UNIVERSITY
OF OSLO

Master Thesis

Dynamic Assessment of Diverse Learners' Language Abilities: A Literature Review

Elizabeth B. Talbott

Master's Thesis in Special Needs Education

40 Credits

University of Oslo

Department of Special Needs Education

Faculty of Educational Science

Spring 2023

2023

Dynamic Assessment of Diverse Learners' Language Abilities: A Literature Review

Elizabeth B. Talbott

http://www.duo.uio.no/

Abstract

The assessment of culturally and linguistically diverse (CLD) children's language abilities is an area that requires greater consideration and development in the field of speech and language pathology. CLD students are often misrepresented as having a language disorder based on their standardized assessment performance. Instead, language difficulties in the classroom may reflect differences in language experience and not impairment in underlying linguistic ability. Dynamic assessment, which integrates teaching into the testing procedure, has been proposed as a less culturally and linguistically biased form of assessment. This study used a mixed-method systematic review to examine dynamic assessment methodology and the role of learner responsiveness for evaluating diverse learners' language skills.

Following the PRISMA framework, ten studies were identified which met the inclusion criteria through keyword searches on ERIC, PsycInfo, and PubMed. The final articles were assessed using the QualSyst appraisal checklist. The results supported the integration of teaching into the assessment procedure through a test-teach-retest approach. Two main variations were present in the current research: the graduated prompting approach and the mediated learning experience approach. Findings supported the significance of children's responsiveness to teaching as an important indicator for identifying CLD children with and without language impairment, although in-depth information regarding how mediators elicited this responsiveness was largely missing. Further qualitative research in the form of case studies and longitudinal research were recommended. The results of this review can contribute to the development of new methods of language assessment to better represent diverse populations.

Acknowledgments

This work is dedicated to my past students who shared their beliefs and cultures with me and inspired me to think beyond standard practice.

Thank you to my husband, Stian, for encouraging me to pursue this master's. And to my two children, Talina and Henrik, for putting it all into perspective.

Finally, thank you to my thesis supervisors, Jørgen and Luca, for the support along the way.

Table of Contents

Total Words: 17,526

List of Tables and Figures

List of Abbreviations

BAS II	British Ability Scales II
BPVS	British Picture Vocabulary Scale
CLD	Culturally and Linguistically Diverse Learners
DA	Dynamic Assessment
ERIC	Education Resources Information Center
EAL	English as an Additional Language
E/ROWPVT	Expressive/Receptive One-Word Picture Vocabulary Test
MLE	Mediated Learning Experience
MLO	Mediated Learning Observation
PRISMA	Preferred Reporting items for Systematic Reviews and Meta-Analysis
SES	Socioeconomic Status
SLI	Specific Language Impairment
SLP	Speech and Language Pathologist
ZPD	Zone of Proximal Development

1. Introduction

Prior to beginning my studies in Norway, I worked as a speech-language pathologist (SLP) at an international secondary school in London. As a white, English-speaking American, my cultural and linguistic background was often different from the students I worked with, who came from countries from all over the world and whose mother tongue was not English. Many of these students were referred for speech and language assessment due to concerns regarding their use and understanding of language in the classroom. One of my tasks was to determine whether language difficulties in the classroom were due to a lack of underlying language skills or due to a lack of experience with the majority language. For students from linguistically and culturally diverse backgrounds, this issue of differentiation can lead to misdiagnosis of language difference as language disorder (Moore-Brown et al., 2006; Peña et al., 2001). The challenge I faced was how to assess a child's language ability when their first language and cultural background was different from my own.

Professional bodies such as the American Speech-Language-Hearing Association (2004) and the Royal College of Speech Language Therapists (2007) advise the assessment of all the languages spoken by a multilingual child to evaluate for language impairment. While assessment of the child's first language would occur in an ideal setting, the reality that most SLPs face is limited access to translators and appropriate assessment materials for culturally and linguistically diverse (CLD) learners. As a result, CLD learners are often tested in the society's majority language (Williams & McLeod, 2012), which is typically a language that the child is in the process of developing. My own experience using standardized language assessment was that the results on paper did not match the learning profile of the student in front of me. While I knew to interpret these test scores with caution, it left me eager to find new methods of assessment to capture the language abilities of my students more accurately. This is what ultimately led me to discover dynamic assessment as an alternative form of assessment for CLD students.

1.1 The Research Problem

Speech-language pathologists are trained to use standardized language assessments, which measure an individual's current level of language ability by comparing their performance on an assessment to their peers' performance. The problem, however, is that these assessments are influenced by cultural and linguistic bias (Bedore & Peña, 2008; Paradis et al., 2011).

While countries like the United States continue to grow in ethnic and cultural diversity, standardized testing practices have not grown alongside these changing demographics to include culturally and linguistically diverse (CLD) populations in their normative samples (Laing & Kamhi, 2003). Multilingual and multicultural students may perform poorer on these tests due to lack of educational opportunity or because they have different learning experiences compared to their majority peers (Gutiérrez-Clellen & Peña, 2001). As a result, CLD students are vulnerable to misdiagnosis by professionals in special education (Hunt et al., 2021; Peña et al., 2001).

In the field of speech-language pathology, there is need for alternative assessment methods for evaluating language skills in student populations that continue to grow in cultural and linguistic diversity.

1.2 Aim and Research Questions

Current research suggests that dynamic assessment, which integrates teaching into assessment, is a clinically useful tool for diagnosing language disorder in children with CLD backgrounds (Hunt et al., 2021, Orellana et al. 2019, Peña et al., 2001). The aim of this research is to investigate further the methodology behind dynamic assessment as a means of differentiating between language difference and language disorder in CLD children. It is hoped that this research will provide speech-language pathologists (SLPs) with practical information on how instruction and assessment can be integrated to derive meaningful information about a child's learning potential and minimize misdiagnosis of language disorder in diverse learners.

The following research questions will be addressed:

1. What are the methods and procedures used by the current dynamic assessment studies for assessing CLD children's language abilities?
2. How is learner responsiveness during dynamic assessment procedures elicited and observed?
3. How can SLPs use dynamic assessment methods in their decision-making process when assessing children from CLD backgrounds?

1.3 Terminology

Key concepts will be explained in this section to ensure that the reader has a clear understanding of the different terminology that will be used throughout the paper.

1.3.1 Dynamic Assessment

Dynamic assessment is an umbrella term for alternative approaches to testing which incorporates teaching into assessment through some form of mediation (e.g., teacher feedback, prompting, etc.). Most dynamic approaches to testing follow a pretest-intervention-posttest format, while some may move directly into intervention (Lidz, 2002). Dynamic assessment focuses on the learning process rather than the learning outcome and may provide useful information about the child's potential for learning when optimal learning conditions are provided (Elliott et al., 2010).

1.3.2 Language Disorder

A language disorder is defined as "impaired comprehension and/or use of spoken or written language" (American Speech-Language-Hearing Association, 1993). The disorder may affect the form (i.e., phonology, morphology, or syntax), content (i.e., semantics), or function (i.e., pragmatics) of communication (American Speech-Language-Hearing Association, 1993). A child who has a true language disorder will have a core language deficit in all of the languages they are exposed to and use (Hasson et al., 2012). The term *language impairment* is also used in this paper synonymously.

1.3.3 Language Difference

A language difference is a variation in language used by a group of individuals that reflects shared regional, social, or cultural factors (American Speech-Language-Hearing Association, 1993). A child whose language learning is different due to cultural or linguistic factors should not be viewed as having a language disorder, but rather "the result of the normal process of second language acquisition, and its impact on the development of the second language" (Prezas & Jo, 2017, p. 5).

1.3.4 Culturally and Linguistically Diverse Learners

With regard to this review, culturally and linguistically diverse (CLD) learners will refer to children whose home environment language is different from the language spoken in school and whose cultural background may differ from the mainstream culture (Laing & Kamhi,

2003; Gutiérrez-Clellen & Peña, 2001). The terms *bilingual* and *multicultural* are also used in current research to describe this population.

1.4 Outline of the Thesis

This literature review will follow a mixed-methods systematic approach to investigate the current dynamic assessment methodology for evaluating CLD children's language abilities, with emphasis on learner responsiveness during assessment, and how this can guide the diagnostic decision-making process. The first chapter begins with an introduction and explains the aims and research question as well as key terminology. Chapter 2 provides an overview of dynamic assessment, including an introduction to the topic, theoretical background, dynamic testing methods, and a summary of the current literature reviews on this topic. Chapter 3 describes the methodology through a detailed description of the research procedure and data extraction. Results are presented in Chapter 4 in accordance with the three research questions. Chapter 5 provides a discussion of the major findings and limitations. Finally, a conclusion is presented in Chapter 6.

2. Dynamic Assessment- Introduction, Theory, and Methodology

This chapter is divided into five parts. Issues of standardized assessment are first presented before introducing dynamic assessment and its core principles. Next, dynamic assessment's theoretical background is discussed followed by a discussion of its methodology. Finally, previous literature reviews on the topic are considered.

2.1 Issues of Standardized Assessment

This section will present some of the issues of standardized assessment for the evaluation of culturally and linguistically diverse (CLD) children, including issues of testing bias and the identification of language disorders within the CLD population.

2.1.1 Testing Bias

Speech-language pathologists (SLPs) are trained to use standardized language assessments, which measure an individual's current level of language ability by comparing their performance on an assessment to their peers' performance. The problem, however, is that standardized tests are typically norm-referenced with monolingual children (Paradis et al., 2011) and reflect the linguistic and cultural experiences of the majority. Language development milestones do not follow the same order across different languages; therefore, direct translation of standardized language tests from the majority language to another language affects the classification accuracy of the test (Bedore & Peña, 2008). Unfortunately, many SLPs continue to use standardized assessment tools for CLD learners due to a lack of other means of assessment for determining language disorder in this population.

A child's performance on a standardized test is largely dependent on their cultural familiarity with the material as well as their understanding of the task, which may be unfamiliar to them (De Lamo White & Jin, 2011). If the child has limited experience with individualized standardized testing then the test becomes a measure of their test-taking abilities as opposed to an accurate measure of their language abilities (Chamberlain, 2005). Furthermore, language skills such as vocabulary and narrative skills are connected to children's previous learning experiences (Lidz & Peña, 2009). For example, a child who is asked to tell a story about a day at the beach who has limited experience with the beach is likely to perform poorer on this task as they lack background knowledge and content-specific vocabulary. This leads to "test scores reflecting life experiences and socio-economic status rather than

language ability" (Hasson et al., 2012, p. 58). As a result, standardized assessments, which only reflect the educational experiences of the majority population, are inherently biased and lead to poorer outcomes for CLD students.

2.1.2 Diagnosis Issues in CLD Population

Overidentification of language disorders is a well-known issue in the assessment of multilingual and multicultural learners (Cummins, 2000; Gutiérrez-Clellen et al., 2008; Mennen & Stansfield, 2006). Overidentification is "when a dual language child is inappropriately diagnosed with a language or learning disability and receives unnecessary services and/or is inappropriately placed in special education classes" (Paradis et al., 2011, p. 213). As previously stated, CLD students are likely to perform poorer on standardized language assessments due to lack of educational opportunity or because they have different learning experiences compared to their majority peers (Gutiérrez-Clellen & Peña, 2001). As a result, CLD students are at greater risk of misdiagnosis of language disorder when what they require is greater educational opportunity to learn the majority language.

Culturally and linguistically diverse learners are also vulnerable to under identification of a language or learning disability. This occurs when a CLD learner has a true language disorder, but it goes undiagnosed because it is assumed that their poor performance in school is the result of learning two languages (Paradis et al., 2011). Speech-language pathologists and other education professionals may have difficulty recognizing the difference between a child learning English as an additional language versus having a core language impairment in all of the languages that they speak (Hasson et al., 2012). Without appropriate language assessments for multilingual learners, CLD students will continue to be vulnerable to misdiagnosis by professionals in special education (Hunt et al., 2021; Peña et al., 2001).

2.2 Dynamic Assessment

Dynamic assessment has been used as an alternative form of assessment for differentiation between language differences and language disorder in culturally and linguistically diverse populations (Camilleri & Botting, 2013; Hasson et al., 2012; Laing & Kamhi, 2003; Peña et al., 2006). In this section, an overview of dynamic assessment will be given, including the integration of teaching and assessment, the role of learner responsiveness, and how dynamic assessment may be used to differentiate between language disorder versus a difference in language experience.

2.2.1 Integrating Teaching and Assessment

Dynamic assessment incorporates teaching into the assessment procedure in order to optimize the performance of the learner. According to Sternberg and Grigorenko (2002), dynamic assessment is not limited to any single domain (e.g., speech-language pathology, psychology), content (e.g., science, history), activity (e.g., teaching, testing) or age. Rather it is a "family" of procedures that share a set of principles (Sternberg & Grigorenko, 2002). The core principle of dynamic assessment challenges conventional views about the relation between teaching and assessment by arguing that the two should not be regarded as separate activities but rather fully integrated (Lantolf & Poehner, 2004; Poehner, 2008). As Lidz (2003) noted, "there is a certain logic to the assertion that the best way to assess learning is to involve the child in a learning situation" (p. 112).

The integration of teaching and assessment occurs as intervention is embedded within the assessment procedure in order to better understand a child's abilities and guide them to higher levels of functioning (Lidz & Gindis, 2003). By integrating teaching into assessment, the examiner's focus shifts away from the learning product and towards the learning process. In doing so, they are able to better understand a child's level of performance and guide the child from there by supporting and extending existing strategies.

The two most common dynamic assessment formats are described as the "sandwich" format and the "cake" format (Sternberg & Grigorenko, 2002). In the sandwich design, the integrated teaching is given as a separate session in between a standardized pretest and posttest. In the cake design, teaching or mediation is provided during the assessment whenever the learner encounters problems. The cake format is particularly effective for one-to-one testing environments as the examiner can support the learner more readily (Sternberg & Gigorenko, 2002). As previously stated, there are a wide range of dynamic approaches which will be discussed in greater detail in the dynamic assessment methods section.

2.2.2 Learner Responsiveness

One of the key characteristics of dynamic assessment is that it is designed to reinforce learner responsiveness by engaging the child in a shared learning experience:

> As was observed by Lidz (1995), traditional standardized assessment follows the child's cognitive performance to the point of "failure" in independent functioning, whereas dynamic assessment in the Vygotskian tradition leads the child to the point of achievement

of success in joint or shared activity. Dynamic assessment begins where standardized

testing ends. (Lindz & Gindis, 2003, p. 103)

During the intervention or mediation phase of dynamic assessment, the child's learning behaviors (e.g., attention, responsiveness to feedback) are observed carefully by the examiner as this helps inform and guide future instruction (Lantolf & Poehner, 2004). By focusing on learner responsiveness, the examiner is able to identify areas of strength, emerging learning strategies, and strategies the child needs for future learning (Moore-Brown et al., 2006). The concept of learner responsiveness is connected to Vygotsky's zone of proximal development and Feuerstein's mediated learning experience, which will be presented in the theoretical background section of this paper.

