Use Authorization

In presenting this dissertation in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my dissertation for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this dissertation for financial gain shall not be allowed without my written permission.

Signature _____

Date _____

A COMPARISION OF NARRATIVE CONTENT ANALYSIS METHODOLOGIES USING THE ORAL AND WRITTEN NARRATIVES OF FOURTH-GRADE CHILDREN

by

Julie M. Spencer

A thesis

Submitted in partial fulfillment

of the requirements for the degree of

Master of Speech Language Pathology

in the Department of Communication Sciences and Disorders

Idaho State University

Spring 2018

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Julie M. Spencer find it satisfactory and recommend that it be accepted.

_____
Dr. Diane Ogiela
Major Advisor

_____
Dr. Kristina Blaiser
Committee Member

_____
Dr. Judith Crews
Committee Member

Research Approvals Page

Dec 1, 2017

Julie Spencer
Comm Sci Disorders/Deaf Educ
MS 8116

RE: regarding study number IRB-FY2018-129: A Comparison of Narrative Content Analysis Methodologies in the Oral and Written Narratives of Fourth-Grade Children

Dear Ms. Spencer:

I agree that this study qualifies as exempt from review under the following guideline: Category 4. Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.
This letter is your approval, please, keep this document in a safe place.

Notify the HSC of any adverse events. Serious, unexpected adverse events must be reported in writing within 10 business days.

You are granted permission to conduct your study effective immediately. The study is not subject to renewal.

Please note that any changes to the study as approved must be promptly reported and approved. Some changes may be approved by expedited review; others require full board review. Contact Tom Bailey (208-282-2179; fax 208-282-4723; email: humsubj@isu.edu) if you have any questions or require further information.

Sincerely,

Ralph Baergen, PhD, MPH, CIP
Human Subjects Chair

Dedication

I dedicate my thesis to my son, Nolan, who used to spin yarns for me about dragons who came over for peanut butter and jelly sandwiches, giant flying spiders that drove cool cars around Australia, and sheriffs that put villains in jail for the crime of making them wear pants. That little boy, now a much bigger boy, and his fantastical tales were my inspiration to study narratives and to find new ways for all children to find the joy in their own stories.

Acknowledgements

I would like to express my deep appreciation for Dr. Diane Ogiela's wise guidance in shepherding me through the process of writing this thesis. She sagely tried to deter me from taking the road less traveled, knowing better than I how much more difficult the process would be, then joyfully skipped down the path with me when she saw that I was determined. I'd like to thank my committee members, Dr. Kristina Blaiser, and Dr. Judith Crews for their support, kindness, patience and insightful questions, all of which spurred me on my way. My professors and faculty have been excellent examples of the frustrating joy to be found in research and for that I am very grateful. The excellent efforts of Angela Parkinson, Jeana Hansen, Taylor Maloney, and Jackie Thurman in collecting, transcribing and coding the data made coding a much easier process and they deserve much gratitude. Hannah Cassim not only performed the reliability check for this study but she was supportive and patient throughout the experience. It is impossible to summarize everything that my parents, Scott and Susan Johnson, have done to get me to this point, but it would take some grandiose words and several pages to make the attempt. To my son, Nolan Spencer, I am forever grateful because he reminded me to take time out for joy and fun no matter how busy I was. More importantly, he was my inspiration to stay dedicated to my work, so that he might be as proud of me as I am of him.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

INC	The Index of Narrative Complexity
INCs	Index of Narrative Complexity scores
LI	Language impairment
LLD	Language and learning disabled
MAIN	Multilingual Assessment Instrument for Narratives
MLU	Mean length of utterance
PDC	Percent of dyadic constituents
SALT	Systematic Analysis of Language Transcripts
SLP	Speech language pathology or speech language pathologist.
SL	Story level
SLI	Specific language impairment
STWR	The Six Traits Writing Rubric
TD	Typically developing or typical development
TNL	The Test of Narrative Language

Abstract

A COMPARISION OF NARRATIVE CONTENT ANALYSIS METHODOLOGIES IN THE ORAL AND WRITTEN NARRATIVES OF FOURTH-GRADE CHILDREN
Thesis Abstract—Idaho State University (2018)

This study compared narrative content assessments to explore differences of information offered by each method and to describe differences of oral and written narratives generated by typically developing children. The oral and written narratives of 21 typically developing fourth-grade children were evaluated using three measures: The Index of Narrative Complexity (INC), Story Level (SL), and Percent of Dyadic Constituents (PDC). All three measures demonstrated a normal distribution, indicating that each offered meaningful descriptive capability. There were no significant correlations between the complexity measures in the oral narratives and only SL and PDC demonstrated significant correlation in the written narratives, indicating fundamental differences in the type of information yielded. Oral narratives scored higher on all measures of complexity and productivity. The results support further research into research to describe the narratives of a wider age range and to determine whether the methods can be usefully applied to individuals with language impairment.

Introduction

Narrative, or telling stories, is integral to the human experience. Stories are part of how individuals interrelate, create and strengthen bonds with one another, and receive emotional support (Dautenhahn, K., 2001; Kali, 2004; Schickedanz, Schickedanz, Forsyth, & Forsyth, 2001; Silliman, 1989). When an individual tells a story to relay their experiences to another, the listener can connect with that story because of their similar experience or ability to imagine that experience. A person's sense of the world is constructed from stories about what is happening, stories about who other people are, and stories about what motivates the actions of others (Bruner, 1991; Heinriques, 2015). Stories can serve as a scaffold for academic success because stories teach children the formal structures that can be used in language while still using words familiar to the child; learning the structure of stories can prepare the child for less familiar academic language structures (Westby, 1991). A developing reader relies upon their knowledge of story structure for text comprehension which is a crucial skill for academic success (Catts & Kamhi, 2005).

With the recognition of the importance of narrative, it is natural that narrative would be the source of extensive research by speech language pathologists (SLPs). There is wide recognition in the field that assessment and treatment of narrative deficits has value and may improve academic outcomes for children with language impairments (Alves, Kennedy, Brown, & Solis, 2015; Catts, Fey, Tomblin, & Zhang, 2002; Gillam, Olszewski, Fargo, & Gillam, 2014; Justice, 2006; Karanski, 2013; Petersen, Gillam, Spencer, & Gillam, 2010; Snow et al, 2014; Stetter & Hughes, 2010; Swanson, Fey, Mills, & Hood, 2005). When children are taught to identify the basic elements of a story, it often assists them in comprehending the stories of others, both oral and written (Alves et al, 2015; Stetter & Hughes, 2010). Knowledge of the

structure of stories appears to assist the listener to pick out salient details of the story which improves overall comprehension of the story. A grounding in the basic elements of a story also provides more structure for the children to craft their own stories, improving the overall quality of the stories produced (Gillam et al, 2014; Koutsoftas & Gray, 2012; Petersen et al, 2010; Swanson et al, 2005). The Functional Language Intervention Program for Narratives, which scaffolds story-telling with pictures and verbal prompts, may increase overall usage of story grammar elements and causality within narratives (Petersen et al. 2010). The benefits of narrative instruction by a speech language pathologist within a classroom may improve the skills, not only of at-risk students, but of typically developing (TD) children as well (Gillam, Olszewski, Fargo, & Gillam, 2014). Such classroom-based instruction may improve reading comprehension, narrative production and increase lexical diversity—the number of different words used (Carretti et al., 2016; Green, Stockholm, Cearley, & Sheffield-Anderson, 2015; McNamara et al., 2011; Petersen, et al., 2010).

Narrative Content Analysis

A variety of means have been employed to assess and treat the production of narratives. Readily quantifiable measures of narrative productivity and complexity often used are: the total number of words used, mean length of utterance (MLU), lexical diversity, or clauses per T-unit, among others (Hughes, McGillivray, & Schmedek, 1997). There is debate in the field about how each measure should be employed, under which circumstances, and with which populations (Burns, de Villiers, Pearson, & Champion, 2012; Gillam, Olszewski, Fargo & Gillam, 2014; Justice et al, 2006). Objective ratings of the quality of narrative content is less readily employed due to the subjective nature of content evaluation. For instance, some approaches rate a narrative to be of higher quality if characters are named rather than unnamed

(Petersen, Gillam & Gillam, 2008), but this assumption can be called into question in the face of many age old and beloved children's stories that designate many characters by description only, such as "the wolf," "the pig," and "the woodcutter." When assessing a skill, an SLP often seeks higher tally counts of certain items to designate a skill as "better," but identifying more content areas used in a story may not be helpful to identify the quality of narratives when the skillful use of allusion may make the use of some story grammar components superfluous or ambiguous.

The ambiguities of narrative content make the objective evaluation of narrative quality problematic at best, and yet ratings among teachers and untrained laypersons can demonstrate significant consistency when evaluating the quality of narratives produced by children (Liles, 1985; Newman & McGregor, 2006; Petersen et al, 2008). This consistency would indicate that there is some objective and measurable standard by which the content of narratives may be evaluated, although difficult to operationalize for research purposes. To that end, there have been many different content analysis approaches, although the most commonly used is the story grammar approach (Applebee, 1978; Liles, 1985; Petersen et al, 2008). Story grammar has its roots in the work of Vladimir Propp (1928), a Russian linguist who engaged in the analysis of folktales. Since the same folktale could be told in many regions but with characters interchanged depending upon the region where tale was told, Propp rejected the idea that folktales should be analyzed based primarily upon what characters were in the story. Instead, he proposed that folktales should be studied by examining the functions of the characters with the following two central ideas: "1. Functions of characters serve as stable, constant elements in a tale, independent of how and by whom they are fulfilled. They constitute the fundamental

components of a tale. 2. The number of functions known to the fairy tale is limited" (Propp, 1928, Chapter 2, para. 9).

A number of linguists adapted Propp's, (1928) theory of story elements to create "story grammar" schemas (Black & Wilensky, 1979). Rumelhart (1975) was the first to modify Propp's ideas into a story grammar schema most closely aligned with those used by SLPs in current practice. Rumelhart postulated that a simple story consists of two elements. The first element is the setting, which includes place, time, and characters. The second element is the episode, which is defined as an event and reactions to the event. He then added a number of "syntactic categories" that functioned within the basic story grammar: action, application, attempt, change-of-state, consequence, desire, emotion, episode, event, internal response, overt response, plan, preaction, reaction, setting, state, story, and subgoal. Applebee (1978) applied a simplified version of Rumelhart's approach to assessing the narratives of children. Applebee asserted that a true narrative must contain an initiating event, an attempt and a consequence and at least two other story grammar elements. Rumelhart and Applebee's work has since been used and adapted by several researchers when studying the quality of narratives and also for teaching story-telling to children. An example of a story grammar teaching tool is the popular Story Grammar Marker® (Moreau, 2011, 2017) often employed by SLPs when teaching story grammar. Moreau's Story Grammar Marker® simplifies Rumelhart's (1975) syntactic categories into seven story grammar elements that are more accessible to children: character, setting, kick-off, internal response, plan, attempts or actions, direct consequence, and resolution.

