‘There's plenty of opinion that supports guided reading, though few supportive studies’

Kerry Hempenstall.

**Guided reading and Levelled Literacy Intervention*(*LLI)**

**Dr Kerry Hempenstall on the DDOLL Network**

Senior Industry Fellow, School of Education, RMIT University, Melbourne, Australia.

The issue of guided reading and Levelled Literacy Intervention*(*LLI) is complicated because it is also often associated with other practices, such as levelled readers and silent reading. The problem is not whether there is a killer disconfirming finding for guided reading, but rather the dearth of supporting evidence to recommend the approach. Theoretically, it is inconsistent with the consensus among empiricists about what constitutes effective reading instruction. For example, contrary to evidence-based practices, it favours minimally guided instruction, levelled readers, the three cueing system, and silent reading. However, the guided reading approach is favoured over initial explicit phonics by a large number of teachers and teacher organisations in the US and Australia (particularly the LLI model espoused by Fountas and Pinnell). Obviously, many students do manage to learn to read when exposed to guided reading. The question is better framed in terms of overall effectiveness and efficiency differences between the two approaches, the grade(s) at which it is to be used, and the cohorts of students with whom it is to be implemented. Unfortunately, most of the supportive studies I've seen do not compare LLI with other worthwhile programs. Further, the lack of careful attention to detail is disadvantageous to those students at risk of becoming, or already are, low progress readers. I found the Murray et al. analysis particularly instructive.

**Some relevant quotes:**

“The Problem with “Leveled” Books Most striking in our classrooms is the popularity of “leveled” book libraries that are often the main tool for teaching students how to read. Leveled texts originated in New Zealand to accompany Reading Recovery lessons (Clay, 1991). They are the centerpiece of Guided Reading (Fountas & Pinnell, 1999). Leveled texts are assigned a rank (level) on a difficulty scale, such as A–Z, according to four major characteristics judged by a panel of experts: a) book and print features; b) content, themes, and ideas; c) text structure; and d) language and literary elements. These judgments are necessarily subjective, because readability formulas that calculate sentence length and word frequency cannot be meaningfully applied to beginning texts. Several studies have shown that primary grade students in the bottom 40% of reading skill often cannot read these leveled texts because they have not acquired the requisite phonic decoding skills (Cunningham et al., 2005; Foorman, Francis, Davidson, Harm, & Griffin, 2004; Hoffman, Roser, Patterson, Salas, & Pennington, 2001). The leveled texts do not control for phonic patterns and do not follow a scope and sequence of decoding skill instruction, so there are few opportunities for students to apply and solidify phonics skills through cumulative practice. Further, leveled texts do not consistently increase in word-level demands as their levels increase and the need for consistent phonic patterns is ignored when new words are introduced. Phonic patterns are introduced at too fast a rate for beginning readers who need repeated exposures to words to store them in memory. Therefore, when leveled texts are used for instruction, students must rely on (and are often encouraged to rely on) memorization, pictures, and guesswork to read. If phonic correspondences have not been taught, then contextual guessing is the only strategy available to a student who cannot figure out print-to-speech correspondences through incidental exposure and self-teaching. It is well established that guessing from context is the default strategy of students who cannot decode and who read poorly (Gough & Tunmer, 1986; Seidenberg, 2017).” (p. 18)

Moats, L. (2017). Can prevailing approaches to reading instruction accomplish the goals of RTI? *Perspectives on Language and Literacy, 43*(3), 15-22.

