



What Do “Levels” Really Mean?

A Closer Look at Text Leveling

By Heather Koons, MetaMetrics,
Elfrieda (Freddy) H. Hiebert, TextProject

What happens in beginning reading instruction matters, since children’s initiation into successful reading predicts their later progress (Hernandez, 2010; Juel, 1988). Different types of texts, such as decodable or leveled texts, are advertised in the marketplace (and mandated by policy-makers) as evidence-based reading programs that will produce efficacious reading acquisition. But, until recently, assessment tools that provide information on aspects of reading development supported by different text types (or at different levels within a text program) have not been available to answer critical questions about the demands made by various text types on beginning readers. A recently developed tool, derived from theory and empirically validated with student performances and teacher ratings, makes it possible to examine both the within-level consistency and the across-level patterns of texts within beginning reading programs. In this study, this tool is applied to the two most prominent text types currently used in many beginning reading classrooms—decodable and leveled texts.

This content was presented at the annual conference of the American Educational Research Association, Philadelphia, PA, April 7, 2014



Theoretical Framework

As text types such as decodable and leveled texts replaced highly controlled texts epitomized by the “Dick and Jane” readers, systems of assessing text difficulty that rely on vocabulary and syntax (e.g., Spache readability formula; Lexiles) have proven less than reliable (Hiebert & Pearson, 2010). Qualitative measures of the complexity of beginning texts such as Fountas & Pinnell’s (2012) guided reading levels have provided a single indicator of text complexity, without an indication of how text features such as decodability, familiarity, and frequency of words influence rater’s assignment of levels. A newly developed tool makes it possible to examine the opportunities for learning afforded by different text types in a manner not previously possible. This tool was derived from theory and research of beginning reading acquisition (e.g., Mesmer, Cunningham, & Hiebert, 2012) and also empirically validated by teachers’ ratings and by students’ comprehension of the texts.

This text complexity tool for early reading texts—the Early Literacy Indicators (ELI)—was developed with the specific aim of addressing unique features of beginning texts (e.g., decodable but infrequent words; additive sentences in predictable texts). The tool was developed in a five-step process that has been described elsewhere (Elmore, 2013; Fitzgerald, 2013). In brief, 238 composites representing features of words (structure and meaning), sentences, and discourse in beginning texts were identified. The analysis was applied to a set of texts that represented prominent and unique beginning reading texts. The comprehension of approximately 1,200 first- and second-graders on a randomly stratified sample of the texts was used to establish which composite variables best accounted for students’ performances. Teachers’ ordering of texts was also used in identifying the composites. The features that best predicted student performance are represented in four composite descriptors as well as an overall scale score: (a) decoding: complexity of patterns of monosyllabic words and number of multisyllabic words in a text; (b) semantic: abstractness, rareness, and age-of-acquisition of words in a text; (c) structural: degree of concept overlap and density of concepts within and across sentences; and (d) syntactic: degree to which word, phrase, and letter patterns are repeated between adjacent sentences. The scores for the composites are presented as percentiles, which were established relative to the domain of beginning reading texts. A “low” percentile on a composite means that the text is “easier” on that feature.

A framework for the design and study of beginning texts recently proposed by Mesmer et al. addresses the three components represented in the ELI. In addition, Mesmer et al. argue that the tasks represented

by texts within and across levels of a program require consideration if the influences of text features on students' reading acquisition are to be understood. In this study, the ELI is used to describe the text opportunities within and across texts in two programs: leveled and decodable texts. The questions that this study addressed were: How similar/different are the features of texts within a level of a particular program? How similar/different are the features of texts across the levels of a particular program?

Methods

Data Sources

Two text sets developed to span the reading range from beginning of kindergarten to end of grade 1 were selected for analysis. Each has a distinctly different instructional purpose and both types are widely used in K-1 reading instruction. The leveled set (Fountas & Pinnell, 2008), ordered according to Fountas and Pinnell's (2012) text gradient, is designed to provide texts in a difficulty order based on ten factors that, in combination, are intended to provide early readers with the incremental increases in challenge needed to develop into independent readers.

The texts from the first 10 levels—those intended for K-1—of a decodable program (Juel, Paratore, Simmons, & Vaughn, 2008) formed the decodable sample. The program does not use the Fountas and Pinnell (2012) levels but, rather, uses a numbering system. For purposes of comparison, this study uses letters rather than numbers to designate the levels of the decodable program.

The unit of analysis for the present study was the set of text. The leveled program has a total of 105 books with levels A—D having an average of 13-14 texts per level and 8-9 texts per level for E—J. The decodable program consists of 90 texts with 9 texts at each of 10 levels.