The Test of Narrative Language (TNL; Gillam, & Pearson, 2004) is possibly the most commonly used standardized test to measure the narrative content of children's stories. The

TNL was normed on children between the ages of 5;0 – 15;11 using oral narratives generated by a restricted set of specific visual and verbal story prompts. The TNL demonstrated interscorer reliability above .90 and test-retest reliability above .80. Validity of the TNL was demonstrated by high correlation between TNL scores and language sample analysis. Petersen, Gillam and Gillam (2008) developed the Index of Narrative Complexity (INC) as a companion to the TNL. The INC is based upon the same schema as the TNL but allows for analysis of narratives other than the story prompts provided in the TNL. The INC may also be used to assess written narratives in addition to oral narratives. The rubric design of the INC assigns points to the narrative for including characters, setting, initiating events, internal responses, planning, action, complications, and other story grammar elements. The narrative can be assigned more points for multiple uses of a particular story grammar element or a more skilled use of a story grammar element. For instance, more points are assigned if one or more characters are named than if no characters are named. This allows a narrative to receive a score between 0 – 31 which allows for a greater differentiation in narrative quality.

Petersen, Gillam and Gillam (2008) conducted a narrative intervention study of 12 children with diagnosed language impairment (LI), between the ages 6;4-9;1 (mean = 8;2). The children were administered two pretests, one month apart using the INC to demonstrate reliability prior to administering treatment. Then the children received narrative intervention treatment over four 90-minute sessions per week for one month. This treatment was offered using a story grammar approach, highlighting character, setting, initiating event, character response, plan, actions, consequence, complication and resolution. Causality, temporal concepts and dialogue were also addressed during treatment. The children practiced narrative skills using wordless picture books and single pictures. After the treatment, the children were

again assessed using standard procedures of the TNL and the INC to demonstrate correlation between INC and TNL scores. The final administration of the INC was a story retell from three wordless picture books. Interrater reliability was above 90% for all stories scored using the INC. Reliability between the two pretests of the INC was significant and there were no significant differences in narrative skill scores between pretests 1 and 2. Scores on the INC were significantly higher after treatment than before treatment.

Colozzo and colleagues (2011) adapted the TNL scoring system to examine the content of stories produced by children with specific language impairment (SLI) versus children with typically developing language (TD). In the first stage of the study, 13 children diagnosed with persistent SLI (mean age 9;0) were paired with 13 age-matched TD children that attended the same school as the child with SLI. The TNL was administered to all participants which was then transcribed using Systematic Analysis of Language Transcripts (SALT; Miller & Iglesias, 2006) conventions and scored using TNL protocols. The TNL scores were then converted into a 4-point scoring system that was used to differentiate the content from the grammaticality measures of the TNL. The second stage of the study used 20 narratives generated by children with SLI and 20 narratives from age and gender-matched pairings (mean age 7;6). Again, the narratives were transcribed using SALT conventions and scored using the TNL scoring as well as the aforementioned 4-point system. Children with SLI scored significantly below their age-matched TD peers when evaluating the content of their narratives, although it was noted that children with SLI were more liable to fall into "extremes of the distribution, obtaining relatively higher scores on either content or form. In contrast, 18 of the 20 TD children earned similar portions of their points from the content and the form items" (p. 1617). Colozzo and colleagues transcribed and coded the narratives in accordance with Merritt and Liles' (1987)

story grammar schema to identify story grammar elements as well as coding grammatical errors. The stories were ranked separately for grammatical errors and for total number of story grammar elements and ranking differences were compared. Using this system, it was determined that the stories of children with SLI were shorter, contained fewer story grammar elements and were less grammatically accurate than their TD peers. The ranked scores using Merritt and Liles' schema correlated strongly with the TNL scores.

Blom and Boerma (2016) designed a study to explore the reasons why children with SLI perform more poorly on story grammar measures than their TD peers, and how production of story grammar can change over time in this population. The participants for this study were 84 children with SLI (21 female and 63 males; mean age 5.97 years old) and 45 TD children (18 female and 27 male; mean age = 5.90 years old) for a total of 129 participants. All participants were monolingual Dutch speakers. Narrative content was measured using the Multilingual Assessment Instrument for Narratives (MAIN; Garinga et al, 2012). The MAIN is similar to the TNL, wherein the child has stories read to them, about which the child answers comprehension questions. The children are then presented with pictures and are prompted to produce a narrative using the stimulus pictures. The narrative is scored via a rubric for the inclusion of story grammar elements with most categories in the rubric being scored for the presence or absence of the story element; only "setting" had a rating of 0 - 2. The children were also evaluated for nonverbal IQ, verbal short-term memory and attention. Although all children had an IQ within normal limits, a higher IQ score was positively correlated with higher scores on story grammar content. Higher story grammar scores were also positively correlated with age, MLU, narrative comprehension and attention. Narratives were gathered from the participants a year later and evaluated by the same means as the first narratives. The gap in

story grammar scores between the children with LI and the TD children narrowed as the children aged. Blom and Boerma hypothesized that this gap narrowed because the children with LI were able to improve their narratives with maturation, which was positively correlated with narrative production, while the children with TD language were already at the ceiling of the comprehension measures during the first test.

Paul and colleagues (1996) also documented the narrowing of the gap in content scores by second grade when examining the narrative skills of early school age children who were diagnosed with slow expressive language development as toddlers, as compared to TD peers. Narratives were elicited on three occasions from the participants using a wordless picture-book prompt. The first narratives were gathered while the participants were in kindergarten and again during their first and second grade school years. The content of the stories was analyzed using an adaptation of Applebee's (1978) story types (Klecan-Aker & Kelty, 1990; Klecan-Aker, McIngvale, & Swank, 1987). During kindergarten, there were significant differences in narrative quality documented between the groups but the differences narrowed during first grade. By the second grade, the majority of the participants demonstrated no significant gap in content scores. Paul and colleagues posited that the children with slow expressive language development were slower to develop narrative skill than their peers but were able to resolve those deficits over time.

Fey and colleagues (2002) conducted a longitudinal study comparing children with LI to TD children in both oral and written narrative skill. During kindergarten, oral narratives were elicited from the participants. Both oral and written narratives were gathered during second-grade and fourth-grade. The narratives were measured for productivity, grammatical accuracy and complexity. Story content was measured for the inclusion of story grammar elements,

language sophistication and plot complexity. The differences in the narrative content scores between the children with LI and TD was observed to narrow between kindergarten and second grade. However, by fourth grade, narrative quality differences in narrative quality score re-emerged. There was no significant difference in mean length of C-unit (the main clause and its subordinate clauses) in words or total number of C-units per story, so it would appear that the TD children were producing narratives of greater complexity without increasing length appreciably over their peers with LI. It is unclear whether the inclusion of the additional elements of language sophistication and plot complexity to story grammar when rating narrative quality was a factor in the observation of these narrative differences.

The gap in story grammar scores may close by second grade because story grammar schemas are less effective in differentiating the performance of older children. Koutsoftas and Gray (2011) argued that the story grammar measures typically employed by speech language pathologists are not sufficient to assist children with language-learning disabilities to meet curricular writing standards. Koutsoftas and Gray compared the narratives of 56 fourth and fifth-grade children using a story grammar analysis versus the scoring system employed by the local school system. Of the participants, 26 were language and learning disabled (LLD) and the remaining participants were TD with no history of receiving special education services. The children were provided with a narrative story prompt and then instructed to write a story with a beginning, middle and end. The children also completed an expository essay about a topic with a verbal prompt. The narratives were transcribed using SALT software conventions and scored for story grammar content using Merritt and Lile's (1987) story grammar schema that identifies the story grammar elements of: setting, initiating event, internal response, attempt, direct consequence, and reaction. The narratives were also evaluated using The Six-Traits Writing

Rubric (STWR; Education Northwest, 2006) employed by the local school district for comparison. The content portions of the STWR focus upon the domains of: main idea, details and support, reasoning/thinking, evidence selection and acknowledgement, awareness/engagement of reader, lead and conclusion, transitions, sequencing, packing, purpose/text structure, title, tone, commitment, and fit with audience/purpose. The TD group produced a mean of 3.81 story grammar elements per episode and the LLD group produced a mean of 3.54 story grammar elements per episode but the difference was not statistically significant. In contrast, the STWR scores demonstrated significant differences in narrative quality between the TD and LLD groups.

The findings of Koutsoftas and Gray (2011) may explain why the aforementioned research has documented an elimination of the gap in narrative content quality between children with LI and TD children as they mature while, in contrast, educators have been documenting what has been termed the "4th grade slump," (Chall & Jacobs, 1996, p.33). During the fourth grade, some children who had been performing adequately in school begin to fall behind (Blom & Boerma, 2016; Carretti, Motta, & Re, 2016; Fey et al 2002; McNamara, Ozuru, & Floyd, 2011; Paul et al, 1996). Children who may have be able to competently decode words may not have acquired the necessary skills to comprehend the content of what they read. Such metacognitive skills can include: understanding causality, perspective taking, and sequencing, which are the very skills necessary to the production of competent narratives (McNamara et al., 2011). It is possible that the children encounter the ceiling of what story grammar schemas are capable of measuring, while curricular demands for narrative production extend beyond the expectations of story grammar.

Botvin and Sutton-Smith (1977) developed an alternate adaptation of Propp's (1928) content analysis approach, which focused upon Propp's assertion that function was of primary importance when evaluating the content of narratives. Botvin and Sutton-Smith noted that significant actions in competent narratives typically occurred in pairs, such as an injury followed by a recovery or a deception that was later resolved by the revelation of the deception. The pairs could be embedded within other pairs to create greater complexity within the narrative. For instance, in the course of nullifying an act of villainy, the hero could be captured and then escape, which creates a subordinated dyad of action within the larger dyad. Botvin and Sutton-Smith labeled the actions of stories "plot units" and the dyads occurring in the stories as "primary plot units." Secondary actions could also occur that were not dyadic in nature and these were labeled "secondary plot units." By reducing the actions of stories to these simplified plot units, Botvin and Sutton-Smith were able to graph out the actions of stories to show their structural complexity in a visual format. To describe the complexity of the resultant graphed narratives and demonstrate a progression of narrative development, Botvin and Sutton-Smith devised a series of story levels as summarized below and depicted in Figure 1:

