and as a resource for new teacher reference. Shelly listed ABC (Antecedent-Behavior-Consequence) charts and FBA's (Functional Behavior Assessments) as particularly valuable assessment tools from coursework. The district in which Shelly teaches offers regular training to special education staff about these items and their use in forming Behavior Intervention Plans, which suggests the district experience that entering special education teachers do not have adequate knowledge in this area without additional training. Behavior consultants within this district also regularly supervise the teacher use of these tools. Within the district, being able to use these particular tools in an effective way is considered highly desirable but the level of district guidance in this area suggests that many teachers do not have this skill. This

would seem to indicate that although this component of assessment is present in SEDLP

coursework it may be lacking for other special education teachers hired by the school

district from other preparation programs.

Lynn also explained the effect of preparation on her current assessment

organizational practices: "they made it sound like it…doesn't have to be that takes a ton

of time out of your day it's really not once you learn…to get different binders or

whatever you have to do." This captures the idea that systematicity as described by the

NRC (2008) is presented in SEDLP coursework. The physical details of how to collect

and store assessment data are not the most interesting component of assessment

instruction, but they are important in that if beginning teachers do not have knowledge of

how to set up such systems they may not be able to implement their broader knowledge

of assessment concepts in an effective way.

Participants also discussed negative or limited aspects of their assessment

preparation. These tended to vary greatly by participant with some overlap.

Carmen presented somewhat contradictory statements about her experience:

"When I took the assessment class our teacher wasn't the best. She was good, she wasn't

the best" and "If I retook it I'm hoping I'd get a better teacher that knows what they're

doing." The contradiction in these statements is present in that even if a teacher is "good"

but not the best she should "know what she is doing"- only a not-good teacher would lack

knowledge of what she is doing. This contradiction may reflect Carmen's unwillingness

to criticize the SEDLP program in an interview by someone who is connected to it and to

other people involved in it; nevertheless, an expression of dissatisfaction is intended, as is

demonstrated by her continued comments. She states that "I felt like it was more

repetitive stuff that we already knew" while also stating a need for more knowledge in this area: "I've thought about retaking that class" and "hopefully they'll give us more I guess assessments 'cause we only got a few…and I know there's a bunch more like I know there is." Carmen was able to provide many examples of assessments, but as is described in the challenges section, she may not see these as usable options when actually assessing. This problem is also explained by Cathy:

> We focused a lot on how to do informal assessments and sort of you know the hindrance of the like basic test exam thing but that doesn't make it not a part of school and I feel like maybe that part was you know skipped over because a lot of people have very adverse feelings about it but…it's still a requirement you know it's still something I have to do.

Carmen's and Cathy's concern about this suggests that addressing this issue in some way is necessary for preservice teachers who will be working in a middle school environment.

A different type of preparation concern was discussed by Dawn. As she explained, "I felt like we had a strong literacy background I guess in math I felt more weak" because "I don't think we had near as many um like having the opportunity to create a math assessment as much as we did the literacy." The list of required SEDLP coursework (Appendix A) tends to support this idea. All coursework contains an assessment component, and literacy coursework includes five classes (Children's Literature, Teaching Reading, Teaching Reading in Elementary Education, Teaching Oral and Written Language in Elementary Education, and Differentiating Reading Instruction in Inclusive Settings) while mathematics coursework includes one class (Teaching Math

in Elementary Education). It is important to consider whether or not assessment for

mathematics is being adequately addressed.

Member Check

I emailed each participant a summary of her individual main points presented in

the combination of both interviews and asked for her to examine if it sounded like what

she said and intended to say. Four of five participants responded to the member check.

Three participants stated "Yes, that matches everything I said!" "Yes, this all looks

accurate!" and "went over the points of the interview and it seems to be accurate to what

I said." One participant asked that I remove one summary sentence but did not provide an

explanation of why. I deleted the line. Perhaps it was important to the participant not to

emphasize that particular component of assessment. Summaries, including this revision,

are included in Appendix E.

CHAPTER 5

Discussion and Conclusions

The purpose of this study was to examine experiences of recent graduates of the SEDLP program in order to answer three research questions:

1. In what ways do recent SEDLP graduates characterize their level of Competence (theoretical understanding and practical application) in assessing the progress of students with disabilities in the classroom?

2. In what ways do SEDLP graduates report they use classroom assessment to inform classroom instruction?

3. What features of the SEDLP teacher preparation program do graduates identify as having positively or negatively impacted their ability to effectively use classroom assessments?

Level of Competence in Assessing Progress of Students with Disabilities

Competence was not intended to be directly measured in this qualitative study, but participants do describe what they understand and can do to assess progress of students with disabilities. As discussed in Chapter 4, this description includes providing complementary definitions of assessment across participants and the ability of each participant to list or describe many assessment types and strategies. The definitions provided by participants easily fit into more formal definitions provided in the literature, including the Ysseldyke and Algozzine (2006) definition of assessment as "a process of collecting data for the purpose of making decisions about students" (p.74) and the NRC (2008) description of assessment as requiring "purposefulness" and "systematicity." Included in "systematicity" is the idea that assessments must be valid and reliable. As

discussed in Chapter 4, Cathy and Shelly particularly discussed concerns about ensuring

that their informal assessment methods were valid assessments.

Regarding assessment types, there are some differences between this study and

the limited research available in this area. Kohler, Henning, and Usma-Wilches (2008)

found that preservice teachers used primarily two types of formative assessments:

"listening to student talk" and "observing student behavior" (p.2113). Although Cathy,

Carmen, Dawn, Lynn, and Shelly did discuss using these methods frequently, they were

also able to provide specific examples of using other methods, such as exit tickets,

journal entries, games, posters, paper folds, and task completion. The level of detail they

were able to provide suggests that these were actual practices used in their classes rather

than simply an understanding of options. For example, Dawn states that she has students

using physical objects as a formative assessment of one to one correspondence: "having

so many objects, bears for instance, and can they move them and count them correctly so

getting an understanding of where they are." Carmen was able to explain a formative

assessment of geologic concepts:

> For my kids that are doing rocks and minerals and how the land is formed I did a
>
> where they made one picture of like the mountains and stuff and they had to cut
>
> and glue all the other like where the spring water comes into the river water to the
>
> ocean water you know and it forks off into different rivers and stuff like that you
>
> know I had them draw that but instead of having them draw it I had them cut with
>
> different colors of paper and they made shapes so that was an assessment I
>
> thought for them knowing how the land was formed and where the water comes

Cathy discussed reading journal entries, similar to how an exit ticket would be used, and Lynn and Shelly provided examples of formative behavioral and social assessments using targeted written observations based on task analysis for situations such as learning to use the restroom and student self-assessment of "appropriate or negative reactions" (Shelly). What extends these behavioral observations beyond the Kohler et al. (2008) conclusions is the focus on measurable behavioral criteria, even in formative assessment. As Shelly explains, "it has to measurable if you're collecting data."

