

Cultural Responsiveness and Formative Reading Assessments: Retellings, Comprehension Questions, and Student Interviews

SUSAN V. PIAZZA
Western Michigan University

Abstract

How do formative reading assessments influence educators' ability to assess readers' understandings in culturally responsive ways? This study examines three formative reading assessments to explore the capacity of each measure to fairly represent readers' understandings without being influenced negatively by social and cultural diversity. The guiding question is "How do these three formative assessments inform and support culturally responsive literacy instruction?" Participants in this study include 10 young adolescent African American male readers. Data collection and analysis took place in a Midwestern urban university in the United States and makes use of a cross case comparison format. Interviews reveal that readers are the best informants regarding their own understandings about texts. Comprehension questions and retellings reveal discrepancies across readers' understandings. It is crucial that students are given the benefit of responsive assessments in order to accurately demonstrate academic strengths and areas of instructional need.

When I say begin, start reading aloud at the top of this page (point). Read across the page (point). Try to read each word. If you come to a word you don't know, I'll tell it to you. Be sure to do your best reading. Ready, begin. At the end of 1 minute, place a (|) after the last word read, and say "Stop" (University of Oregon Center for Teaching and Learning, 2012).

If you took a cross-section of classrooms in the United States, you would likely observe reading assessments like the one cited above. This example was retrieved from the official website of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessment. This assessment is widely used in the United States, and has gained influence internationally as a quick way to measure oral reading fluency, among other sub-skills, based on number of words read accurately within one minute. It is recommended that this assessment be followed up with a retelling fluency assessment to measure comprehension, which quantifies words retold in one minute from the same passage. The resulting scores are compared to pre-established benchmarks and provide a hierarchy of performance levels. Sometimes these benchmarks are posted in classrooms or school hallways to show students' weekly progress. Many of my graduate students who are full-time teachers express uneasy feelings about this display of progress monitoring and the limitations of the data in terms of informing their instruction. Research shows that the predictive validity of the retelling portion of DIBELS does not provide accurate information representing students' abilities to comprehend text (Billinger, 2011), and the oral reading fluency is not a valid measure of the fluency construct (Samuels, 2007).

This kind of assessment is more likely to be implemented in low performing schools that have diverse populations and high levels of poverty. If you were in an affluent school district, you might observe different kinds of formative assessment such as student-teacher conferences to discuss questions, comments, and personal connections. There are many different ways to formatively assess students' understandings of texts. Formative reading assessments are most valuable to classroom educators because they provide continuous feedback loops that reveal students' needs while allowing for adjustments in instruction to ensure continued reading development (Roskos & Neuman, 2012). This article will examine several forms of assessments based on previous research and experience in a K-12 clinical reading research center. It is likely that most educators, both those who believe in text-centered skills-based approaches, and those who support student-centered sociocultural approaches, would agree that formative reading assessments are intended to inform instruction and serve students. Assessments that are most responsive to diverse learners provide information that helps educators design instruction based on differences and individual needs.

Skills-based approaches use narrow definitions of reading that rely primarily on the alphabetic code and word recognition as precursors to reading comprehension and vocabulary development (Adams, 1990). Narrow definitions of reading are arguably neater because they focus on text and process rather than the increasingly diverse student backgrounds in today's classrooms (García & Bauer, 2009). Skills-based approaches to reading receive a great deal of funding and scientific recognition, and provide opportunities to conduct experimental studies that measure reading skills by examining speed and accuracy, such as the DIBELS example above.

A sociocultural student-centered approach, by definition, is a more complex perspective that considers the diverse background experiences, cultures, and languages of readers. Researchers have documented the growing nature of student populations who are culturally diverse and do not fit social or cultural norms, and who receive labels such as at-risk, trainable, or low-level functionaries when evaluated from a skills-based definition of reading (Luke, 2005; Strauss, Goodman, & Paulson, 2009). Notwithstanding the United States' well-intentioned No Child Left Behind (NCLB) Act (2002), students would benefit tremendously if our policy makers, educators and the general public would think more deeply about how children might be inappropriately positioned when assessments do not elicit and respond to differences in race, gender, class, language, and the like.

For example, African-American boys have a greater chance of experiencing cultural conflict with texts in school settings (Delpit, 1995; Piazza & Duncan, 2012; Tatum, 2008). The National Center for Education Statistics (2011) reports pervasive achievement gaps in reading regardless of the countless attempts to address the issue. African American boys sometimes perceive literacy activities as feminine and not appropriate for boys (Newkirk, 2002). They may also perceive some mainstream texts to be irrelevant to their lives (Piazza & Duncan, 2012; Tatum, 2005). A sociocultural and student-centered definition of reading would include all of these factors when evaluating the abilities of students. Since there are no formal or standardized assessments designed to account for these factors, ultimately teachers must use multiple approaches in order to benefit diverse learners.

Cultural responsiveness is addressed here in an analysis of three formative reading assessments, examining which measures have the capacity to fairly address reading comprehension without being negatively influenced by the readers' cultural and linguistically diverse backgrounds. The research question guiding this study is "How do these three formative assessments inform and support culturally responsive literacy instruction?"