“Elementary school students are often placed into groups with peers of similar reading ability in a practice called within-class ability grouping for guided reading instruction. Through this practice, students are differentially exposed to reading skills, strategies, and texts that are presumed to match their current level of ability. This widespread practice is particularly problematic given that (1) current notions of matching early readers to texts for reading instruction are based on traditional instructional practice rather than empirical evidence, (2) poor, minority students are overrepresented in the lowest ranked groups, (3) students in higher ranked groups make greater academic gains than those in lower ranked groups, and (4) teacher perceptions of students’ abilities are often inaccurate. Conversely, several studies have shown that when students are presented with texts of increased difficulty and given appropriate instructional support, they are able to make accelerated reading progress. (p.ii)

Young, T.T. (2019). *Redesigning guided reading instruction: Achieving equity through heterogeneity.* The University of Nebraska - Lincoln, ProQuest Dissertations Publishing, 2019. 13861841. Retrieved from [https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1105&context=teachlearnstudent](https://protect-au.mimecast.com/s/UK5GCq71jxflLE8DhZExIM?domain=digitalcommons.unl.edu)

“School texts, at least those in grades one and three (Gamson, Lu, & Eckert, 2013), appear to have increased in complexity since the 1960s. Therefore, first- through third-grade texts in Bormuth’s sample may not represent current primary-level texts. However, since text complexity beyond the primary grades appears to have remained fairly stable over past decades (Gamson et al.), higher-level texts in the Bormuth criterion may be relatively representative of current texts. … When readability researchers began to use traditional cloze rather than multiple-choice questions as the criterion measure, correlations between predictors and criterion increased (Miller, 1975), usually into the .80–.89 range (Klare, 1974). For these two reasons, in this study we expected each text tool to predict Bormuth’s criterion variable with a correlation of at least .80. This anticipated level was not attained by either the Lexile Framework (r = .70; r2 = .49) or the Flesch–Kincaid (r = - .76; r 2 = .57). These results suggest that the two widely used tools, especially Lexile, may lack adequate validity for their current high-stakes uses in schools.” (p. 828-829)

Cunningham, J.W., Hiebert, E.H., & Mesmer, H.A. (2018). Investigating the validity of two widely used quantitative text tools. *Reading and Writing, 31*(4), 813-833.

“In LLI phonics lessons, the sounds associated with vowel letters are not always explicitly taught. For example, although letter “o” is introduced in Lesson 19 in the word *got*, the sound that letter “o” represents is not explicitly taught. In this lesson, students are instructed to watch the teacher write letters “g,” “o,” and “t” in a sequence of three boxes. The suggested instructional language includes the teacher only saying the sound made by letter “g”—not the sounds of letters “o” and “t.” In these instances, we still considered the vowel to be “taught,” and later words featuring the short “o” sound were counted in LTTMs, provided the other letters and sounds in the words were also taught.

 … The LLI program appears to be in alignment with a meaning-emphasis philosophy. Comparisons at the word and text levels revealed that LLI features more highly frequent words and more multisyllable words, thus confirming our initial hypotheses. This makes sense in a program emphasizing meaning, semantic cues, natural language patterns, predictable syntactic patterns and word repetition, and a philosophical foundation underlying texts promoting “wholebooksuccess” (Martin & Brogan, 1971). …

[There is a] repetition of syntax and words in LLI, which matches the characteristic of texts with a foundation in whole language traditions. …

A particular contribution of the present study is the attention to multisyllable words, a feature of beginning texts that is infrequently addressed in theory and research on intervention for at-risk readers. LLI texts had a relatively high percentage of multisyllable words at beginning levels and lower percentages at the final levels (33%to10%). This shift is explained by early LLI texts’ inclusion of many multisyllable words that are matched to pictures such as *pancakes, rooster*, and *blueberries*. The high percentage of multisyllable words in LLI texts may serve to assist students in the earliest stages of reading (e.g., Ehri’s prealphabetic phase) when they do not yet know alphabet letters or sounds (Ehri, 1998, 2005). This is because visual aspects of the words, such as their shapes, may provide a beneficial scaffold. An alternative explanation is that the pictures serve as the scaffold in supporting beginning readers, rather than the length or appearance of the word. …