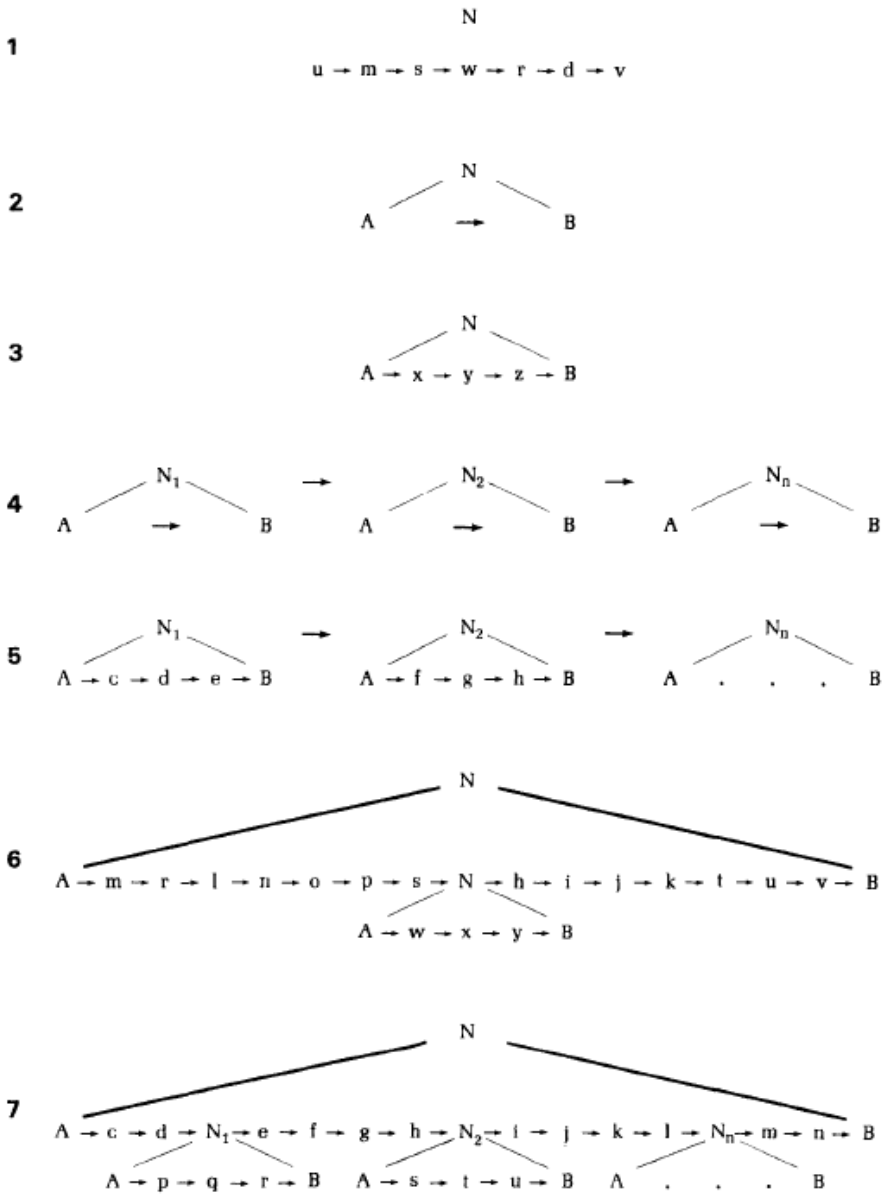
- Level 0: The narrative contains no explicit action and relationships between characters are not established.
- Level 1: The narrative lacks coherence or structural unity. Action is implied.
- Level 2: The narrative is short and symmetrical and contains only one dyad. Action before and after the dyad may occur, but none in the middle of the dyad.
- Level 3: The single dyad is expanded with secondary plot units occurring in the middle of the dyad.
- Level 4: Two or more dyads are concatenated with no secondary plot units.

- Level 5: Two or more dyads are concatenated and secondary plot units are used.
- Level 6: A subordinated dyad occurs within the primary dyad (main action interrupted by subplot).
- Level 7: The narrative contains multiple dyads with subordination.

Botvin and Sutton-Smith (1977) used a series of experiments to test their dyadic approach. The initial study analyzed narratives from 80 children between the ages of 3 – 12. The children were given verbal prompts to generate the story of their choice. Several stories were gathered from each participant but only one fantasy narrative was included from each participant. The stories were coded for primary and secondary plot units and then categorized by story level, using levels 1-7. There was a high correlation between age and narrative complexity level, as well as between narrative length and complexity level, which suggested that the approach was valid to measure narrative development. A second study of 140 children between the ages of 3 – 12 was conducted using the same procedure to gather and score narratives as the first study, with the addition of Level 0 to the schema. The second study also demonstrated positive correlations between age, narrative level and narrative length. The researchers noted that, while the order of narrative level acquisition appeared to be the same for all of the children, there was variability in rate of acquisition by age, indicating that the approach may be of value when differentiating the quality of narratives produced by children of similar age.

Gillam and Johnston (1992) adapted Botvin and Sutton-Smith's (1977) framework for the content portion of their narrative analyses. Gillam and Johnston retained Botvin and Sutton-Smith's primary plot units, renaming them *dyadic constituents* and gave credit for each action in the dyad, the initiating event and the resolution. This provided scoring of two dyadic constituents for each dyad. Single process constituents were story actions that did not

Figure 1
Schematization of levels of structural complexity



Narrative N is composed of primary, upper case, and secondary, lower case, plot units which may be formed into single narratives, coordinated episodes, or hierarchized into plot and subplot or subplots. Reprinted with permission from "The Development of Structural Complexity in Children's Fantasy Narratives," by G. Botvin and B. Sutton-Smith, 1977, *Developmental Psychology, 13*, p. 380. Copyright 1977 by the American Psychological Association.

require pairing with a resolution, similar to Botvin and Sutton-Smith's secondary plot units. Gillam and Johnson also included contextual constituents which are often considered "setting" in many narrative evaluation schemas. Contextual constituents included participants (characters), location, time and contextual events. Measuring the total number of constituents and the percentage of dyadic constituents allowed for greater statistical analysis to examine the differences in the narratives. When this schema was applied to the oral and written narratives of children between the ages of 9-12, oral narratives contained more constituents per story than written narratives, but written narratives appeared to have a significantly higher percentage of dyadic constituents per T-unit (the shortest grammatically allowable sentence) than oral narratives. The researchers compared children with language/learning impairment (LLI) to 3 groups: matched for age, matched for reading level, and matched for language age. No significant interaction effects between group and mode (oral vs. written) were found. Gillam and Johnston did not present simple group differences by constituents per story or dyadic constituents per T-unit, so the utility of this approach in measuring narrative complexity warrants further examination.

Oral Versus Written Narratives

Writing demands also increase in the fourth grade and, although written production is supported by oral language, significant differences have been documented between oral and written language (Carretti, et al., 2016; Erftmeyer & Dyson, 1986; Gillam & Johnston, 1992; Justice, et al.; 2006; Kroll & Vann, 1981; Moyle, et al., 2011; Pinto, Tarchi & Bigozzi, 2015; Thordardottir, 2008). While it may seem intuitive to equate spoken and written languages with one another, their distinctiveness becomes evident when the written word is spoken aloud or the spoken word is written down verbatim. Mechanical production considerations aside, the

lexical choices, grammar, rhythm and tone can differ widely. Prior research indicates that this differentiation between production of oral and written narratives begins to occur between 9 – 12 years of age, with written language typically becoming more cohesive, complex and organized than spoken language (Gillam & Johnston, 1992; Kroll & Vann, 1981).

We know little about the content differences between oral and written narratives during this period of differentiation because the majority of research performed has focused upon narrative microstructure differences (Cain & Patson, 2005; Dockrell & Connelly, 2015; Ebert & Scott, 2014; Klecan-Aker & Kelty, 1990; Losh & Gordon, 2014; Nippold, 1995; Nippold et al, 2014; Scott & Windsor, 2000). An exclusive focus upon microstructure may be of limited utility when evaluating narratives because both teachers and laypeople appear to use content to differentiate the quality of narratives as much as microstructure (Koutsoftas & Gray, 2011; Newman & MacGregor, 2006). The few studies that have compared the content differences of oral and written narratives have tended to take a broad view. For instance, Fey and colleagues (2004) assigned quality scores to the narratives but they only had a 0-3 point scale based on rating characters, setting, ending and language sophistication. Such a limited scale for differentiation between narratives allows only the broadest categorization of the narratives and may be of limited utility when differentiating the complex narratives of older children. Carretti, Motta, and Re (2016), only evaluated reader's general impression of the text by rating their general impressions of the narrative on a 1 – 5 point scale while the bulk of their analysis focused upon microstructural features.

The aforementioned study by Gillam and Johnston (1992) compared the content of oral and written narratives in children between the ages of 9 – 12 in detail and found that the communication mode did make a significant difference in content. Both the LI and control

groups produced oral narratives were longer than the written narratives while the written narratives had a higher percentage of dyadic constituents. No significant interaction effects between group and modality were found. Gillam and Johnston proffered two possible theories for the content differences. First, that the difference in content organization was a pragmatic choice made by the participants, based upon the different needs of the audience in each mode. Since oral language is ephemeral, it places greater demands upon the working memory of the listener to integrate prior information while listening to the current information, requiring more cohesive ties for the listener to relate the prior with the incoming information. In contrast, a reader may stop reading to reflect and can review previously read text at any time, allowing broader, more global relationships between ideas and information. A secondary theory for the differences in content by modality is that the participant's language production skill was the main driver for the difference in content organization. Since oral narration requires greater on-line processing, the participants were only capable of creating cohesion in close proximity, while they had more processing time to create cohesion across the text in the written condition.

Kroll (Kroll & Vann, 1981) described research with children in the third, fourth and sixth grades where in the children were taught how to play a game, then prompted to create an instruction set for the game. Both oral and written instructions were collected from each child. The results indicated that the third and fourth-grade children produced oral instructions that were superior to written. The sixth-grade children produced oral and written instructions of equal quality. However, these results may not apply to narratives due to the significant differences in language structure.

Purpose

The current literature does not adequately address the development of narrative content in older school-age children, particularly in understanding content differences between oral and written narratives. Current methods of assessing narrative content, while effective when assessing younger children, may not be effective in addressing the deficits of older, typically developing children. However, teachers and laypersons both differentiate the quality of narratives based as much upon the content as microstructure. Understanding the development of narrative content production in typically developing children is important for diagnosis of deficits and the development of treatment targets for older school-age children with language deficits that affect narrative quality. Additionally, understanding the differences in content between oral and written narratives will allow a more complete picture of typically developing narrative skills.

This research seeks to augment current content analysis procedures by exploring the efficacy of a dyadic approach to when examining narrative content analysis to determine whether additional information about the quality and complexity of narratives may be ascertained beyond traditional story grammar approaches. Additional methodologies for narrative content analysis are warranted since story grammar analysis appears reach an upper limit when scoring the more complex narratives produced by older children. The primary purpose of this research is to determine whether a closer examination of function and how functions are structured within the narrative may provide additional information about the complexity of the narratives produced by children in the fourth grade beyond a story grammar analysis.

A secondary goal of this research is to understand differences in narrative content complexity when comparing oral and written forms in a typically-developing fourth-grade population. Since curricular demands for written narrative production increase during fourth-grade, it is important to understand how oral and written narrative content is differentiated in typically developing children. Understanding how written narrative content may differ from oral narratives in a typically developing population may assist clinicians in developing appropriate treatment targets for children at risk for academic deficiencies in narrative production. Since story grammar approaches may be insufficient to describe the more complex narratives produced by typically-developing fourth-graders, additional methods must be explored to provide this differentiation.

Hypotheses and Predictions

It is hypothesized that the dyadic analysis methods: both that developed by Botvin and Sutton-Smith (1977) as well as that of Gillam and Johnston (1992), will yield different information about narrative quality and complexity than a story grammar approach. To compare these approaches, three analysis methods will be used: 1) story grammar will be assessed using the INC, 2) story levels will use Botvin and Sutton-Smith's story levels (SL), and 3) The percent of dyadic constituents (PDC) will be calculated using Gillam and Johnston's procedure. It is expected that all three of the content analysis methods will have a significant, positive correlation since all are analyzing the same narratives and there are parallels in how each measure defines and evaluates narrative content. The correlation between the story level scores and the percent of dyadic constituents is anticipated to be more highly correlated with one another than with INC scores since they are both based upon plot units. Prior research has indicated that story grammar methods may reach an upper limit when measuring the stories of

older children. Therefore, we would expect the INC scores to be more negatively skewed than the dyadic methods which are expected to demonstrate a more normal distribution.