2.2.3 Language Disorder or Language Difference

Gathering information about learner responsiveness during dynamic assessment can yield clinically useful information for distinguishing between language disorder or language difference in culturally and linguistically diverse (CLD) children (Hasson et al., 2012; Lidz & Peña, 2009; Moore-Brown et al., 2006). Characteristic of the definition of a language disorder is the "intransient nature of the difficulty", suggesting that "it is not readily modifiable without an investment of intervention" (Hasson et al., 2012). Consequently, CLD learners with language disorder are expected to show less responsiveness to a short teaching session during dynamic assessment as they require more intensive intervention. On the other hand, language difficulties that come from cultural and linguistic difference, or limited experience with the majority language, are more likely to be remediable following short exposure to appropriate language models (Hasson et al., 2012).

Peña and colleagues (1992, 1993) early work with dynamic assessment in CLD populations suggested that children's modifiability- how they responded to teaching and transferred new learning- were important indicators for identifying children with and without language impairment. By examining a child's ability to learn when given appropriate assistance, clinicians may be more likely to effectively differentiate between the effects of language disorder versus language experience when working with diverse groups of students.

2.3 Theoretical Background

In order to understand dynamic approaches to assessment it is important to understand its' origins. The most often recognized roots for dynamic assessment stems from Vygotsky's

sociocultural theory and the zone of proximal development (ZPD), as well as Feuerstein's mediated learning experience (MLE). The concept of intersubjectivity (Rommetveit, 1976) is also presented as it relates to language acquisition.

2.3.1 Vygotsky's Sociocultural Theory and the Zone of Proximal Development

The synthesis of assessment and instruction is grounded in Vygotsky's understanding of development (Poehner, 2008). Vygotsky's (1986) sociocultural theory maintains that it is through social interaction that the child learns to master cognitive function. The learning process is mediated through signs and tools, language as the most powerful tool, to internalize higher psychological functions. According to sociocultural theory, "cognitive, language, and social functioning in educational settings are not innate abilities or disabilities but are sociocultural formations resulting from the interactions of a child with culture" (Lidz & Gindis, 2003, p.100). An essential feature of learning is the concept of the zone of proximal development (Vygotsky, 1978) where the child develops higher psychological functions during shared activities with a more competent partner (Lidz, 1991).

According to Vygotsky (1978), standard IQ testing measures only the "fruits" of development, that which is already matured or learned, it does not test the "buds", what is not yet ripe or learned. The "fruits" that Vygotsky refers to are a child's *actual* level of development, the psychological functions that have matured, whereas the "buds" are their *proximal* level of development, the psychological functions that are in the process of maturing. The ZPD is the distance between the actual level of development and the child's potential level of development. The child's actual level of development can be observed in independent problem-solving tasks, but his or her *proximal* level of development can only be seen through collaboration and guidance from an adult or more capable peer (Vygotsky, 1978).

It is of particular value to acknowledge that Vygotsky's notion of the ZPD has been misinterpreted and oversimplified by many educators as short-term, collaborative activities between teacher and student to encourage independent problem-solving (Smagorinsky, 2018). Instead, it has been argued that the ZPD represents a much deeper and complex long-term process of maturation: "These higher mental functions take many years to develop through extensive experience with a culture's values and practices" (Smagorinsky, 2018, p. 72). What has been lost in translation is the significance of the social system in which the

child and adult collaborate. Moreover, it is the understanding that it is both the teacher and the student who actively and mutually create this social system (Moll & Whitmore, 1993).

2.3.2 IQ Testing Through a Vygotskian Lens

Intelligence tests provide children with a mental age that does not account for the dynamic versatility in their mental functioning. That is, two children who score the same on an IQ test are likely to have notably different capacity for learning. In order to illustrate this point Vygotsky (1978) used the example of two children who had an 8-year-old mental age derived from an IQ test, but who responded differently to adult guidance in problem-solving situations: "Under these circumstances it turns out the first child can deal with problems up to a twelve-year-old's level, the second up to a nine-year-old's. Now, are these children mentally the same" (Vygotsky, 1978, p.86)? Not only are these two children's mental stages of development different, but their paths of learning will also be different. In this regard, IQ tests provide little information about a child's learning potential.

Vygotsky argued that standardized IQ tests wrongly balanced the natural and cultural processes, and, consequently, are "unable to make the differentiation of impaired functioning that can be due to cultural deprivation or can be the result of organic damage" (Gindis, 1999, p. 337). That is, standardized assessments may wrongly attribute a child's poor performance on a standardized test to an organic disability without consideration for the child's sociocultural background. A child who has received less exposure to the ideal form of language will likely score lower on an IQ test in comparison to his or her peers who have received more exposure to the ideal form of language. Who is to say that this child's potential for learning is any less than the child who comes from an enriched language background? Vygotsky's work developing our understanding of the zone of proximal development was cut short after his death in 1934.

2.3.3 Mediated Learning Experience

Feuerstein and his colleagues (1979, 1981) are known for their contributions regarding the specific nature of interactions within the ZPD that facilitate student development, following the assumption that not all interactions are equal (Lidz & Gindis, 2003). Feuerstein's mediated learning experience (MLE) stemmed from his work helping socialize and integrate children who immigrated from Europe and North Africa to Israel. He found that there was a discrepancy between what he believed the children were capable of academically and socially

versus what his test results showed (Lidz, 2002). Feuerstein adapted static testing measures to include an interactive component, known later as MLE, which attempted to guide the child in their understanding while observing their responsiveness to the interactions (Lidz, 2002).

Feuerstein's findings led to his MLE theory that poor cognitive functioning is the result of a lack of mediation, marking a departure from the idea that intelligence is fixed or static (Deutsch, 2003; Grosser & Waal, E., 2008). Following in Vygotsky's footsteps, Feuerstein shifted focus away from the numeric measurement of a child's performance to an exploration of *how* children learn to use and apply cognitive strategies when learning is mediated. Feuerstein argued that cognitive abilities are modifiable and can be developed in a range of ways depending on the quality and type of instruction the individual receives (Poehner, 2008). This idea is now known as Feuerstein's Structural Cognitive Modifiability theory.

Mediation is not simply offering the learner assistance, but rather providing appropriate assistance: assistance that focuses not on getting to the correct solution, but on assisting the learner toward independent problem-solving (Poehner & Lantolf, 2010). Feuerstein and colleagues (1991) described three key attributes that are necessary to create an MLE:

> (1) *Intentionality and reciprocity*: The mediator makes an intentional effort to mediate and guides the child's attention e.g., "I want you to hear this, therefore I say it loudly". (Feuerstein et al., 1991).
>
> (2) *Transcendence*: The mediator moves beyond the immediate goal of the interaction to help the child make connections between ideas and events e.g., relating a story-telling activity to the child's home and school activities (Peña et al., 2006).
>
> (3) *Mediation of meaning*: The mediator emphasizes the parts of the activity that are significant, mediating the "why" and "what for" (Feuerstein et al., 1991).

Within a mediated learning experience, the roles of examiner-examinee shift to teacher-student in which both individuals are working towards a common goal of development (Feuerstein et al., 1979).

2.3.4 Rommetveit and Intersubjectivity

The Norwegian psychologist, Ragnar Rommetveit, maintained that language is a purely social phenomenon which must be examined within the context of human interaction (Rommetveit, 1976). Rommetveit argued that a child's verbal communication can only be appropriately assessed "if we venture to explore the architecture of intersubjectivity within

which it is embedded" (Rommetveit, 1976, p. 93). Communication relies on different "states of intersubjectivity" in which the speaker and the listener have shared knowledge and pre-made assumptions about the social world they live in (Rommetveit, 1976). That is, a word or phrase can carry different meaning depending on the social context. By establishing joint attention, intersubjectivity between two people can be attained (Rommetveit, 1998 as cited in Kowal & O'Connell, 2016).

The establishment of shared perspective between teacher and student is of particular importance when working with culturally and linguistically diverse children. Rommetveit (1980) emphasized that a fundamental part of communicative competence in a diverse world is "our capacity to adopt the perspective of *different* 'others'" (Rommetveit, 1980, p. 108 as cited in Kowal & O'Connell, 2016). Rommetveit maintained that this "sharing of consciousness" relies upon both speaker and listener making reciprocal adjustments in their perspective-setting and perspective-taking which are culturally and linguistically bound (Rommetveit, 1992 as cited in Kowal & O'Connell, 2016). Creating reciprocity and intersubjectivity within the assessment environment is, therefore, a dynamic process which depends greatly on the quality of the relationship between the teacher and learner.

2.4 Dynamic Assessment Methods for Language Assessment

Speech-language pathologists carry out individual language assessments to create intervention plans and then monitor the child's progress. Therefore, dynamic assessment methods, which link assessment and intervention, are particularly applicable to speech-language pathology. In this section, an overview of the approaches to dynamic assessment that have been used for language assessment will be discussed.

2.4.1 Test-Teach-Retest Approach

The test-teach-retest approach is the predominate dynamic assessment approach for evaluating language abilities and has been used, in particular, for the identification of language impairment in CLD students (Lidz & Peña, 2009). This approach originated in the field of cognitive and educational psychology (Budoff, 1974; Feuerstein, 1979). Lidz (1991) adapted Feuerstein's (1979) mediated learning experience (MLE) to develop a pretest-mediation-posttest format, which Peña (1993) used as a guide to evaluate language learning differences between mediated and non-mediated preschool children with varying language abilities.

Following the test-teach-retest approach, the examiner first administers a pretest (standardized or non-standardized) with no support to identify target skills that the child requires more mediation from an adult to develop (Elliot et al., 2010). Next, the examiner provides a teaching phase (mediated learning experience) to target areas of difficulty and modify the child's level of functioning. Through teaching the principles of the task, "the test situation changes from an *evaluative* interaction (typical of traditional test situations) to a *teaching* interaction where the examiner maximizes the child's feelings of competence" (Gutiérrez-Clellen & Peña, 2001, p. 215). Finally, the examiner administers a posttest, either similar to or identical to the pretest, to determine how much improvement the child made as a result of mediation (Poehner, 2008).

2.4.2 Graduated Prompting Approach

The graduated prompting approach (See Brown & Ferrara, 1985; Campione & Brown, 1987) has been used in assessment to gather information about children's learning potential using a hierarchy of prompts from least to most supportive. During the administration of a test, the examiner uses a "menu of standardized hints and leading questions" after each item or problem (Poehner, 2008, p. 51). The child's learning efficiency is then measured by "the number of hints required for the attainment of the learning criterion" (Brown & Ferrara, 1985, p. 82). After identifying the target skill, a hierarchy of prompts are created alongside a scoring system to track progress. Whereas the test-teach-retest approach focuses on maximizing task performance, the graduated prompting approach focuses on the amount of assistance required for the learner to efficiently achieve a prespecified outcome (Elliott et al., 2010).

The use of Brown's graduated prompting procedure has been used in language assessment (Bain & Olswang, 1995; Olswang & Bain, 1996). Bain and Olswang (1995) explored whether a graduated prompting procedure could help determine young children's readiness for producing two-word utterances. The graduated prompting procedure followed a hierarchy of six verbal cues, from least supportive (e.g., general statements) to most supportive (e.g., direct language models). If the child failed to respond to the least supportive cue, the examiner used a more explicit cue until the child correctly responded. The child's responses were scored on a 1 to 6 scale based on the amount of support required (Bain & Olswang, 1995).

Results showed that those children who demonstrated responsiveness to the cuing system exhibited greater mean length utterance increase over the 9-week treatment period. This finding indicates that children's responsiveness to a cuing hierarchy may have good predictive value for determining their readiness for language intervention and the selection of appropriate target skills (Bain & Olswang, 1995). Graduated prompting may also provide information regarding which instructional techniques are most effective for the child (Laing & Kamhi, 2003), which can guide future teaching and learning.

2.4.3 Feedback and Verbalization

Standardized testing procedures have been modified to include feedback during the testing experience (Carlson & Wiedl, 1978, 1992). Carlson and Wiedl (1978, 1992) modified cognitive and academic testing measures by incorporating elaborate feedback and verbalization techniques during testing with children from diverse cultural and linguistic backgrounds. Elaborate feedback was defined as feedback about the accuracy of the child's answers in combination with an explanation of the correct response. Verbalization techniques had the child describe the test question in their own words and then explain how they found their answer. The use of feedback and verbalization strategies during testing of CLD children has led to improved performance on posttest measures and has been connected to a decrease in testing anxiety (Carlson & Wiedl, 1978, 1992).

Verbalization techniques have also been incorporated into dynamic testing procedures following clinical interview methods (Ginsburg, 1997). Ginsburg (1997) argued that standardized tests are inadequate for understanding children's underlying cognitive processes. In particular, "they usually do not elicit subject's verbalizations concerning strategies and methods of solution – verbalizations which might provide considerable insight into children's thinking" (Ginsburg, 1997, p.15). During the clinical interview a child is presented with a problem or a set of tasks and are simultaneously asked questions by the examiner to encourage their understanding of the task. Children are asked questions such as "How did you know that?" or "what would happen if?" (Gutiérrez-Clellen & Peña, 2001). In contrast to the elaborate feedback approach of Carlson & Wiedl (1978, 1992), the interviewer is not concerned with the correctness of the child's answer, but rather the thought processes behind the child's response. This method has been used with CLD learners as an alternative method for demonstrating their language knowledge (Peña, 2001, Ginsburg, 1997).

2.4.4 Modifiability Scales

Dynamic assessment proponents have argued that pretest to posttest change alone cannot account for the individual differences between learners (Lidz & Peña, 2009). Namely, it is difficult to establish how much of a posttest gain is significant as "a very small positive change may be highly significant for an individual child" (Lidz, 2002, p. 123). Research has found that a child's modifiability- their metacognitive skills (i.e., self-awareness) and flexibility (i.e., response to feedback)- is a good indicator of their future language development (Peña et al, 2007).

Lidz (1991, 2002) created modifiability scales using Likert-type ratings to examine the child's overall responsiveness to mediation. The Response to Mediation Scale (Lidz, 1991) is designed to capture the learner's behavior during dynamic assessment related to the areas of (1) self-regulation, (2) strategic problem-solving, (3) active learning, and (4) representational thinking. This scale can be used in conjunction with The Mediated Learning Experience Rating Scale (Lidz, 1991 based on Feuerstein et al., 1979, 1980) which captures the examiner's experience. These scales have been modified and used in dynamic assessment research within the field of speech-language pathology (Gutiérrez-Clellen & Peña, 2001; Kapantzoglou et al., 2012; Peña et al., 2014).

The Response to Mediation Scale and The Mediated Learning Experience Rating Scale (Lidz, 1991) are presented in Appendix A and B.

2.5 Previous Literature Reviews

There are two previous literature reviews that have examined the topic of dynamic assessment's use for diagnosing language disorder in culturally and linguistically diverse (CLD) children.