It is interesting to note that similar to previous descriptive studies comparing the LTTM of meaning- and code-emphasis first-grade classroom reading programs (Beck & McCaslin, 1978; Chall, 1967/1983; Foorman et al., 2004; Meyer et al., 1987; Stein et al., 1999), our findings show the meaning-emphasis program that included phonics instruction (LLI) featured much less of a systematic match to the words in students’ texts than the code-emphasis program (MS) did. It appears that philosophical foundations continue to exert influence on interventions, much like they influenced the construction and content of first-grade whole-class reading programs in previous decades. …

A diet of LLI may be inadequate in reference to the following: percentages of multisyllable words remain high even after students have presumably moved into the alphabetic stages of reading, which may encourage guessing at words or overreliance on picture cues. A low percentage of phonetically regular words (and LTTM) may discourage the use of decoding strategies, and the high percentage of singletons may cause difficulty for at-risk beginners who cannot “rely on either semantic supports or repetition to identify or retain meanings of words in texts” (Mesmer et al., 2012, p. 241). …

Since LLI texts are not highly decodable, at-risk students may not gain sufficient experience in the application of the lettersound knowledge they are taught, particularly if decodable texts are also not used in their classrooms.” (p.490, 493)

Murray, M. S., Munger, K. A., & Hiebert, E. H. (2014). *An analysis of two reading intervention programs: How do the words, texts, and programs compare? Elementary School Journal, 114*(4), 479-500.

“LLI had no impact on students’ reading comprehension and a negative impact on their mastery of English language arts/literacy standards.” (p.1)

Gonzalez, N., MacIntyre, S., & Beccarvarela, P. (2018). Challenges in adolescent reading intervention: Evidence from a randomized control trial. Working Paper 62. Oakland, CA: Mathematica Policy Research, June 2018. pp. 1-20. Retrieved from [https://www.mathematica-mpr.com/-/media/publications/pdfs/education/2018/adolescent-reading-intervention-wp62.pdf](https://protect-au.mimecast.com/s/HVozCr810kCowk8pszN_BT?domain=mathematica-mpr.com)

“Hanover identified five major studies, all published within the past decade, examining the impact of LLI on student outcomes. The methodological rigor of these studies varies. Some researchers employ randomized student selection, demographic matching, and control groups to determine a causal relationship between LLI participation and student outcomes, whereas other studies can only determine a correlational relationship between LLI participation and student outcomes. Likewise, the assessments used to measure outcomes vary, including assessments such as the Fountas and Pinnell benchmarks, the Dynamic Indicators of Basic Early Learning Skills (DIBELS), and the Gates-MacGinitie Reading Test (GMRT). These differences notwithstanding, all seek to determine the efficacy of LLI in increasing early elementary literacy achievement.

Overall, results of both causal and correlational studies suggest that LLI has a neutral to positive impact on multiple measures of literacy among young, struggling readers. The two most methodologically rigorous, independent studies of LLI found that students in LLI intervention groups made statistically significant improvements on the Fountas and Pinnell benchmarks, DIBELS subtests, and the Developmental Reading Assessment (DRA2) as compared to students in control groups who receive no literacy interventions.25 A less rigorous, independent study of LLI likewise found that students who received 12 to 15 weeks of LLI instruction made statistically significant progress on the GMRT.26 Finally, both of Heinemann’s internal studies found that, for each month students participated in LLI, they made roughly two months of reading progress.27 In a 2013 Heinemann study, the discrepancy between students’ observed and expected reading scores was also significantly lower after participation in LLI.28

Although these findings indicate that LLI has a positive impact on student literacy, they do not provide any evidence regarding the impact of LLI as compared to other interventions. Even the most rigorous studies compared the progress of students participating in LLI against the progress of students not participating in any interventions—not against the progress of students participating in other interventions.” (p.11-12)

Hooper, J (2016). *Literature review: Reading intervention programs*. Hanover Research. Retrieved from [http://gssaweb.org/wp-content/uploads/2016/04/Literature-Review-Reading-Intervention-Programs.pdf](https://protect-au.mimecast.com/s/LkZKCvl1g2SKOY7Gczo8q0?domain=gssaweb.org)