Current Dilemmas

Current assessment practices are grounded in the National Reading Panel's (NRP) report (2000) that sought to make a final determination about which is the best and most reliable way to teach reading. Upon reviewing a limited list of experimental studies, the panel identified the "big 5" components of reading instruction for all children, towards which many student evaluations are explicitly and quantitatively directed. These five big ideas focus on a skills-based approach to reading and include the alphabetic principle, phonemic awareness, oral reading fluency, vocabulary knowledge, and reading comprehension (both literal and inferential).

An understanding of reading as a social practice contrasts with the traditional view of reading that defines it as a product resulting from the accumulation of hierarchical skills and sub-skills. In keeping with a view of reading as a social practice, assessments tend to be more qualitative in nature and tend to ask students themselves to describe their thinking before, during, and following reading events. This paper focuses on three such formative assessments: 1) retellings, 2) comprehension questions, and 3) interviews. The first two assessments examined are adapted from the Qualitative Reading Inventory-5 (QRI-5) (Leslie & Caldwell, 2010), which offers assessment tools that cover both retellings and comprehension questions. The QRI-5 is an oral reading assessment that includes concept questions for pre-reading, oral reading analysis, retellings, comprehension questions, and optional think alouds. This kind of assessment might be considered a more balanced measure of comprehension because it emphasizes accuracy and provides quantifiable data, but also offers some qualitative information with the pre-reading questions, retellings, and optional think-alouds. The third assessment examined is open-ended interviews with students in response to textual understandings.

Sociocultural Theoretical Framework

Sociocultural theories of literacy and literacy as a social practice inform this article (Au, 2011; Lewis, Enciso, & Moje, 2007; Luke, Woods, & Dooley, 2011; Vygotsky, 1978). From these perspectives, readers construct meaning from texts through lenses of social and cultural lived experiences (Lankshear, 1997). Readers' prior knowledge and backgrounds serve as filters through which they transact with ideas in the texts they encounter and each reader will construct nuanced understandings of each text they read (Bruner, 1986; Rosenblatt, 2005).

Sociocultural literacy research has demonstrated the importance of race, class, gender, and the affective influence of these factors. For example, boys typically gravitate to books with humor, high levels of activity, and books with male characters (Newkirk, 2002). As well, boys often perceive reading as a feminine activity (Hall & Piazza, 2008; Newkirk, 2002; Smith & Wilhelm, 2002). Tatum (2005) advocates changes to skills-

based instruction and assessments so that the relationship between literacy and black males is strengthened and equal opportunities for success are provided. Even though the relationship between sociocultural factors and reading is well established, it is rare to see this discussion occur in connection with classroom assessments in the contemporary American context.

Instead of viewing reading assessments as a measurement device, they are better viewed as an inquiry process in which teachers take a curious stance toward students' abilities to use language and represent their understandings of texts (Delandshere, 2002; Serafini, 2005). Assessment can be a form of inquiry presented as a socially contextualized and interpretive activity in which "teachers and students are viewed as active creators of knowledge rather than as passive recipients" (Wells, 1994 cited in Serafini, 2005, p. 248). This sociocultural framework for conceptualizing the role of reading assessments is grounded in social learning theories of Vygotsky (1978) and Halliday's (1977) language-based learning theories. When we look closely at how students think during reading, it is necessary for teachers to look at multiple data sources, in multiple ways, to gain a credible perspective on how students understand text. I hold that assessment should be situated as a form of inquiry to help educators begin connecting sociocultural theories and reading assessments.

According to Brown and Cambourne (1990), retelling a story or experience is a natural linguistic behavior and an authentic way to assess. The documentation of oral language traditions (Delpit, 1995; Smitherman, 2000) in many African-based, and Aboriginal or Native American cultures supports the use of oral communication as a valid way to assess understandings of text. Therefore, oral retellings might serve as socially and culturally responsive ways to assess some diverse learners' thinking in relation to text. Another important piece of responsive evaluation is the triangulation of multiple sources of data. Assessment as inquiry provides a framework to look at the reading process from a position of genuine curiosity, rather than as a construct to be measured. Given a framework of assessment as inquiry, this study synthesizes retellings, comprehension questions, and student interviews collected over a period of three years.

Methods

This study examines the use of retellings, comprehension questions, and interviews with adolescent males considered at-risk of academic failure due to institutions and assessment practices that do not account for their differences, or meet their needs, related to issues such as race, class, gender, language, general learning disabilities, parental incarceration, and poverty. This study extends an earlier data set (Piazza, 2006) which looked closely at three participants. The data set has been extended to ten. The guiding question in the current study focuses on how three formative reading assessments compare to one another in their ability to provide culturally responsive assessments of readers' instructional needs.

Through the lens of assessment as inquiry within a sociocultural framework, a cross-case comparative case study design (Merriam, 1998; Wolcott, 2001) was used to collect, analyze, and report data. The participants' understandings of texts in relation to their lived experiences were gained primarily through interviews. Cross-case analysis reveals ways in which assessments may or may not offer a window on diverse learners'

understandings and interactions with texts. Looking across ten cases revealed themes related to the use of each assessment (Yin, 1994). This follow-up study to an earlier project is an effort to extend the database and confirm earlier understandings about ways that readers think about texts. However, this study focuses on the utility of the assessments themselves, rather than detailed case analyses of each participant. This perspective should facilitate the examination of whether these assessments offer culturally responsive measures of reading comprehension.