Orellana and colleagues (2019) performed the first meta-analysis of dynamic assessment accuracy for English-speaking bilinguals. They reviewed six quantitative studies from the USA with a majority of Spanish-English speaking bilingual children. Their aim was to evaluate the diagnostic accuracy of dynamic assessment for language impairment within bilingual populations. Orellana et al. (2019) findings showed that gain scores between pre- and posttest were not generally useful for differentiating between bilingual children with and without language impairment. However, their analysis did show that bilingual children's

modifiability ratings during the teaching phase of dynamic assessment "consistently yielded significant group effects favoring the typically developing children" (Orellana et al., 2019). That is, children who showed greater responsiveness to mediated learning were less likely to present with language impairment. Issues of methodological quality were acknowledged regarding participant selection and risk of bias.

Following Orellana et al. (2019), Hunt and colleagues (2021) performed the second meta-analysis to evaluate the evidence for the use of dynamic assessment for identifying language disorder in multilingual and multicultural children. Their aim was to include students with a wider range of cultural and linguistic backgrounds than Orellana et al. (2019) in order to reflect a more global population (Hunt et al., 2021). The authors reviewed 10 quantitative studies from the U.S.A, U.K., and Switzerland which used dynamic assessment procedures to identify language disorder in children whose home language was different from the majority language. Their findings (Hunt et al., 2021) showed that nine out of ten studies indicated dynamic assessment as a suitable method of diagnosis of language disorder in CLD students based on improvement in pretest to posttest measures and modifiability scores. Hunt et al. (2021) acknowledged that dynamic assessment of language skills is an emergent area of research and, consequently, there are design issues and small sample sizes found in these studies.

While the above systematic reviews have contributed valuable information regarding the effectiveness of dynamic assessment practices, the primary focus is on the quantitative results. In order for more practitioners to confidently utilize dynamic assessment methods there needs to be reviews which focus on *how* teaching and assessment are integrated to evaluate the language skills of CLD learners. In addition, the role of learner responsiveness needs to be explored more deeply as this is fundamental to dynamic assessment. To close this gap, I will focus on the methodology and procedures of current dynamic assessment studies and examine how learner responsiveness is elicited and observed.

3. Methodology

The methodology chapter is divided into five parts. First, the PRISMA method is described as it is followed to gather data for the literature review. The second section details the literature selection procedure, including the inclusion and exclusion criteria, as well as databases and keywords for searches. Third, methodological quality and risk of bias are presented. Following this, the data extraction procedure is described. Finally, the data analysis procedure is detailed.

3.1 PRISMA Method

A mixed-methods systematic review was selected in order to investigate what the current literature shows regarding dynamic assessment procedures that can be used by speech-language pathologists (SLPs) for identifying language disorders in culturally and linguistically diverse (CLD) populations. The selection of literature and analysis followed the Preferred Reporting items for Systematic Reviews and Meta-Analysis (PRISMA) Statement 2020 (Page et al., 2021). PRISMA is a meta-analyses framework first developed by David Moher and his colleagues (2009) intended for systematic reviews in the field of healthcare. Now PRISMA is widely used to evaluate interventions in the fields of psychology and education and allows for mixed-methods systematic reviews of quantitative and qualitative studies (Page et al., 2021).

Following PRISMA (2020), articles are first identified through relevant database searching using keywords. Duplicate articles are removed from the search results. Next, the articles' abstracts are screened based on the inclusion and exclusion criteria, and articles which do not meet the inclusion criteria are eliminated (Moher et al., 2009). In the third step, eligibility, a full text screening is performed to further exclude articles which do not adhere to the inclusion criteria (Moher et al., 2009). The final studies to be included in the review are then determined.

Please refer to the PRISMA flow diagram in the Results section for further illustration of the PRISMA framework.

3.2 Literature Selection

This section will detail how the literature was selected for this review. The first part presents the inclusion and exclusion criteria, followed by a detailed explanation of the search procedure which includes (1) databases, (2) search keywords and (3) selection of articles.

3.2.1 Inclusion and Exclusion Criteria

Studies which met all of the following criteria were included:

1. Participants:
    a. Are under the age of 12, corresponding to primary or elementary education in the U.S., U.K., or Australia.
    b. Speak a language other than the majority language (English) at home.
    c. Do not have a developmental disorder diagnosis (e.g., autism, intellectual disability, cerebral palsy, hearing loss, etc.).
2. The assessment procedure:
    a. Fits the description of dynamic assessment (Lidz, 2002, 2009; Elliot et al., 2010) and is used to measure any aspect of the child's language ability (e.g., vocabulary, comprehension, narrative skills, etc.).
    b. Includes an intervention, or teaching/learning component, and the change after intervention is evaluated. Comparison groups are not required.
    c. May include quantitative and/or qualitative measures.
3. Studies that have English as the majority language.
4. Studies are from a peer-reviewed journal.
5. Studies are published in 2000 or later.

Studies that met any of the following criteria were excluded:

1. Participants:
    a. Are over the age of 12.
    b. Speak only the majority language (English).
    c. Have a developmental disorder diagnosis, precluding the diagnosis of a specific language impairment or disorder.
2. The assessment procedure:
    a. Does not fit the description of dynamic assessment (Lidz, 2002, 2009; Elliot et al., 2010).

       b.   May fit the description of dynamic assessment but does not measure the child's language ability. Reading/literacy skills are excluded as they relate to a reading disorder diagnosis.

       c.   Does not include an intervention or teaching/learning component.

3. Majority languages other than English.

4. The article is a review or theoretical paper.

5. The studies were done before 2000.

3.2.2 Search Procedure

Databases

A total of three databases, ERIC, PsycInfo, and PubMed were used based on their well-established connection to education research. The search took place over a two-month period from January to February 2023 using the keywords shown in Table 1.

Search Keywords

Three keyword categories were selected from the topic in order to answer the research questions: *culturally and linguistically diverse (CLD) learners, dynamic assessment,* and *language disorder*. After closely examining the definition and scope of the three categories, synonyms and associated terms were identified and used as search keywords. These terms are shown in Table 1.

Table 1. *Search Terms by Category*

Categories	Search Terms
CLD Learners	CLD learners, bilingual, multilingual, multicultural, minority language, minority culture, diverse learners, English language learners (ELLs), English as an additional language (EAL), second language learners
Dynamic Assessment	Dynamic assessment, dynamic testing, dynamic teaching, alternative assessment, mediated learning, sociocultural theory, zone of proximal development, test-teach-retest
Language Disorder	Language disorder, language impairment, learning disability, communication disorder, language delay, language difference, communication difference

3.2.3 Selection of Articles

After identification of articles from the database search, articles were sorted and duplicates were removed. Following PRISMA (2020) framework, article abstracts were screened against the inclusion and exclusion criteria. During this stage articles were removed which came from a country whose majority language was not English. Information regarding participant age and language status was also used to exclude further studies. Articles that passed this stage of screening were then read in full and screened using the inclusion criteria. Following full text screening, a final list of studies was generated to be further analyzed.

3.3 Methodological Quality

The articles selected for this review were assessed for quality using guidelines from QualSyst (Kmet et al., 2004) quality assessment of quantitative and qualitative studies. The included quantitative studies were assessed using Kmet (2004) *Checklist for Assessing the Quality of Quantitative Studies* (refer to Appendix E), and the included qualitative studies were assessed using Kmet (2004) *Checklist for Assessing the Quality of Qualitative Studies* (Refer to Appendix F). Each study was reviewed and scored: a strong quality score was a score > 80%, good quality was a score between 70 to 79%, fair quality was a score between 50 to 69%, and poor quality was a score <50%. Quality assessment of the included articles is shown in Table 2 of the results section.

3.3.1 Risk of Bias

The included articles were rated by one reviewer which increases the risk of personal bias. Ideally, the selected articles should be reviewed by two or more researchers to minimize the risk of bias and increase the strength of the findings.

3.4 Data Extraction

Data extraction procedures were followed on the ten articles selected. Data regarding the participants, aims, methodology (including dynamic assessment strategies), learner responsiveness, and findings were extracted. Participant data is important as it relates to the research questions and whether the selected participants are representative of CLD populations. Information regarding methodology, dynamic assessment strategies, and learner responsiveness corresponded directly to the three research questions. Data related to the studies' aims and findings were also extracted as they provide further information related to

how SLPs may utilize dynamic assessment in their decision-making process (research question 3). The results from data extraction are presented in Table 3 of the Results chapter.
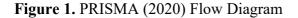
3.5 Data Synthesis

The aim of this review is to synthesize qualitative data in order to investigate the dynamic assessment methods that may be used to guide speech-language pathologists' diagnostic decision-making process when working with children whose first language is different from their own. In qualitative synthesis, the findings of qualitative studies, and sometimes mixed-methods and quantitative research, are combined (Bearman & Dawson, 2013). This review followed a mixed-methods approach, using qualitative synthesis to extract data from qualitative and quantitative research studies in order to answer the three research questions. The results from data synthesis were organized into a table in order to provide a comprehensive summary of the current work in the field (Koons et al., 2019).
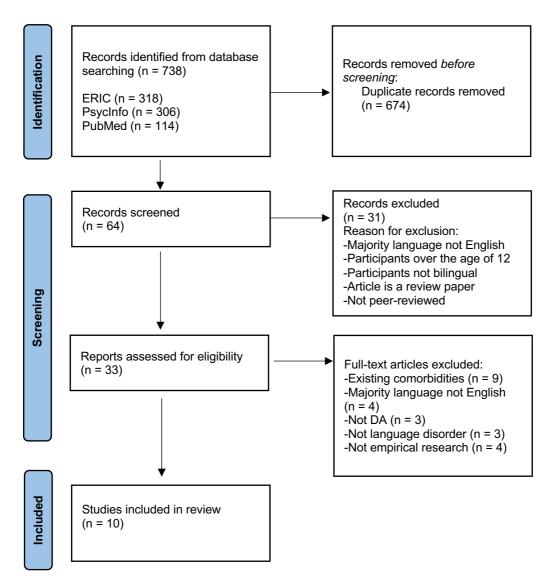
4. Results

Research findings will be presented in this chapter, which is divided into six parts. First, the PRISMA flow chart of research article selection is presented. Next, the results of Kmet (2004) methodological quality assessment for the included studies is shared followed by a summary of the studies' research designs. Finally, the results are presented in accordance with the three research questions.

4.1 Flow Diagram of Study Selection

This literature review followed the PRISMA (2020) framework as the research design. Using the preselected search terms, 738 sources were initially identified from databases ERIC, PsycInfo, and PubMed. Following the PRISMA (2020) framework, duplicates were removed and the remaining 674 articles' abstracts were screened using the inclusion and exclusion criteria. Abstracts which showed majority languages other than English and which did not include bilingual participants or participants under the age of 12 were excluded. Full-text reviews were performed on the remaining 33 articles. It was determined that ten articles met the inclusion criteria and were analyzed for this review article. Refer to Figure 1. PRISMA (2020) Flow Diagram for an illustration of the research article selection procedure.

Figure 1. PRISMA (2020) Flow Diagram

Note. This flowchart is adapted from: Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

4.2 Methodological Quality

Table 2 presents the results of the Kmet (2004) methodological quality assessment for the included research studies. Six of the articles scored between 70 to 79% and are of good methodological quality. Four articles scored below 70% and are of fair quality. No articles scores above 80% and were considered of strong methodological quality. Methodological quality was generally low due to issues of participant selection bias (e.g., lack of random sampling, small sample size, or lack of control group).

Table 2. *Methodological Quality for Included Research Studies*

Reference	Kmet Score (%)	Kmet Methodological Quality
1. Camilleri & Law, 2007	20/28 (71%)	Good
2. Gutiérrez-Clellen & Peña, 2001	15/20 (75%)	Good
3. Hasson et al., 2012	19/28 (68%)	Fair
4. Hemsley et al., 2014	15/20 (75%)	Good
5. Kapantzoglou et al., 2012	18/28 (64%)	Fair
6. Lazewnik et al., 2019	19/28 (68%)	Fair
7. Peña et al., 2001	21/28 (75%)	Good
8. Peña et al., 2014	22/28 (79%)	Good
9. Petersen et al., 2017	19/28 (68%)	Fair
10. Petersen et al., 2020	20/28 (71%)	Good

4.3 Study Design

A total of ten research articles were included in this literature review. A majority (8 articles) were quantitative studies while two were qualitative case studies. All of the quantitative studies applied case-control designs to compare students with suspected or pre-diagnosed language disorder to groups of students with typical language development. Of the children diagnosed with language disorder, none of them had comorbid diagnoses (e.g., hearing loss, autism) which would preclude the diagnosis of a language disorder. Control groups were made up of bilingual students of similar age and language experience. Camilleri and Law

(2007) also included a monolingual group of English-speaking children as a second control group.

Regarding the qualitative case studies, both studies analyzed dynamic assessment techniques for evaluating language skills in culturally and linguistically diverse (CLD) children. One study examined the language skills of two bilingual children with a previous diagnosis of language impairment (Hemsley et al., 2014) and the other evaluated two bilingual children without previous diagnosis (Gutiérrez-Clellen & Peña, 2001).

4.3.1 Participants and Location

Two studies were located in the United Kingdom (Camilleri & Law, 2007, Hasson et al., 2012) and one study was located in Australia (Hemsley et al., 2014). The remainder of the studies were located in the United States. Therefore, the majority language in all of the studies was English. The total number of participants in each study ranged from two (Gutiérrez-Clellen & Peña, 2001; Hemsley et al., 2014) to 79 (Peña et al., 2001). All participants were between the ages of three and nine, with the exception of one 11-year-old participant (Hemsley et al., 2014).

All of the studies observed the use of dynamic assessment for children who spoke more than one language. The participants from the seven studies located in the U.S.A were Spanish-English bilinguals, while the participants from the studies in the U.K. (Camilleri & Law, 2007; Hasson et al., 2012) represented a diverse group of children with a range of home languages. Home languages included Bengali, French, Gujarati, Lingala, Polish, Portuguese, Spanish, Turkish, Twi, and Yoruba (Camilleri & Law, 2007; Hasson et al., 2012). English was also considered a home language for these participants. The case study located in Australia (Hemsley et al., 2014) analyzed the language skills of two children, one of Samoan heritage and the other of Vietnamese heritage.

4.3.2 Study Aims and Language Targets

All of the included studies investigated the use of dynamic assessment approaches as a method to evaluate language skills and identify impairment in CLD learners. Six of the studies assessed vocabulary learning skills (see Table 3), with five of these studies (Gutiérrez-Clellen & Peña, 2001; Kapantzoglou et al., 2012; Lazewnik et al., 2019; Petersen et al., 2020) using similar protocols based on the mediated learning experience procedures of

Peña and colleagues (2001). Camilleri & Law (2007) utilized a graduated prompting approach to assess vocabulary skills.

Narrative language skills were assessed by two studies, using dynamic assessment story retell procedures (Peña et al., 2014; Petersen et al., 2017) and the remaining two studies assessed multiple language skills, including phonology, sentence structure, and vocabulary (Hasson et al., 2012; Hemsley et al., 2014). Further details regarding participants, study aims, and language targets are shown below in Table 3.

4.4 Dynamic Assessment Methods and Procedures

The first research question asks what methods and procedures are being used by current dynamic assessment studies for the assessment of CLD children's language abilities. To answer this question, information regarding methodology and dynamic procedures was extracted and analyzed.