Results

Within-Level Consistency of Text Programs

To understand within-level consistency of texts within the two programs, box-and-whiskers plots are used to illustrate patterns, with the minimum composite percentile, the 25th percentile of the distribution of composite percentiles, the median (50th percentile), the 75th percentile, and the maximum for each of the four composites (decoding, semantic, structure, and syntax) on the horizontal axis. Composites for the texts in all levels were examined as part of this research, but, for purposes of brevity, this report focuses on Levels B and J.

Decodable texts. The box-and-whiskers plot in Figure 1a indicates a low level of decoding challenge for texts at Level B (minimum, 1.3, maximum 12.9). Variation across texts in syntax is similarly not substantial but syntax (minimum, 14.6, maximum, 29.4) is somewhat more demanding than decoding. For structural and semantic features, demands differ as a function of the particular text: semantic—minimum 5.7, maximum 40.7 and structure—minimum 9.7, maximum 57.9. Although texts at Level B consistently present a low level of decoding demands and a relatively low level of syntax demands, the amount of support provided by structure and semantics vary considerably among the texts at this level.

The pattern of the composites represented by Level B is typical for sets of decodable texts at other levels. The overall degree of challenge generally increases with each level. Whatever the level, however, decoding demands continue to be low and the structure demands high. Figure 1b shows the composite profile for Level J, which follows this general pattern. However, in contrast to Level B, decoding, semantic, and syntactic present more variability than structure.

Leveled texts. For the majority of levels, within-level variability for most composites was considerable. As evident in the box-and-whiskers plots in Figure 2a for Level B of the leveled text program, structural demands are low (minimum .1, maximum 5.9) but other features of the texts indicate more challenge and variability: decoding—minimum 3.9, maximum 62.2; semantic—minimum 2.9, maximum 43.3; and syntactic demand—minimum 6.8, maximum 26.7. The demands vary considerably from text to text within levels.

Unlike the decodable program, the pattern of composites across levels changes dramatically in the leveled set. Figure 2b shows that Level J has an opposite profile to Level B: structure presents by far the greatest demand (minimum 68.3, maximum 76.0). Demands for decoding (minimum 34.4, maximum 67.7), semantic (minimum 39.2, maximum 57.0) and syntactic (minimum 40.8, maximum 60.9) are less. The minimum value for structure is higher than the maximum value for the other composites.

Variation across Levels within Text Type

The median percentile for each text level was used to represent changes in the challenge presented by each composite across the text levels.

Decodable texts. The pattern of challenge for the four composites in the decodable set follows a fairly consistent pattern across levels, as shown in Figure 3a. The striking exception is seen in Level A, which demonstrates

Figure 1a
Percentile Distributions: Decodable Text Set B

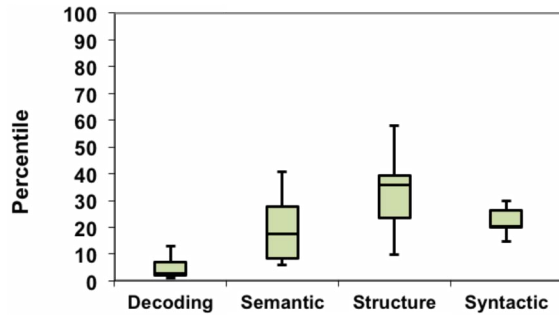


Figure 1b
Percentile Distributions: Decodable Text Set J

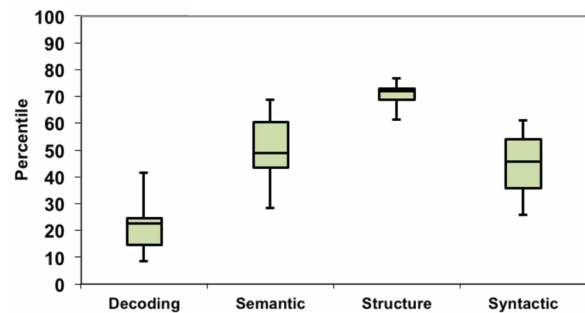


Figure 2a
Percentile Distributions: Leveled Text Set B

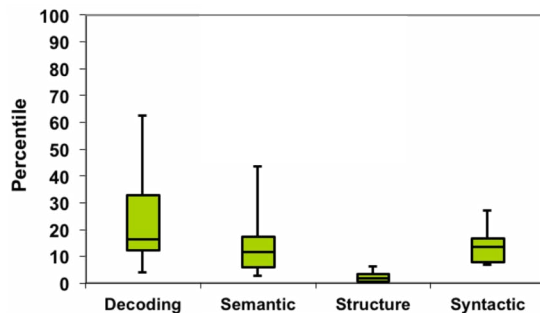
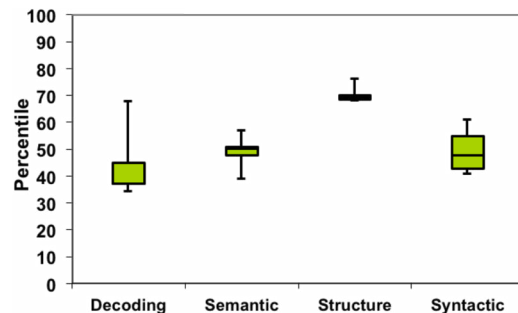


Figure 2b
Percentile Distributions: Leveled Text Set J



a low level of challenge in all four composites. Levels B to J generally show a pattern of highest challenge in the structure composite, with lower challenge in semantic, slightly lower in syntactic challenge, and the lowest challenge in decoding.