Procedures and Participants

Ten African American male readers between the ages of 9-11-years-old were purposefully selected for this study from a larger pool of reading program attendees. These participants attended an after-school reading program at the university, where they received individualized instruction based on detailed assessments. The data was collected over a period of five sessions. The first four sessions consisted of listening to the text on audio tape while they read along in the book. Following the advice of J. Hale (personal communication, April 3, 2001) each text was audio recorded by an African American male to provide relevance, to control for oral reading differences, and to ensure the text presented in dialect was delivered authentically. This study does not aim to measure fluency, but places primary emphasis on understanding texts. The audio recordings minimize the effects of oral reading difficulties.

Two of the 10 case descriptions, Tony and Steve (pseudonyms), are used to provide context alongside of the assessment data. First, Tony was ten years old and lived with his single mother who was observably supportive of his academic achievements. He struggled between engaging academically, and his social status of being cool. His grades were falling below average in school, and while he was slightly uneasy during the interviews and assessments, he provided candid remarks, making his thinking more transparent than many of the other participants.

The second student, Steve, was 11-years-old during his participation in the reading program. His father was incarcerated at the time and his mother brought him to the reading program with regular attendance and displayed a strong commitment to ensure his academic achievement in school. She also attended classes at the university, and struggled financially to make ends meet. Steve had no difficulty engaging in the reading program and interacting with texts, however, his mother reported that he was not getting good grades in school and wanted additional support.

Over a period of four sessions, each participant engaged in the following assessments: 1) retellings which were analyzed using two different kinds of analyses (Feathers, 2002; Goodman, Watson, & Burke, 2005), 2) 10 comprehension questions analyzed using the pre-determined answers modeled after the QRI-5 (Leslie & Caldwell, 2010), and 3) open-ended informal student interviews analyzed for thematic patterns (Merriam, 1998; Schensul, Schensul, & LeCompte, 1999). Lastly, participants engaged in a fifth and final interview in which they compared their thinking across four texts and made overall reflective statements. The retelling guides and comprehension questions were created for each text in the original study by 30 graduate students as outside raters for reliability and validity. There were four retelling narratives collected per participant in

this study, which resulted in 40 retellings. Each session ended with open-ended interviews focused on intersection between the text and the reader’s lived experiences.

Data Analysis

The significance of this cross-case comparative design is in the combination of data sources that provide multiple lenses from which to analyze participants’ textual understandings and how the assessments may or may not reflect their thinking. Attempting to measure participants’ thinking always raises concerns about validity, and one solution to this is through triangulation of multiple data sources (Miles & Huberman, 1994). Pappas (1987) and Sulzby (1985) caution researchers that social contexts of retelling assessments might influence outcomes; therefore, the study was designed to incorporate retelling procedures that students were already accustomed to from their after-school tutoring experiences at the university.

The first data source, open-ended retellings, were analyzed using two different protocols. First, retelling guides (Goodman, Watson & Burke, 2005) provided percentage scores. There are typically six sections in retelling guides: a) character identification, b) character development, c) theme statement, d) plot statement, e) list of events, and f) subtleties that credit readers with additional information recalled that was not accounted for in the guide. Graduate students provided 99% inter-rater reliability on the scoring of retelling guides.

The retellings were analyzed a second time using the Text-Related Thinking (TRT) analysis (Feathers, 2002). This tool provides a framework that helps identify the various kinds of thinking used to retell a passage. The analysis consists of parsing units of the retelling narrative into categories of meaning, in some ways similar to discourse analysis (Gee, 1999). Table 1 provides a brief sampling of categories used in the analysis of text-related thinking. There are thirteen kinds of thinking in addition to seven different levels of summarization. For further details, see Feathers (2002).

Table 1

Examples of Text-Related Thinking (Feathers, 2002)

Summarization	Statements that condense text information, i.e. single sentences, multiple sentences, single paragraphs, multiple paragraphs, single episodes, or macro level.
Character Description	Statements that describe physical, mental, or emotional features of a character.
Character Thoughts/Feelings	Statements that describe the thoughts or feelings of a character.
Causality	Statements that identify character intent or motive and/or other reasons for events.
Comparison	Statements that identify similarities or differences between characters, objects, or events in the text.
Conclusion	Statements that synthesize multiple text items and use reader logic to infer something about character or events.

Condition	Statements that identify the conditions under which an action will or has taken place.
Evaluation	Statement that present a judgment of characters, events, ideas, or thoughts in the text, or of the text itself, or portions of the text.

The second data source, comprehension questions, consisted of five implicit and five explicit questions. The questions were generated by 30 graduate students who were studying assessments, and they served as raters for the validity of the comprehension questions. The questions were scored based on pre-determined answers just as the QRI-5 assessment (Leslie & Caldwell, 2010). The data generated percentage scores based on accuracy. This was the most rigid, and least responsive, assessment used in the study, however it provides useful information for the overall comparisons.