4.4.1 Testing Format and Duration

All of the studies, including the two case studies, used a test-teach-retest design that involved a pretest of the target skill, an intervention phase to develop the language skill(s), and a posttest of the target skill (see Table 3). English was the language used for all dynamic procedures in five of the studies (Camilleri & Law, 2007; Hasson et al., 2012; Hemsley et al., 2014; Peña et al., 2014; Petersen et al., 2017). For the studies that involved Spanish-English bilingual children, the speech-language pathologists (SLPs) were also bilingual and used Spanish and English throughout dynamic procedures by responding in whichever language the child used to communicate (Gutiérrez-Clellen & Peña, 2001; Kapantzoglou et al., 2012; Lazewnik et al., 2019; Peña et al., 2001; Petersen et al., 2020).

All components of the test-teach-retest procedure were completed in one or two sessions lasting 30 to 45 minutes in seven of the studies (see Table 3). The remaining three studies (Lazewnik et al., 2019; Peña et al., 2014; Petersen et al., 2017) were completed over three sessions which included a separate session for posttests.

4.4.2 Pre- and Posttest Measures

Subtests from standardized language assessments were used in eight studies (see Table 3) to establish a pretest score of the target skill e.g., use of the Expressive One-Word Picture

Vocabulary Test (EOWPVT). Standardized assessments were administered in the traditional static manner of testing where the examiner is not allowed to engage with the student beyond what is written in the test script. A majority of these studies also used the same standardized assessment as a posttest, except for three of the studies (Camilleri & Law, 20017; Kapantzoglou, 2012; Petersen et al., 2020) which followed dynamic procedures for post testing of concepts taught during the intervention phase. For example, Camilleri and Law (2007) administered a posttest of vocabulary which imitated the same word activity the child participated in during the dynamic teaching phase of the assessment.

Two studies which assessed children's narrative skills (Peña et al., 2014; Petersen et al., 2017) created their own story telling measures for the pre- and posttest using guidelines from Systematic Analysis of Language Transcripts (SALT). Peña et al. (2014) asked children to tell a story from a wordless picture book for their pretest measure and scored the audio recordings based on three subsections: (1) story components (e.g., setting, character description, etc.), (2) story ideas and language (i.e., vocabulary, grammar, complexity of ideas), and (3) story structure (i.e., problem, climax, resolution). After the teaching phase, the same task was repeated for the posttest except with a different wordless picture book. Petersen and colleagues (2017) followed a similar pre- and posttest storytelling procedure, except the examiner first read a story aloud to the child and then asked them to retell the story using the wordless picture book.

4.4.3 Variations in Teaching Phase

While all of the studies followed some form of a scripted teaching phase, there were two main variations that emerged: (1) the graduated prompting approach and (2) the mediated learning experience (MLE) approach. In the graduated prompting approach, focus is placed on the number of cues or prompts required by the child to reach a learning target; whereas the MLE approach focuses on the child's underlying cognitive behaviors during the teaching phase (Feuerstein, 1991; Lidz, 1991).

Graduated Prompting

Four studies (Camilleri & Law, 2007; Hasson et al., 2012; Petersen et al., 2017; Petersen et al., 2020) followed a graduated prompting approach using a standardized hierarchy of cues to support and guide the child towards achieving the learning goal. Of these studies, two of them followed similar protocols for teaching vocabulary and incorporated a three-level

prompting hierarchy with associated scoring system (Camilleri & Law, 2007; Hasson et al., 2012). During the intervention phase, children were asked to post pictures of vocabulary words (e.g., tractor) into a mailbox after correctly identifying the word alongside two distractor words (e.g., ladder, nest). The child received a score based on the level of independence: *independent identification* (i.e., child able to identify word without SLP prompting) received three points, *implicit identification* (i.e., SLP guides child to find distractor words first) received two points, and *explicit identification* (i.e., SLP models explicit connection between target word and picture) received one point. This graduated prompting approach was also replicated in a phonology and syntax structure activity (Hasson et al., 2012).

Petersen and colleagues incorporated aspects of graduated prompting for the dynamic assessment of narrative skills (2017) and vocabulary skills (2020). Their cue system followed "least-to-most verbal prompting" (Petersen et al., 2017, p. 989), ranging from open-ended questions to direct modelling of the language target. In their narrative skills study (2017), Petersen et al.'s (2017) teaching phase followed four preset steps (See Table 3) which the examiner repeated with the participant as many times as possible during the 15-to-20-minute session. The prompting was systematically faded with each repetition of the lesson and student responsiveness was observed and recorded. Petersen et al., (2020) followed a similar scripted prompting approach to teach children how to use context clues to infer word meaning (see Table 3).

MLE Approach

Four studies (Hemsley et al., 2014; Lazewnik et al., 2019; Peña et al., 2001; Peña et al., 2014) utilized the MLE approach using a mediated learning script based on previous research by Peña (2001) and Lidz (1991). The mediated learning script included principles from the key concepts of MLE instruction (Lidz, 1991): *intentionality, meaning, transcendence, and competence* (See Table 3). Lazewnik et al., (2019) and Hemsley et al. (2014) created mediated learning scripts to develop bilingual children's word learning and syntax ability based on the previous work of Peña and colleagues (2001). Pena et al. (2014) followed a similar MLE framework using "story intervention scripts" to teach narrative skills to bilingual children. They also included the principle of *transfer* to help children focus on how they could use learned strategies in the future.

In the above-mentioned studies, the content and materials used were the same for each participant but the mediated learning script allowed the SLP to respond flexibly to the individual learner. At the beginning of the teaching session, the mediator (SLP) clearly stated the goal of the session (i.e., mediation of *intentionality*) and the purpose (i.e., mediation of *meaning*). The SLP then connected the planned activity to the child's school and home activities (i.e., mediation of *transcendence)*. Next, the SLP helped the child develop their own plan for the learning activity and guided them as they carried out the plan (i.e., mediation of *competence*). At the end of the activity, the mediator (SLP) reviewed the key learning principles with the child and described what they had observed the child do during the session (Peña et al., 2001).

Finally, the remaining two studies (Gutiérrez-Clellen & Peña, 2001; Kapantzoglou et al., 2012) utilized a combined approach in their teaching phase. This included a mediated learning script and the use of graduated prompting.

4.5 Learner Responsiveness

The second research question concerns how learner responsiveness is elicited and observed during mediated learning experiences. To answer this question, information was extracted regarding strategies or actions on behalf of the mediator to elicit student responsiveness. Data related to how learner responsiveness is observed, including modifiability measures (Lidz, 1991, Peña, 1993) and descriptions of children's behavior during dynamic assessment, was also obtained.

4.5.1 Mediator Strategies

Use of Feedback

The majority of studies emphasized the importance of feedback throughout the teaching phase of dynamic assessment. However, there were some differences noted in the type of feedback given to the individual child. For the studies that followed a graduated prompting approach (Camilleri & Law, 2007; Hasson et al., 2012; Petersen et al., 2017; Petersen et al., 2020) feedback was given in a predetermined, structured manner following a cuing hierarchy. If the child's responses matched the correct or desired response then they received no feedback. Feedback was given in the form of open-ended questioning (least supportive feedback) or modelling the desired response (most supportive feedback). In addition, feedback was given immediately following an incorrect response. For example, if the child

incorrectly labelled a picture of a tractor, the examiner immediately stopped the child and said, "No, that's not the tractor; that was hard work; let's try and find the easier ones first" (Camilleri & Law, 2007; p. 316).

In five of the studies which followed the principles of MLE (Gutiérrez-Clellen & Peña, 2001; Lazewnik et al., 2019; Kapantzoglou et al., 2012; Peña et al., 2001; Peña et al., 2014), feedback was given flexibly based on each child's individual response. A majority of these studies utilized bilingual English-Spanish SLPs who provided feedback in whichever language the child used in their response (Gutiérrez-Clellen & Peña, 2001; Kapantzoglou et al., 2012; Lazewnik et al., 2019; Peña et al., 2001). In addition, the goal of feedback was not to correct the child's response, but to explore the child's thought process during problem-solving. Open-ended questions were used, such as "How did you know that?" or "What would happen if?" (Gutiérrez-Clellen & Peña, 2001; Peña et al., 2001). Researchers noted that open-ended questions like these create opportunities for the child to engage in problem-solving.

Reinforcing what children had learned during the teaching session is another way in which feedback was utilized (Gutiérrez-Clellen & Peña, 2001; Kapantzoglou et al., 2012; Peña et al., 2001). Peña and colleagues (2001) finished each activity by reviewing the key principles of the target skill (e.g., labelling pictures) and describing the changes the mediator observed in the child's planning skills (e.g., 'you compared the pictures') and their self-regulation skills (e.g., 'you waited for directions and asked questions'). Gutiérrez-Clellen and Peña (2001) provided positive feedback at the end of each task, for example, "Good! At first you didn't use special names, but we worked on that and now you know that special names are important" (p. 220). Another study (Kapantzoglou et al., 2012) commented on the positive changes observed by the examiner, e.g. "You are learning new words!" (p. 87), in order to reinforce the child's learning.

Repetition
Several studies (Hasson et al., 2012; Kapantzoglou et al., 2012; Petersen et al., 2017) highlighted the importance of repetition and increasing opportunities for learning during dynamic assessment tasks. Petersen et al. (2017) used an "overcorrection procedure" to have the child produce the target skill multiple times (e.g., repetition of target word in different contexts). They also utilized task repetition by having the participants cycle through the same

story retell task as many times as they could during the 20-minute session (Petersen et al., 2017). In a word-learning category task, examiners repeated aloud the items named by participants (Kapantzoglou et al., 2012). Repetition of instructions was also noted as a teaching strategy (Hemsley et al., 2014).

Regulating Attention

Regulating and refocusing the child's attention throughout the teaching session was highlighted in several studies (Gutiérrez-Clellen & Peña, 2001; Peña et al., 2001; Kapantzoglou et al., 2012), although details regarding how this was done was limited. Peña and colleagues (2001) used strategies to focus the "visual and verbal attention of the learner" (p. 220). One way in which they achieved this was by clearly stating the goal and expectations of the task. In another study which assessed word-learning (Kapantzoglou et al., 2012), students were reminded to pay attention to the names of the objects prior to starting each new phase of the activity.

Establishing Rapport

Overall, there was limited data regarding how the examiner (i.e., mediator) established the relationship between mediator and child during dynamic asessment. One study acknowledged the importance of establishing rapport with the child (Kapantzoglou et al., 2012) prior to beginning the dynamic assessment procedure. It was noted that "affective involvement and enthusiasm [on behalf of the examiner] were critical elements" (Kapantzoglou et al., 2012, p. 87). At the end of each task the examiner commented on the positive change they observed (e.g., "you are learning new words!") before moving on to the next task (Kapantzoglou et al., 2012). Positive reinforcement was mentioned as a strategy in three studies (Gutiérrez-Clellen & Peña, 2001; Kapantzoglou et al., 2012; Peña et al., 2001).

4.5.2 Modifiability Measures

Nine of the studies (see Table 3) included measures of modifiability which observed and scored the child's response to mediation (i.e., teaching phase of dynamic assessment procedure). The remaining study (Lazewnik et al., 2019) did not include any measures or mention of learner responsiveness or modifiability measures during dynamic assessment procedures. Two of the studies which followed a graduated prompting approach (Camilleri & Law, 2007; Hasson et al., 2012) gave students a "mediational score" depending on how they

responded to the hierarchy of prompts. This mediational score was based on the correctness of the child's response in connection with the level of prompting required.

Student's responsiveness to mediation was considered using the Modifiability Scale and Learning Strategy Checklist (refer to Appendix C) in three studies (Gutiérrez-Clellen & Peña, 2001; Kapantzoglou et al., 2012; Peña et al., 2001). The Modifiability Scale (Peña et al., 2001 based on Lidz, 1991) used a Likert scale to summarize the child's overall change based on three areas: (1) *examiner effort* (i.e., amount of support provided to child), (2) *child responsivity* (i.e., flexibility when learning new information), and (3) *transfer* (i.e., applying learnt strategies across tasks). In addition, examiners used the Learning Strategy Checklist (Lidz, 1991; Peña, 1993) to observe and record the child's learning behavior during mediated learning activities. This checklist is made up of 13 three-point items (scored 0-2) which examine learning behaviors such as attention, self-awareness, planning, and motivation (see Appendix C for full list).

Two studies (Hemsley et al., 2014; Peña et al., 2014) used a Mediated Learning Observation (MLO) form created by Peña and colleagues (2007) to examine children's behavior in regard to four main areas: (1) *affect* (e.g., anxiety, motivation), (2) *behavior* (e.g., attention, reaction to feedback), (3) *arousal* (e.g., self-awareness), and (4) *elaboration* (e.g., problem-solving). Refer to Appendix D for a complete MLO form.

Petersen and colleagues (2017, 2020) modified the MLO form to produce a 7-point modifiability scale for their dynamic assessment studies. Questions one through six observed how frequently behaviors occurred during the teaching phase of dynamic assessment using a 3-point Likert scale. Target behaviors included paying attention to tasks, responsiveness to prompts, displaying frustration, and disruptive behavior (see Table 3 for complete list). The final question considered the child's potential for future learning of the target skill based on the level of difficulty observed during the teaching phase.

4.5.3 Descriptions of Behavior

Overall, there was minimal qualitative description of children's responsiveness and their behavior during dynamic assessment procedures. The majority of studies followed a quantitative design and evaluated student responsiveness through quantitative measures as discussed above. Qualitative descriptions of children's behavior during dynamic assessment

procedures were documented in three studies (Gutiérrez-Clellen & Peña, 2001; Hemsley et al., 2014; Peña et al., 2001), although even these descriptions were abbreviated. Of the three research articles, two presented case studies of dynamic assessment of two bilingual children's language skills (Gutiérrez-Clellen & Peña, 2001; Hemsley et al., 2014). Children's testing behavior was described in terms of their attention, motivation, self-regulation, planning, and their ability to transfer learned skills (see Table 3 for full descriptions). It is important to note that only one research study (Hemsley et al., 2014) provided direct quotes from the students during mediated learning experiences.

4.6 Key Findings

The third research question asks how speech-language pathologists can use dynamic assessment methods in their decision-making process when assessing children from culturally and linguistically diverse backgrounds. This question has, in part, been answered through information presented in the previous results sections (4.4 Dynamic Assessment Methodologies and Procedures; 4.5 Learner Responsiveness). To answer this question further, information related to key findings and outcomes of the included studies is presented here.

4.6.1 Pretest to Posttest gains

While the ten research studies captured a range of findings, four studies (Gutiérrez-Clellen & Peña, 2001; Peña et al., 2001; Hemsley et al., 2014; Petersen et al., 2020) showed that pretest to posttest gains (following an intervention phase) was a significant predictor of language disorder in CLD children. That is, CLD children who made small improvement between pre- and posttest were more likely to have an underlying language disorder. On the other hand, CLD children who made significant improvement following the teaching phase of dynamic assessment, were representative of CLD learners whose language difficulties are due to limited experience with the majority language and culture (Gutiérrez-Clellen & Peña, 2001; Hemsley et al., 2014).