It is possible that the understanding and use of a greater variety of formative assessment methods is a product of participation in the SEDLP. There is a gap in the research in this area, and this study contributes to an understanding of what new dually licensed teachers understand about how to implement formative assessment.

Competence in implementing summative assessment was an area in which two participants described concerns. As Cathy and Carmen explained, they conduct summative assessment of middle school science units or chapters using tests provided by the textbook publishing company and adapt these tests to meet student needs. As explained in Chapter 4, Cathy expressed concern about her ability to write a good formal assessment and both participants failed to express that their students experienced success when using these tests. Although both participants were able to describe authentic assessments, both stuck to the use of the tests described above as summative assessments. It is possible that both participants perceive limited options or personal control in summative assessments. Cathy provided support in an inclusive science classroom. Her role appears to be constructed as one of support, which means that there is unequal power in making instructional decisions, including assessment decisions. Carmen taught in a

segregated science setting but still is a part of the community of science teachers within her school and school district. The pressure to assess in a particular manner, although perhaps not directly stated, is present.

There were additional areas of knowledge across participants. As demonstrated in the results section, participants described the ability to meet district assessment requirements, to attend to student affect issues during assessments, to broaden the application of an assessment by individualizing or differentiating, and to frequently conduct informal assessments. This study did not measure participant ability to apply the items listed above, but participant knowledge of the existence of these items is an indicator of the practical relevance of particular assessment ideas to participants. Participant identification of these items as relevant to their practice may also demonstrate the impact of preparation with the SEDLP.

It is interesting to notice the assessment terms used by participants as additional descriptors of competence in assessment knowledge. Participants regularly used the terms "formal" and "informal" in a manner consistent with their field-defined definitions, but did not use the terms "formative" and "summative" except in one case, where one participant used "formative" one time only, in a manner inconsistent with the field-defined definition. She used it to mean "standardized." From this we cannot generally conclude that participants did not know the concepts that these terms describe, just that they described their assessment behavior rather than using these labels. Additionally, the current data showed some confusion or inconsistency about the meanings of the terms "individualize" and "differentiation" related to assessment. Additional probing questions

would be needed to determine participant belief about the precise meanings of these terms.

Classroom assessment is a broad field of study. Within this broad field there were many aspects of assessment knowledge that all or most participants did not describe. For example, participants did not describe historical, legal, or human rights' understandings of the role of assessment. One participant (Shelly) commented one time that her students "deserved" appropriate assessments, but other than this instance, these concepts were either assumed and not mentioned or not part of participants' assessment constructs. Participants also did not widely discuss evidence-based assessment understandings, although concerns about using measurable criteria and assessment validity of purpose were addressed, and these are evidence-based considerations. It would have been difficult for participants to discuss evidenced-based assessment understandings when, as the literature demonstrates, there is no current consensus on what practices are evidence-based in the special education field (Cook, Tankersley, & Landrum, 2009), though even without consensus it is important for beginning educators to be able to differentiate between better supported and less supported practice claims (Emmons et al., 2009). Curriculum based assessment, curriculum based measurement, authentic assessment, and assessment of embedded IEP goals were also absent from participant discussion although some components of authentic assessment were discussed, such as the basic intent of finding a way for students to demonstrate knowledge of the particular skill rather than knowledge of reading about the particular skill. The literature review explains gaps in teacher knowledge in these areas, and these gaps appeared to present for the participants in this study as well.

Ways of Using Classroom Assessment to Inform Instruction

Participants did consistently present the idea of using informal classroom assessment to know generally what to do next. Other important uses were not described consistently across participants but included instructional grouping, knowing more specifically what to teach next, and determining effectiveness of interventions. Second round interview questions attempted to get more data about, for some participants, the unaddressed topics of student behavioral and social skills and how to use assessment in these areas. The result was greater detail for the initial participant responses, but not an increase of information on this topic across participants. For example, Carmen and Cathy presented minimal comments about assessing behavioral and social needs, and even after additional probing questions, limited material was added to these understandings. This reinforces the idea that assessment in these areas was not relevant to these particular participants and therefore was not used to inform their practice. This may be because Carmen and Cathy taught middle school students in a district in which students with behavioral or social skill issues at the middle school level would be more likely to be taught in a class specifically labeled as an environment in which to address these needs. Students with these difficulties would likely be labeled as having an emotional disorder or behavioral disorder and placed in a more restrictive behavior "program." It is possible that addressing these issues within a less restrictive setting is more of a norm for elementary schools.

SEDLP Program Features that Affected Graduate Assessment Skills

College of Education faculty were indicated as a program feature that affected assessment understanding, both positively and negatively, as explained in more detail in

Chapter 4. Only two faculty members were mentioned, so there isn't a way to gauge the

perceived impact of this program feature within the context of this study. The positive

faculty member description included a range of comments about the variety of

assessment methods modeled and the overall positive view of assessment presented.

Additionally, the negative description of one faculty member's instructional practice

seems to be entangled with dissatisfaction with participant summative assessment skills.

This is an issue that may not be a direct criticism of the quality of instruction provided by

the faculty member, but instead is a criticism of the situation in which participants are not

able to cause success for their students by adapting grade level summative science

assessments. In this case providing more examples of types of summative assessment

would not have increased student success unless participants were free to actually use a

greater variety of these assessments. Both middle school science teachers expressed a

desire for greater skill in creating formal, summative assessments, with one participant

describing this as needing to know more types of assessment.

Positively described program features in the context of course instruction included

the opportunity to see many examples of types of assessments, appropriate modeling of

assessment techniques, and frequent practice in understanding or implementing informal

assessments. All participants were able to describe an array of informal assessment

options and give examples of how they had used them in daily practice. Individual

participants also mentioned coursework that provided assessment resources for later

reference and that provided ideas about how to organize a classroom for assessment.

Negatively described program features included gaps in some content coverage

expressed by individual participants, including additional examples of types of formal

and summative assessments, instruction and practice in designing formal assessments, and instruction and practice in designing assessments for mathematics.

Limitations

A limitation of this study is its small number of participants. Additionally, participants were all working in elementary school preK-5 and middle school (6-8) settings. This could limit generalizability across other grade levels and settings. None of the participants who were teaching in segregated settings were working in classrooms identified as primarily serving students with more severe disabilities and multiple disabilities, which makes this study potentially less applicable for students with severe disabilities. All participants were female, and although caution must be used in making assumptions based on gender definitions, including male participants as well would have been preferable. Four of five participants identified themselves as Caucasian, which does not reflect the ethnic and racial composition of the people who live in the geographic region in which this study was conducted. A larger number and more diverse group of participants may have provided a more thorough understanding of the topic due to greater variance in perspective. A further limitation of this study is that data is self-reported by participants and these reports may differ from actual participant practice. Conclusions about actual practice must therefore be limited.

