

Critical Review:
Investigating the Classification Accuracy of Dynamic Assessment with Culturally and Linguistically Diverse Children

Devan Gingrich
M.Cl.Sc SLP Candidate
University of Western Ontario: School of Communication Sciences and Disorders

This critical review examined published literature investigating the classification accuracy of the dynamic assessment of narratives in culturally and linguistically diverse children. A literature search yielded five case-control studies which met inclusion criteria. Overall, findings indicated that dynamic assessment of narratives could accurately classify CLD children with and without language impairment with high sensitivity and specificity.

Introduction

Language impairments in children have historically been identified using standardized language assessments. These assessments rely on comparisons between the child's level of performance on a given test to the performance of norm-referenced peers. Traditional assessment tools, such as standardized assessments, tend to reflect the cultural values, knowledge and communication strategies of their culture of origin (Carter, et al., 2004). Culturally and linguistically diverse (CLD) children are often under-represented or absent from normative samples (Qin Teoh, Brebner, & McCormack, 2012). Therefore, a CLD child's limited performance on standardized assessment may not accurately reflect their language learning potential (Gutiérrez-Clellen & Pena, 2001).

Over the past 20 years, it has been proposed that dynamic assessment represents a culturally and linguistically fair alternative to traditional standardized assessments for identifying the presence of language impairment in CLD children (Gutiérrez-Clellen & Pena, 2001). Dynamic assessment has been influenced by Vygotsky's sociocultural theory and Feuerstein's theory of mediated learning experiences (MLE) (Pena et al., 2006). Vygotsky, in his research regarding the zone of proximal development, proposed that learners could successfully complete a previously difficult to perform cognitive task when provided with appropriate support through mediation and/or teaching support (Gillam, Pena & Miller 1999; Petersen et al., 2017). Feuerstein's MLE theory further extended Vygotsky's findings to include a focus on the child's behaviour during mediated learning experiences (Pena et al., 2006).

Driven by these theories, current dynamic assessment in speech-language pathology uses a test-teach-retest format to evaluate the child's response to intervention. The child's language learning ability is

quantified by measuring gains in performance from pre-test to post-test and the evaluation of the child's use of cognitive strategies to modify outcomes, referred to as modifiability. In examining a child's language learning potential, rather than their existing language knowledge, dynamic assessment avoids many of the biases and classification issues inherent in norm-referenced assessment tools (Petersen et al., 2017).

Although a large body of evidence supporting the use of dynamic assessment exists, there is substantial variability among the types used in practice, and among which methods are most effective and efficient at accurate identification (Gutiérrez-Clellen & Pena, 2001). The dynamic assessment of narratives has recently shown considerable potential as a diagnostic indicator differentiating language disorder from language difference in CLD children. The current model for the dynamic assessment of narratives uses the traditional test-teach-retest format. In both the test and retest phases, the child creates stories that correspond to wordless picture books. Following the pretest phase, the child's narrative structures are evaluated and targeted areas of narrative structure are established as goals for the teach phase. The child's language learning abilities are then quantified during the teach and retest phases of assessment by evaluating their response to intervention, examiner effort required during mediation and overall pretest to post-test changes in targeted narrative structures. It has been proposed that narratives are a universal genre that provide an appropriate medium for assessing language cross-culturally (Kramer et al., 2009). It has further been proposed that early narrative language skills offer a strong predictor of later language and literacy skills (Bishop & Edmundson, 1987; Fazio, Naremore, & Connell, 1996; Wetherell, Botting, & Conti-Ramsden, 2007).

As the multicultural landscape of Canada expands, it is imperative that Canadian clinicians develop a stronger understanding of the evidence-base surrounding the dynamic assessment of narratives. In evaluating the evidence, clinicians can make informed decisions regarding its clinical utility when assessing and creating intervention goals for CLD children.

Objectives

The purpose of this critical review was to evaluate the existing literature on the classification accuracy of the dynamic assessment of narratives with culturally and linguistically diverse children.