Two studies (Kapantzaglou et al., 2012; Petersen et al., 2020) found that pretest to posttest gains were significant when combined with modifiability scores. The remaining four studies (Camilleri & Law, 2007; Lazewnik et al., 2019; Peña et al., 2014; Petersen et al., 2017) did not indicate pretest to posttest gains as a significant predictor of language disorder in CLD learners.

4.6.2 Significance of Modifiability

Participants' modifiability scores, which evaluated how children responded to the teaching phase of dynamic assessment, represented the most significant finding for the current dynamic assessment studies. Six studies (Gutiérrez-Clellen & Peña, 2001; Hemsley et al., 2014; Kapantzoglou et al., 2012; Peña et al., 2001; Peña et al., 2014; Petersen et al., 2017) indicated that modifiability ratings (i.e., scores taken from Modifiability Scale, Learning Strategies Checklist, etc.) were the best predictor of language ability for CLD children. Petersen et al. (2017) highlighted that children who showed higher levels of frustration, inattention, and disruption were more often part of the language impaired group. Three studies (Hemsley et al., 2014; Peña et al., 2001, Peña et al., 2014) pointed towards issues of planning, self-regulation, use of strategies, and carryover between tasks as common areas of need for bilingual children with language impairment. See Table 3 for further details.

4.6.3 Further Findings

Two studies which followed a graduated prompting approach (Hasson et al., 2012; Camilleri & Law, 2007) indicated that CLD children with language impairment required more prompting during teaching sessions in order to achieve the target goal. Another study (Petersen et al., 2017) showed that the duration of the teaching session, following modifiability scores, was the second-best predictor for differentiating language disorder from language difference in bilingual children. Specifically, bilingual children who required longer teaching sessions to reach the predetermined goal were more likely to have a language impairment (Petersen et al., 2017).

A summary of the data in relation to the three research questions is shown in Table 3.

Table 3. *Summary of Research Question Findings*

Study	Participants	Aim/Target Area	Methodology/Procedure	Mediator Strategies	Learner Responsiveness	Findings
Camilleri & Law (2007)	N = 54 CLD children 14 children with typical language 40 children referred to SLP, 12 English as an additional language (EAL), 28 English only Ages 3;05-5;0 (years; months)	Investigated use of dynamic assessment (DA) alongside standardized assessment measures to evaluate receptive vocabulary skills of children with suspected language disorder.	Test-teach-retest design in a single 45-min session: *Administered in English* 1. Block-building task from BAS II (Elliott, 1996) 2. Static pre-test phase: BVS II 3. DA phase: activity targeting vocabulary items that were difficult for child from pre-test 4. The Picture Similarities task from the BAS II (Elliott, 1996) 5. Re-test phase: testing vocabulary targeted in intervention phase	Graduated prompting following a standardized hierarchy of cues (least to most assistive) Providing feedback to child's responses and guiding child to correct response ("that's not quite right, let's try and find the easier ones first") Creating opportunities for problem-solving	Limited information regarding the child's response to mediation or their engagement in the dynamic task Children were given a 'mediational' score based on how they responded to prompting, ranging from a score of 1 to 3, with independent identification of words receiving maximum points	Bilingual and monolingual children showed similar abilities to make new word-referent matches during DA session, despite differences in their static assessment vocabulary scores. DA provided additional information about vocabulary learning not captured by BPVS II (e.g., captured responses to learning opportunities) Limits regarding predictive value of DA procedure for future vocabulary skills.
Gutiérrez-Clellen & Peña (2001)	N = 2 English/Spanish bilingual children Ages 4;0 and 4;6 No existing diagnoses	To explore how a DA protocols for vocabulary learning may be used to differentiate between language disorder and language difference in bilingual children Hypothesized that when CLD learners receive appropriate MLE, typical language learners should show improvement between pre-test and post-test	Case study design: two cases taken from a previous DA study of CLD children (Peña et al., 2001) *Administered in English and Spanish* Refer to Peña et al., 2001 for details	Examiner followed script with prompting hierarchy to make MLE sessions consistent (Refer to Peña et al., 2001 for details) Providing feedback during testing: simple yes/ no feedback to explanations for why answer was correct or not Use of graduated prompting	Modifiability Scale and Learning Strategies Checklist (See Peña et al., 2001) Child A exhibited: -Good motivation and attention -Inconsistent use of strategies, minimal carryover -Difficulties with planning and self-regulation Child B exhibited: -Strengths in attention, planning, self-regulation, motivation -Ability to self-correct -Transferred learned skills across tasks and sessions	Student A demonstrated little improvement between pre-test and post-test following MLE. *Indicative of language disorder.* Student B demonstrated significant improvement between pre-test and post-test following MLE. *Not indicative of language disorder, but need for more language exposure.* Limitations of reliability of DA procedures.

Study	Participants	Aim/Target Area	Methodology/Procedure	Mediator Strategies	Learner Responsiveness	Findings
Hasson et al. (2012)	N = 26 bilingual children (EAL) from diverse linguistic/ cultural backgrounds 12 referred to speech language services; 14 not receiving services Ages 3;0-5;0 years	Evaluated whether DA procedures for phonology, sentence structure, and vocabulary resulted in significantly different performances for bilingual children referred to SLP compared to bilingual children with typical language	Test-teach-retest design carried out in a 40-min session *Administered in English* DAPPLE administered: 1. Block-building task from BAS II (Elliott,1996) 2. DA of vocabulary: pre-test, teaching phase, post-test (Camilleri & Law, 2007) 3. DA of expressive language, static pre-test followed by teaching phase 4. Goodenough Draw a Man Test (Goodenough, 1926) 5. DA of phonology 6. Post-test: assessment of expressive language repeated	Children encouraged to use process of elimination strategies Graduated prompting following a standardized hierarchy of cues (least to most assistive) SLP feedback (e.g., "that's not quite right") Modelling correct use of language for child to repeat	Limited information regarding how children responded to testing or how the examiner assessed the child's responsiveness to mediated learning. Children were given a 'mediational' score based on how they responded to prompting (same as described in Camilleri & Law, 2007)	Children referred to SLP services required more prompting for vocabulary and sentence structure tasks and demonstrated lack of retention of learning, compared to bilingual students not flagged for SLT services. The group design of the study obscured individual differences of participants such as age and English language exposure. A case-study design is suggested for further investigation of these factors.
Hemsley et al. (2014)	N = 2 bilingual children (EAL) Both diagnosed with specific language impairment (SLI) from standardized language assessment. Ages 11;0 and 8;0	To examine whether CLD children's language skills were consistent with SLI or whether they were characteristic of typical bilingual language development Receptive and expressive language tasks administered	Case study design, w/wide range of assessment techniques: *Administered in English* 1. Interviews with students and parents to obtain language background 2. Peer-child comparative analysis to compare child's language abilities to a peer from the same CLD background w/typical LD 3. DA of target area. Intervention/teaching phase over two 45-minute sessions. 4. Post-test using standardized measures	DA intervention did not "teach to the test", instead real-life examples were used Goal of intervention explained to child in relation to classroom and real-life situations. Mediated learning experience protocol was followed (See Gutiérrez-Clellen & Peña, 2001).	Mediated Learning Observation (MLO) form (see Peña et al., 2014) Student A exhibited: - Attentive/cooperative - Good recall for purpose of intervention - Good retention of information - Able to apply learned strategies to new tasks Student B exhibited: - Lack of awareness of performance/purpose of intervention - Difficulty planning and identifying problems - Restricted use of strategies w/o prompting	Findings show the importance of using a range of assessment techniques for diagnosing SLI in CLD learners. Student A showed positive change from pretest to posttest following intervention and their learning behaviors demonstrated good cognitive modifiability. *Performance not consistent with SLI.* Student B showed limited change from pretest to posttest after intervention and significant difficulties were noted regarding his learning behaviors. *Performance consistent with SLI.*

Study	Participants	Aim/Target Area	Methodology/Procedure	Mediator Strategies	Learner Responsiveness	Findings
Kapantzog-lou et al. (2012)	N = 28 Spanish/English bilingual children 15 w/typical language development; 13 w/language impairment Ages 4;0-5;0	To investigate whether DA of vocabulary learning skills is an effective way to identify bilingual children with underlying language disorder	Test-teach-retest design carried out in one 30-40 min session * Administered in Spanish and English Parent and teacher questionnaire to gather background information 1. Pretest: Standardized vocabulary test (EOWPVT); story retell task 2. Three-phase DA task including pre-test, teaching, and post-test of target words. 3. Learning Strategies Checklist and Modifiability Scale administered	Scripted MLE approach with graduated prompting (See Gutiérrez-Clellen & Peña, 2001) Examiner established rapport with children before beginning activity Use of positive feedback Connecting activities with children's previous experiences Use of repetition and opportunities for learning	Learning Strategies Checklist and Modifiability Scale (Lidz, 1991; Peña, 1993). See Gutiérrez-Clellen & Peña (2001) for details. Only quantitative data for modifiability, no qualitative description of how children responded to MLE	Suggest that DA of vocabulary skills is a useful tool for screening language ability in CLD learners and differentiating between language difference and language disorder. Modifiability scores were strongest indicator for separating typical language learners from impaired. Classification accuracy improved when modifiability scores were included with pretest-posttest change.
Lazewnik et al. (2019)	N = 30 Spanish/English bilingual children 15 w/typical language development; 15 with language impairment Ages 4;0-5;0	To explore whether adding informal measures (including DA) to standardized assessment improves the classification accuracy for language ability in second language learners. Targeted vocabulary skills	Followed test-teach-retest DA procedure (Peña et al., 2001) over 3 sessions *Administered in English and Spanish 1. Interviews with parents and teachers to gather language background information 2. Pre-test: tested vocabulary using EOWPVT 3. DA: two 20-minute sessions following Mediated Learning Script to teach single-word labels 4. Post-test: re-administered EOWPVT	Followed Mediated Learning Script (See Peña et al., 2001). Individualized feedback given in whichever language the child used in their response (i.e., Spanish or English).	No measures or discussion of student responsiveness or modifiability during DA procedure.	Pretest to posttest gains were not significant and did not differentiate between language impaired and typical development groups. DA was not suggested as a method to identify language impairment in CLD learners. Inclusion of language sample to identify language impairment in CLD learners is recommended.

Study	Participants	Aim/Target Area	Methodology/Procedure	Mediator Strategies	Learner Responsiveness	Findings
Peña et al. (2001)	N = 79 Spanish/English bilingual children 17 children w/low language ability (parent report, teacher report, observation); 45 children w/typical language ability Ages 3;9-4;9	To study whether DA approaches help distinguish between typical and atypical word learning in culturally and linguistically diverse children.	Followed a test-teach-retest design *Administered in English and Spanish* 1. Pre-test: vocabulary subtests from standardized language assessments (E/ROWPVT, CSSB, PLS) 2. DA: children taught vocabulary strategies in two 30-minute sessions 3. Post-test: re-administer standardized measures. 4. Analysis of modifiability ratings and children's responses to the tests after MLE.	Mediated Learning Script following four concepts: -<u>Intentionality</u>: communicates goal/purpose of intervention to focus learner's attention -<u>Meaning</u>: encourages learner's awareness of learning target by putting meaning to specific behavior -<u>Transcendence</u>: creates connections b/w learning tasks to child's experiences to carryover skills across contexts -<u>Competence</u>: provides positive feedback to reinforce and encourage use of new strategies	Learning Strategies Checklist (Lidz, 1991; Peña, 1993) to rate attention, self-awareness, planning, motivation, etc. Modifiability Scale (Lidz, 1991; Peña, 1993): <u>High performers:</u> -Good motivation, planning, focus, initiation -Able to transfer strategies -Good problem-solving skills -Good task understanding <u>Low performers:</u> -Needed more redirecting, encouragement, praise -Became frustrated when task was too difficult -Less perseverance	Compared to pretest static measures, DA results (posttest scores of single-word labelling and modifiability ratings) predicted the language ability groups more accurately. Children in the low language ability group showed less pretest-posttest change after intervention compared to bilingual children with typically developing language. The "modifiable child" who shows more improvement in the target area after mediation is likely to have typical language learning ability.
Peña et al. (2014)	N = 54 Spanish/English bilingual children 18 with language impairment (LI); 36 with typical language (TL) M= 5;9	To explore the classification accuracy of a DA procedure to evaluate narrative ability for bilingual children learning English as a second language.	Test-teach-retest design over 3 sessions *Administered in English* 1. Pretest: Story-telling task with wordless picture book 2. DA intervention: two MLE sessions 30 min each. Focused on increasing length and complexity of narratives 3. Mediated Learning Observation (MLO) form administered 4. Posttest: Story-telling task (different story than pre-test)	Scripted intervention sessions incorporated MLE strategies (i.e., meaning, transcendence, transfer, etc.) See Gutiérrez-Clellen & Peña (2001) for MLE strategies. Modelling and practicing complex and complete sentences.	MLO form (Peña, 2007) rated 4 areas based on support required: 1. Affect (e.g., anxiety, motivation) 2. Behavior (e.g., attention, reaction to feedback) 3. Arousal (e.g., task orientation) 4. Elaboration (e.g., problem-solving) Only quantitative data provided, no qualitative description of how children responded to MLE.	No significant differences among groups (LI, TL) for pretest to posttest gains. Possible that language demands of narrative task exceeded children's English ability. MLO form significant for differentiating between LI and TL groups. Children with LI needed more support during MLE. Group differences were greatest for attention, problem-solving and flexibility.

Study	Participants	Aim/Target Area	Methodology/Procedure	Mediator Strategies	Learner Responsiveness	Findings
Petersen et al. (2017)	N = 42 bilingual Spanish/English children 32 children were not receiving SLP services; 10 children receiving SLP services Ages 6;4-9;6	To investigate the validity of a narrative DA procedure for identifying language impairment in bilingual children. Hypothesized that modifiability ratings would be greatest predictor of language ability in Spanish-English bilinguals.	Test-teach-retest design carried out in 3 sessions *Administered in English* 1. Pre-test: Frog Story Retell. Children listened to story in English, then asked to retell 2. DA: (1) Examiner models narrative; (2) Retell w/pictures and icons; (3) Retell w/icons only; (4) Retell w/o visual supports 3. Post-test: Story retell task with a new story 4. Modifiability rating scale Audio recordings of DA teaching were timed	Graduated prompting strategies: 1. Correct/prompt immediately upon error 2. Least to most verbal prompting (e.g., open-ended questions to modelling target) 3. Create opportunities for child to produce target skill multiple times across contexts 4. Foster independence (i.e., use least amount of verbal prompting)	7-item modifiability scale to assess students' responsiveness to instruction (based on Peña 2006, 2007) Rating form evaluated: 1. Responsiveness to prompts 2. Showing transfer of learning across lessons 3. Paying attention 4. Ease of teaching 5. Level of frustration 6. Disruptive behavior 7. Learning potential	Findings were in contrast to previous DA studies. No significant difference in gain scores between children with typical language and children with impairment. Best predictor of language performance was modifiability rating. Children who showed greater levels of frustration, disruption, and inattention were part of language impaired group. Next best predictor was teaching duration. Results were only for a short period of DA administration (two 25-minute sessions), longitudinal data advised.
Petersen et al. (2020)	N = 31 bilingual Spanish/English school-age children 21 children with typical language, 10 children with language disorder Ages 5;9-9;7	Examined whether the use of DA of inferential word learning was more accurate at identifying bilingual children with a language disorder compared to a standardized measure of vocabulary	Test-teach-retest design carried out over two sessions *Static measures in Spanish and English, DA English only* 1. Pre-test: E/ROWPVT vocabulary test in Spanish/English 2. DA: children taught context clue strategies to infer word meaning 3. Post-test of inferential word learning (similar to DA task) 4. Modifiability rating scale	Examiner modelled think-aloud strategies and had the child imitate the strategies out loud Use of open-ended questions Use of graduated prompting (see Petersen et al. 2017)	7-item modifiability scale to assess students' responsiveness to instruction (See Petersen et al. 2017). Provided quantitative scores for modifiability but no qualitative description of children's responsiveness to DA	DA differentiated between children with and without language impairment more accurately than standardized vocabulary tests. There may be age-related differences in word learning. Bilingual children's response to DA may be related to differences in their English ability, not in their underlying language ability.