Implications for Practice

The greatest implication of this study is that recent SEDLP graduates do have assessment knowledge and gaps in knowledge that they are able to articulate and that can therefore be used to consider changes in the SEDLP. Assessment knowledge is a critical component of teacher education that must continue to be addressed. SEDLP graduates

have assessment knowledge that will directly impact students with disabilities in the area of learning, which is a property right provided in the United States and a human right provided internationally.

It is also interesting that participants consistently raised the issue of standardized testing, even though it was outside of the overall stated study goal and the semi-structured interview questions did not address this topic. This provides contextual implications; teachers are working in a climate that may be highly focused on standardized testing. Appropriate assessment practices may be processed through a filter of relation to standardized assessment rather than a relation of what is needed for a particular assessment situation.

Recommendations for Teacher Preparation

This study highlights the understanding that the context of high stakes standardized assessment is the omnipresent reality of teacher's classroom experience; therefore, it is important to address this issue in teacher preparation programs. It is tempting to resist the idea of addressing this issue because it does not seem relevant to the actual important work that occurs in classrooms, but it is an issue important to new teachers because of the potential professional consequences. New teachers are now being evaluated and potentially penalized based on the level of growth their students achieve on these tests but do not enter the profession with the level of experience needed to cause optimum change. It is unlikely that even experienced teachers of students who receive special education services will be able to cause their students to achieve the level of student growth required to label the teacher "effective" because these are assessments that are often not valid in determining authentic student knowledge. Under this

assessment system, groups of students, including students with disabilities, will continue to score poorly even when provided with high quality instruction. In this environment it is unlikely that new teachers will want to be told to focus on authentic classroom instruction and assessment rather than standardized testing, but this is critical. Teachers only have instructional opportunity with students for a limited period of time, making it wasteful to spend time focusing on a system in which it is unlikely that a positive impact can be made. Standardized tests are partially tests of content but largely tests of reading comprehension skills. It is usually not possible to raise comprehension levels to the degree required within these standardized tests because the tests are not sensitive enough to measure, for example, an eighth grader's growth in reading from second grade level to fourth grade level. The tests instead continue to measure lack of growth at eighth grade level.

Participants in this study struggled with this type of issue on a smaller scale when they tried to adapt grade level summative assessments for classroom use. Although providing accommodations and modifications is more likely to allow students to demonstrate knowledge than not providing accommodations and modifications, participants did not report more than minimal success using these methods.

There isn't really any comforting statement that can be made to new teachers about high stakes testing. Perhaps the best that can be done is to acknowledge the issue and focus on what is more likely to be within teacher control- actual classroom instruction and assessment. Even in the classroom the issue of control remains. Teachers who provide special education services are part of a collaborative group at all levels and at times may not have equal power or control of the curriculum, instructional methods,

and assessments. It is important for teacher preparation programs to provide preservice teachers with an understanding of methods to use what power they do have to positively affect outcomes for students with disabilities. This means that they will need to be able to communicate in a positive and assertive way about what they can offer regarding instruction and assessment options. New dually licensed teachers need to be able to collaborate effectively with general education teachers, which is a complex process. In order to do this they need to not only have knowledge about teaching but also have knowledge about how to communicate in a way that allows them to accomplish what they have been trained to do. More direct instruction and practical experience may be needed in this area and it would be worth considering increasing its inclusion in existing or new coursework.

An additional recommendation would be to investigate whether or not adequate opportunity is present for SEDLP students to learn and practice math assessment. It is possible that such practice is adequate and is provided across a variety of courses, but it is important to know whether or not this is the case so that adjustments can be made if needed.

A final recommendation would be to further investigate SEDLP students' knowledge of the relationship between classroom instruction, assessment, and IEP goals. Only one participant discussed this topic but knowing how to use assessment in a way that connects both instruction and IEP goals is essential for meeting the intent of assessment practices as described in IDEA (2004). This knowledge would fit into the "systematicity" assessment description from the NRC (2008).

Future Research

Although this study does provide some beginning insight into the assessment knowledge of recent graduates of the SEDLP program, it is not a comprehensive analysis and more research in this area is needed to determine strengths and areas of improvement for the program. Studies that examine actual implementation of assessment knowledge, such as planning for assessment, assessing, and use of assessment results, are needed to understand success in not only the SEDLP program but in the broader field of teaching beginning teachers to use classroom assessment appropriately. At present, there is an inadequate research base in the field of assessment. We do not know how teachers implement their plans for assessments, what opportunities are present for teacher individual assessment choices, and what models of success are available to illustrate teaching beginning teachers to asses appropriately. We also do not know if there is a difference or the extent of difference in assessment knowledge among general education, special education, and dually educated beginning teachers. There is a significant gap in this area across the field of higher education as a whole, as demonstrated in the literature review.

Expansion of assessment knowledge is needed across a range of areas. Particular areas of research need include classroom assessment for students with disabilities, classroom assessment for students with disabilities in inclusive environments, and classroom assessment in specific disability categories, such as multiple disabilities. Gaps also exist in understanding the impact of specific assessment strategies on long-term student achievement.

Classroom assessment is an exciting area in which to increase academic and professional knowledge. Taking advantage of the newly implemented use of common

core standards as a research context may aid in generalizing research results across areas of the United States, thus increasing the quantity and quality of classroom assessment research available. This is an opportune time to increase classroom assessment research across the field of education.

Reflection

There are many gaps in assessment research, and this study contributes specifically to the understanding of what new dually licensed teachers understand about how to implement formative assessment. This study is also worthwhile in that it uses interviewing as a data collection method rather than simply analyzing artifacts. Interviewing is a method that is respectful of the complexity of human beings in constructing the world around them. The study would have been stronger using a combination of methods to triangulate results, including a greater number and a more diverse group of participants and including a deeper analysis of the environments in which participants were teaching. It may also have been helpful to directly analyze SEDLP coursework using syllabus descriptions and faculty interviews.

More deeply analyzing teaching environments could have led to greater understanding of the broader context in which graduates of the SEDLP are teaching. At the school level, it is important to understand differences and similarities among the populations of students who were being provided with special education services. Power relationships, viewed using a critical lens, would be more apparent with greater information about each school. It is important to also use a critical lens in considering the context of the districts in which teachers taught. District decisions can affect teacher practice. Even the most robust teacher efforts at implementing assessment practices are

not effective if the structure that has been created by the district does not allow such practices and preserves the systemic status quo. It would have been interesting to compare stated district assessment practices and non-stated assessment practices as demonstrated by other artifacts, perhaps including IEPs.

Finally, it is important to consider the broader context of assessment research and this study in relation to what is important in human lives. In human lives, education as a property right is important because it is what allows some level of equity and prevents the abuses that have historically and sometimes modernly occurred for individuals with disabilities. Assessment is a necessary and interesting component of education as a property right and has been addressed within both NCLB (2001) and IDEA (2004). Preserving only appropriate use of power is an important component of assessing and remains relevant in continuing assessment research.