Based upon limited prior research (Gillam & Johnston, 1992), written narratives are anticipated to demonstrate greater complexity than oral narratives. The written narratives are expected to be scored at higher story levels, have a higher percentage of dyadic constituents and receive higher INC scores than the oral narratives. However, this prediction should be anticipated with caution due to the limited research base in comparing the content or oral and written narratives. It is predicted that oral narratives will contain more total words, more T-units and more constituents per story. These differences are expected because prior research has shown that oral narratives tend to be longer than written narratives and should, therefore, contain more constituents per story.

Methods

Participants

The data for this study was obtained from a larger data set from research that is currently in progress. Fourth-grade teachers from a rural Idaho school sent explanations of the study with permission slips to parents and asked for signed consent if the parents were willing to have their children participate. A total of 35 children participated in the initial study. All children were invited to participate as long as their parents signed an informed consent form. Data was only analyzed for the 21 children who were native English speakers, did not have an individualized educational plan, and completed all segments of data collection. Out of the 21 participants, there were 11 males and 10 females. Their ages, as calculated on the first day of testing, were between 9;1 and 10;7 years of age.

Procedure

Signed informed consent forms were obtained from the parents/guardians for all participants. Prior to data collection, each child listened to an assent statement which was read aloud to them. The child was given the opportunity to choose to participate or decline to participate. All children chose to participate and signed an assent form to this effect. Data collection occurred over two sessions that occurred on different days, but no more than two weeks apart. The testing was conducted in an unoccupied classroom of the participants' school during the school day.

Two pictures from the Story Retell Fun Deck (Ducworth, 2006) were used to elicit the narratives analyzed in this study. The first, which will subsequently be described as *Show and Tell*, depicts a classroom with a teacher and three students. The chalkboard in the background says "Show & Tell." One of the students is attempting to catch a jumping frog. The second picture, which the researchers designated as *No Girls Club*, depicts five children in an outdoor setting. Two boys sit inside of a homemade tent while reading comic books and playing with a frog. The tent has the signs "No Girls Club" and "Boys Only!" posted on the outside. Three girls stand outside of the tent, looking on at the boys. See Appendix A for a copy of the stimulus pictures used.

The oral narratives were gathered in a one-on-one format with the participant and the researcher. The examiner modeled a narrative and then prompted the participant to produce an original narrative. This narrative was digitally recorded for later transcription. The written narratives were gathered in groups of 4 – 24 students at a time and their papers were scanned for later transcription. All narratives were produced within a 30-minute time limit. See Appendix B for the details of the modeled narratives and prompts used.

Half of the participants of the original study completed the oral narrative on the first day and the written on the second day. The other half completed the written narrative on the first day and the oral narrative on the second day. To provide additional counterbalancing, half of the oral narratives were elicited using the *No Girls Club* picture and the other half were elicited with the *Show and Tell* picture. Had all original data been included, this would have resulted in equal groups of four conditions. Due to the exclusionary criteria listed above, not all narratives collected were included in the final study, resulting in inequalities of group size: 12 oral *Show and Tell* narratives, 9 written *Show and Tell* narratives, 9 oral *No Girls Club* narratives, and 12 written *No Girls Club* narratives.

Following the narrative elicitation task, the sentence repetition subtest of the Clinical Evaluation of Language Fundamentals, fifth edition (Wiig, Semel, & Secord, 2013) was administered to each participant. The intent of the administration of the sentence repetition task was to screen any children with an unidentified language impairment (Wiig, Semel & Secord, 2013). All participants scored at or above one standard deviation below the mean and were thus determined to have language skills within the normal range.

Transcription and Coding

The narratives that were previously gathered using the above procedures were also transcribed using SALT software conventions (Miller & Chapman, 2012), as well as project specific conventions by trained undergraduate and graduate research assistants, prior to the initiation of this study. The lead researcher for this study analyzed the narratives and obtained a narrative complexity score using The Index of Narrative Complexity (INC; Petersen et al, 2008; see Appendix C for the measure). Each narrative was also analyzed for primary and secondary plot units. Since Botvin and Sutton-Smith (1977) did not provide their operational

definitions for the plot units, the operational definitions of primary and secondary constituents as created by Gillam and Johnston (1992) was used to define the dyadic and non-dyadic story actions. However, some of the constituent definitions were modified and additional constituents were added because the story actions provided by Gillam and Johnston did not adequately define all of the story actions present in the narratives for the current study. Dyadic and non-dyadic story actions as well as the contextual story elements of character, time and place were coded in the transcripts using SALT software.

Table 1

Non-Dyadic Story Actions

Announcement	A verbal statement.
Command	A directive verbal statement.
Inquiry	A question intended to seek information.
Response	A mental or verbal acknowledgement of a previous event, announcement, or inquiry.
Reaction	A physical activity that has been influenced by a previous event, announcement or inquiry.
Consequence	The outcome or effect of a previous process or set of processes.
Fortuitous occurrence	An accidental event that is advantageous to a participant.
Arrival	A participant arrives at a location.
Departure	A participant leaves a location.
Return	A participant comes back to a former location.
Narrator elaboration	An explanation of a participant's behavior or circumstance that lies outside the story.
Conclusion	Narrator's statement which brings the story to a logical end.

Contextual Story Elements

Character: named	A human or animal in the story that behaves, senses or exists in the story and has been assigned an individual name.
Character: unnamed	A human, animal, or group of human/animals that behaves, senses or exists in the story but is identified only by a common noun. Experiences identified only by pronouns are not included. For example, "the girl," "the dog," or "the herd" would be acceptable.
Place	Specific reference to a site.
Time	Temporal reference to the period of existence or duration of a contextual event, dyadic process or single process.

Dyadic Story Actions

Initiation of the dyad	**Positive dyad completion**	**Negative dyad completion**
Villainy: An evil or wicked act.	Villainy nullified: Either the wicked act was negated or the perpetrator was punished.	Villainy not nullified: The villainous act was not rectified or the perpetrator was not punished.
Lack: A problematic shortage, deficiency or lack of access.	Lack liquidated: The item(s) needed or access to a resource obtained or a substitute secured.	Lack not liquidated: The needed/desired items/access were not obtained.
Deception: A misleading act or statement that is intended to cause a participant to think or act wrongfully.	Deception revealed: The nature of the misleading act/statement is revealed.	Deception not revealed: A statement must be made to indicate that necessary parties to the misleading act/statement continue to think/act wrongly.
Threat: A potential source of danger, harm, or distress.	Threat nullified: The potential harm/distress was addressed to stop or prevent harm/distress from occurring.	Threat not nullified: Harm or distress was inflicted by the threatening character, event or situation.
Plan: A scheme (thought, intention) for doing something.	Plan carried out: The intended scheme/thought/intention was acted upon.	Plan not carried out: The character(s) abandoned the plan or were unable to carry out the plan.
Attack: To strike or shoot at a participant with intent to harm.	Counterattack/wound/kill/flee: The protagonist counterattacked or fled the attacker.	
Injury: Physical harm or damage to a participant.	Recovery: The physical harm or damage was healed or rectified in some manner	Not recovered/death: The injury is not healed or the participant died.
Pursue: To chase in order to capture or harm.	Capture: The pursued participant was captured/harmed.	Not captured: The pursued participant was not captured/harmed.
Escape: A participant has escaped confinement.	Capture: The escaped participant was captured.	Not captured: The escaped participant was not captured.
Search: To look for a participant/item that is needed or has been lost.	Found: The participant/item in question is located.	Not found: A statement occurs in the narrative that, despite searching, the participant/item is not located.
Problem: A difficulty encountered by the participant(s) that does not correspond to one of the above categories.	Resolution: The problem has been solved.	Problem not resolved: A statement occurs in the narrative that attempt to solve a problem was not successful.

Adapted with permission from "Spoken and Written Language Relationships in Language/Learning-Impaired and Normally Achieving School-Age Children," by R. Gillam and J. Johnston, 1992, *Journal of Speech and Hearing Research, 35,* p. 1315. Copyright 1992 by the American Speech-Language-Hearing Association.

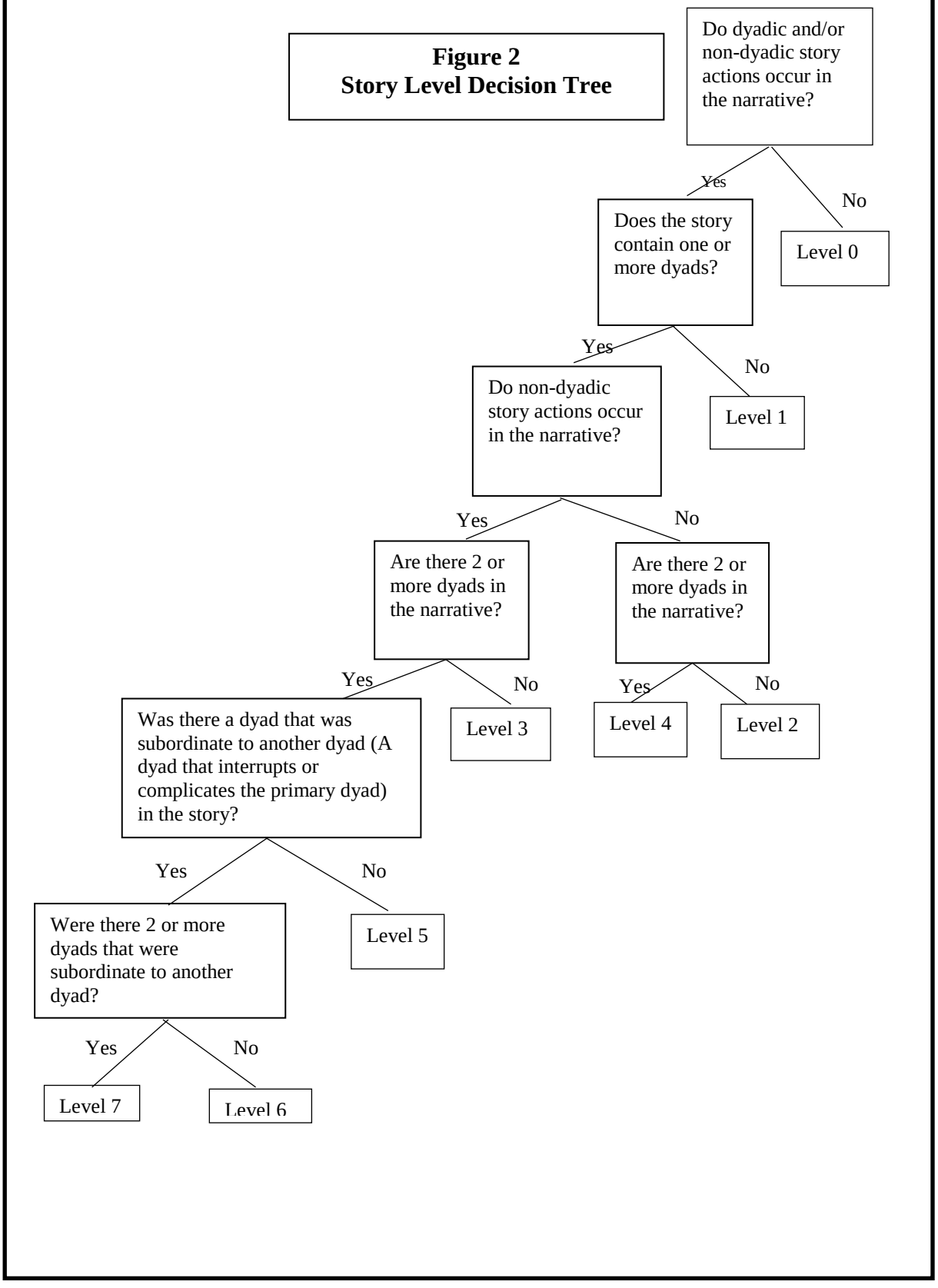
A decision tree, based upon the operationalized story level definitions that were provided by Botvin and Sutton Smith (1977), was used to determine the story level of each narrative, as depicted in Figure 2.