Methods

Search Strategy

Online databases (PubMed, Google Scholar, Web of Science, CINAHL) were searched using the following search terms: (dynamic assessment) AND (narrative) AND (bilingual) OR (multicultural) OR (culturally diverse) OR (linguistically diverse). Reference lists of related or included articles were further reviewed for additional resources.

Selection Criteria

Studies selected for inclusion were required to investigate the classification accuracy of the dynamic assessment of linguistically and/or culturally diverse participants. Studies were required to use narratives as the only form of dynamic assessment. Studies were limited to those which evaluated participants (under the age of 18), and were written in English.

Data Collection

Results of the literature search yielded five case-control studies that met selection criteria.

Results

Case-Control Studies

Case-control studies clearly define two identical groups of participants which are differentiated only by the presence or absence of a condition. Participant group allocation is determined based on a particular condition, in this critical review the presence of a language impairment, and is therefore not a randomized study design. Case-control studies are appropriate for questions regarding the classification accuracy of alternative forms of assessment, such as dynamic assessment. The study design retrospectively compares classification outcomes on one, or several traditional assessment approaches to those reached using dynamic assessment. Therefore,

allowing researchers to draw conclusions on the power and validity of the alternative assessment approach.

Henderson, Adelaida and Aiken (2018) conducted a case-control study to determine whether the dynamic assessment of narratives in English could accurately identify language impairment in Navajo preschoolers. Dynamic assessment of narratives was completed using the Language subtest of the Predictive Early Assessment of Reading and Language (PEARL; Petersen & Spencer, 2014). The PEARL uses the traditional dynamic assessment framework of test-teach-retest to evaluate the vocabulary and story grammar of preschool-age children. The PEARL measures modifiability (the child's response to intervention) in the areas of language production and comprehension. The Language subtest of the PEARL evaluates three areas of narrative production: story grammar, language complexity and episode.

Participants included a case group of 45 language impaired preschoolers, and a control group of 45 typically developing preschoolers. Participant inclusion criteria was well described, along with an extensive a priori classification system. The PEARL was administered in English over 1 session consisting of 2 assessment and 4 mediation phases. In both the test and retest phases the examiner read a short story and the participant was asked to retell the story. Narratives were scored online using PEARL protocol, which evaluated story grammar, language complexity and episode. Teach phases consisted of the examiner targeting missed narrative elements and the participant retelling the pretest story with accompanying pictures representing story content and icons representing story grammar. Following each teach phase, the participants' responsiveness to mediation and behaviour was rated using the PEARL's modifiability rating scale. Behaviours evaluated included number of prompts needed, the child's confidence, the amount of child disruption, rate of task completion, quantity of assistance required by the examiner and the child's level of comfort and frustration with the task. Appropriate blinding and interrater agreement were described for all phases of testing.

Appropriate statistical analysis indicated that the PEARL accurately identified Navajo preschoolers with language impairment with high sensitivity and specificity. Modifiability scores combined with post-test narrative scores were the best indicators of accurate classification.

Strengths of the study included a large sample size, strong fidelity and well described methods for replicability. Weaknesses included the limited geographic representation of participants from the overall Navajo population, along with lack of detail regarding how case and control groups were matched. Overall, this study provides compelling evidence that the dynamic assessment of narratives using the PEARL is an effective and efficient measure in identifying Navajo preschoolers with language impairment.

Kramer, Mallett, Schneider and Hayward (2009) conducted a case-control study to determine whether the dynamic assessment of narratives in English was accurate in identifying language learning difficulties in First Nations students. Dynamic assessment of narratives was completed using the Dynamic Assessment and Intervention Tool (DAI) (Miller et al., 2001). The DAI focuses on a child's narrative abilities to determine language learning potential using a test-teach-retest format.

Participants included a case group of 5 third grade children identified by school personnel as having a possible language learning difficulty, and a control group of 12 third grade children identified as normal language learners. The DAI was administered in English over 2 assessment and 2 mediation sessions. In the test phase participants were asked to create an oral narrative from a wordless picture book without any assistance from the examiner. In the teach phase, participants received supported mediation from the examiner directly targeting missed narrative elements. In the retest phase participants produced a second oral narrative from a new wordless picture book without any assistance from the examiner.