“Fountas and Pinnell advocate for a system of text placement that has been widely and long accepted in the field of reading (I've previously written about the sources of those ideas). F&P add to that a philosophical position that maintains students learn best from figuring things out themselves from reading, rather than from the explicit instruction a teacher might provide. In their plan, much of the teacher’s work is devoted to accomplishing an appropriate placement of students in texts, and they strive to minimize the distance between what a text demands and what students can now do current so that students can scale these small challenges with minimum teacher input. … the book leveling schemes that are being used are pretty dubious. I’m not talking about Lexiles or other well-validated readability schemes, but the book-leveling schemes for guided reading are pretty shaky. … Studies, quoted in previous blogs, show that students can make real learning progress while matched to a variety of text levels, though they tend to do best when matched with more challenging texts than guided reading advocates recommend. Thus, placing students in easy text CAN lead to learning, but placing students in more challenging texts and then making sure they can successfully negotiate them (through rereading, analysis of information, etc.) may lead to even greater success.”

Shanahan, T. (2012). *Common core or guided reading.*Retrieved from [http://www.shanahanonliteracy.com/search/label/Guided%20reading](https://protect-au.mimecast.com/s/Eg4aCyoj8Puk68rgcZAW-cN?domain=shanahanonliteracy.com)

“ … one of the most widely implemented approaches to early reading instruction in the primary grades in schools across the United States is Guided Reading (GR; Fountas & Pinnel, 1996). This approach deemphasizes decontextualized instruction and practice of reading-related skills in favor of extended time spent reading text under the guidance of a teacher who supports the development of effective reading strategies (Fountas & Pinnell, 2012–13). Despite its widespread implementation (Ford & Opitz, 2008), **GR has rarely been empirically validated**. … GR teachers teach and prompt students to use reading strategies that involve three sources of text information: meaning cues from background knowledge and text context (including cues from illustrations), cues derived from students’ understanding of English syntax, and visual information derived from print, including sound–symbol relationships and sound-spellings associated with larger orthographic units such as onsets and rimes. As described by Fountas and Pinnel, word study instruction is primarily embedded in text reading and does not follow a predetermined scope and sequence.” (p. 269-70)

“Explicit instruction approaches most often emphasize synthetic phonics instruction (i.e., teaching individual sound-spelling correspondences and encouraging children to “sound out” words), although any approach to word identification could be taught explicitly. Published early reading programs that incorporate highly explicit instruction often include application of skills in text that is decodable using sound-spelling patterns and intact words that have been previously taught. … Thus, the ability to decode unknown words using sound–symbol correspondences appears to be best supported through explicit instruction. This is not trivial and may become increasingly important as students encounter more challenging multisyllable words in the intermediate grades. … EX [explicit instruction] is more likely to substantially accelerate student progress in phonemic decoding, text reading fluency, and reading comprehension than GR [guided reading].” (p.268, 285, 269)

“We had expected that phonemic decoding and passage reading fluency would be better supported by the explicit phonics instruction and structured repeated reading practice students received in the EX group, but the finding that EX was associated with **stronger effects on comprehension** was unexpected, as GR [guided reading] students spent more time engaged in text reading and discussion of text.” (p.287)

Denton, C.A., Fletcher, J.M., Taylor, W.P., Barth, A.E., & Vaughn, S. (2014). An experimental evaluation of guided reading and explicit interventions for primary-grade students at-risk for reading difficulties*. Journal of Research on Educational Effectiveness*,*7*(3), 268-293. Retrieved from [http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4712689/pdf/nihms723846.pdf](https://protect-au.mimecast.com/s/RtkuCxngGkfwm71yuRZiL-?domain=ncbi.nlm.nih.gov)