Data Analysis

To obtain a measure of complexity for each narrative, in accordance with Gillam and Johnston (1992), the number of dyadic constituents was divided by the total number of constituents to obtain the percent of dyadic constituents (PDC). A Nonparametric Spearman's **ρ,** compared the complexity measures of the oral narratives for the INC scores (INCs), percent of dyadic constituents (PDC) and story levels (SL) to determine whether there were correlations between the three measures of narrative complexity and how correlated the measures were with one another. This same Nonparametric Spearman's **ρ** analysis was conducted separately for the written narratives. A distribution analysis was performed individually for each of the measures of complexity to determine whether the measures had a normal distribution; a Shapiro Wilkes Test was conducted on each complexity measure to test for goodness of fit.

Oral and written narratives were compared for differences in productivity and complexity. The productivity of each narrative was measured by total number of words, total number of T-units, and total number of constituents. Paired-samples t-test compared oral and written narratives in each of the productivity measures. To measure differences in complexity, a Wilcoxon Signed Rank test was used to compare oral and written (SL) differences while t-tests were used to compare oral and written differences for PDC and INCs.

Figure 2
Story Level Decision Tree



Reliability

To test the reliability for the measures of complexity, another SLP graduate student was trained in the procedure to code dyadic and non-dyadic story elements as well as contextual narrative information. She was then trained in how to assign a story level to the narrative and the decision tree depicted in Figure 2 was provided to determine the story level of each narrative. The student was provided with the procedural information from both the Gillam and Johnston (1992) study as well as the Botvin and Sutton-Smith (1977) article for additional reference to clarify questions when coding. The student was provided with the INC scoring rubric and the information on scoring the INC that was provided by the authors of the measure (Petersen, Gillam & Gillam, 2008) but no additional training was provided on the INC. Five participants were randomly selected for reliability testing. The 10 narratives produced by the 5 participants were independently coded and scored by the lead researcher and the trained graduate student.

There were separate procedures to compare the reliability of each of the complexity measures. To score the reliability of the INC, the scores on each of the thirteen measures of complexity were compared for agreements between the raters. For both the oral and written narratives, point by point agreement between the raters was 69.2%. To test for PDC reliability, point by point agreement on dyadic and non-dyadic story elements was obtained. Reliability for dyadic story elements was 64.3% in the oral narratives and 68.8% in the written narratives. Reliability for the oral total constituents was 76.7% in the oral narratives and 76.9% in the written narratives. To rate story level agreement, the decisions on the aforementioned story level decision tree were compared for each narrative for the number of same and different

decisions each rater made on the decision tree. Agreement on the story level decision tree was 63.1% for oral narratives and 79.3% for written narratives.

Results

To determine whether the story prompt—*No Girls Club* or *Show and Tell*—affected the productivity or complexity of the narratives, a least squares fit was performed for the measures of total words, total T-units, total constituents, INCs, SL and PDC. No significant differences were found between the stories on any of the measures. Distribution analyses were performed for the measures of complexity—INC, PDC and SL—in both oral and written stories. The oral INC distribution appeared to be negatively skewed but a Shapiro-Wilkes test yielded a value of $p = .11$ which would indicate that there was no statistically significant difference from a normal distribution. All other complexity measures in both oral and written narratives yielded a normal distribution.

A Nonparametric Spearman's **ρ** comparing INCs, PDC and SL for the oral narratives did not demonstrate any significant correlations, while written narratives only demonstrated a significant correlation between SL and PDC.

The means and standard deviations of the measures of complexity and productivity were obtained, as reported in Table 2.

Table 2

Means and Standard Deviations

Complexity Measures: Oral Narratives

	N	Mean	Standard Deviation	Min	Max
Oral INCs	21	20	4.7	9	26
Oral Dyadic Constituents	21	7.4	3.7	2	16
Oral PDC	21	0.2	0.1	0.1	0.4
Oral SL	21	6	1.5	2	7

Complexity Measures: Written Narratives

	N	Mean	Standard Deviation	Min	Max
Written INCs	21	15.7	4.3	7	24
Written Dyadic Constituents	21	3.5	2.3	0	10
Written PDC	21	0.1	0.1	0	0.3
Written SL	21	4.3	2	1	7

Productivity Measures: Oral Narratives

	N	Mean	Standard Deviation	Min	Max
Oral Total Words	21	234.3	114.4	85	572
Oral Total T-Units	21	26.1	12.6	9	61
Oral Total Constituents	21	35.9	14.3	16	74

Productivity Measures: Written Narratives

	N	Mean	Standard Deviation	Min	Max
Written Total Words	21	134.3	67.5	56	316
Written Total T-units	21	18.4	9	7	38
Written Total Constituents	21	26.9	11.7	11	53

Table 3

Correlations Between Measures of Complexity

Nonparametric: Spearman's ρ: Oral Narratives

Variable	by Variable	Spearman ρ	Prob>|ρ|
Oral Percent Dyadic	Oral Inc Score	0.2334	0.3085
Oral Story Level	Oral Inc Score	0.3548	0.1146
Oral Story Level	Oral Percent Dyadic	0.3798	0.0895

Nonparametric: Spearman's ρ: Written Narratives

Variable	by Variable	Spearman ρ	Prob>|ρ|
Written Percent Dyadic	Written Inc Score	0.1166	0.6149
Written Story Level	Written Inc Score	0.3589	0.1101
Written Story Level	Written Percent Dyadic	0.6806	0.0007*

The productivity measures of total words, total T-units and total constituents per story were compared in the oral versus written narratives using paired-samples t-tests. Oral narratives were found to be significantly more productive in all three measures. The mean total words used in oral narratives was 234.3, SD = 114.4, while the mean total words of the written narratives was 134.3, SD = 67.5, t (20) = 4.0, p = 0.0007. The mean total T-units of oral narratives was 26.1, SD = 12.6 while written narratives had a mean total T-units of 18.4, SD = 9.0, t (20) = 3.04, p = 0.0064. The mean total constituents per story was 35.9, SD = 14.3 in the oral stories and 26.9, SD = 11.7 in the written stories, t (20) = 2.65, p = 0.0154.

Complexity measures in the oral and written narratives yielded results that indicate greater complexity in the oral narratives. A paired-samples t-test was conducted to compare the INCs of the oral and written narratives; oral narratives obtained significantly higher scores (M = 20.0, SD = 4.7) than written narratives (M = 15.7, SD 4.3), t (20) = 3.69, p 0.0014. A Wilcoxon Signed Rank test indicated that oral narratives (M = 6.0, SD = 1.5) were rated at higher story level than written narratives (M = 4.3, SD = 2.0), S = 73.0, p = 0.007. A paired-samples t-test comparing the PDC of the narratives indicated a significantly higher score for

oral narratives (M = 0.206, SD = 0.1) than written narratives (M = 0.133, SD = 0.1), t (20) = 3.86, p = 0.0010.

Complexity measures were also compared to measures of productivity to determine if there were correlations between productivity and complexity. The oral narratives demonstrated positive correlations of the INCs with all of the productivity measures: TW, t (20) = 0.663, p = 0.0020; TTU, t (20) = 0.603, p = 0.0038; and TCS t (20) = 0.607, p = 0.0035. There were no significant correlations when comparing either PDC or SL with productivity measures. The written narratives also demonstrated positive correlations between INCs and productivity measures: TW, t (20) = 0.554, p = 0.0091; TTU, t (20) = 0.645, p = 0.0016; and TCS t (20) = 0.695, p = 0.0005. Written SL demonstrated positive correlations with all productivity measures: TW, t (20) = 0.438, p = 0.0470; TTU, t (20) = 0.5057, p = 0.0194; and TCS t (20) = 0.4496, p = 0.0409. No significant correlations were found between the written PDC and productivity measures.

Table 4

Complexity and Productivity Relationships

Complexity and Productivity Relationships: Oral

	Total Words	**Total T-Units**	**Total Constituents**
INC Score	t (20) = 0.633, p = 0.0020	t (20) = 0.603, p = 0.0038	t (20) = 0.607, p = 0.0035
Percent Dyadic	t (20) = 0.633, p = 0.0020	t (20) = 0.022, 0.9262	t (20) = 0.176, p = 0.4458
Story Level	t (20) = 0.328, p = 0.1463	t (20) = 0.384, 0.0861	t (20) = 0.317, 9 = 0.1621

Complexity and Productivity Relationships: Written

	Total Words	**Total T-Units**	**Total Constituents**
INC Score	t (20) = 0.554, p = 0.0091	t (20) = 0.645, p = 0.0016	t (20) = 0.695, p = 0.0005
Percent Dyadic	t (20) = 0.001, p = 0.9960	t (20) = 0.069, p = 0.7674	t (20) = 0.0504, p = 0.8254
Story Level	t (20) = 0.438, p = 0.0470	t (20) = 0.506, p = 0.0194	t (20) = 0.450, p = 0.0409

Discussion

Measures of Complexity

The purpose of this research was to explore new avenues of narrative content measurement beyond traditional story grammar approaches, with the hope that these methods may yield novel information to augment our current understanding about narrative development in fourth-grade children. Prior to the initiation of this study, it was expected that the three methods of narrative complexity would be positively correlated because they were measuring the same narratives and the three methods appeared to contain parallels in how they evaluated narratives. While not absolutely equivalent, dyads that have been decomposed into "action or potential action" and "resolution" would appear to have a surface similarity to "initiating event" and "consequence" as measured within the INC, which would suggest a potential correlation of the methods. It was also expected that the INC would demonstrate a negatively skewed distribution, as suggested by research conducted by Blom and Boerma (2016), indicating that story grammar measures may reach a ceiling of descriptive capability in older children. This was contrasted by a prediction that PDC would not reach a ceiling of descriptive capability and would maintain a normal distribution as was seen in the research of Gillam and Johnston (1992).

This study also sought to increase current knowledge about content differences in the content complexity and productivity of oral and written narratives produced by TD fourth-grade children. Content complexity differences were not established in the literature for this population. It was proposed with trepidation due to the lack of literature in the subject, that written narratives would be more complex than oral narratives. The greater productivity of oral narratives over written in narratives produced by fourth-grade children has been previously

demonstrated, leading to a prediction that oral narratives would be more productive in this study as well.