Narratives were scored online by examiners assigning a numeric value between one and seven for story components, story ideas and language, and episode elements. Appropriate blinding of examiners was reported, and interrater agreement was reached following consensus between coders on all narrative structures. Following each teach session, responsiveness to mediation was rated using two Likert scales to quantify modifiability. Scale one described the amount of effort required to teach the participant, and scale two described the participants' responsiveness to teaching.

Appropriate statistical analysis revealed that the DAI accurately classified participants into similar groups as the a priori classification made by school personnel with acceptable sensitivity and specificity. Average change in DAI story score on targeted

components of mediation between pre-and post-test was the best predictor of overall classification.

Strengths of the study included appropriate research design and sufficient detail for replicability. Weaknesses included the small sample size, lack of inclusion criterion, the use of school personnel for a priori classification and lack of information regarding interrater reliability for modifiability scales. Overall, this study provides somewhat suggestive evidence that the dynamic assessment of narratives using the DAI is an effective and accurate measure in identifying First Nations students with language impairment.

Pena, Gillam, and Bedore (2014) conducted a case-control study to determine the identification accuracy of dynamic assessment of narrative ability in English for English language learning children. Dynamic assessment of narratives was completed using the Dynamic Assessment and Intervention Tool (DAI).

Participants included a case group of 18 children with language impairment, a control group of 18 age, sex, language experience and IQ matched children with normal language development, and a compare-control group of 18 age and language experience matched children with normal language development. Participant inclusion criteria was well described, along with an acceptable a priori classification system. High interrater agreement was reported for a priori classification.

Participants were assessed in English over three sessions using two wordless picture books. At the pretest and post-test phases, participants were asked to create an oral narrative from a wordless picture book with minimal cuing from the examiner. Narratives were scored using DAI protocol to produce a total story score consisting of 10 qualitative items. Items were divided into the three subsections of story components, story ideas and language, and episode structure. High interrater reliability was reported in both transcription and scoring. Participants modifiability (responsivity to mediation) during the first teach phase of assessment were observed and rated using the 12 item Mediated Learning Observation (MLO) form. The MLO reflects examiner judgements of participants' learning performance in the areas of affect, arousal, elaboration and behaviour. Examiners assigned a numeric value between one and five for each item on the MLO form, with one indicating the participant required little examiner support and five indicating the participant required maximum examiner support.

Appropriate statistical analysis revealed that the dynamic assessment of narratives in English with English language learning children could accurately classify children with and without language impairment. A combination of modifiability scores and post-test narrative measures yielded strong sensitivity and specificity.

Strengths of the study included the use of a compare-control group to cross-validate findings, an appropriate study design, and the analysis of specific narrative and modifiability measures that best differentiated children by language learning ability. Weaknesses of the study included a small sample size of language impaired children, and lack of information regarding interrater reliability for modifiability scales. Overall, this study provides highly suggestive evidence that the dynamic assessment of narratives using the DAI is an effective measure in identifying language impairment in English language learning children.

Pena et al. (2006) conducted a case-control study to determine the classification accuracy of dynamic assessment of narrative ability in culturally diverse school-age children. Dynamic assessment of narratives was completed using the Dynamic Assessment and Intervention Tool (DAI). Participants included a case group of 14 children with language impairment, a control group of 27 children with typical language development and a no-treatment control group of 30 children with typical language development. An acceptable a priori classification system was well described.

Participants were assessed in English over four sessions using two wordless picture books and a test-teach-retest approach. In the test and re-test phases of assessment, participants were instructed to generate a story that corresponded with the pictures from one of the wordless books. The participants spontaneous oral narratives were recorded, transcribed and scored for 10 aspects of narrative language according to DAI protocol. The protocol divides narrative language among the three subcategories of story components, story ideas and language, and episode structure, which combine to form the DAI total story score. During the teach phases of assessment, examiners administered the same scripted 30-minute mediated learning (MLE) sessions to all participants. MLE sessions focused on increasing length and complexity of narratives. Participants modifiability (responsivity to mediation) during both MLE sessions was assessed using a 5-point Likert scale. Examiners used the scale to make judgements on the amount of examiner effort

and support required by each participant during the teach phases of assessment.