Appendix A: Required Coursework for SEDLP students

Courses	Credits
ENGL 101	3
ENGL 102	3
ENGL 219 or 220	3
LING 101, C&J 130, 220, 321, 331, THEA 418, 415	3
HIST 101 or 102	3
HIST 161 or 162	3
HIST 260	3
HIST Elective (HIST, AMST, RELG, CLST)	3
Science NTSC 2611, 2621, 2631, CHEM 111L, 121 L, 122L, 131L, 132L, BIOL 110/112L, 122L, 123L, 124L, GEOG 101/105L, PHYC 102/102L, 151/151L, 152/152L, 160/160L, 161/161L, ENVS 101/102L, EPS 101/ 105L, 201L, ASTR 101, ANTH 150/151L, PSY 105 & 220/260	12
Math 111	3
Math 112, 121, 129, 150, 162, 163, 180, 181, 215, STAT 145	9
SOC 101, ANTH 101, 130, ECON 105, 106, POLS 110, 200, 240, PSY 105	6
ARTH 101, 201, 202, DANC 105, MA 210, MUS 139, 140, THEA 122	6
Second Language	3
SPCD 201	3
SPCD 204	2
SPCD 420 Introduction to ID	3

LLSS 443 Children's Literature	3
EDPY 310 Learning and the Classroom	3
EDPY 303 or PSY 220 or 260	3
EDUC 330L Teaching Reading	3
EDUC 353L Teaching Science in EL Ed	3
EDUC 361L Teaching Math in El Ed	3
SPCD 303 Methods for Learners w/Disabilities	3
SPCD 495 Field Experience	3
EDUC 331L Teaching Reading in El Ed	3
EDUC 333L Teaching Oral & Written Language in El Ed	3
SPCD 319 Classroom Organization and Management	3
SPCD 481 Assistive Technology	1
SPCD 304 Practicum	2
EDUC 321L Teaching Social Studies in El Ed	3
EDUC 400 Student Teaching in El Ed	6
SPCD 486 Differentiating Rdg Inst. In Inclusive Settings	3
SPCD 313 Curriculum for Learners w/Disabilities	3
SPCD 462 Student Teaching in SPCD	7
SPCD 464 Classroom Assessment and Program Planning	3
SPCD 481 Assistive Technology	1
EDUC 400 Student Teaching in El Ed	3
EDUC 493 T/Professional Seminar	2

Appendix B: Demographic Questionnaire

Please fill in the blanks or circle the most appropriate choices to describe yourself.

Pseudonym Choice	_____		Male	Female
Year of Graduation	2010	2011	2012	
Age Range	18-25	26-36	37-47	47+
Ethnicity	African American	American Indian	Asian	
	Caucasian	Hispanic	Other	
Current Assignment	Preschool	Elementary	Middle	
	High	None/Other		
Current Assignment Type	Special Ed.	General Ed.	Neither/Other	
	Co-teacher	Sole teacher	Other	

If employed as a Special Education Teacher:

Is your class entirely self-contained for academics?	Yes	No
Is your class self-contained only in math, reading, and writing?	Yes	No
Is your class 80% or more included with a general education class?	Yes	No
Is your class a resource room?	Yes	No

Please describe any self-contained/inclusive arrangement not listed above:

_____

If employed as a General Education Teacher:

Do you have any students who receive special education services?   Yes   No

Prior Adult Employment, if any, other than teaching:

_____

For how many years?   0-1   2-3   4-6   6+

Appendix C- Color Assignments for Data Processing and Analysis

Name	Color
Carmen	White
Cathy	Pink
Dawn	Green
Lynn	Yellow
Shelly	Blue

Appendix D- Initial Coding and Categorization

P- purpose
        PB baseline
        PGP general progress
        PS specific
        PGIEP goals
        PC competence
        PG groups
        PSB non-academic content
        PP compare to peers

C-challenges
        CKU keeping up
        CI improve teaching
        CIK implement already known
        CO need other options
        CNV not valid

T-types
        TV verbal/Oral
        TG games
        TPF paper folds
        TPH projects, hands-on
        TO observation
        TPres presentations
        TE exit ticket
        TPos posters
        TJ journal
        TF formal (tests/quizzes)
        TI (informal/hidden)
        TM movement
        TTask task

B-best practice
        BF frequency
        BI interpreting results
        BU using other ways
        BT type
        BAF accessing formal
        BTec technology
        BK keeping up to date
        BAd adjusting methods
        BC creating

        BMI matching/individualizing

        BColl collaboration

O-other

        OCR competence/resolved challenges

        OA student affect

        OTC teacher personal connection

        OCou assessment coursework

        OS standardized testing

Appendix E- Member Checks

Summary Points of Carmen's Interviews

1. Assessment is an activity, project, or test to see what students learned. I prefer projects but also do activities and have to do tests.

2. Teachers can use assessments to determine student skill level in one area (ex. reading) while assessing knowledge of a different topic.

3. Teachers use assessment to determine a class's beginning skill level.

4. I have to adapt the curricular tests I give to match my student's reading levels and in other ways (i-pad, students who don't write in English well, length).

5. Students' reading and writing levels really affect how they do on tests (they understand the content but can't show their understanding using reading and writing on the test).

6. My informal assessments include checking in as we read, activities and games like jeopardy, calling on students to answer in class, and exit tickets.

7. Some activities I give students to assess include trifolds, posters, and other visual presentations.

8. It is important to consider how students' feelings affect testing accuracy.

9. Multiple choice tests are useful for my students- I've taught them how to eliminate some choices and get the correct answers.

10. My younger students (6th grade) are not ready for some of the assessments I can give my older students.

11. I would have liked to know more about assessment from my classes. I especially wanted to know more types of assessments and how to give those assessments correctly.

Summary Points of Cathy's Interviews

1. Assessment gives you a reading of student knowledge.

2. I give mostly informal assessments and help students navigate formal assessments from their general education teachers.

3. I do curriculum based assessments for IEPs.

4. My informal assessments include projects, posters, models, science and history journals, oral question and answer, check-in's, and looking at their responses on assignments.

5. I can understand what students have learned by gauging responses to informal assessments. I had to get to know my students before being able to do this.

6. I can understand what students have learned based on both correct answers and incorrect answers- and whether or not they're in the "ballpark".

7. I do frequent checking for understandings- perhaps 10 times in a class period.

8. It's important to ask the right questions to students to gain a real understanding of whether or not they have learned. It can take some time to figure out how to ask questions that give the specific information you want about what students have learned.

9. I would have liked to have had more instruction in how to write good formal assessments, including appropriate formats for students with disabilities and a broad range of skills.

10. Informal assessments tell me where to go next in supporting my students.

11. Behavioral and social informal assessments can be done by monitoring disciplinary forms, checking in with teachers, checklists, and daily observations.

12. It's useful to start assessing by taking what assessments are already available and adapting them for what is needed. Later you can make your own assessments.

Summary Points of Dawn's Interviews

1. Assessments tell you where students are at. They give you a basis for where to start teaching and a baseline of student skills.

2. I started the year with two days of baseline skill assessments and then did assessments about every two weeks. These assessments were useful in grouping students for instruction.