The lack of correlation between the three complexity measures, despite that they were used to measure the same narratives, was unexpected. This may indicate that the three methods measure fundamentally different information about narrative content complexity from one another. As mentioned previously, the INC contains measures for action that indicated an initiating event and conclusion, similar to the action or potential action and resolution of the dyads used in SL evaluation. Since the INC contains 13 total categories to measure content, it is possible that the two categories of the INC that contained information similar to dyadic schema were insufficient to establish a correlation between the measures. Comparison between the individual items of the INC to SL may demonstrate a correlation between the INC items most similar to the information used to determine SL. This comparison may be conducted in future research. It is also possible that INC and SL provide divergent information about narratives. SL is a measure that provides detail about content structure; INC addresses whether structures are present but may not do so with the same depth as SL. For instance, the categories of "plan," "action/attempt," and "complication," may allude to subordinate dyads within a larger dyad, but they do not describe such structures as concretely and with as much detail as SL.

Although both PDC and SL are based in a dyadic approach, they only demonstrated the predicted positive correlation in the written narratives and not the oral narratives. It is possible that the wider distribution of SL (SD = 2.0) than in the oral narratives (SD = 1.5) allowed for greater opportunity to demonstrate a correlation. It is also possible that the correlation in the written condition could be due to the participant's ability to review what they have written, reminding them of the need to complete the dyad; this review process during production is

more difficult to perform during an oral narrative due to the real-time processing demands and ephemeral nature of oral narratives. However, this possibility is unlikely since oral narratives contained significantly more complete dyads than written narratives.

PDC did not correlate with INC in either the written or oral condition. In the course of scoring the narratives, a potential drawback to the PDC became apparent, which called into question the utility of PDC as a valid measure of narrative complexity. Figure 3 demonstrates the difference in PDC using two level 2 narratives: one with zero non-dyadic story elements and the other containing four non-dyadic story elements. The first narrative received a PDC of 100% and the second scored 33%, even though the second narrative contained more story information and actions and the same dyadic structure. It was observed on several occasions in the study that when two narratives of the same story level with the same number of dyads occurred, one narrative would receive a higher INC score while simultaneously receiving a lower PDC than the other.

Figure 3:

Differences in PDC of Narratives of the Same SL

To further examine this phenomenon, the oral narratives were arranged into matched pairs with identical story level and number of dyadic constituents, which yielded 16 paired narratives. Some narratives were included in more than one pair because there were multiple narratives of the same story level and number of dyadic constituents. Narratives that did not

have a matched pair for story level and number of dyadic constituents were excluded. The PDC and the INC scores for each narrative in the pair were compared to one another. Of the 16 pairs, 12 demonstrated an inverse relationship between INC scores and PDC. Of the 4 pairs that did not demonstrate an inverse relationship between INC scores and PDC, 3 narratives had INC scores that were within 2 points of difference, which may indicate negligible differences in the complexity of the narratives. The same procedure was conducted for the written narratives. One pair was excluded because neither narrative contained dyadic constituents and two pairs were excluded because they had the same INC score and PDC scores that only demonstrated negligible difference. Of the remaining 20 pairs, 18 demonstrated a negative relationship between INC score and PDC and the remaining 5 a positive relationship. More variability was observed in the INC scores of written narratives that had a positive relationship between INC score and PDC; between 2 – 5 points of difference were observed in INC scores between these pairs, which contrasts with the results of the oral narratives. This inverse relationship between INC score and PDC may indicate that PDC is not a valid measure of narrative complexity because PDC appears to penalize the addition of story information and actions that enrich the story. However, these concerns are anecdotal and there is insufficient data to determine whether PDC is or is not a valid measure of complexity with reasonable certainty. It should also be considered that oral narratives had a higher mean PDC than written narratives, which is consistent with the other two complexity measures which may indicate that PDC is a valid measure of complexity.

Both SL and PDC demonstrated normal distributions as predicted but INC also demonstrated a normal distribution despite the anticipation that it would reach a measurement ceiling when rating the narratives of older children, as other story grammar measures appear to

do. This may be because the INC included 13 different measures with each category able to yield a score between 0 – 3. This allows the INC to have a potential composite score between 0 – 31. This wide range of possible scores is in contrast to other story grammar measures, such as that used by Merritt and Liles (1987) which used an episodic analysis combined with six story grammar categories: setting, initiating event, internal responses, attempts, consequences, and reactions. While this schema counted the total number of occurrences of a story element in the narrative which would provide a greater score range than a simple register of presence or absence of the occurrence in the narrative, the reduced number of categories may not be adequate to measure narrative complexity in older children. The INC also included measures for temporal markers and causal adverbial clauses which may be considered word-level rather than narrative content measures. However, the use of causal adverbial clauses would indicate the introduction of causality to the narrative which is a measure of content. Temporal markers can also be used to indicate the setting by placing the narrative in a point in time or by indicating the passage of time in the narrative, which are also content-level functions. Whether these measures of complexity are measuring the content or the form of the narratives, they are still measures not present in other measures of narrative content complexity.

In summary, both the INC and SL may be valid and valuable methods for evaluating the content complexity of narratives in fourth-grade children. Both have a normal distribution but each appears to offer distinctly different information about the complexity of narratives which could indicate that they could be used in a complimentary fashion to evaluate narratives. PDC may not be a valid measure of complexity since it does not appear to quantify the dyadic structure of stories as intended.

Complexity Differences: Oral and Written Narratives

Contrary to the prediction, oral narratives scored significantly higher than written narratives in all measures of complexity. The challenges of the acquiring of writing skills could be a determining factor in the difference. During the fourth grade, students may still be learning spelling, sentence structure, writing conventions, and even may continue to work on acquisition of legible handwriting. The narratives were not analyzed for the effects of age due to the narrow age range of the participants. Were the study to include a larger population with a wider range of ages, the emergence of more complex written narratives with development may become apparent. It is also possible that the educational sequence taught in the schools could be different from what may be taught in other schools, or other schools may place greater emphasis upon learning narrative writing skills at this age. The school that the participants were drawn from had a high percentage of students of low socioeconomic status which has been persistently linked to lower academic performance. It may be of benefit to administer the measures to a more heterogeneous population or in schools in other geographic locations, to determine whether oral narratives are consistently more complex than written narratives across various populations and educational conditions.

Fourth-grade may be an early period of narrative writing development with the expectation that written narratives would be less complex than oral narratives. It is possible that written narratives would develop parity or even greater complexity than oral narratives with maturation.

Productivity Differences: Oral and Written Narratives

As predicted, oral narratives were significantly more productive than written narratives in all measures of productivity. The strongest effect was found in the mean total words of each

narrative. It is possible that a stronger effect was found for total words than for total number of T-units because written T-units may have been longer. The results of Gillam and Johnston (1992) may also illuminate the productivity differences between oral and written narratives; their study examined a wide range of microstructural narrative features in addition to content measures; it was noted that oral narratives had more unconnected constituents which may have increased the productivity of oral narratives. This may explain the difference in total constituents between oral and written narratives. Since those unconnected constituents would be composed of additional words and T-units, this explanation may also extend to the differences in total words and T-units. This study did not examine cohesion in conjunction with productivity differences, which may be a fruitful direction for further study. As also previously mentioned, fourth-grade students may be acquiring narrative writing skills which could have hampered productivity which may be ameliorated or reversed with maturation.

Oral narrative productivity appears to be consistently higher in oral narratives than written narratives. Although this may be related to the acquisition of writing skills, the use of superfluous, unconnected statements in oral narrative may contribute to oral productivity differences.

Complexity and Productivity Comparisions

Relationships between productivity and complexity were not predicted in this study, but these relationships were analyzed to contribute data to the body of literature upon the subject, since there is little research in this area. PDC was not correlated with productivity in oral or written measures but the nature of how PDC is calculated—dyadic constituents divided by total constituents—would nullify the effect of greater productivity. However, INC scores were significantly positively correlated with all productivity measures in both oral and written

narratives. Since INC scores measure whether a variety of elements are present within the narrative, it would logically follow that the more narrative material is present, the higher the chance that the narrative would have of meeting the criteria of the INC rubric. Story level was significantly correlated with all productivity measures in the written narratives but there was no significant correlation in the oral narratives. It is possible that the greater standard deviation of written narratives (SD = 2.0) than oral narratives (SD = 1.5) allowed greater opportunity to demonstrate a correlation.

Reliability

The reliability of all three complexity measures was lower than is generally considered acceptable. Since PDC is calculated by dividing dyadic constituents by the total number of constituents, reliability for both was obtained separately. Reliability for dyadic constituents was 64.3% in the oral narratives and 68.8% in the written narratives. Reliability for the oral total constituents was 76.7% in the oral narratives and 76.9% in the written narratives. Reliability for SL was 63.1% in the oral narratives and 79.3% in the written narratives. Agreement on INC scoring was 69.2% for both oral and written narratives.

A possible explanation for the lack of agreement was due to the method used to train the raters. Some training was provided and narratives that were not from this data set were used for practice analysis. However, the majority of direction in how to score the narratives was provided directly from the articles that detailed the various measures because the intent was to examine the measures for potential clinical application, which does not allow for extensive hours of training and practice to obtain maximal interrater reliability in a clinical setting. This, as a result, led to differing interpretations of the instructions between the raters when scoring. When the two raters compared their identification of dyads, they often both identified the same

potential dyads. However, one rater decided that inferred dyads should not be included, while the other decided that inferencing was a valid literary method and included inferred dyads if the intent was clear. Neither of the studies that used a dyadic approach to assess narratives addressed inferencing in their schema, so it was not included in the methodology of this study (Botvin & Sutton-Smith, 1977; Gillam & Johnston, 1992). Reliability may improve if the inclusion or exclusion of inferencing were addressed in the operational definitions of the dyads and the scoring training.

Higher SL agreement was obtained on the written narratives than the oral. Both raters perceived a difference in the ease of identifying dyads in the oral versus the written narratives: oral narratives were perceived to use more inferencing to initiate or complete dyads and written narratives were perceived to have more concrete, readily identifiable dyads. It is possible that participants engaged in more planning before completing their narratives, or that participants relied upon pragmatic aspects of language when face to face with the examiner which they could not do when writing. This could explain the difference in reliability for written story levels versus oral story levels. While this perception in oral versus written narratives is a qualitative evaluation, it could bear further study because it may indicate a difference in the use of oral and written language when producing narratives. It is recommended that future research include ratings regarding whether the dyad was concrete or required inferencing to identify. The number of total constituents per story may have been affected by this lack of reliability in identifying dyads. However, Gillam and Johnston (1992) noted significant reliability differences in the identification of secondary constituents, despite providing the second rater with six hours of training. This lack of reliability may call into question the efficacy of using PDC to rate the complexity of stories, since intensive training may be insufficient to obtain

reliable scores using this method. Botvin and Sutton-Smith (1977) also noted low reliability when identifying non-dyadic plot units but did not find this to be of concern for their purposes because the reliability of identifying the overall structure of the stories obtained higher agreement scoring.