Appropriate statistical analysis revealed that the dynamic assessment of narratives in English with culturally diverse children could accurately classify language impairment with high sensitivity and specificity. The best independent indicator of classification was the child's response to learning during MLE sessions (modifiability). Results indicated that a combination of modifiability, number of different words, total number of words and the total story scores at post-test yielded the highest classification accuracy.

Strengths of the study included the use of scripted MLE sessions allowing for standardization, strong fidelity and replicability of the study, and the use of a no-treatment control group. Weaknesses included lack of details regarding interrater agreement for a priori classification and modifiability scales, and lack of details regarding matching between control and case groups. Overall, this study provides highly suggestive evidence that the dynamic assessment of narratives in English accurately identifies culturally diverse children with and without language impairment.

Petersen et al. (2017) conducted a case-control study to determine the classification accuracy of a condensed version of dynamic assessment of narratives with bilingual school-age children. Participants included a case group of 10 children with language impairment, and a control group of 32 children with typical language development. Participant inclusion criteria and demographic information was well described, along with an extensive a priori classification system.

Participants were assessed in English over two short (25-30 min) sessions using a narrative retell format. Each session included a single test and retest narrative retell task, and multiple teaching cycles. In the test and re-test phases, the participant listened to a story modeled by the examiner and was asked to the retell the story supported by a wordless picture book. Narratives were scored for the presence and quality of story grammar elements and complexity of episodic structure in real-time by the examiners during the story retell. In the teach phase, the examiner cycled one to four times through a brief set of structured steps targeting story grammar, including any of the narrative elements omitted or poorly represented by the child's narrative in the test phase. Participants responsivity to mediation following each teach cycle was assessed using a 7-item modifiability

rating scale. The items focused on how frequently child behaviours occurred during the teach phase including, responsivity to prompts, transfer of targeted skills between cycles, attending to teaching, ease of teaching, level of frustration, level of disruption and overall potential to learn narrative language. High inter-and-intra-rater reliability was reported for all phases of testing, including qualitative judgements of modifiability made by examiners.

Appropriate statistical analysis revealed that a condensed form of dynamic assessment using a retell narrative format in English with bilingual children could accurately classify language impairment with high sensitivity and specificity. The best independent indicator of classification was the participants modifiability (response to intervention), as judged by examiners, using an overall modifiability rating.

Strengths of the study included the use of a structured 4-step teaching cycle allowing for replicability of the study, strong fidelity, and the overall clinical feasibility of the proposed assessment procedure. Weaknesses included the small sample size and lack of details regarding matching between control and case groups. Overall this study provides compelling evidence that the dynamic assessment of narratives, using a condensed test-teach-retest structure in English accurately identifies bilingual children with and without language impairment.

Discussion

This critical review analyzed five studies to determine the classification accuracy of the dynamic assessment of narratives with culturally and linguistically diverse children. Although there is some variation in the clinical importance and validity of the studies reported, the overall data suggests that the use of dynamic assessment of narratives can accurately classify CLD children with and without language impairment with high sensitivity and specificity. The dynamic assessment of narratives provides clinical insight into the learning process of CLD children. All five reviewed studies reported that children with a language impairment had difficulty learning and incorporating new information into the narrative structures even when provided with structured mediation sessions. Whereas, typically developing CLD children made more rapid changes in their narrative structures and were more responsive during structured mediation sessions.

The clinical indicators that best predicted classification varied between studies; however,

modifiability was the single best indicator of classification in four of the five reviewed studies (Henderson et al., 2018; Pena et al., 2014; Pena et al., 2006; Petersen et al., 2017). Despite the use of three unique modifiability rating scales, all four of these studies reported that CLD children with a language impairment demonstrated lower responsivity to change during mediation sessions and required higher examiner effort to elicit change. Post-test narrative scores were also a strong classification indicator in three of the five reviewed studies (Henderson et al., 2018; Pena et al., 2014; Pena et al., 2006). Post-test narrative scores in these three studies indicated that children with a language impairment consistently achieved lower scores at post-test when compared with typically developing children even after receiving higher levels of support during mediation sessions.