The third data source consists of open-ended interviews analyzed with Spradley's (1980) thematic analysis in which the transcripts are read and re-read to identify categories of meaning. Initially, I began with three overall themes: (a) people/characters, (b) places, and (c) activities. I sorted each participant's statements into these three areas. Within these larger themes, sub-themes began to emerge and used to compare and contrast their thinking across the multiple assessments. Following the interview analyses, the assessments became contextualized in ways that attended to sociocultural and linguistic nuances emerging during their interactions and responses with each text.

Text Selections

Children's literature was not the intended focus of this study. However, four authentic picture books were selected to provide comparisons of participants' transactions with the books. The text selections provided a backdrop of social, cultural, and linguistic variations. In order to uncover differences in each reader's understandings, it was important to minimize the differences between texts so that responses were not markedly different due to textual features. Therefore, the texts were carefully selected with consideration to difficulty, character gender, racial and cultural features, and variations across linguistic features, setting, readability, length, theme, genre, and story structure (see Table 2).

Table 2

Texts Selected for Retellings and Comprehension Questions

Enemy Pie
(Munson, 2002)

Two cartoonish white male friends, formerly enemies, set in a suburban neighborhood. It contains a complex plot with high interest, outdoor activities, conflict and resolution.

Three Wishes

Male-female African-American friendship

(Clifton, 1992)

story told through dialogue, written in a form of dialect, and contained conflict and resolution.

The Best Friends Club
(Winthrop, 1989)

Male-female white best friends formed a clubhouse. The story introduces another male character, has outdoor activities, conflict and resolution.

Heroes
(Mochizuki, 1995)

Japanese-American male is bullied by his mostly white friends who do not believe his father is a war hero. It contains outdoor activities, group conflict and resolution.

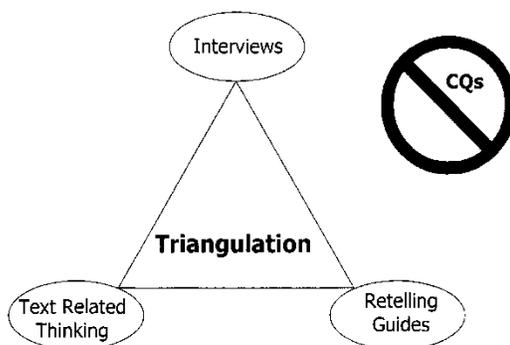
Findings

The three kinds of formative reading assessments produced different outcomes regarding the ways in which participants' understandings of texts are portrayed. Findings across the four data sources include: (a) comprehension questions that were inconsistent with the retelling and interview findings, (b) participants' personal connections with texts were consistent with retelling and interview findings, and (c) the majority of participants felt that retellings are the most effective way to demonstrate how they understand texts. The following sections will present evidence within each of these areas.

Interviews

Overall, Figure 1 represents how the two retelling data sources and interviews triangulated, and comprehension question scores were the outliers. Interview data revealed that most personal connections occurred with texts that elicited higher scores on retelling guides and higher levels of text-related thinking (see Tables 3 and 4 in next sections). Themes that emerged from the interview data were identified and include overarching categories such as people, places, activities, and questions or comments to the researcher. 17 interview questions were used to elicit the ways that students were thinking about and understanding the texts.

Figure 1. Triangulation of Data Sources



Sample interview questions included: (a) Tell me what interested you (or not) about this story, (b) Tell me what you thought about the language in the story, (c) Tell me about the message in this book, and (d) Tell me about any experience you may have had similar to the events in this story, and (e) Describe why or why not you think this story represents how kids in real life, your own neighborhood, get along. How is getting along similar or different in your life? A thematic analysis (Spradley, 1980) was used to determine themes and sub themes from the fifty interviews. Examples from the interview themes and sub-themes are displayed in Table 3.

Table 3

Operational Definitions of Themes and Sub Themes

Places		
• Neighborhoods	Descriptions and evaluations about neighborhoods.	"...because I live in a nice neighborhood because it's friendly." (P3,I1).
Activities		
• Conflict Resolution/ Violence	Ways to resolve conflicts and talk about violence.	"You might not know somebody, but they keep on messin' with you..." (P4, I1).
• Interests	Activities and events that are referred to as interesting and fun.	"And you could make his tree house fall out of the tree just to make it a funny story." (P5, I1).
• Popular Culture	References made to current social and/or cultural interests, fashions, or games.	"Well, I saw it on Law and Order." (P3, I4).

The overall trend in the boys' responses to the books at first seemed counter-intuitive because they ranked the texts in the following order from most interesting to least: 1) *Enemy Pie*, 2) *The Best Friends Club*, 3) *Heroes*, and 4) *Three Wishes*, which ranks the two mainstream texts with highest interest. There were only three exceptions to these rankings where preferences were in a different order. For example, Tony ranked *Three Wishes* as his second favorite and *The Best Friends Club* as his least. *Enemy Pie* was by far the most engaging text according to all ten participants because of the interesting plot and outdoor activities. *Three Wishes* was, for the most part, the least preferred text in this study.