Leveled texts. As shown in Figure 3b, the pattern of challenge for the four composites in the leveled set is much less consistent, with many shifts occurring in the first few levels. In Level A, students encounter texts with very low structure demand and relatively high decoding demand, but by Level D, this pattern has reversed: the decoding demand is low and the structure demand is high. From Levels B through E the increase in structure demand is sharp, in contrast to semantic demand, which increases slightly with each level. From Levels F to J, only the semantic demands increase steadily. Decoding, structure, and syntactic demands are fairly consistent across the higher levels with structure demands high. That is, texts do not have the repetitions of episodes or clusters of sentences that characterize earlier levels.

Significance and Contributions

Understanding the opportunities that texts of different types provide for students during the reading acquisition phase is among the most critical endeavors in the school enterprise. To this point, the selection of texts for lessons has largely depended on teachers' expertise. The data from the

6 Text Findings

Figure 3a
Medians of Text Levels: Decodable Texts

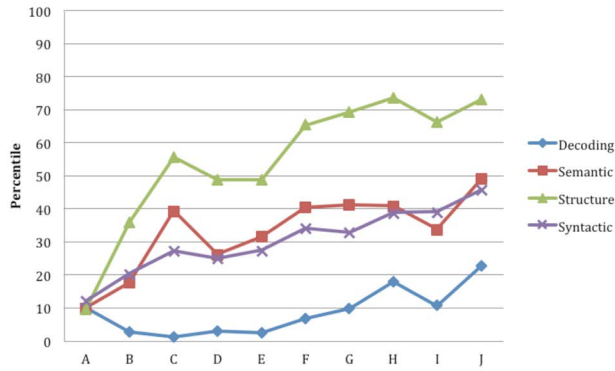
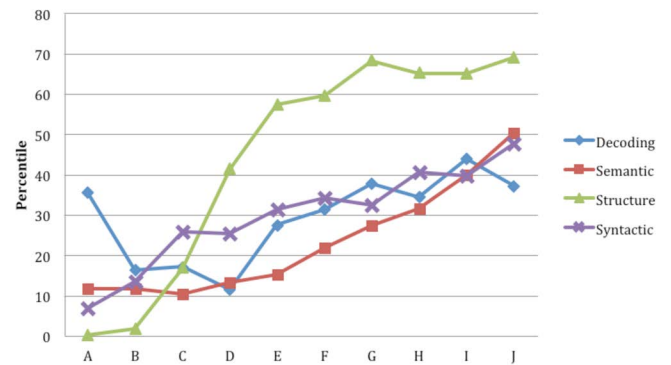


Figure 3b
Medians of Text Levels: Leveled Texts



ELI can provide teachers and program developers with information on how texts may be appropriate for different readers or different purposes in a beginning reading program. With such knowledge, teachers may be able to select from different levels or even programs of texts as they work to foster particular literacy proficiencies and with particular students.

References

- Elmore, J. & Baker, R. (2013, April). *Modeling computer-based text characteristics as potential predictors of text complexity*. Paper presented at the American Educational Research Association Annual Meeting, San Francisco, CA.
- Fitzgerald, J. & Elmore, J. (2013, April). *Text-complexity predictors*. Paper presented at the American Educational Research Association Annual Meeting, San Francisco CA..
- Fountas, I. C., & Pinnell, G. S. (2008). *Leveled literacy intervention*. Portsmouth, NH: Heinemann.
- Fountas, I. C., & Pinnell, G. S. (2012). *The Fountas and Pinnell text level gradient: Revision to recommended grade-level goals*. Portsmouth, NH: Heinemann.
- Hernandez, D.J. (2001). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Baltimore, MD: The Annie E. Casey Foundation.
- Hiebert, E.H., & Pearson, P. D. (2010). *An examination of current text difficulty indices with early reading texts* (Reading Research Report #10-01). Santa Cruz, CA: TextProject. Retrieved from http://textproject.org/assets/publications/TextProject_RRR-10.01_Text-Difficulty-Indices.pdf
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first to fourth grades. *Journal of Educational Psychology*, 88, 3-17. doi: 10.1007/978-1-4612-4282-6
- Juel, C., Paratore, J. R., Simmons, D., & Vaughn, S. (2008). *My sidewalks on reading street*. Glenview, IL: Scott Foresman.
- Mesmer, H. A., Cunningham, J. W., & Hiebert, E. H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading Research Quarterly*, 47, 235-258.