“Much of the research and practice of the past 20 years has dealt with how to support students with instructional level reading, consistent with Betts’ (1946) view, described earlier. Researchers have addressed instructional level reading as part of intervention (e.g., Schwartz 2005) as well as more general classroom instruction (e.g., Iaquinta 2006), but there is little research, and certainly no consensus, on the best ways to support students in reading more challenging texts (i.e., frustration level, according to Betts 1946). At the same time, there is little evidence to suggest that Betts’ guidelines––or various adaptations of Betts, like those used by Fountas and Pinnell (1996), Leslie and Caldwell (2011), or Morris (2008)—hold any standing (e.g., Cunningham 2013). While there is a long history of employing thresholds for accuracy in considering text difficulty, no definitive word recognition percentage exists to guide matching readers with texts. Instead, research on reading instruction should likely consider how aspects of the text, (such as its structure, cohesion, or narrativity) might interact with the reader’s word recognition and comprehension skills.” (p.145)

Amendum, S.J., & Conradi, K., & Hiebert, E. (2018). Does text complexity matter in the elementary grades? A research synthesis of text difficulty and elementary students’ reading fluency and comprehension. *Educational Psychology Review,*30(1), 121–151.

“Research has not been kind to the idea of mechanical “instructional level” criteria like 90-95% accuracy (e.g., Jorgenson, Klein, & Kumar, 1977; Kuhn, Schwanenflugel, Morris, Morrow et al., 2006; Morgan, Wilcox, & Eldredge, 2000; O’Connor, Swanson, & Geraghty, 2010; Powell, & Dunkeld, 1971; Stahl, & Heubach, 2005; Stanley, 1986).”

Shanahan, T. (2016). *Further explanation of teaching students with challenging text.* Retrieved from [http://www.shanahanonliteracy.com/search/label/text%20complexity](https://protect-au.mimecast.com/s/2zeYCzvkmpf98kM3I4BviRT?domain=shanahanonliteracy.com)

“Several teachers also stated that there were inconsistencies between materials (e.g., the lesson did not “match” the written materials), that the system was too fast-paced for at-risk students, and that the amount of information presented in each lesson could be overwhelming to students. Several teachers also recommended improving the lesson sequence or “flow” and the word work, which they stated should be consistent and derive directly from the text. Some teachers also disagreed with the examples used for certain concepts in the curriculum, such as using “ear” for the long “e” sound, presenting “bread” and “read” at the same time (i.e., one is an example of the “ea” rule while one is an exception), and including words like “moon” and “spoon” in the same word ladder as words like “book.” Further, some teachers suggested providing more skill review so that mastery can be obtained before progressing to the next skill, providing a more specific vocabulary introduction for each book, and reducing the amount of “runoff” paper that is wasted when printing out CD resources.” (p.50)

Ransford-Kaldon, C.R., Flynt, E.S., Ross, C.L., Franceschini, L., Zoblotsky, T., Huang, Y., Gallagher, B.. (2010). *An empirical study to evaluate the efficacy of Fountas & Pinnell’s Leveled Literacy Intervention System (LLI) (2009-2010).*Center for Research in Educational Policy (CREP).

“The results of this study imply that the implementation of Fountas and Pinnell Leveled Literacy Intervention does not significantly affect the growth of reading achievement by primary students who are identified as Targeted Title One students. While growth was achieved by the majority of the students in the treatment groups, those not in the treatment groups achieved similar growth whether examined by grade level, by race (except for White), or by special service program such as Special Education and ELL.” (p.18)

Waldera, J. (2017).  Study of the efficacy of Leveled Literacy Intervention on meeting reading achievement goals for targeted Title One Students in Grades 1 and 2. Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Education, May 2017. Graduate Programs in Education, Goucher College. Retrieved from [https://mdsoar.org/bitstream/handle/11603/3909/Final%20Action%20Research%20Paper.docx?sequence=1&isAllowed=y](https://protect-au.mimecast.com/s/l7jlCANpnDC2rwNghGORbC6?domain=mdsoar.org)