In contrast, Petersen and colleagues (2008) obtained considerably higher reliability when using the INC to rate narratives but that reliability was not replicated in this study. A possible explanation for this difference could be in the training procedure of the raters. Petersen and colleagues trained the raters using narratives generated by a narrow set of prompts. Although the narratives rated for reliability were not the same narratives used in training, they were generated by the same set of prompts. For this study, the second rater was trained using a variety of narratives but none of the narratives used in training employed the same narrative prompts as the narratives in the study. This may indicate that the level of reliability reported for the INC by Petersen and colleagues (2008) may not be generalizable to clinical settings that lack the same level of training or use a broader set of prompts than was used in the study.

Clinical Implications and Future Research

Both INC and SL information provided descriptive capability that appeared to be adequate to assess the narratives produced by fourth grade students which may be useful in a clinical setting after the use of the measures receive further operationalization to clarify the scoring. To be of maximal benefit to clinicians, the use of the measures should be explored with students of a wider range of ages and abilities—and disabilities—to determine their efficacy for assessing a range of school-age populations. It is anticipated that both oral and written SL and INCs will attain parity with maturation in a typically developing population.

Prior research indicates that this parity may take longer to achieve in individuals with LI and may not be achieved in some individuals.

Each measure appeared to measure different aspects of narrative complexity which may indicate that combining the measures will offer greater descriptive capability when assessing narratives. To determine whether these measures will offer sufficient measurement capability to assess narratives effectively in an academic setting, it would be advantageous to compare INCs and SL to academic scoring procedures such as the STWR. If the INCs and SL provide comparable or complementary results to academic scoring procedures, then clinicians may rely upon them to assist in assessing student narrative deficits and to determine if students' narrative skills have improved sufficiently.

References

Alves, K., Kennedy, M., Brown, T., & Solis, M. (2015). Story grammar instruction with third and fifth grade students with learning disabilities and other struggling readers. *Learning Disabilities: A Contemporary Journal, 13,* 73-93.

American Speech-Language-Hearing Association. (2001). Roles and responsibilities of speech-language pathologists with respect to reading and writing in children and adolescents. Retrieved from: ww.asha.org/policy

Applebee, A. N. (1978). *The child's concept of story*. Chicago: The University of Chicago Press.

Black, J., & Wilensky, R. (1979). An evaluation of story grammars. *Cognitive Science, 3*, 213-229.

Blom, E., & Boerma, T. (2016). Why do children with language impairment have difficulties with narrative macrostructure? *Research in Developmental Disabilities, 55,* 301-311. Doi: 10.1016/j.ridd.2016.05.001

Botvin, G. & Sutton-Smith, B. (1977). The development of structural complexity in children's' fantasy narratives. *Developmental Psychology, 13,* 377-388.

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry, 18,* 1-21.

Cain, K., Patson, N., & Andrews, L. (2005). Age and ability-related differences in young readers' use of conjunctions. *Journal of Child Language, 32,* 877-892. doi: 10.1017/S0305000905007014

Caretti, B., Motta, E., & Re, A. M. (2016). Oral and written expression in children with reading and comprehension difficulties. *Journal of Learning Disabilities, 49,* 65-76. doi: 10.1177/0022219414528539

Carnes, R. (2014). A perceptual model of the Whorfian Thesis. *ETC: A Review of General Semantics, 71,* 263-271.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairment. *Journal of Speech, Language, and Hearing Research, 45,* 1142–1157.

Catts, H., & Kamhi, A. (2005). *Language and Reading Disabilities*. Boston, MA: Pearson Education, Inc.

Chall, J., Jacobs, V., & Baldwin, L. (1996). The reading crisis: Why poor children fall behind. In Shimron, J. (eds.) *Literacy and Education: Essays in Memory of Dina Feitelson*. Cresskill, NJ: Hampton Press, Inc.

Colozzo, P., Gillam, R. B., Wood, M., Schnell, R. D., & Johnston, J. R. (2011). Content and form in the narratives of children with specific language impairment. *Journal of Speech, Language and Hearing Research, 54,* 1609-1627. doi: 10.1044/1092-4388(2011/10-0247)

Dautenhahn, K., (2001). The narrative intelligence hypothesis: In search of the transactional format of narratives in humans and other social animals. *Cognitive Technology: Instruments of Mind, Proceedings,* 248-266.

Dockrell, J. E., & Connelly, V. (2015). The role of oral language in underpinning the text generation difficulties in children with specific language impairment. *Journal of Research in Reading, 38,* 18-34. doi: 10.1111/j.1467-9817.2012.01550.x

Ducworth, D. (2006). *Story Retell Fun Deck*. Super Duper© Publications.

Ebert, K. D., & Scott, C. M. (2014). Relationships between narrative language samples and norm-referenced test scores in language assessments of school-age children. *Language, Speech and Hearing Services in Schools, 45,* 337-350. doi: 10.1044/2014_LSHSS-14-0034

Education Northwest. (2006). 6 + 1 trait writing. Retrieved from http://www.educationnorthwest.org

Erftmier, T., & Dyson, A. (1986). 'Oh, ppbbt!': Differences between the oral and written persuasive strategies of school-aged children. *Discourse Processes, 9,* 91-114.

Fey, M. E., Catts, H. W., Proctor-Williams, K., Tomblin, B. J., & Zhang, X. (2004). Oral and written story composition skills of children with language impairment. Journal *of Speech, Language, and Hearing Research, 47,* 1301-1318. doi: 1092-4388/04/4706-1301

Garinga, N., Klop, D., Kunnari, S., Tantele, K., Valimaa, T., Balciuniene, L., & Walters., J. (2012) MAIN—Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics, 56,* 1-155.

Gillam, R., & Johnston, J. (1992). Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. *Journal of Speech and Hearing Research, 35,* 1303-1315.

Gillam, R. B., & Pearson, N. (2004). *Test of Narrative Language*. Austin, TX: PRO-ED.

Gillam, S. L., Olszewisi, A., Fargo, J., & Gillam, R. B. (2014). Classroom-based narrative and vocabulary instruction: Results of an early-stage, nonrandomized comparison study. *Language, Speech, and Hearing Services in Schools, 45,* 204-219. doi:10.1044/2014_LSHSS-13-0008

Green, L., Stockholm, M., Cearley, J., & Sheffield-Anderson, L. (2015). Direct vocabulary instruction with two 5th-grade English-language learners with language-learning disabilities: A treatment study. *Contemporary Issues in Communication Science and Disorders, 42,* 191-201.

Heinriquez, G. (2015). Your version of reality, and mine: We construct our own sense of the way things are—including our relationships. *Psychology Today*. Retrieved from https://www.psychologytoday.com/blog/theory-knowledge/201504/your-version-reality-and-mine

Hughes, D., McGillivray, L., & Schmidek, M. (1997). *Guide to Narrative Language*. Austin, TX: PRO-ED.

Justice, L. (2006). Evidence-based practice, response to intervention, and the prevention of reading difficulties. *Language, Speech, and Hearing Services in Schools, 37,* 284-297. doi:10.1044/0161-1461(2006/033)

Justice, L. M., Bowles, R. P., Kaderavek, J. N., Ukrainetz, T. A., Eisenberg., S. L., & Gillam, R. B. (2006). The Index of Narrative Microstructure: A clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech-Language Pathology, 15,* 177-191. doi:1058-0360/06/1502-0177

Kali, R. (2004). *Children and their development*. Upper Saddle River, NJ: Pearson:Prentice Hall.

Karanski, C. (2013). Behavior problems and social functioning in adolescents with language impairment. *Perspectives on Language Learning and Education, 20,* 36-43.

Klecan-Aker, J., & Kelty, K. (1990). An investigation of oral narratives of normal and language-learning disabled children. *Journal of Childhood Communication Disorders, 13,* 207-216.

Klecan-Aker, J., McIngvale, G., & Swank, P. (1987). Stimulus considerations in narrative analysis of normal third grade children. *Language and Speech, 30,* 13-24.

Koutsoftas, A., & Gray, S. (2011). Comparison of narrative and expository writing in students with and without language-learning disabilities. *Language, Speech and Hearing Services in Schools, 43,* 395-409.

Kroll, B., & Vann, R. (1981). *Exploring speaking-writing relationships: Connections and contrasts*. Urbana, IL: National Council of Teachers of English.

Liles, B. Z. (1985). Cohesion in the narratives of normal and language-disordered children. *Journal of Speech and Hearing Research, 28,* 123-133.

Liles, B. (1987). Episode organization and cohesive conjunctions in narratives of children with and without language disorders. *Journal of Speech and Hearing Research, 30,* 185-196.

McNamara, D. S., Ozuru, Y., & Floyd, R. G. (2011). Comprehension challenges in the fourth grade: The roles of text cohesion, text genre, and readers' prior knowledge. *International Electronic Journal of Elementary Education, 4,* 229-257.

Merritt, D. & Liles, B. (1987). Story grammar ability in children with and without language disorder: Story generation, story retelling, and story comprehension. *Journal of Speech and Hearing Research, 30,* 539-552.

Merritt, D., & Liles, B. (1989). Narrative analysis: Clinical applications of story generation and story retelling. *Journal of Speech and Hearing Disorders, 54,* 438-447.

Moreau, M. (2011). Making connections: The Critical Thinking Triangle® of the Story Grammar Marker®. Retrieved from www.asha.org/Events/convention/handouts/2011/Rooney-Moraeu/

Moreau, M. (2017). About MindWing Concepts. Retrieved from https://mindwingconcepts.com/pages/about-us

Moyle, M. J., Karasinski, C., Weismer, S. E., & Gorman, B. K. (2011). Grammatical morphology in school-age children with and without language impairment: A discriminant function analysis. *Language, Speech & Hearing Services In Schools, 42*(4), 550-560. doi:10.1044/0161-1461(2011/10-0029)

Newman, R. M., & McGregor, K., K. (2006). Teachers and layperson discern quality differences between narratives produced by children with or without SLI. *Journal of Speech, Language, and Hearing Research, 49*, 1022-1036. doi: 1092-4388/06/4905-102.

Nippold, M. A. (1995). Language norms in school-age children and adolescents: An introduction. *Language, Speech, and Hearing Services in Schools*, *26*(4), 307-08. doi:10.1044/0161-1461.2604.307

Nippold, M. A., Frantz-Kaspar, M. W., Cramond, P. M., Kirk, C., Hayward-Mayhew, C., & MacKinnon, M. (2014). Conversational and narrative speaking in adolescents: Examining the use of complex syntax. *Journal of Speech, Language, and Hearing Research, 57*(3), 876-886. doi:10.1044/1092-4388(2013/13-0097)

Parkinson, A. (2016). Grammatical accuracy and syntactic complexity in the oral and written narratives of fourth grade typically developing children. Department of Communication Sciences and Disorders, Idaho State University, Meridian, ID.

Paul, R., Hernandez, R., Taylor, L., & Johnson, K.(1996). Narrative development in late talkers; Early school age. *Journal of Speech & Hearing Research, 39,* 1295-1303.