3. I had to adapt the way I was assessing to meet the needs and abilities of some students, and particularly because of attention span.

4. It was easier for me to assess my younger students, perhaps because of the team I was working with and the way we set things up.

5. It was easier to give reading assessments, using for example DRA's and running records.

6. It was harder for me to keep up with third and fourth grade math assessments. It was hard at times to be sure the assessment I was using matched the current content taught and the math assessments were more teacher created than the reading were.

7. In my teacher preparation program I would have liked to have had more learning opportunities in math assessment- it seemed that much more was available in literacy assessment.

8. I had to create many of my own informal assessments, but I used more formal than informal to assess knowledge about specific content standards.

9. I did behavioral assessments, particularly I timed on-task behavior for one student.

10. Social skills assessment addressed individual learning goals for each child. I did these using oral language centers.

11. Documentation is essential in assessing social and behavioral growth.

12. Collaboration with peers is important in gaining the right assessment resources and individualizing for students.

Summary Points of Lynn's Interviews

1. Assessment is used to understand what students know about a topic so you will know what to teach next.

2. I use informal assessment often (checking in constantly) and less often I use formal assessment, like the Brigance.

3. For students who do not communicate verbally I put them in a situation where they can show what they know by completing a task.

4. To complete formal assessments a baseline and lots of documentation is needed.

5. I document observations using data sheets that we keep in relevant places around the classroom. I keep trying to make better data sheets.

6. Task analysis is important in assessing my students and helping them to progress.

7. It is important to individualize assessments. Collaboration is a good way to keep yourself being creative in finding what works for students.

8. It is important not to give up too early and to allow students time and varied opportunities to show what they know.

9. After assessing I do an intervention to help the student (ex. social story) and then reassess to see if the intervention worked. Sometimes it takes many attempts to figure out what the problem is and to help the student find a better way to meet their needs.

10. I gained a positive view of classroom assessment from my teacher preparation coursework. I learned enjoyable ways to collect data and saw many examples.

11. I wish I better understood how the SBA is used.

12. It is important to differentiate instruction to meet the needs of all students. Assessments should also be differentiated.

Summary Points of Shelly's Interviews

1. Assessment involves being able to determine whether or not students have learned what was taught. It is measurable both formally and informally.

2. I assess academics, social skills, behavior, and oral and nonverbal responses.

3. To assess I use a scale to get a baseline grade level and take a lot of written data and samples to get a full picture. I create my own assessments and record my own data to see progress levels.

4. I have had to be creative in making assessments for creativity, critical thinking, and other issues for twice exceptional students.

5. Informal assessments are less detailed but prepare me for how students will handle formal assessments.

6. It is important to individualize assessments. All students need and deserve this. One way to do this is to assess using modes in which students are stronger (ex.

written or oral depending on the student). Assessments need to be adjusted for the individual. It is critical to know your students in depth in order to accomplish this.

7. It can be difficult to determine what was done with a child in the past. BIPs and IEPs may not have adequate information about how assessment was conducted.

8. It is important to be very thorough in assessments. Detailed criteria need to be written out. Actions need to be thoroughly documented.

9. It can be useful to discuss assessment results and next steps with students.

10. Collaboration with parents, other teachers, and the school social worker can be useful in behavioral assessments.

11. I would like to better assess progress in interactions with peers in general education classes and with general education content.

12. I regularly use materials from my assessment coursework when I need more ideas or something isn't working. I use the ABC charts and FBA information.

13. I would have liked to know more about assessments in writing IEPs.

14. Some formal assessments don't work well for my students. For example, the DRA had material that did not match some students' developmental levels. Students could read at low levels but had interest/conceptual levels that were much higher than what the DRA offered.

References

Allen, D., Ort, S. W., & Schmidt, J. (2009). Supporting classroom assessment practice:

Lessons from a small high school. *Theory Into Practice, 48*(1), 72-80.

Allinder, R., Bolling, R. M., Oats, R. G., & Gagnon, W. A. (2000). Effects of teacher

self-monitoring on implementation of curriculum-based measurement and

mathematics computation achievement of students with disabilities. *Remedial and*

Special Education, 21(4), 219-226.

Allsopp, D. H., Kyger, M. M., Lovin, L., Gerreison, H., Carson, K. L., & Ray, S. (2008).

Mathematics dynamic assessment: Informal assessment that responds to the needs

of struggling learners in mathematics. *TEACHING Exceptional Children, 40*(3),

6-17.

Alonzo, J., Ketterlin-Geller, L. R., & Tindal, G. (2007). Curriculum-based measurement

in reading and math: Providing rigorous outcomes to support learning. In L.

Florian (Ed.), *The SAGE Handbook of Special Education* (pp.307-318). Thousand

Oaks, CA: SAGE Publications.

Anderson, K. T., Zuiker, S. J., Taasoobshirazi, G., & Hickey, D. T. (2007). Classroom

discourse as a tool to enhance formative assessment and practice in science.

International Journal of Science Education, 29(14), 1721-1744.

Begeny, J. C., & Martens, B. K. (2006). Assessing pre-service teachers' training in

empirically-validated behavioral instruction practices. *School Psychology*

Quarterly, 21(3), 262-285.

Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An*

introduction to theories and methods (5th ed.). Boston: Pearson.

Bourke, R., & Mentis, M. (2007). Self-assessment as a lens for learning. In L. Florian

 (Ed.), *The SAGE Handbook of Special Education* (pp. 319-330). Thousand Oaks,

 CA: SAGE Publications.

Bricker, D. (2000). Inclusion: How the scene has changed. *Topics in Early Childhood*

 Special Education, 20(1), 14-19.

Bricker, D. (2001). The natural environment: A useful construct? *Infants and Young*

 Children, 13(4), 21-31.

Bricker, D., & Gumerlock, S. (1988). Application of a three-level evaluation plan for

 monitoring child progress and program effects. *The Journal of Special Education,*

 22(1), 66-81.

Brookhart, S. M., Andolina, M., Zuza, M., & Furman, R. (2004). Minute math: An action

 research study of student self-assessment. *Educational Studies in Mathematics,*

 57, 213-227.

Brookhart, S. M., & Bronowicz, D. L. (2003). 'I don't like writing. It makes my fingers

 hurt': Students talk about their classroom assessments. *Assessment in Education,*

 10(2).

Brookhart, S., Moss, C., & Long, B. (2008, November). Formative assessment.

 Educational Leadership, 52-57.

Campbell, C., & Collins, V. L. (2007). Identifying essential topics in general and special

 education introductory assessment textbooks. *Educational Measurement: Issues*

 and Practice, 9-18.

Campbell, C., & Evans, J. A. (2000). Investigation of preservice teachers' classroom

 assessment practices during student teaching. *The Journal of Educational*

Research, 93(6), 350-355.

Cannon, C. (2006, April/May). Implementing research practices. *The High School Journal, 8-13.*

Catone, W. V., & Brady, S. A. (2005). The inadequacy of individual education program (IEP) goals for high school students with word-level reading difficulties. *Annals of Dyslexia 55*(1), 53-78.