Note. N = number of participants in study.

5. Discussion

This chapter will present a discussion of the three research questions based on the findings of the dynamic assessment studies included in the literature review. Implications for speech-language pathologists (SLPs) and educational practitioners will be discussed as part of the third research question. Following a discussion of the research questions, limitations of the dynamic assessment research studies included in this review will be discussed as well as limitations of the literature review itself. Finally, suggestions for further research will be presented.

5.1 Dynamic Assessment Methods and Procedures

The first research question examined what methods and procedures the current dynamic assessment studies employ for the assessment of culturally and linguistically diverse (CLD) children's language abilities. The findings of this literature review indicated a number of similarities in the methodology of the selected dynamic assessment studies. All of the studies followed a test-teach-retest approach in which children participated in a teaching session between assessments to address the target skill (e.g., vocabulary, sentence structure). This finding supports the idea that a key aspect of dynamic assessment is the integration of teaching into the assessment procedure (Lidz & Gindis, 2003) in order to gather valuable information regarding students' learning processes.

While all studies incorporated an intervention or teaching session into their testing procedure, there were some differences regarding the methodology of the teaching session. Studies followed either the mediated learning experience (MLE) approach (Feuerstein et al., 1979, 1980; Lidz, 1991) or the graduated prompting approach (Brown & Ferrara, 1985; Campione & Brown, 1987), with two studies incorporating aspects of both approaches. Within these two approaches, there was little variability in the assessment procedures as most studies replicated the methodology of previous studies. Furthermore, there appeared to be a correlation between dynamic assessment methodology and the location of the study. That is, the studies which took place in the U.K. followed the same graduated prompting procedure, while most of the studies from the U.S. followed the MLE approach based on the original work of Peña and colleagues (2001). Given that dynamic assessment is a relatively new approach for assessing the language skills of bilingual students, it is unsurprising that current research replicates the

methodology of previous dynamic assessment studies in order to validate its use as an assessment tool.

In nine of the studies the teaching session was given in between non-dynamic pretest and posttests, which is consistent with the sandwich design described by Sternberg and Gigorenko (2002). The cake format, which provides the child with immediate mediation when problems arise during testing (Sternberg & Gigorenko, 2002), was found in only one study despite all studies utilizing one-to-one testing environments. One possible explanation for this is that all of the studies followed some form of a teaching script. The scripted approach is used to standardize the dynamic assessment process and allow for scoring (Lidz & Gindiz, 2003) and makes it easier for examiners to learn how to administer the procedure (Poehner, 2008). While this structured approach allows for replication of studies and the statistical analysis of findings, it may not be suitable for "depicting the kinds of dynamic, emergent abilities that are of interest in dynamic assessment" (Poehner, 2008, p. 20).

Standardized assessment was used as a tool for pretest and posttest comparison in the majority of studies, excluding the narrative skill studies (Peña et al., 2014; Petersen et al., 2017) which used their own story retell activities. This indicates that current dynamic assessment procedures for CLD learners still rely on standardized assessment as a tool for measuring growth in children's language skills. In order to establish the effectiveness of dynamic assessment, the majority of research within the field has followed standardized approaches to quantify their results (Poehner, 2008). Of the ten studies in this literature review, eight followed quantitative research designs. The two case studies (Gutiérrez-Clellen & Peña, 2001; Hemsley et al., 2014) also followed test-teach-retest procedures and included standardized measures for pre- and posttest assessment. This finding represents the influence and value of traditional testing methods even within an area of research that challenges standardized assessment practices.

Overall, dynamic assessment procedures were carried out within similar timeframes. The majority of studies occurred within one or two sessions lasting 30 to 45 minutes. While critics of dynamic assessment have argued that it is too time-intensive (Petersen et al., 2017), this finding suggests that dynamic assessment procedures may be carried out in comparable timeframes as standardized assessment. In addition, the length of dynamic assessment sessions may also provide useful information for distinguishing between language

impairment and language difference. Petersen et al. (2017) showed that bilingual children who required longer teaching sessions to reach the predetermined goal were more likely to have a language impairment than those who took less time.

Finally, an important difference in methodology related to the language of instruction during mediated learning or teaching sessions. While all of the studies took place in English majority countries, only five of the studies used English only for dynamic assessment procedures. The other five studies included bilingual Spanish-English SLPs who shared the same language and culture as the participants in the study and were able to respond to the participant in either language. In the included studies it is unclear to what extent language of instruction or shared cultural background affected dynamic assessment outcomes, as there was no discussion of this in the research. Whether there is shared language and culture between mediator and child is an important factor to consider when analyzing the research, especially as it supports the establishment of intersubjectivity (Rommetveit, 1976) between the mediator and the learner.

5.2 Learner Responsiveness

The second research question explored the role of learner responsiveness in the current dynamic assessment studies. Learner responsiveness has been previously described as a key characteristic of dynamic assessment which aims to engage children in a shared learning experience during assessment. Two aspects are considered: (1) how the examiner or mediator elicited children's responsiveness and (2) how the child's responsiveness to testing was observed by the mediator.

5.2.1 Mediator Strategies

In general, the findings lacked in-depth information regarding how examiners elicited participants' responsiveness during language tasks. Of specific interest was how mediators engaged children during the teaching and learning portion of dynamic assessment. In all of the studies the examiner followed a script which was either based off of the principles of graduated prompting (Brown et al., 1985; 1987) or the mediated learning experience (MLE) (Feuerstein et al., 1979, 1980; Lidz, 1991). The graduated prompting studies provided information in the form of lists and tables about the hierarchy of prompts used by the examiner, ranging from general statements to direct language modelling. In these studies, the

focus was on the amount of assistance the mediator provided, as opposed to the quality of assistance.

Mediators demonstrated a flexible approach to learning tasks in the MLE studies, in contrast to the studies which took a graduated prompting approach. Feedback was adjusted to the pace of the testing session and depending on the needs of the individual child (Kapantzoglou et al., 2012). Another important difference between the graduated prompting studies and the MLE studies was the emphasis placed on guiding the child to independent problem-solving by building their self-awareness throughout the learning process. Asking the child open-ended questions, such as "How did you know that?", encouraged the participants to use self-reflection as a learning tool. When asked these types of open-ended questions, children often changed their answers, demonstrating more knowledge than was first revealed from the original assessment question (Gutiérrez-Clellen & Peña, 2001). This finding emphasizes the importance of the role of the adult during dynamic assessment, not as a traditional examiner but as a teacher or mediator guiding the child towards future learning.

Research studies which reflected the values of MLE (Feuerstein et al., 1979, 1980) provided more information related to how the mediator engaged the child in dynamic assessment tasks. Mediators followed a mediated learning script based on the key attributes of MLE (i.e., intentionality, meaning, transcendence, competence) to uncover information about the child's learning process. Of particular significance throughout all of the studies were mediator strategies which focused and maintained the child's attention and engagement with the task. Strategies such as drawing the child's attention to key information, clearly explaining the goal of a task, offering suggestions, using non-verbal cues, and asking questions were noted as common strategies. Positive feedback was also noted in some of the studies as a strategy for encouraging the child's engagement and building their competence.

Overall, mediated learning strategies highlighted the dynamic nature of dynamic assessment in which the mediator engages the child in reciprocal learning tasks in order to transform them from a passive responder to an active participant. The literature supports the importance of establishing joint attention between mediator and child as the first step towards creating a mediated learning experience (Feuerstein, 1991). Use of feedback during assessment was an important tool across studies, but how the examiner used feedback was dependent upon the different approaches to dynamic assessment.

5.2.2 Measures of Responsiveness

The second part of the research question examined how learner responsiveness was observed by the researchers and mediators. Overall, children's responsiveness to intervention sessions was predominantly captured through quantitative scores and not through qualitative description. In nine of the ten studies learner responsiveness was measured through behavioral checklists and modifiability scales (Lidz, 1991; Peña 1993), which examined how the child responded to teaching (e.g., attention, self-regulation) and transferred new learning across tasks. Studies which used a graduated prompting approach evaluated children's responsiveness based on a rating of how much support the child required from the examiner to achieve the learning target. Another study viewed student responsiveness as a measure of how long it took a child to achieve the learning goal (Petersen et al., 2017).

The dynamic assessment case study of Hemsley and colleagues (2014) provided a more in-depth look at how CLD student's behaviors and responsiveness to instruction may be used to differentiate between impaired language and typical language development. This case study observed two bilingual students who had been previously flagged for a language disorder based on below average performances on standardized language assessment. After engaging in dynamic methods of assessment, significant differences emerged based on how these two children responded to teaching.

Of particular interest was differences between the two children's ability to retain information from the first intervention session to the second session. While one child demonstrated little change in retention of information, the other child was able to recall key information as well as explain the purpose of intervention by the second session (Hemsley et al., 2014). This child also demonstrated "a systematic and efficient approach to problem solving…in his ability to talk tasks through using a range of strategies to achieve success" (Hemsley et al., 2014, p. 106). It was decided that this child's learning behaviors were not consistent with language impairment, as the child was able to effectively interpret information from the environment to develop the targeted language skill (Hemsley et al., 2014). Without looking deeper beyond static measures of assessment, this child would likely have received a language disorder diagnosis.

These findings support MLE theory which argues that cognitive abilities are modifiable and can be developed depending on the quality of mediation or teaching the child receives

(Feuerstein, 1993). However, it is important to note that while modifiability scores were found significant for differentiating between language impaired and typically developing CLD students in a number of studies (Gutiérrez-Clellen & Peña, 2001; Hemsley et al., 2014; Kapantzoglou et al., 2012; Peña et al., 2001; Peña et al., 2014; Petersen et al., 2017), there was inadequate information regarding how clinicians were trained to use modifiability scales which rely heavily on personal judgment. Clinical judgments about student's behavior during testing are influenced by cultural expectations of "typical" student behavior. This represents an area of cultural bias that should be considered when interpreting the modifiability findings of the current research.

5.3 Implications for Professionals Working with CLD Learners

The final research question considered the implications of the findings for speech-language pathologists (SLPs) working with CLD populations. Of specific interest is how SLPs can utilize dynamic assessment methods to aid their decision-making when assessing the language skills of children who have different cultural and linguistic backgrounds from their own. The following implications may also apply to other educational specialists and teachers working in the field.

First, this literature review has revealed the value of incorporating teaching into the assessment process for the evaluation of children's language abilities. In doing so, the testing environment is transformed into a teaching environment which maximizes the child's feelings of accomplishment (Peña et al., 2001) and has the power to uncover their true learning potential. This is particularly important when working with children from CLD backgrounds as they are often at a disadvantage due to issues of testing bias or lack of opportunity to learn the assessment content (Bedore & Peña, 2008; Chamberlain, 2005). Incorporating dynamic methods into assessment procedures is particularly achievable for SLPs who often work with children one-to-one or in small group environments.

Second, the findings of this literature review have emphasized the importance of learner responsiveness and measures of modifiability during the assessment process. While pretest to posttest gain scores were not always significant, the included studies did support previous findings (Pena et al., 1992, 1993) which suggest that modifiability measures are reliable predictors of language disorder (Kapantzoglou et al., 2012; Peña et al., 2014; Petersen et al., 2017). That is, bilingual children who scored higher in areas such as attention, problem-

solving, and ability to transfer learned strategies across tasks were less likely to have a true language impairment. By focusing on how a child uses cognitive strategies during intervention, SLPs may be better equipped to identify CLD children with good potential for language learning versus those who require more intensive language intervention.

In addition, the amount of prompting required during dynamic testing may also help SLPs differentiate between language disorder and language difference (Camilleri & Law, 2007; Hasson et al., 2012), such that CLD children who require greater levels of prompting during dynamic assessment tasks may be more likely to have a language disorder. Teaching effort in combination with teaching duration to reach a predetermined goal may also be important factors for SLPs to consider when making clinical judgments. These findings further emphasize the significance of observing the teaching and learning process alongside learning outcomes when assessing children's language skills.

Another finding of interest is the value of assessing CLD children's narrative skills as it is an area that may represent less testing bias (Peña et al., 2014, Petersen et al., 2017). Dynamic assessment of narrative skills required only wordless picture books to generate language samples from CLD children, whereas other studies relied on standardized assessment measures for pre and posttest comparison following intervention. The use of picture books to stimulate language production, instead of formal testing measures, is more likely to create a safer and friendlier learning environment for the child to perform to the best of their ability.

Finally, speech and language therapists must consider how their own linguistic and cultural background affects their clinical judgment and expand their perspective-taking skills to create more channels of communication. It is important that clinicians and educators try to understand the perspectives and experiences of their students as this contributes to the quality of the relationship between the teacher and learner. Gathering background information through parent and child interviews about CLD children's language use and home environment is an essential first step in language assessment (Hemsley et al., 2014; Lazewnik et al., 2019). In addition, incorporating opportunities for children to self-reflect during the assessment process (Peña et al., 2001) can unveil information about the learning process from the child's point-of-view, which may be different from the clinician's interpretation.

5.4 Limitations

In this section, limitations of the current research studies will be addressed as well as limitations of the literature review itself. Following this, critical remarks are presented.

5.4.1 Limitations of Included Research Studies

First, a majority of the participants in the studies were Spanish-English bilinguals from the United States which limits the generalizability of the findings to a larger multilingual and multicultural population. In addition, the language used for dynamic assessment was not consistent throughout studies as half of the studies used Spanish and English for assessment while the others used only English. It was unclear in the research to what extent using multiple languages during assessment affected the findings.

Second, most of the participants were recruited from the preschool age group (three- to five-year-olds); therefore, the findings may not generalize to older school-age children. There were also issues of participant selection bias as most studies recruited children who had known language impairment. The majority of studies divided participants into groups based on typical language development and atypical language development, but it's important to note that not all children clearly fall into these two categories (Peña et al., 2001).