In response to *Three Wishes*, the other boys shared, "I didn't like it very much...the part where they always talk and stuff. And they didn't do anything fun!" and "I'd change that boy to a girl 'cause he look like one...the light eyebrows and eyes, how they look (pointing to the curve in the main character's eyebrows)...and you know how girls be hangin' out together and stuff." Multiple comments were made about too much talking and too little activity in *Three Wishes*.

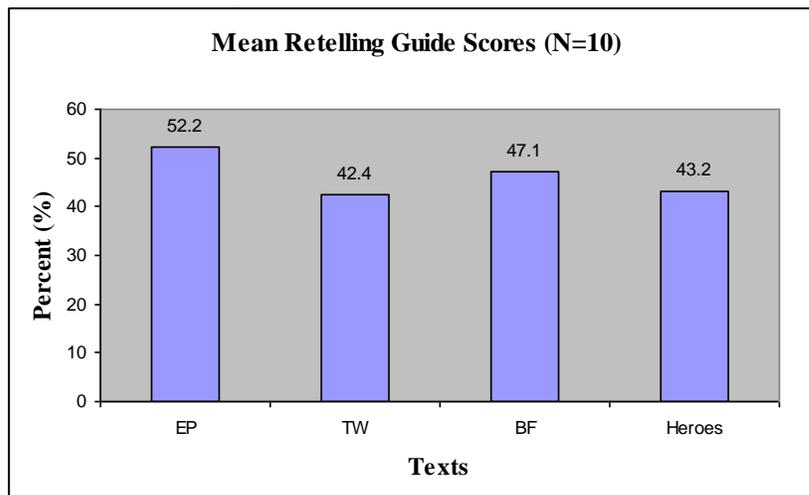
Tony explained the complexity of his preferences with regard to race, “It [books with dark-skinned characters] don’t make the story better...I just don’t like seein’ light skinned people all the time, I wanna see my color, too”. Even though Tony expressed that the story was not that exciting, he made it clear that the racial identities represented in *Three Wishes* were valued. Steve shared similar views in which he acknowledged race, but referred to *Three Wishes* as a boring text.

In sum, the interview data revealed that the two mainstream texts, *Enemy Pie* and *The Best Friends Club* elicited more personal connections and, therefore, richer conversations about ideas presented in those texts. The two texts chosen for the diversity represented in settings, characters, and even language did not elicit as many personal connections. In fact, many comments across all ten participants revealed a distancing from the ideas presented in those texts because of the lack of masculinity displayed by the male characters. Some commentary regarding the bullying in *Heroes* included, “nobody ever bullied me”, and “he started cryin’ and...then he stopped cryin’”, and one participant referred to the character being bullied as a “wimp”. In response to *Three Wishes*, comments about the male character’s feminine features, hanging out with a girl, and wearing the color pink were abundant. The majority of participants expressed prototypical views of masculinity and rejected the alternative perspectives presented in the two non-mainstream texts. For further details that elaborate on personal connections with texts, see the original study (Piazza, 2006).

Retelling Guides

The mean retelling guide scores aligned well with participants’ personal connections to characters and ideas found in the texts. Note the relatively lower mean scores in *Three Wishes* and *Heroes*. Interviews revealed that participants paid a great deal of attention to the characters, events, and plot details in *Heroes* because of the outdoor war games and plot structure, but did not attend to details in *Three Wishes*, which further supports their self-reported lack of interest and lower scores. As well, the mean retelling guide scores in *Enemy Pie* were higher across all participants (see Figure 2).

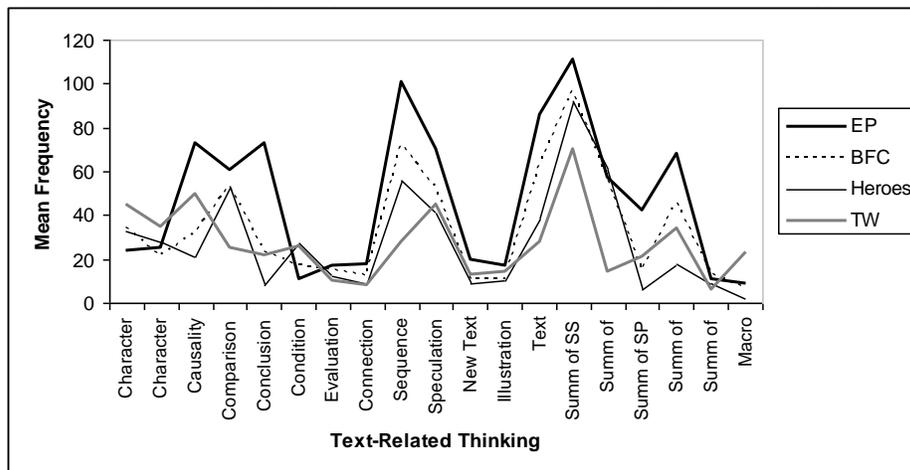
Figure 2. Retelling Guides



Retellings: Text-Related Thinking (TRT)

The TRT analysis revealed trends, rather than actual scores. In this analysis, retellings were parsed into units representing different kinds of thinking. Figure 3 displays the frequency of each text-related-thinking category across the books. The significance of examining data across books in this study is that it allows the sociocultural contexts and readers' connections to these contexts to surface while examining the reading assessments. The findings from this data source are consistent with the interview and retelling guide findings that show higher levels of connection and interest in the two mainstream texts *Enemy Pie* and *The Best Friends Club*. A line graph format visually highlights the higher and lower frequency in the TRT across books. It is fascinating to compare the ways that the boys examined character traits and thinking about characters and events given the categories in this analysis.