Codding, R. S., Skowron, J., & Pace, G. M. (2005). Back to basics: Training teachers to interpret curriculum-based measurement data and create observable and measurable objectives. *Behavioral Interventions, 20,* 165-176.

Cole, D. J., Ryan, C. W., Kick, F. K., & Mathies, B. K. (2000). *Portfolios across the curriculum and beyond* (2nd ed.). Thousand Oaks, CA: Corwin Press, Inc.

Conderman, G., Katsiyannis, A, & Franks, D. (2001). Program assessment practices in special education teacher preparation programs. *Preventing School Failure, 45*(4).

Conderman, G., Morin, J., & Stephens, J. T. (2005). Special education student teaching practices. *Preventing School Failure, 49*(3), 5-10.

Cook, B., Tankersley, M., & Landrum T. (2009). Determining evidence-based practices in special education. *Exceptional Children, 75*(3), 365-383.

Cook, R. S., Rule, S., & Mariger, H. (2003). Parents' evaluation of the usability of a web site on recommended practices. *Topics in Early Childhood Special Education, 23*(1), 19-27.

Corbin, J. & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Los Angeles, CA: SAGE Publications.

Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for

kindergarten students: Comparing extended instruction to embedded instruction

and incidental exposure. *Learning Disability Quarterly, 30, 74-88.*

Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five*

approaches (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Cummings, Jack A. (2000). Academic assessment. In *Encyclopedia of special education:*

A reference for the education of the handicapped and other exceptional children

and adults (2nd ed). (Vol. 1, pp. 22-25). New York: John Wiley & Sons.

Cushing, L. S., Carter, E. W., Clark, N., Wallis, T., & Kennedy, C. H. (2009). Evaluating

inclusive educational practices for students with severe disabilities using the

program quality measurement tool. *The Journal of Special Education, 42*(4), 195-

208.

Daugherty, S., Grisham-Brown, J., & Hemmeter, M. L. (2001). The effects of embedded

skill instruction on the acquisition of target and nontarget skills in preschoolers

with developmental delays. *Topics in Early Childhood Special Education, 21*(4),

213-221.

Davidson, C. (2009). Transcription: Imperatives for qualitative research. *International*

Journal of Qualitative Methods, 8(2), 1-47.

Deisinger, J. A. (2007). Administrating discipline, grading, and 504 plans for students

with special needs. *Current Perspectives in Special Education Administration:*

Advances in Special Education, 17, 109-124.

Delano, M. E., Keefe, L., & Perner, D. (2010). Personnel preparation: Recurring

challenges and the need for action to ensure access to general education. *Research*

and Practice for Persons with Severe Disabilities, 33-4(4-1), 232-240.

Deno, S. L., Reschly, A. L., Lembke, E. S., Magnusson, D., Callender, S. A., Windram, H., & Stachel, N. (2009). Developing a school-wide progress-monitoring system. *Psychology in the Schools, 46*(1), 44-55.

Dodd, E. E., & Scheetz, N. A. (2003). Preparing today's teachers of The Deaf and Hard of Hearing to work with tomorrow's students: A statewide needs assessment. *American Annals of the Deaf, 148*(1), 25-30.

Duncan, A. (2009, July 24). Education Reform's Moon Shot. *The Washington Post.* Retrieved from http://www.washintonpost.com

Emmons, M., Keefe, E.B., Sanchez, R.M., Mais, M.M., & Neely, T.Y. (2009). Teaching information literacy skills to prepare teachers who can bridge the research-to-practice gap. *Reference & User Services Quarterly, 49*(2), 140-150.

Foegen, A. (2008). Progress monitoring in middle school mathematics: Options and issues. *Remedial and Special Education, 29*(4), 195-207.

Foegen, A., Espin, C. A., Allinder, R. M., & Markell, M. A. (2001). Translating research into practice: Preservice teachers' beliefs about curriculum-based measurement. *The Journal of Special Education, 34*(4), 226-236.

Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics: A review of the literature. *The Journal of Special Education, 41*(2), 121-139.

Fuchs, L. S., & Fuchs, D. (1986). Curriculum-based assessment of progress toward long-term and short-term goals. *The Journal of Special Education, 20*(1), 69-82.

Fuchs, L. S., & Fuchs, D. (1991). Curriculum-based measurements. *Preventing School Failure, 35*(3).

Ganley, K. (2000). Quality goals, objectives, benchmarks. *ASHA Leader, 5*(18), 1-3.

Gansle, K. A., & Noell, G. H. (2008). Consulting with teachers regarding academic

skills: Problem solving for basic skills. *International Journal of Behavioral*

Consultation and Therapy, 4(2), 199-211.

Gansle, K. A., VanDerHeyden, A. M., Noell, G. H., Resetar, J. L., & Williams, K. L.

(2006). The technical adequacy of curriculum-based and rating-based measures of

written expression for elementary school students. *School Psychology Review,*

35(4), 435-450.

Garcia, T. (2007). Facilitating the reading process: A combination approach. *TEACHING*

Exceptional Children, 39(3), 12-17.

Gartin, B. C., & Murdick, N. L. (2008). Individualized education program. In E.L.

Grigorenko (Ed.), *Educating individuals with disabilities: IDEIA 2004 and*

beyond (pp. 337-359). New York: Springer Publishing Company.

Glesne, C. (2006). *Becoming qualitative researchers: An introduction* (3rd ed.). Boston:

Pearson.

Hammond, H., Casteneda, R., & Ortega, R. (2006). Taking a child's IEP goals home. *EP*

Magazine, 26-32.

Hawkins, L. J., & Riley, M. N. (2008). Local educational authorities and IDEIA. In E.L.

Grigorenko (Ed.), *Educating individuals with disabilities: IDEIA 2004 and*

beyond (pp. 403-420). New York: Springer Publishing Company.

Hessler, T., & Konrad, M. (2008). Using curriculum-based measurement to drive IEPs

and instruction in written expression. *TEACHING Exceptional Children, 41*(2),

28-37.

Hexom, D., Menoher, J., Plummer, B. A., & Stone, M. (2008). Increasing the academic

    performance of struggling students and students with disabilities in

    underperforming schools. In E. L. Grigorenko (Ed.), *Educating individuals with*

    disabilities: IDEIA 2004 and beyond (pp.383-402). New York: Springer

    Publishing Company.

Horn, E., Lieber, J., Li, S., Sandall, S., & Schwartz, I. (2000). Supporting young

    children's IEP goals in inclusive settings through embedded learning

    opportunities. *Topics in Early Childhood Special Education, 20*(4), 208-223.

Hosp, M. K., & Fuchs, L. S. (2005). Using CBM as an indicator of decoding, word

    reading, and comprehension: Do the relations change with grade? *School*

    Psychology Review, 34(1), 9-26.