Finally, there was the absence of children's voices during mediated learning experiences. This in part due to the quantitative nature of the current dynamic assessment research, but also reflects an overarching issue of children's perspectives and personal experiences not being represented in educational research.

5.4.2 Limitations of Literature Review

A main limitation of this literature review is the risk of personal bias. Given that the current study only had one researcher it increased the risk of personal judgment affecting decisions such as inclusion and exclusion criteria, key search words, article ratings, and the decision of which articles to include in the final review. Two researchers would have minimized issues of personal bias and reduced the chances of missing dynamic assessment studies which should have been included in this review. In addition, the findings of this literature review are highly dependent on the researcher's interpretations and personal judgment. By clearly describing the methodology process and using the PRISMA guidelines it is hoped that this literature

review can be repeated by other researchers interested in the topic of dynamic assessment to increase the reliability of the findings.

A second limitation of the current study is the number and quality of articles which were reviewed. It was judged that only ten of the studies from the database search met the inclusion and exclusion criteria for the topic of dynamic assessment of language skills in CLD populations. The inclusion criteria of only examining dynamic assessment studies which had English as the majority language limited the total number of articles for this study and was less representative of the global population. If the inclusion criteria had been expanded to majority languages other than English, more dynamic assessment studies may have been identified and incorporated into this review to represent a larger multilingual population. Furthermore, of the ten studies which were included, none of the studies were considered "strong" based on the Kmet (2004) methodological quality assessment due to issues of participant selection bias, small sample size, or lack of control group.

Finally, there were limitations related to the mixed methodology approach used by the researcher. While the research questions were qualitative in nature, a majority of the current dynamic assessment research that fulfilled the inclusion criteria adhered to quantitative research methodology. Consequently, there was less qualitative data to pull from to answer the research questions, especially with regard to the second research question (learner responsiveness). It is possible that the inclusion and exclusion criteria were too narrow or there have been too few case study designs done on this topic within the last twenty years.

5.4.3 Additional Critical Remarks

Given that dynamic assessment of bilingual children's language skills is an emergent area of research, there were often more questions than answers generated by the current research studies. One question I often found myself asking was how researchers and clinicians knew that they were working within a child's zone of proximal development (ZPD)? There appeared to be an assumption in the current research that by incorporating shared learning activities into the assessment procedure, children were automatically accessing and working within their ZPD. As previously discussed, co-construction of the ZPD is revealed through close collaboration between learner and teacher, and yet the current research studies provided little to no information regarding the relationship between the mediator and child.

In addition, there was no description of the testing environment, leading to the question of how researchers created safe learning environments to foster reciprocity between adult and child? Furthermore, what strategies did they utilize when children became unresponsive during teaching sessions? While the MLE studies focused on the quality of the interaction or support provided by the mediator, they largely overlooked the child's contributions in this process. As a researcher as well as a clinician in the field, I was interested in learning more about how children responded to teaching within assessment, not just as a quantitative result but as a meaningful description. The quality of the interaction between the mediator and child is fundamental to unlocking the child's learning potential, therefore, both perspectives should be reflected in dynamic assessment research.

5.5 Suggestions for Future Research

It is recommended that future research incorporate longitudinal or follow-up studies which evaluate the predictive value of dynamic assessment procedures. From the current research studies, it is unknown whether bilingual children who were identified as having an underlying language disorder continued to show language learning difficulties in the future, or whether those children who were identified as typical language learners made good progress in their schooling. This is especially relevant for studies which used graduated prompting as it has been previously noted that there may not be a direct relationship between a child's responsiveness to graduated prompting and their future learning curve (Long & Olswang, 1996). Further research is also suggested regarding how modifiability judgements can be used clinically for speech-language pathologists as well as other specialists who assess CLD children.

As shown in this literature review, the majority of research studies have used quantitative research designs to investigate the use of dynamic assessment as a tool for evaluating the language skills of CLD learners. The research methodology for studying dynamic assessment should reflect a similarly diverse and innovative approach to assessment. It is recommended that future research builds on the case study or qualitative approach in order to derive in-depth information regarding how mediator and learner must work together to co-construct the zone of proximal development.

As previously acknowledged, more research is needed which investigates the child's perspective and contributions during dynamic assessment procedures. It is believed that this

is a missing piece in current dynamic assessment research that needs to be explored. One way that this could be achieved in the current research is by incorporating clinical interview methods (Ginsburg, 1997) during dynamic assessment as well as student interviews before and after participation in a mediated learning experience.

Finally, further research should investigate the role of cultural bias in assessment. In some studies (Gutiérrez-Clellen & Peña, 2001; Peña et al., 2001; Peña et al., 2014) the clinicians shared the same linguistic and cultural background as their participants, which is a factor to consider when comparing research results to studies where the examiners do not share the same cultural and linguistic background. It is recommended that future research compares the use of multiple languages during dynamic assessment to the use of only the majority language to see how this factor impacts the findings. A significant part of the challenge SLPs face when assessing CLD learners is not being able to assess the child in their first language; therefore, this is an area that needs further exploration.

6. Conclusion

The purpose of this literature review was to investigate the use of dynamic assessment as a means for differentiating between language disorder and language difference in culturally and linguistically diverse (CLD) learners. Previous systematic reviews on this topic (Hunt et al., 2021, Orellana et al. 2019) have suggested that dynamic assessment is an effective method for identifying language impairment in bilingual populations, yet there has been little progress for its use within the field of speech and language pathology. This literature review has focused on dynamic assessment methodology and the role of student responsiveness, with the goal of providing useful information to speech-language pathologists (SLPs) and other educational practitioners who are responsible for assessing the language skills of CLD students. PRISMA guidelines were followed for literature screening and selection of the ten studies included in this literature review.

The information presented in this literature review has provided insight into the topic of dynamic assessment and has supported its use as a less culturally and linguistically biased approach to language assessment. Key findings emphasized the value of integrating teaching into assessment in order to access children's language learning potential. In addition, this literature review has reinforced the significance of observing children's responsiveness to teaching, including their ability to transfer learned skills across tasks, as an important indicator for identifying CLD children with and without language impairment.

Finally, it is hoped that the findings of this literature review can contribute to the development of new methods of language assessment for CLD children within speech language pathology. Further research on this topic is required to better understand the child's experience during dynamic assessment and how the relationship between mediator and child is linked to the co-construction of the zone of proximal development. More research in the form of case studies and longitudinal designs is recommended to achieve this. As the world continues to diversify, practitioners and educators must grow alongside their students and look beyond standard practice towards innovative and dynamic means of assessment.

References

American Speech-Language-Hearing Association. (1993). *Definitions of Communication Disorders and Variations*. https://www.asha.org/policy/rp1993-00208/

American Speech-Language-Hearing Association (2004). Preferred practice patterns for the profession of Speech-Language Pathology. https://www.asha.org/policy/pp2004-00191/#sec1.3.19

Bain, B. J., & Olswang, L. B. (1995). Examining Readiness for Learning Two-Word Utterances by Children With Specific Expressive Language Impairment. *American Journal of Speech-Language Pathology*, *4*(1), 81–91. https://doi.org/10.1044/1058-0360.0401.81

Bearman, M., & Dawson, P. (2013). Qualitative synthesis and systematic review in health professions education. *Medical Education*, *47*(3), 252–260. https://doi.org/10.1111/medu.12092

Bedore, L. M., & Peña, E. D. (2008). Assessment of Bilingual Children for Identification of Language Impairment: Current Findings and Implications for Practice. *International Journal of Bilingual Education and Bilingualism*, *11*(1), 1–29. https://doi.org/10.2167/beb392.0

Brown, A. and Ferrara, R. A. (1985) Diagnosing zones of proximal development. In J. V. Wertsch (Ed.) *Culture, Communication and Cognition. Vygotskian Perspectives.* (pp. 273-305). Cambridge: Cambridge University Press.

Budoff, M. (1974). *Learning potential and educability among the educable mentally retarded.* (Final Report Project No. 312312). Cambridge, MA: Research Institute for Educational Problems, Cambridge Mental Health Association.

Camilleri, B., & Law, J. (2007). Assessing children referred to speech and language therapy: Static and dynamic assessment of receptive vocabulary. *Advances in Speech-Language Pathology*, *9*(4), 312–322. https://doi.org/10.1080/14417040701624474

Camilleri, B., & Botting, N. (2013). Beyond static assessment of children's receptive vocabulary: the dynamic assessment of word learning (DAWL). *International Journal of Language &Amp; Communication Disorders*, *48*(5), 565–581. https://doi.org/10.1111/1460-6984.12033

Campione, J., & Brown, A. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82–115). New York: Guilford.

Carlson, J. S., & Wiedl, K. H. (1978). Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. *American Journal of Mental Deficiency*, *82*(6), 559–564.

Carlson, J. S., & Wiedl, K. H. (1992). The Dynamic Assessment of Intelligence. *Springer EBooks*, 167–186. https://doi.org/10.1007/978-1-4612-4392-2_6

Chamberlain, S. C. (2005). Recognizing and Responding to Cultural Differences in the Education of Culturally and Linguistically Diverse Learners. *Intervention in School and Clinic*, *40*(4), 195–211. https://doi.org/10.1177/10534512050400040101

Cummins, J. (2000). *Language, Power and Pedagogy: Bilingual Children in the Crossfire*. Multilingual Matters Limited.

De Lamo White, C., & Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language & Communication Disorders*, *46*(6), 613–627. https://doi.org/10.1111/j.1460-6984.2011.00049.

Deutsch, R.M. (2003). Mediation from the Perspective of the Theory of Mediated Learning Experience. *Journal of Cognitive Education and Psychology,* 3(1)29-45.

Dollaghan, C. A., & Horner, E. A. (2011). Bilingual Language Assessment: A Meta-Analysis of Diagnostic Accuracy. *Journal of Speech Language and Hearing Research*, *54*(4), 1077–1088.

Elliott, C.D. (1996) *British Ability Scales II*. Windsor: NFER-Nelson.

Elliott, J.G., Grigorenko, E.L., & Resing, W.C.M. (2010). Dynamic assessment. In P. Peterson, E., Baker, & B. McGaw (Eds.), *International Encyclopedia of Education* (Vol. 3, pp. 220-225). Oxford: Elsevier.

Feuerstein, R., Rand, Y., & Hoffman, M. B. (1979). *The dynamic assessment of retarded performers: the learning potential assessment device, theory, instruments, and techniques*. Baltimore, MD: University Park Press.

Feuerstein, R., Miller, R., Rand, Y., & Jensen, M. H. (1981). Can Evolving Techniques Better Measure Cognitive Change? *Journal of Special Education*, *15*(2), 201–219. https://doi.org/10.1177/002246698101500209

Feuerstein, R., Rand, Y., Jensen, M. R., Kaniel, S., & Tzuriel, D. (1987). Prerequisites for assessment of learning potential: The LPAD model. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 35–51). The Guilford Press.

Feuerstein, R., Klein, P.S. *& Tannenbaum, A.J. (1991). *Mediated Learning Experience: Theoretical, Psychological and learning Implications*. England: Freund Publishing House Ltd.

Gindis, B. (1999). Vygotsky's Vision: Reshaping the Practice of Special Education for the 21st Century. *Remedial and Special Education*, *20*(6), 333–340. https://doi.org/10.1177/074193259902000606

Ginsburg, H. (1997). *Entering the Child's Mind: The Clinical Interview In Psychological Research and Practice*. Cambridge University Press.

Goodenough, F.L. (1926) *Measurements of Intelligence by Drawings*. New York: World Book Co.

Grosser, M., & De Waal, E. (2008). Recentering the teacher: from transmitter of knowledge to mediator of learning. *Education as Change*, *12*(2), 41–57. https://doi.org/10.1080/16823200809487206

Gutiérrez-Clellen, V.F., & Peña, E. (2001). Dynamic Assessment of Diverse Children. *Language Speech and Hearing Services in Schools*, *32*(4), 212–224. https://doi.org/10.1044/0161-1461(2001/019)

Gutierrez-Clellen, V. F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics*, *29*(1), 3–19. https://doi.org/10.1017/s0142716408080016

Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2012). Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Language Teaching and Therapy*, *29*(1), 57–75. https://doi.org/10.1177/0265659012459526

Hemsley, G., Holm, A., & Dodd, B. (2014). Identifying language difference versus disorder in bilingual children. *Speech, Language and Hearing*. https://doi.org/10.1179/2050572813y.0000000027

Hunt, E., Nang, C., Meldrum, S., & Armstrong, E. (2021). Can Dynamic Assessment Identify Language Disorder in Multilingual Children? Clinical Applications From a Systematic Review. *Language, Speech, and Hearing Services in Schools*, *53*(2), 598–625. https://doi.org/10.1044/2021_lshss-21-00094

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic Assessment of Word Learning Skills: Identifying Language Impairment in Bilingual Children. *Language Speech and Hearing Services in Schools*, *43*(1), 81–96. https://doi.org/10.1044/0161-1461(2011/10-0095

Koons, G. L., Schenke-Layland, K., & Mikos, A. G. (2019b). Why, When, Who, What, How, and Where for Trainees Writing Literature Review Articles. *Annals of Biomedical Engineering*, *47*(11), 2334–2340. https://doi.org/10.1007/s10439-019-02290-5

Kowal, S., & O'Connell, D. L. (2016). Ragnar Rommetveit's Approach to Everyday Spoken Dialogue from Within. *Journal of Psycholinguistic Research*. https://doi.org/10.1007/s10936-015-9404-0

Kmet, L. M., Lee, R. C., Cook, L.S., & Alberta Heritage Foundation for Medical Research. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields.* Alberta Heritage Foundation for Medical Research. https://www.deslibris.ca/ID/200548

Laing, S. P., & Kamhi, A. (2003). Alternative Assessment of Language and Literacy in Culturally and Linguistically Diverse Populations. *Language, Speech, and Hearing Services in Schools*, *34*(1), 44–55. https://doi.org/10.1044/0161-1461(2003/005)

Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: bringing the past into the future. *Journal of Applied Linguistics*, *1*(1), 49–72. https://doi.org/10.1558/japl.1.1.49.55872

Lazewnik, R., Creaghead, N. A., Smith, A. D. B., Prendeville, J., Raisor-Becker, L., & Silbert, N. H. (2019). Identifiers of Language Impairment for Spanish–English Dual Language Learners. *Language Speech and Hearing Services in Schools*, *50*(1), 126–137. https://doi.org/10.1044/2018_lshss-17-0046

Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment*. New York, NY: Guilford Press.

Lidz, C. S. (2002). *Early Childhood Assessment* (1st ed.). Wiley.

Lidz, C. S., & Gindis, B. (2003). Dynamic Assessment of the Evolving Cognitive Functions in Children. *Vygotsky's Educational Theory in Cultural Context*, 99–116. https://doi.org/10.1017/cbo9780511840975.007

Lidz, C. S., & Peña, E. D. (2009). Response to Intervention and Dynamic Assessment: Do We Just Appear to Be Speaking the Same Language? *Seminars in Speech and Language*, *30*(02), 121–133. https://doi.org/10.1055/s-0029-1215719

Luria, A. R. (1961). Study of the abnormal child. *American Journal of Orthopsychiatry*, *31*, 1–16.