Figure 3. Mean Frequency of Text-Related Thinking in Retellings



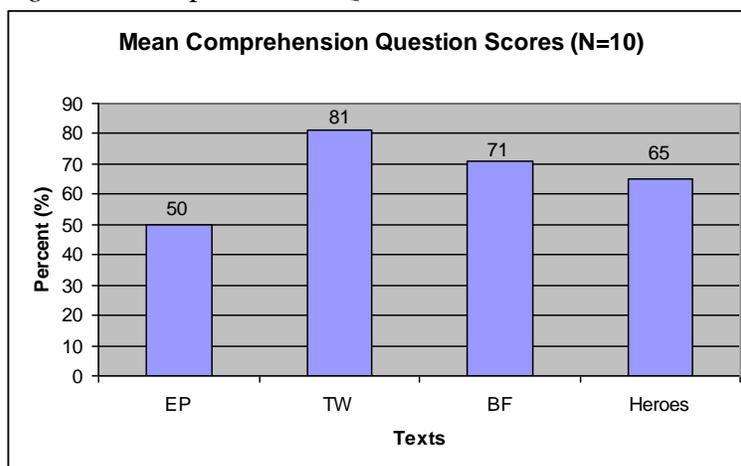
For example, *Enemy Pie* elicited a high level of personal connections evidenced by the higher number of units of causality, conclusions, sequencing, and speculation (more complex thinking). Also, the TRT analysis showed that *Heroes* and *Three Wishes* elicited slightly more thinking in the first two categories, character description and character thoughts and feelings, which is supported during interviews when readers reported these characters to be “wimpy” and “like a girl” and “not cool”. Descriptive thinking is often considered a lower level of thinking. Nine of 10 participants pointed out character flaws and details about those characters during interviews. The characters’ behaviors contradicted the boys’ own versions of their conceptions of male identities.

Comprehension Questions

The comprehension questions across all four texts were not consistent with the two retelling analyses or the interviews. As previously noted, all of the data sources triangulated with the exception of comprehension questions. The mean comprehension question scores were lowest in *Enemy Pie*, which was the text in which a majority of participants made strong personal connections. On the other hand, mean comprehension

scores were higher in *Three Wishes*, which proved to be a simplistic story according to multiple comments found in the interviews (Figure 4). Overall, the boys expressed a disinterest in *Three Wishes* because it lacked action, focused mostly on dialogue between a male and female, and the most notable concern expressed by eight of the 10 participants, was that the male character acted and looked like a girl. Even though the books only varied slightly in text structures such as length, readability, and grade level, the comprehension question scores were slightly higher in the least complex stories, and slightly lower in the more complex stories. One of the features that made *Three Wishes* a bit more simplistic was that it presented a sequence of events between two characters primarily through dialogue.

Figure 4. Comprehension Questions



Discussion of Findings

Interviews, retelling guides, and the TRT analyses appear to be more culturally responsive forms of reading assessments because they offer a window into the participants' understandings and connections with texts. The comprehension questions seemed to elicit a narrower view of the readers' understandings and connections to the texts. The comprehension questions position the readers as less capable of understanding the very texts that they found most interesting and connected with most. The comprehension questions as formative reading assessment might lead educators to draw inaccurate conclusions when looking at student comprehension, and are the least culturally responsive tool of the assessment strategies used in this study. These findings—that demonstrated the limitations of the comprehension questions in providing a culturally responsive assessment strategy – have confirmed the results of the original study.

Another significant benefit in comparing two kinds of retelling approaches was that the TRT compared nicely with retelling guides in the areas of character description and character thoughts and feelings. The two retelling approaches together revealed detailed opportunities to target instructional planning regarding character analysis,

alternative perspectives, and critical stances with texts (Hall & Piazza, 2008). However, a limitation with retellings is that they typically value quantity of recall, because lengthier retellings are rewarded, and brief summaries result in lower scores. It was the TRT analysis that revealed brief, but efficient, summary statements produced by the readers. For example, one participant's scores were higher on retelling guides due to his attention to detail, while another had lower scores because he provided higher levels summary statements in the TRT analysis.

The comprehension questions themselves were the outliers among the four different data sources. The questions are reliable and valid within a skills-based definition of reading because questions were developed using inter-rater protocols to reflect the same constructs as the Qualitative Reading Inventory-5 comprehension questions (Leslie & Caldwell, 2010). The contradiction in findings in the comprehension questions may have arisen due to the underlying skills-based definitions of reading that assume the questions will result in a single, most common, correct answer. These assumptions disregard the sociocultural lived experiences of readers and the context in which they are reading (Gee, 1999; Rosenblatt, 2005). Other researchers have also documented bias in mainstream approaches that assume a standardized social and cultural knowledge base (Au, 2011; Delpit, 1995; Heath, 1983; Klinger & Edwards, 2006).