Individuals with Disabilities Educational Improvement Act of 2004, PL108-446, 20

    U.S.C. § 1400 *et seq.*

Jameson, J. M., McDonnell, J., Johnson, J. W., Riesen, T., & Polychronis, S. (2007). A

    comparison of one-to-one embedded instruction in the general education

    classroom and one-to-one massed practice instruction in the special education

    classroom. *Education and Treatment of Children, 30*(1), 23-44.

Jenkins, A. A., Pateman, B., & Black, R. S. (2002). Partnerships for dual preparation in

    elementary, secondary, and special education programs. *Remedial and Special*

    Education, 23(6), 359-371.

Jenkins, J. R., Graff, J. J., & Miglioretti, L. (2009). Estimating reading growth using

    intermittent CBM progress monitoring. *Exceptional Children, 75*(2), 151-163.

Johns, B. H., Crowley, E. P., & Guetzloe, E. (2002). Planning the IEP for students with

emotional and behavioral disorders. *Focus on Exceptional Children, 34*(9), 1-12.

Johnson, J. W., McDonnell, J., Holzwarth, V. N., & Hunter, K. (2004). The efficacy of

embedded instruction for students with developmental disabilities enrolled in

general education classes. *Journal of Positive Behavior Interventions, 6*(4), 214-

227.

Jones, C. J. (2001). Teacher-friendly curriculum-based assessment in spelling.

TEACHING Exceptional Children, 34(2), 32-38.

Kanne, S. M., Randolph, J. K., & Farmer, J. E. (2008). Diagnostic and assessment

findings: A bridge to academic planning for children with autism spectrum

disorders. *Neuropsychological Review, 18*, 367-384.

Kaufman, A. K. (2008). Policy and law of Individuals with Disabilities Education

Improvement Act of 2004: Attempting no student left behind to the extent

enforceable. In E. L. Grigorenko (Ed.), *Educating individuals with disabilities:

IDEIA 2004 and beyond* (pp.39-59). New York: Springer Publishing Company.

Keenan-Takagi, K. (2000). Embedding assessment in choral teaching. *Music Educators

Journal, 86*(4), 4-12.

Keilty, B., LaRocco, D. J., & Casell, F. B. (2009). Early interventionists' reports of

authentic assessment methods through focus group research. *Topics in Early

Childhood Special Education, 28*(4), 244-256.

Kohler, Henning, & Usma-Wilches (2008). Preparing preservice teachers to make

instructional decisions: An examination of data from the teacher work sample.

Teaching and Teacher Education 24, 2107-2117.

Kowalski, E., Lieberman, L., Pucci, G., & Mulawka, C. (2005). Implementing IEP or 504

goals and objectives into general physical education: What does the physical

educator do when a student with a disability has goals and objectives that differ

from the class goals and objectives? *The Journal of Physical Education,*

Recreation, & Dance, 33(5), 1-8.

Kowalski, E., McCall, R., Aiello, R., Lieberman, L. (2009). Effectively using IEP goal

banks: A web-based goal-bank program can significantly improve the creation

and maintenance of an IEP. *The Journal of Physical Education, Recreation, &*

Dance, 44(6), 1-9.

Lapadat, J. C., & Lindsay, A. C. (1998, April). *Examining transcription: A theory-laden*

methodology. Paper presented at the Annual Meeting of the American

Educational Research Association, San Diego, CA.

Levine, E. L., & Wexler, E. M. (1981). *PL94-142: An act of congress.* New York:

Macmillan Publishing Co, Inc.

Lincoln, Y. S. & Guba, E. G. (1985). *Naturalistic inquiry.* Newbury Park, CA: SAGE

Publications.

Lipson, M. Y., & Wixson, K. K. (2003). *Assessment & instruction of reading and writing*

difficulty: An interactive approach (3rd ed.). Boston, MA: Pearson Education, Inc.

Luckasson, R. (2006). The human rights basis for student personal empowerment in

education. In E. B. Keefe, V. M. Moore, & F. R. Duff (Eds.), *Listening to the*

experts: Students with disabilities speak out (pp. 11-20). Baltimore, MD: Paul H.

Brookes Publishing Co.

Maclellan, E. (2004). Initial knowledge states about assessment: Novice teachers

conceptualizations. *Teaching and Teacher Education 20,* 523-535.

Macy, M. G., & Bricker, D. D. (2007). Embedding individualized social goals into

routine activities in inclusive early childhood classrooms. *Early Child*

Development and Care, 177(2), 107-120.

Madelaine, A., & Wheldall, K. (2004). Curriculum-based measurement of reading:

Recent advances. *International Journal of Disability, Development, and*

Education, 51(1), 57-82.

Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.).

Thousand Oaks, CA: SAGE Publications.

McBride, B., & Schwartz, I. S. (2003). Effects of teaching early interventionists to use

discrete trials during ongoing classroom activities. *Topics in Early Childhood*

Special Education, 23(1), 5-17.

McCombes-Tolis, J., & Feinn, R. (2008). Comparing teachers' literacy-related

knowledge to their states standards for reading. *Reading Psychology, 29,* 236-

265.

McDonnell, J., Thorson, N., Disher, S., Mathot-Buckner, C., Mendel, J., & Ray, L.

(2003). The achievement of students with developmental disabilities and their

peers without disabilities in inclusive settings: An exploratory study. *Education*

and Treatment of Children, 26(3), 224-236.

McGuinn, P. (2012). Stimulating reform: Race to the top, competitive grants and the

Obama education agenda. *Educational Policy, 26*(1), 136-159.

McMaster, K., & Espin, C. (2007). Technical features of curriculum-based measurement

in writing: A literature review. *The Journal of Special Education, 41*(2), 68-84.

McMillan, J.H., & Schumacher, S. (2006). *Research in education: Evidence-based*

inquiry (6th ed.). Boston: Pearson.

McNicholas, J. (2000). The assessment of pupils with profound and multiple learning

    difficulties. *British Journal of Special Education, 27*(3), 150-153.

Mertler, C. A. (1999). Assessing student performance: A descriptive study of the

    classroom assessment practices of Ohio teachers. *Education, 120*(2), 285-296.

Mertler, C. A. (2005). Secondary teachers' assessment literacy: Does classroom

    experience make a difference? *American Secondary Education, 33*(2), 76-92.

National Commission on Excellence in Education. (1983). *A nation at risk: The*

    imperative for education reform. Retrieved from http://www2.ed.gov

    /pubs/NatAtRisk/risk.html

National Research Council of the National Academies. (2008). *Early childhood*

    assessment: Why, what, and how. (C. E. Snow & S. B. Van Hemel, Eds.).

    Washington, D.C.: The National Academies Press.

No Child Left Behind Act of 2001, PL 107-110, 20 U.S.C. § 6301 *et seq.*

Ochs, E. (1979). Transcription as theory. In E. Ochs & B.B. Schieffelin (Eds.)

    Developmental Pragmatics (pp.43-72). New York: Academic Press.