Mennen, I., & Stansfield, J. (2006). Speech and language therapy service delivery for bilingual children: a survey of three cities in Great Britain. *International Journal of Language & Communication Disorders*, *41*(6), 635–652. https://doi.org/10.1080/13682820600623911

Miller, L., Gillam, R. B., & Peña, E. (2001). Dynamic Assessment and Intervention: Improving Children's Narrative Abilities. Austin, TX: PRO-ED

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine*, *151*(4), 264. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Moll, L., & Whitmore, K. (1993). Vygotsky in educational settings: Moving from individual transmission to social transaction. *Contexts for learning: Sociocultural dynamics in children's development*, 19-42.

Moore-Brown, B. J., Huerta, M. J. D., Uranga-Hernandez, Y., & Peña, E. D. (2006a). Using Dynamic Assessment to Evaluate Children With Suspected Learning Disabilities. *Intervention in School and Clinic*. https://doi.org/10.1177/10534512060410040301

Olswang, L. B., & Bain, B. J. (1996). Assessment Information for Predicting Upcoming Change in Language Production. *Journal of Speech Language and Hearing Research*, *39*(2), 414–423. https://doi.org/10.1044/jshr.3902.414

Orellana, C., Wada, R., & Gillam, R. B. (2019). The Use of Dynamic Assessment for the Diagnosis of Language Disorders in Bilingual Children: A Meta-Analysis. *American Journal of Speech-Language Pathology*, *28*(3), 1298–1317. https://doi.org/10.1044/2019_ajslp-18-0202

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T., Mulrow, C. D., Shamseer, L., Tetzlaff, J., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Paradis, J., Genesee, F., & Crago, M. B. (2011). *Dual Language Development and Disorders: A Handbook on Bilingualism and Second Language Learning*. Van Haren Publishing.

Peña, E., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language assessment: A non-biased procedure. *Journal of Special Education, 26*(3), 269–280.

Peña, E. (1993). *Dynamic assessment: A non-biased approach for assessing the language of young children.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses.

Peña, E., Iglesias, A., & Lidz, C. S. (2001). Reducing Test Bias Through Dynamic Assessment of Children's Word Learning Ability. *American Journal of Speech-Language Pathology*, *10*(2), 138–154. https://doi.org/10.1044/1058-0360(2001/014)

Peña, E. D., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C. E., & Sabel, T. (2006). Dynamic Assessment of School-Age Children's Narrative Ability: An Experimental Investigation of Classification Accuracy. *Journal of Speech Language and Hearing Research*, *49*(5), 1037–1057.

Peña, E. D., Reséndiz, M., & Gillam, R. B. (2007). The role of clinical judgments of modifiability in the diagnosis of language impairment. *Advances in Speech Language Pathology, 9*(4), 332–345. https://doi.org/10.1080/14417040701413738

Peña, E. D., Gillam, R. B., & Bedore, L. M. (2014). Dynamic Assessment of Narrative Ability in English Accurately Identifies Language Impairment in English Language Learners. *Journal of Speech Language and Hearing Research*, *57*(6), 2208—2220 https://doi.org/10.1044/2014_jslhr-l-13-0151

Petersen, D. B., Chanthongthip, H., Ukrainetz, T. A., Spencer, T. D., & Steeve, R. W. (2017). Dynamic Assessment of Narratives: Efficient, Accurate Identification of Language Impairment in Bilingual Students. *Journal of Speech Language and Hearing Research*, *60*(4), 983–998. https://doi.org/10.1044/2016_jslhr-l-15-0426

Petersen, D. B., Tonn, P., Spencer, T. D., & Foster, M. C. (2020). The Classification Accuracy of a Dynamic Assessment of Inferential Word Learning for Bilingual English/Spanish-Speaking School-Age Children. *Language Speech and Hearing Services in Schools*, *51*(1), 144–164. https://doi.org/10.1044/2019_lshss-18-0129

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, *9*(3), 233–265. https://doi.org/10.1191/1362168805lr166oa

Poehner, M.E. (2008). *Dynamic Assessment: A Vygotskian approach to understanding and promoting second language development.* Berlin, Germany: Springer.

Poehner, M. E., & Lantolf, J. P. (2010). Vygotsky's Teaching-Assessment Dialectic and L2 Education: The Case for Dynamic Assessment. *Mind, Culture, and Activity*, *17*(4), 312–330. https://doi.org/10.1080/10749030903338509

Prezas, R. F. C., & Jo, A. A. (2017). Differentiating Language Difference and Language Disorder: Information for Teachers Working with English Language Learners in the Schools. *Journal of Human Services: Training, Research, and Practice*, *2*(1), 2.

Rommetveit, R. (1976). On the Architecture of Intersubjectivity. In *Springer eBooks* (pp. 201–214). https://doi.org/10.1007/978-1-4615-8765-1_16

Royal College of Speech & Language Therapists. (2007). *Good practice for speech and language therapists working with clients from linguistic minority communities.* https://www.rcslt. org/members/publications/publications2/slt_assessment_and_ intervention_main

Smagorinsky, P. (2018). Deconflating the ZPD and instructional scaffolding: Retranslating and reconceiving the zone of proximal development as the zone of next development. *Learning, Culture and Social Interaction*, *16*, 70–75. https://doi.org/10.1016/j.lcsi.2017.10.009

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measure- ment of learning potential*. New York: Cambridge University Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* Cambridge, Mass.: Harvard University Press

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Williams, C. J., & McLeod, S. (2012). *Speech-language pathologists' assessment and intervention practices with multilingual children*. International Journal of Speech-Language Pathology, 14(3), 292–305. https://doi.org/10.3109/17549507.2011.636071

Appendix A: The Response to Mediation Scale (Lidz, 1991)

A. Self-regulation of attention:

    1. Unable to maintain attention to task
    2. Fleeting attention to task even with input from adult
    3. Maintains with significant input from adult
    4. Maintains with occasional input from adult
    5. Maintains with no input from adult

    * Does not apply

B. Interactivity with the mediator

    1. Does not engage in turn-taking communication
    2. Minimal engagement in turn-taking communications
    3. Moderate engagement in turn-taking communications
    4. Comfortable, frequent engagement in turn-taking communications
    5. Initiate and responds appropriately and expansively in several chains of conversational interactions

    * Does not apply

C. Responsiveness to initiations of mediator

    1. Resistive to mediator's initiatives
    2. Passive noncompliant
    3. Passive, minimally responsive
    4. Consistently responsive
    5. Enthusiastic and responsive

    * Does not apply

D. Response to challenge

    1. Refuses, cries, or tantrums in response to challenge
    2. Begins, but quickly gives up
    3. Persists, but with significant encouragement from adult.
    4. Persist and completes task, with minimal adult encouragement
    5. Energized by challenge; enjoys the challenge

    * Does not apply

E. Use of adult as a resource when child needs help

    1. Does not refer to adult
    2. Nonverbally, passively signals need for help
    3. Nonverbally actively seeks help
    4. Verbally asks for help
    5. Actively seeks help and seems to appreciate help provided

    * Does not apply

F. Interest in activity materials

    1. Shows dislike of materials
    2. Neutral reaction to materials
    3. Minimal interest in materials
    4. Fluctuating interest in materials
    5. Consistently strong interest in materials

    * Does not apply

Appendix B: The Mediated Learning Experience Rating Scale (Lidz, 1991)

MEDIATED LEARNING EXPERIENCE (MLE) RATING SCALE
(for use with parent-child, teacher-child, examiner-child interactions with preschool children)

Developed by Carol S. Lidz, PsyD
Based on the theory and research of Prof. Reuven Feuerstein

Child: _____   Mediator: _____   Task: _____

Rater: _____   Date: _____   Location: _____

INTENTIONALITY: a conscious attempt by the mediator to influence the behavior of the child. This includes communication to the child of the purpose for the interaction, as well as attempts by the mediator to maintain the child's involvement in the interaction. For children who are already self-regulating and do not require interventions by the mediator to engage them in the activity, rating of intentionality includes the readiness of the mediator to become involved as necessary; therefore, the mediator shows ongoing interest in the activity involvement of the child (in this case, the rating would be a 2, unless a statement of a principle is provided).

0 = no evidence
1 = inconsistently present; loses involvement
2 = consistently in evidence
3 = in evidence, with statement or encouragement of a principle to induce self-regulation in the child; this principle would apply to the child's ability to maintain attention and inhibit impulsivity
NOTES:

TRANSCENDENCE: promotion of cognitive bridges between the task or activity and related but not currently present experiences of the child; these may refer to the past or may anticipate the future. These bridges must promote visual images and help to move the child from the perceptual to the conceptual.

0 = not in evidence
1 = simple, nonelaborated reference to past or future experience
2 = elaborated reference
3 = elaborated reference that includes hypothetical, inferential, or cause-and-effect thinking
NOTES:

MEANING: moving the content from neutral to a position of value and importance; this may be done by affective emphasis or stating that the object or aspect of focus is important and should be noticed (or, in contrast, that it is negative and to be ignored or avoided).

0 = not in evidence
1 = calling up labels or concepts already within the child's repertory; saying that it is important and should be noticed (e.g., "Look at this"), but without elaboration
2 = adding animation or affect to make the activity come alive and provoke interest
3 = elaboration that expands the information about the activity or object; this elaboration addresses information that is perceptible to the child within the situation
NOTES:

COMPETENCE (Task Regulation): manipulation of the task to facilitate mastery by the child.

0 = not in evidence
1 = simple directions or passive manipulation of the task (e.g., holding it, moving pieces toward the child, building a model without elaborated directions)
2 = elaborated directions; nonverbal organization into a kind of conceptual grouping
3 = induction/statement/encouragement of strategic thinking and a planful attitude (e.g., "Where shall we start?" "What should we do first?"), or statement of a principle that the child can use to solve similar problems

NOTES:

Appendix C: Learning Strategies Checklist and Modifiability Scale (Peña, 1993)

Learning Strategies Checklist

	None of the time	Some of the time	Most of the time
Attention/Discrimination			
• initiates focus with minimum cues	0	1	2
• maintains focus with minimum cues	0	1	2
• responds to relevant cues, ignores irrelevant cues	0	1	2
Comparative Behavior			
• comments on features of task	0	1	2
• uses comparative behavior to select item	0	1	2
• talks about same/different	0	1	2
Planning			
• talks about overall goal	0	1	2
• talks about plan	0	1	2
Self-Regulation/Awareness			
• waits for instructions	0	1	2
• seeks help when difficult	0	1	2
• corrects self	0	1	2
• rewards self	0	1	2
Transfer			
• applies strategies within tasks	0	1	2
• applies strategies between tasks	0	1	2
Motivation			
• persists even when frustrated	0	1	2
• shows enthusiasm	0	1	2

Modifiability Scale

	Extreme	High-Moderate	Moderate	Slight
Examiner Effort	3	2	1	0
Child Responsivity	3	2	1	0

	Low	Medium	High
Transfer	0	1	2

Appendix D: Mediated Learning Observation Form (Peña, 2007)

Mediated Learning Observation

	1	2	3	4	5
Internal Social-Emotional (Affect)					
Anxiety	Calm, little to no soothing required	Fidgety, but can be soothed	Uncomfortable, breaks needed to sooth	Distressed, much soothing required	Distraught, crying, cannot be soothed
Motivation	Enthusiastic, engages in tasks readily	Curious, shows interest	Ambivalent, unsure about tasks	Guarded, seems fearful of tasks	Avoidant, does not want to engage
Non-verbal persistence	Persistent, wants to continue despite difficulty	Indicates difficulty non-verbally, but continues	Tentative, appears unsure about continuing	Demonstrates non-verbal frustration, continues under protest	Non-verbal rejecting, cannot continue

Comments

	1	2	3	4	5
Cognitive Arousal					
Task orientation	Completely understands tasks	Mostly understands tasks (75%)	Understands tasks some of the time (50%)	Often does not understand tasks (25% of the time)	Doesn't understand tasks
Meta-cognition	Aware of all errors	Aware of most errors (75%)	Aware of some errors (50%)	Unaware of most errors (25%)	Unaware of any errors
Non-verbal self reward	Positive response to task regardless of difficulty	Positive response related to task difficulty	Demonstrates insecurity, positive & negative responses related to difficulty	Negative response related to task difficulty	Negative response regardless of task difficulty

Comments

	1	2	3	4	5
Cognitive Elaboration					
Problem-solving	Systematic and efficient, used forethought, reflection	Organized, but somewhat inefficient, (less than 25% off task)	Sketchy plan, trail & error	Disorganized, haphazard plan	No plan; unsystematic guessing
Verbal mediation	Elaborates plan clearly	Talks through problem	Talks occasionally	1–2 word utterances only	No verbal mediation
Flexibility	Uses multiple strategies readily	Has preferred strategies, but can change when necessary	Some evidence of more than one strategy and occasionally utilizes them	Recognizes limitations of strategy, but cannot see alternatives	Persists with one strategy, regardless of outcome

Comments

	1	2	3	4	5
External Social-Emotional (Behavior)					
Responsiveness to feedback	Very positive, maintains enthusiasm	Positive, but hesitant; requires some feedback	No response to feedback	Negative, disheartened; requires much feedback	Very negative, rejects feedback
Attention	Attentive and focused	Focused, but distractible at times	Distractible, but can be refocused, needs prompting	Distracted, and difficult to refocus	Distracted and off task
Compliance	Cooperative	Insecure	Hesitant	Uncooperative	Refusing

Comments

Appendix E: Checklist for Assessing the Quality of Quantitative Studies (Kmet, 2004)

	Criteria	Yes (2)	Partial (1)	No (0)	N/A
1	Question / objective sufficiently described?				
2	Study design evident and appropriate?				
3	Method of subject/comparison group selection *or* source of information/input variables described and appropriate?				
4	Subject (and comparison group, if applicable) characteristics sufficiently described?				
5	If interventional and random allocation was possible, was it described?				
6	If interventional and blinding of investigators was possible, was it reported?				
7	If interventional and blinding of subjects was possible, was it reported?				
8	Outcome and (if applicable) exposure measure(s) well defined and robust to measurement / misclassification bias? Means of assessment reported?				
9	Sample size appropriate?				
10	Analytic methods described/justified and appropriate?				
11	Some estimate of variance is reported for the main results?				
12	Controlled for confounding?				
13	Results reported in sufficient detail?				
14	Conclusions supported by the results?				

Appendix F: Checklist for Assessing the Quality of Qualitative Studies (Kmet, 2004)

Criteria		Yes (2)	Partial (1)	No (0)
1	Question / objective sufficiently described?			
2	Study design evident and appropriate?			
3	Context for the study clear?			
4	Connection to a theoretical framework / wider body of knowledge?			
5	Sampling strategy described, relevant and justified?			
6	Data collection methods clearly described and systematic?			
7	Data analysis clearly described and systematic?			
8	Use of verification procedure(s) to establish credibility?			
9	Conclusions supported by the results?			
10	Reflexivity of the account?			