The relative consistency in the findings of all reviewed studies is quite compelling given the use of three different dynamic assessment protocols. Three of the reviewed studies used the DAI tool (Kramer et al., 2009; Pena et al., 2006; Pena et al., 2014). The DAI elicits and assesses narratives using wordless picture books which helped to constrain the task in ways that facilitate comparisons across individuals, between groups and across time. Based on the reviewed studies, the DAI offers highly suggestive evidence in regard to its clinical utility in accurately identifying language impairment with CLD children. The most compelling evidence exists for the dynamic assessment of narratives using narrative retell tasks (Henderson et al., 2018; Petersen et al., 2017). Narrative retell tasks in both the Henderson and Petersen studies allowed the examiner to complete online scoring and goal setting without the timely process of transcribing and scoring narratives between test and mediation/teach sessions. Dynamic assessment using narrative retell tasks offer strong clinical feasibility in an efficient format for busy speech-language pathologists.

Lack of information regarding reliability when quantifying participants response to intervention using modifiability scales was a common weakness among four of the five reviewed studies (Henderson et al., 2018; Kramer et al., 2009; Pena et al., 2014; Pena et al., 2006). Modifiability ratings are subjective ratings based on examiner observations of child effort during mediation. Due to absence of reliability in these studies, the face validity of the dynamic assessment procedure should be interpreted with some caution. However, one study did assess and establish strong intra and inter-rater reliability using a simple modifiability rating scale (Petersen et al.,

2017). It is recommended that future research seeks to investigate reliability in modifiability ratings across examiners to improve the overall validity of dynamic assessment using narratives. In addition, future research should include larger sample sizes and stronger information regarding case-control matching.

Clinical Implications

Overall, the evidence examined in this critical review strongly supports the use of dynamic assessment of narratives as an accurate indicator of language impairment in culturally and linguistically diverse children. The evidence suggests that modifiability ratings, as judged by examiners during the teach and retest phases of testing, are the most robust indicator of classification for children as language impaired or typically developing. Despite this strong body of evidence, the dynamic assessment of narratives as a clinically valid, culturally fair and unbiased tool has largely been overlooked in clinical practice. As speech-language pathologists seek to identify CLD children with and without language impairment, dynamic assessment can and should be used. The three (DAI, PEARL and condensed version of DA) protocols outlined in this review all offer highly suggestive to compelling evidence that dynamic assessment using narratives is an appropriate alternative to standardized assessment tools. Clinicians should strongly consider adopting the dynamic assessment of narratives as an adjunct to more traditional forms of assessment including language sample analysis, direct observation and parent or teacher report.

References

- Carter, J. A., Lees, J. A., Murira, G. M., Gona, J., Neville, B. G. R., & Newton, C. R. J. C. (2005). Issues in the development of cross-cultural assessments of speech and language for children. *International Journal of Language & Communication Disorders*, 40 (4), 385-401.
- Gutiérrez-Clellen, V. F., & Pena, E. (2001). Dynamic assessment of diverse children: a tutorial. *Language, Speech, and Hearing Services in Schools*, 32, 212-224.
- Hendersen, D. E., Restrepo, M. A., & Aiken, L. S. (2018). Dynamic assessment of narratives among Navajo preschoolers. *Journal of Speech, Language, and Hearing Research*, 61, 2547-2560.
- Kramer, K., Mallett, P., Schneider, P., & Hayward, D. (2009). Dynamic assessment of narratives with grade 3 children in a first nations community. *Canadian Journal of Speech-Language Pathology and Audiology*, 33 (3), 119-128.
- Miller, L., Gillam, R., & Pena, E. D. (2001). *Dynamic Assessment and Intervention: Improving children's narrative abilities*. Austin, TX: PRO-ED.
- Pena, 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, 2208-2220.
- Pena, E. D., Gillam, R. B., Malek, M., Ruiz-Felter, M., Resendiz, M., Fiestas, C., & 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, 1037-1057.
- 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, 983-998.
- Petersen, D.B., & Spencer, T. D. (2014). *The Predictive Early Assessment of Reading and Language (PEARL)*. Palmer, AK: Language Dynamics Group.