O'Connor, E. A., & Yasik, A. E. (2007). Using information from an early intervention

    program to enhance literacy goals on the individualized education program (IEP).

    Reading Psychology, 28, 133-148.

Parrish, P. R., & Stodden, R. A. (2009). Aligning assessment and instruction with state

    standards for children with significant disabilities. *TEACHING Exceptional*

    Children, 41(4), 46-56.

Polychronis, S. C., McDonnell, J., Johnson, J. W., Riesen, T., & Jameson, M. (2004). A

comparison of two trial distribution schedules in embedded instruction. *Focus on Autism and Other Developmental Disabilities, 19*(3), 140-151.

de Ramirez, R. D., & Shapiro, E. S. (2006). Curriculum-based measurement and the evaluation of reading skills of Spanish-speaking English language learners in bilingual education classrooms. *School Psychology Review, 35*(3), 356-369.

Riesen, T., McDonnell, J., Johnson, J. W., Polychronis, S., & Jameson, M. (2003). A comparison of constant time delay and simultaneous prompting within embedded instruction in general education classes with students with moderate to severe disabilities. *Journal of Behavioral Education, 12*(4), 241-259.

Roehrig, A. D., Duggar, S. W., Moats, L., Glover, M., Mincey, B. (2008). When teachers work to use progress monitoring data to inform literacy instruction. *Remedial and Special Education, 29*(6), 364-382.

Schenck, S. J. (1980). The diagnostic/instructional link in individualized education programs. *The Journal of Special Education, 14*(3), 337-345.

Schmitt, Carol S. (2000). Assessment, educational. In *Encyclopedia of special education: A reference for the education of the handicapped and other exceptional children and adults (2ⁿᵈ ed).* (Vol. 1, pp. 155-158). New York: John Wiley & Sons.

Shapiro, E. S., & Ager, C. (1992). Assessment of special education students in regular education programs: Linking assessment to instruction. *The Elementary School Journal, 92*(3), 283-296.

Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers and the social Sciences* (3ʳᵈ ed.). Teachers College Press.

Sigafoos, J., O'Reilly, M., Ma, C. H., Edrisinha, C., Cannella, H., & Lancioni, G. E.

(2006). Effects of embedded instruction versus discrete-trial training on self-injury, correct responding, and mood in a child with autism. *Journal of Intellectual & Developmental Disability, 31*(4), 196-203.

Smith, E., & Gorard, S. (2005). 'They don't give us our marks': The role of formative feedback in student progress. *Assessment in Education, 12*(1), 21-38.

Smoot, S. L. (2004). An outcome measure for social goals of inclusion. *Rural Special Education Quarterly, 23*(3), 15-22.

Spinelli, C. G. (2008). Addressing the issue of cultural and linguistic diversity and assessment: Informal evaluation measures for English language learners. *Reading & Writing Quarterly, 24,* 101-118.

Spradley, J. P. (1979). *The ethnographic interview.* Belmont, CA: Wadsworth Group.

Stecker, P. M., Lembke, E. S., & Foegen, A. (2008). Using progress-monitoring data to improve instructional decision making. *Preventing School Failure, 52*(2), 48-58.

Stecker, P. M., & Fuchs, L. S. (2000). Effecting superior achievement using curriculum-based measurement: The importance of individual progress monitoring. *Learning Disabilities Research & Practice, 15*(3), 128-134.

Tomlinson, C.A. (2009). Learning to love assessment: From judging performance to guiding students to shaping instruction to informing learning, coming to grips with informative assessment is one insightful journey. *Annual Editions: Education 09/10, 36,* 15-18.

Trent, J. A., Kaiser, A. P., & Wolery, M. (2005). The use of responsive interaction strategies by siblings. *Topics in Early Childhood Special Education, 25*(2), 107-118.

Tymitz, B. L. (1981). Teacher performance on IEP instructional planning tasks. *Exceptional Children, 48*(3), 258-260.

United Nations. (1948). *Universal declaration of human rights.* Retrieved from http://www.un.org/en/documents/udhr/index.shtml#a26

UNM College of Education. (2013). *Special education (pre K-12) and elementary education (k-8) dual licensure program.* Retrieved November 4, 2013 from http://coe.unm.edu /index.php/departments/alias-7/special-education /special-education-degree-and-certificate-programs.html

U. S. Const. amend. XIV.

U. S. Department of Education. (1991). *America 2000: An education strategy.* (ERIC Document Reproduction Service No. 327009). Retrieved from http://www.eric.cd.gov

VanDerHeyden, A. M., Snyder, P., Smith, A., & Longwell, J. (2005). Effects of complete learning trials on child engagement. *Topics in Early Childhood Special Education, 25*(2), 81-94.

Wallace, T., Espin, C. A., McMaster, K., Deno, S. L., & Foegen, A. (2007). CBM progress monitoring within a standards-based system: Introduction to the special series. *The Journal of Special Education, 41*(2), 66-67.

Walsh, J. M. (2001). Getting the "Big Picture" of IEP goals and state standards. *TEACHING Exceptional Children, 33*(5), 18-26.

Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41*(2), 85-120.

Weishaar, M. K. (2008). The law and reality: Understanding the Individuals With
   Disabilities Education Improvement Act. In E.L. Grigorenko (Ed.), *Educating*
   individuals with disabilities: IDEIA 2004 and beyond (pp. 63-83). New York:
   Springer Publishing Company.

Wertz, M. G., Wolery, M., Venn, M. L., Demblowski, D., & Doren, H. (1996). Effects of
   transition-based teaching with instructive feedback on skill acquisition by
   children with and without disabilities. *The Journal of Educational Research,*
   90(2), 75-86.

Wolery, M., & Anthony, L. (1997). Training elementary teachers to embed instruction
   during classroom activities. *Education & Treatment of Children, 20*(1), 3- 22.

Wolery, M., Anthony, L., Caldwell, N. K., Snyder, E. D., & Morgante, J. D. (2002).
   Embedding and distributing constant time delay in circle time and transitions.
   Topics in Early Childhood Special Education, 22(1), 14-25.

Wolery, M., Brashers, M. S., & Neitzel, J. C. (2002). Ecological congruence assessment
   for classroom activities and routines: Identifying goals and intervention practices
   in childcare. *Topics in Early Childhood Special Education, 22*(3), 131-142.

Yell, M.L. (2006). *The law and special education* (2nd ed.). Upper Saddle River, NJ:
   Pearson.

Ysseldyke, J., & Algozzine, B. (2006). *Effective assessment for students with special*
   needs: A practical guide for every teacher. Thousand Oaks, CA: Corwin Press.

Yun, J., Shapiro, D., & Kennedy, J. (2000). Reaching IEP goals in the general physical
   education class. *Journal of Physical Education, Recreation, & Dance,, 71*(8), 33-
   37.

Zeece, P.D., Graul, S. K., & Hayes, N. (2004). Stories for all children: The use of

    literature in inclusive early childhood classrooms. *Early Childhood Education*

    Journal, 31(4), 255-260.