The participants' preferences for the texts aligned with their demonstrated strengths on the retelling guides and complex ways of thinking. When asked about which reading assessment is the best method for teachers to evaluate students' understandings of texts, seven of the ten participants chose retellings. One participant summed it up with "[I like to] retell because it's like reading the story all over again". Another, "...because I'm good at it" and another, "I get to own the story". Another young man shared, "I really like how they made the illustrations and how they painted and drew [sic] it...the pictures help me to remember the story better and make it more interesting".

Implications for Choosing Culturally Responsive Formative Reading Assessments

Skills-based classroom reading assessments typically place value on attention to details, speed, and accuracy in recalling information, i.e. the DIBELS example provided at the beginning of this paper. As students move into middle and secondary grades, greater emphasis is placed on summarization, paraphrasing, and synthesizing information across disciplines. These tasks are more complex and require a greater understanding of texts and strategy use. Without a detailed feedback loop that informs teachers' efforts to support the comprehension, diverse learners will continue to struggle in upper grades, perhaps in part due to the culturally unresponsive diagnostic assessments used to address their needs in the earlier grades.

This study only touches on the grand conversation that is critical to understanding the concept of culturally relevant reading assessments and why they are so important to diverse populations. The close examination of only three different formative assessments has demonstrated that there are differences and some can indeed help teachers to be more responsive. Tatum (2005) confirms that classroom instruction with texts needs to move beyond comprehension questions if young black males' academic needs are to be met.

He says that understanding during the reading process must be an application of one's own life experiences to new situations in text.

“A large percentage of students are unable to extend the ideas of a text, make inferences, draw conclusions, and connect text to their own experiences” (Tatum, 2005, p. 111). There is an abundance of statistical data that encourages deficit perspectives regarding adolescent black males' academic proficiency (National Center for Education Statistics, 2011). These conclusions do not imply that there is no room for improvement in skills and reading proficiency, but the implications reveal that it is problematic to use formative reading assessments based on skills-only definitions of reading. They are limited for addressing individual and sociocultural differences. This has been particularly salient with the participants in this study who demonstrated their competence in thinking deeply about the texts we read.

The point of this study was to inquire into how formative reading assessments might be combined to explore student thinking in relation to the reading of various texts. In order to answer a question that asks “how” something is done, it is necessary to collect descriptive data from multiple sources, which included comprehension questions stemming from a more traditional paradigm, another in the form of retellings, coupled with interviews that position the boys as informants regarding their own understandings about texts. The findings support the argument for compiling evolving academic profiles of students to inform instruction rather than relying on single measures or benchmarking over time. It is crucial that students are given the benefit of responsive assessments of their understandings, as well as many different opportunities to demonstrate academic strengths and areas of instructional need.

It is the lived experience of each individual reader that determines how a particular text is understood. It also means that each text will take on multiple variations in each reader's mind when it comes to constructing meaning during any literary transaction (Iser, 1978; Rosenblatt, 2005; Smith, 1996). The texts that the boys liked least were the texts in which their performance was lowest in both retellings and interviews. Contrary to the boys' preferences, it was the texts that presented alternative perspectives and challenged their own worldviews that might challenge them to think critically about texts *and* the world they live in. Rather than using stopwatches and standardized protocols, the time has come for teachers to collect data *with* students as informants and compare the findings across time and space in order to find culturally responsive ways to assess diverse learners' reading.

References

- Adams, M. J. (1990). *Beginning to read*. Cambridge, MA: MIT Press.
- Au, K. H. (2011). *Literacy achievement and diversity: Keys to success for students, teachers, and schools*. New York, NY: Teachers College Press.
- Brown, H. & Cambourne, B. (1990). *Read and retell: A strategy for the whole language/natural learning classroom*. Portsmouth, NH: Heinemann.
- Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.
- Clifton, L. (1976). *Three wishes*. New York, NY: Doubleday Books.

- Delandshere, G. (2002). Assessment as inquiry. *Teachers College Record*, 104(7), 1461-1484.
- Delpit, L. (1995). *Other people's children: Cultural conflict in the classroom*. New York, NY: The New Press.
- Feathers, K. (2002). Young children's thinking in relation to texts: A comparison with older children. *Journal of Research in Childhood Education*, 17(1), 69-83.
- García, G. E., & Bauer, E. B. (2009). Assessing student progress in the time of No Child Left Behind. In L. M. Morrow, R. Rueda, & D. Lapp (Eds.), *Handbook of research on literacy and diversity* (pp. 233-253). New York, NY: Guilford Press.
- Goodman, Y., Watson, D., & Burke, C. (2005). *Reading miscue inventory: Alternative procedures*. Katonah, NY: Richard C. Owen.
- Gee, J. P. (1999). *Introduction to discourse analysis: Theory and method*. New York, NY: Routledge.
- Hall, L. A., & Piazza, S. V. (2008). Critically reading texts: What students do and how teachers can help. *The Reading Teacher*, 62(1), 32-41.
- Halliday, M. A. K. (1977). *Learning how to mean: Explorations in the development of language*. New York, NY: Elsevier.
- Heath, S.B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. New York, NY: Cambridge University Press.
- Iser, W. (1978). *The act of reading: A theory of aesthetic response*. Baltimore, MD: Johns Hopkins University Press.
- Klinger, J. K. & Edwards, P. A. (2006). Cultural considerations with Response-to-Intervention models. *Reading Research Quarterly* 41(1), 108-118.
- Lankshear, C. (1997). *Changing literacies*. Philadelphia, PA: Open University Press.
- Leslie, L. & Caldwell, J. (2010). *Qualitative reading inventory-5 (QRI-5)*. New York, NY: Longman.
- Lewis, C. Enciso, P. & Moje, E. B. (2007). *Reframing sociocultural research on literacy: Identity, agency, and power*. Mahwah, NJ: Erlbaum.
- Luke, A. (2003). Literacy and the other: A sociological approach to literacy research and policy in multilingual societies. *Reading Research Quarterly*, (38)1, 132-141.
- Luke, A., Woods, A. & Dooley, K. (2011). Comprehension as social and intellectual practice: rebuilding curriculum in low socioeconomic and cultural minority schools. *Theory into Practice*, 50, 157-164.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.
- Miles, M. & Huberman, A.M. (1994). *Qualitative data analysis* (2nd ed.). Los Angeles, CA: Sage Publications.
- Mochizuki, K. (1995). *Heroes*. New York, NY: Lee Low Books.