Petersen, D. B., Gillam, S. L., & Gillam R. B. (2008). Emerging procedures in narrative assessment: The Index of Narrative Complexity. *Topics in Language Disorders, 28,* 115-130. doi: 10.1097/01.TLD.0000318933.46925.86

Petersen, D. B., Gillam, S. L., Spencer, T., & Gillam, R. (2010). The effects of literate narrative intervention on children with neurologically based language impairments: An early state study. *Journal of Speech, Language, and Hearing Research, 53,* 961-981.

Pinto, G., Tarchi, C., & Bigozzi, L. (2015). The relationship between oral and written narratives: A three-year longitudinal study of narrative cohesion, coherence, and structure. *British Journal of Educational Psychology, 85,* 551-569. doi: 10.1111/bjep.12091

Propp, V. (1968). *The Morphology of the Folk Tale*. [Kindle DX Version]. Retrieved from Amazon.com

Rubin, D. (1982). Adapting syntax in writing to varying audiences as a function of age and social cognitive ability. *Journal of Child Language, 9,* 497-510. Doi: 10.1017/S0305000900004839

Rumelhart, D., (1975). Notes on a schema for stories. In D. Bobrow & A. Collins (Eds.) *Representation and Understanding: Studies in Cognitive Science (pp. 211-235).* New York: Academic Press.

Schickdanz, J., Schickdanz, D, Forsyth, P., & Forsyth, G. (2001). *Understanding children and adolescents* (4th ed). Needham Heights, MA: Allyn & Bacon.

Scott, C., & Windsor, J. (2000). General language performance measures in spoken and written narratives and expository discourse of school-age children with language learning disabilities. Journal of Speech, Language and Hearing Research, 43, 324-339.

Silliman, E. R. (1989). Narratives: A window on the oral substrate of written language disabilities. *Annals of Dyslexia, 39,* 125-139.

Snow, P., Eadie, P., Connell, J., Dalheim, B., McCusker, H., & Munro., J. (2014). Oral language supports early literacy: A pilot cluster randomized trial in disadvantaged schools. *International Journal of Speech-Language Pathology, 16,* 495-506.

Stetter, M., & Hughes, M. (2010). Using story grammar to assist students with learning disabilities and reading difficulties improve their comprehension. *Education and Treatment of Children, 33,* 115-151. doi: 10.1353/etc.0.0087

Swanson, L., Fey, M., Mills, C. & Hood, L. (2005). Use of narrative-based language intervention with children who have specific language impairment. *American Journal of Speech-Language Pathology, 14,* 131-143. doi: 1058-0360/05/1402-0131

Thordardottir, E. (2008). Language-specific effects of task demands on the manifestation of specific language impairment: A comparison of English and Icelandic. *Journal of Speech, Language and Hearing Research, 51*(4), 922-37. doi:10.1044/1092-4388(2008/068)

Westby, C. (1991). "Learning to talk – talking to learn: Oral-literate language differences.' In C. S. Simon (Ed.), *Communication Skills and Classroom Success* (pp. 334–355). San Diego, CA: College-Hill.

Wiig, E. H., Semel, E., Secord, W. A. (2013*). Clinical evaluation of language fundamentals (5th ed.).* Bloomington, MN: Pearson.

Appendix A

Story Prompt Pictures

Appendix B

Narrative Elicitation Examiner Instructions

Written Narrative

1. The examiner reads The Dragon Story from the Test of Narrative Language (TNL; Gillam & Johnston, 2004) aloud to the participant(s).
2. The examiner instructs the participant(s) "Stories have a beginning, things that happen, and an ending. Tell a story that's as good as one in your favorite book."
3. Show the picture prompt to the participant(s).
4. Provide the participant(s) with a lined sheet of paper that has the prompt "Last week Miss Clarkson's fourth grade class…" on the top line of the paper.
5. The examiner encourages the participants to take a few minutes to think about their story before writing.
6. Allowed prompts:
    a. "How does the story begin?"
    b. If the child is erasing frequently, the examiner may say "If you make a mistake while you are writing, it is okay to just cross it out and keep writing."

b. Time limit: 30 minutes

    a. Provide a "three-minute warning" after 27 minutes of writing.

Oral Narrative

1. The examiner reads The Dragon Story from the Test of Narrative Language (TNL; Gillam & Johnston, 2004) aloud to the participant.
2. The examiner instructs the participant "Stories have a beginning, things that happen, and an ending. Tell a story that's as good as one in your favorite book" (Parkinson, 2016)
3. Show the picture prompt to the participant(s).
4. The examiner gives the participants one minute to think about their story before speaking.
5. Allowed prompts:
    a. "How does the story begin?"
    b. "Go ahead and start."
    c. If the child's story is only a few sentences in length, the prompt "Can you tell me anything else about your story?" may be given.

    Time limit: 30 minutes

    . Provide a "three-minute warning" after 27 minutes of speaking.

Appendix C

Index of Narrative Complexity Story Coding Form

Narrative Element	**0 Points**	**1 Point**	**2 Points**	**3 Points**
Character A character is any reference to the subject of a clause in a narrative.	No main character is included, or only ambiguous pronouns are used. Examples: a. They were walking. b. He was walking.	Includes at least one main character with non-specific labels only. Note: Only code each character one time. Examples: a. Once there was a boy. b. The boy was walking.	Includes one main character with a specific name for the character. Examples: a. Once there was a boy named Charles.	Includes more than one main character with specific names. Examples: a. Once there was a boy named Charles and a girl named Mary.
Setting A setting is any reference to a place of time in a narrative.	No reference to a specific or general place. Examples: a. The boy and the girl were walking.	Includes reference to a general place or time. Examples: a. The boy and the girl were outside. b. It was daytime. c. One day, they went to the park.	One or more references to specific places or times. Examples: a. Once there was a boy and a girl walking in Central Park. b. They were walking at 10:00 at night.	

Narrative Element	**0 Points**	**1 Point**	**2 Points**	**3 Points**
Initiating Event An initiating event is any reference to an event or problem that elicits a response from the character(s) in a narrative.	An event or problem likely to elicit a response from a character is not stated. Examples: a. The girl looked at the boy. The boy and girl were walking in the park.	Includes at least one stated event or problem that is likely to elicit a response from a character, but there is no response directly related to that event. Examples: a. The girl was walking in a park and saw a spaceship land (event) and she saw some aliens, and she saw a dog, and a table, and…	Includes at least one stated event or problem that elicits a response from the character(s). Examples: a. The girl was walking in a park and saw a spaceship land and she saw some aliens (IE). The girl started to run away (action).	Two or more distinct stated events or problems that elicit a response from the character(s). Examples: a. The girl was walking in a park and saw a spaceship land and she saw some aliens (IE-1). The girl started to run away (action). But while she was running, her shoe got stuck in a hole. (IE-2). She quickly knelt down and took off her shoe to get unstuck (Action.).

Narrative Element	**0 Points**	**1 Point**	**2 Points**	**3 Points**
Internal Response An internal response is any reference to information about a character's psychological state including emotions, desires, feelings, or thoughts.	No overt statement about a character's psychological state.	One overt statement about a character's psychological state not causally related to an event or problem. Examples: a. The dog was sad, the girl was happy.	One or more overt statements about a character's psychological state causally related to an event or problem. Examples: a. The aliens landed. Sara saw the ship and was terrified.	
Plan A plan is any cognitive verb reference that is intended to act on or solving an initiating event. It must include a "cognitive verb" that indicates a plan. Note: The plan and the action/attempt can share the same clause (see 2 points example b)	No overt statement is provided about the character's plan to act on or solve the event or problem. Examples: a. The girl was very excited and she ran out to meet the aliens.	One overt statement about how the character might solve the complication or problem. Examples: a. The girl thought that it would be neat to go and meet the aliens.	Two overt statements about how the character might act on or solve the event(s) or problem(s). Examples: a. The girl was very excited and she told the boy that she wanted to go meet the aliens. b. The boy was very scared so he decided to sneak away quietly.	Three or more overt statements about how the character might act on or solve the event(s) or problem(s).

Narrative Element	**0 Points**	**1 Point**	**2 Points**	**3 Points**
Action Actions are taken by the main characters but are not directly related to the IE. Attempts are taken by the main character(s) that are directly related to the IE.	No actions are taken by the main character(s). Examples: a. There is a girl. There is a boy. It is sunny.	Actions by main characters are not directly related to the IE. Examples: a. The boy and the girl were walking in a park. b. They saw a boy alien waving.	Attempts by main character are directly related to the IE. Examples: a. The girl thought that it would be neat to go and meet the aliens so she got away from the boy and walked out on the grass.	
Complication A complication is an event that prohibits the execution of a plan or action taken in response to an initiating event. Note: A complication can also be a second initiating event. In this case, code both a complication and an initiating event.	No complications.	One complication that prohibits a plan or action from being accomplished. Examples: a. The spaceship landed. The girl decided to get away from the aliens and started running from the spaceship. While she was running, her shoe got stuck in a hole. She could not get away from the aliens.	Two distinct complications that prohibit plans or actions from being accomplished. Examples: a. The girl was walking in a park and saw a spaceship land and she saw some aliens (IE-1). The girl started to run away (Action-1). But while she was running, her show got stuck in a hole (complication-1/IE-2). She quickly knelt down and took off her shoe to get unstuck (Action 2) but she was shaking too much to get her show off (Complication—2).	

Narrative Element	**0 Points**	**1 Point**	**2 Points**	**3 Points**
Temporal Markers e.g., when, next, then, immediately, instantly, after, again, already, always, before, lately, now, once, presently, rarely, today, weekly, while	No temporal markers.	One temporal marker. Examples: a. The girl walked over to the aliens. Then they all ate some lunch. b. After the aliens landed, the girl screamed.	Two or more temporal markers. Examples: a. When the girl saw the aliens, she ran out to meet them. She already knew they would be nice.	
Causal Adverbial Clauses e.g., because, since, so that, therefore, as a result, consequently, thus, hence, etc.	No causal adverbial clauses.	One causal adverbial clause. Examples: a. The aliens were not nice to the girl <u>because they were scared</u>.	Two or more causal adverbial clauses. Examples: a. The aliens were not nice to the girl <u>because they were scared.</u> <u>Since they were mean,</u> she ran away.	
Knowledge of dialogue Knowledge of dialogue is registered by a comment or statement made by a character or by characters engaging in conversation.	No dialogue.	One character makes a comment or statement. Examples: a. He said “Ow.” b. He said “Don’t come over here!”	Two or more characters engage in conversation. Examples: a. He said, “Oh look, there is an alien!” and she said “Oh, let’s go see them…”	

Narrative Element	**0 Points**	**1 Point**	**2 Points**	**3 Points**
Narrator Evaluations Narrator evaluations are any explanation provided in the story to justify why an action or event took place. e.g. because, since, so, and in order to.	No narrator evaluations.	One narrator evaluation. Examples: a. She ran up to say hello to the alien <u>because she always wanted to meet one.</u>	Two or more narrator evaluations. Examples: a. She knew that it was an alien spaceship <u>because everyone knows about UFOs.</u> b. He wanted to run from the aliens <u>since they were his worst nightmare.</u>	

Reprinted with permission from "Emerging procedures in narrative assessment: The Index of Narrative Complexity," by D. Petersen, S. Gillam and R. Gillam (2008), *Topics in Language Disorders, 28*, pp. 122-125. Copyright 2008 by the American Speech Language and Hearing Association.