- Munson, D. (2000). *Enemy pie*. Vancouver, BC: Raincoast Books.
- National Center for Education Statistics. (2011). *The Nation's Report Card: Reading 2011*. Washington, DC. Retrieved from <http://nces.ed.gov/nationsreportcard/pubs/main2011/2012457.asp>
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington DC: U.S. Department of Health and Human Services, National Institute of Child Health and Human Development.
- Newkirk, T. (2002). *Misreading masculinity: Boys, literacy, and popular culture*. Portsmouth, NH: Heinemann.
- No Child Left Behind*. (2002). Public Law 107, Jan. 8, 2002, a reauthorization of the Elementary and Secondary Education act. Washington, DC.
- Pappas, C. and Brown, E. (1987). Young children learning story discourse: Three case studies. *The Elementary School Journal*, 97(4), 455-466.
- Piazza, S.V. (2006). Case studies of transactions between boys' lived experiences and texts. *Dissertation Abstracts International*, 67(05), 1-235.
- Piazza, S. V. & Duncan, L. E. (2012). After-school literacy engagements with struggling readers. *Reading Writing Quarterly* 28(3), 229-254.
- Rosenblatt, L. (2005). *Making meaning with texts: Selected essays*. Portsmouth, NH: Heinemann.
- Roskos, K. & Neuman, S. B. (2012). Formative assessments: Simply, no additives. *The Reading Teacher*, 65(8), 534-548.
- Samuels, S.J. (2007). The DIBELS Test: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly*, 42(4), 563-566.
- Schensul, S. L., Schensul, J. J., & LeCompte, M. D. (1999). *Essential ethnographic methods: Observations, interviews, and questionnaires*. New York, NY: Rowman & Littlefield Publishers, Inc.
- Serafini, F. (2005). Three paradigms of assessment: Measurement, procedure, and inquiry. In Mason, P. A. & Schumm, J. S. (Eds.). *Promising practices for urban reading instruction* (pp. 243-259). Newark, DE: International Reading Association.
- Smith, F. (1996). *Reading without nonsense*. New York, NY: Teachers College Press.
- Smith, M.W. and Wilhelm, J. D. (2002). *Reading don't fix no chevys: Literacy in the lives of young men*. Portsmouth, NH: Heinemann.
- Smitherman, G. (2000). *Talkin' that talk: Language, culture and education in African America*. New York, NY: Routledge.
- Spradley, J. P. (1980). *Participant observation*. Toronto, ON: Thomson Learning.

- Strauss, S. L., Goodman, K. S., & Paulson, E. (2009). Brain research and reading: How emerging concepts in neuroscience support a meaning construction view of the reading process. *Educational Research and Review*, 4(2), 21-33.
- Sulzby, E. (1985). Children's emergent reading of favorite storybooks: A developmental study. *Reading Research Quarterly*, 20(4), 458-481.
- Tatum, A. W. (2005). *Teaching reading to black adolescent males: Closing the achievement gap*. Portland, MN: Stenhouse Publishers.
- Tatum, A. W. (2008). Toward a more anatomically complete model of literacy instruction: A focus on African American male adolescents and texts. *Harvard Educational Review*, 78(1), 155-181.
- University of Oregon Center on Teaching and Learning. (2012). The official DIBELS website. Retrieved from <https://dibels.uoregon.edu/measures/orf.php#video>
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.). Cambridge, MA: Harvard University Press.
- Winthrop, E. (1989). *The best friends club*. New York, NY: Lothrop, Lee, & Shepard Books.
- Wolcott, H. F. (2001). *Writing up qualitative research*. Thousand Oaks, CA: Sage Publications.
- Yin, R. K. (1994). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.

Author Biography

Susan V. Piazza an Associate Professor of Literacy Studies in the College of Education and Human Development at Western Michigan University. Her research areas include literacy learning and assessment practices as they intersect with sociological influences and issues of equity, culturally responsive teaching, and school and community partnerships.