“A persistent fear among classroom teachers is that some students may not keep their eyes on their text when they are assigned silent independent reading tasks (Donovan, Smolkin, & Lomax, 2000; Fresch, 1995; Hiebert, Wilson, & Trainin, 2010). Guidance within silent reading contexts is key, as students achieving in the bottom quartile of their class frequently attend less well when they read silently in an unguided context as compared to a guided context (Hiebert et al., 2010). … Although previous studies have shown silent reading to be an effective way to improve reading skills, more recent studies have shown that the conditions for silent reading practice in school often result in students acting like they are reading when they are not” (Hiebert & Reutzel, 2010). (p. 123-4)

Hiebert, E.H. (2015). *Teaching stamina & silent reading in the digital-global age.* Reading Essentials Original Series, June 2015. Retrieved from [http://textproject.org/assets/library/resources/Hiebert-2015-Teaching-Stamina-and-Silent-Reading-PRINT.pdf#page=124](https://protect-au.mimecast.com/s/kP8HCANpnDC2rwNghO041z?domain=textproject.org)

“ … the [National Reading Panel (2000)](https://protect-au.mimecast.com/s/3JZeCBNqgBCBPw71U73_4S?domain=ncbi.nlm.nih.gov) report sparked considerable controversy when the panel reported a lack of research supporting independent, silent reading practice as an effective means for developing students’ reading fluency (e.g., Silent Sustained Reading [SSR] or Drop Everything and Read [DEAR]; Allington, 2002; Coles, 2000; J. W. Cunningham, 2001; Edmondson & Shannon, 2002; Krashen, 2002).” (p. 404)

Reutzel, D.R., Spichtig, A.N. & Petscher, Y. (2012). Exploring the value added of a guided, silent reading intervention: Effects on struggling third-grade readers’ achievement. *Journal of Educational Research, 105*(6), 404–415. Retrieved from [http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4557881/](https://protect-au.mimecast.com/s/Gd_LCD1vRkC0jE51snJmUp?domain=ncbi.nlm.nih.gov/)

"The results of this study suggest that children can benefit from reading material well below the 95% accuracy rate traditionally recommended for instruction. In fact, students appeared to benefit from reading stories in the first sampling even though they were reading them with an average accuracy rate of 85%, which would be considered frustration level. Students were able to benefit from reading material at these lower levels of accuracy because of the higher support they were given for the reading through the routines of the program. In this program, students were supported in their reading by having multiple exposures to the same material, by having stories read to them, by exposure to the vocabulary prior to their own reading, by reading the story at home one or more times, possibly by echo reading, and by partner reading…We argue that the instructional reading level for a given child is inversely related to the degree of support given to the reader. That is, the more support given, the lower the accuracy level needed for a child to benefit from instruction. In classroom organizations such as our fluency-oriented instruction, students benefited from reading material of greater relative difficulty because they were given greater amounts of support for that reading."

*Stahl,*S., &*Heubach,* K. M. (*2005*)*.* Fluency-oriented reading instruction. *Journal of Literacy Research, 37*, 25-60.

“The What Works Clearinghouse (WWC) identified**two studies** of LLI that fall within the scope of the Beginning Reading topic area and meet WWC group design standards. Two studies meet WWC group design standards without reservations, and no studies meet WWC group design standards with reservations. Together, these studies included 747 students in grades K–2 in 22 schools in three school districts across three states.

According to the WWC review, the extent of evidence for LLI on the reading achievement outcomes of beginning readers was medium to large for general reading achievement and small for two other student outcome domains—reading fluency and alphabetics. No studies meet WWC group design standards in one other domain, so this intervention report does not report on the effectiveness of LLI for that domain.3” (p.1)

Institute of Education Sciences. (2017). WWC Intervention Report: Leveled Literacy Intervention. US Department of Education. Retrieved from [https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc\_leveledliteracy\_091917.pdf](https://protect-au.mimecast.com/s/iiejCE8wlRCM1v3msK-OIE?domain=ies.ed.gov)

“Thus, the ability to decode unknown words using sound–symbol correspondences appears to be best supported through explicit instruction. This is not trivial and may become increasingly important as students encounter more challenging multisyllable words in the intermediate grades.) (p. 19)

Denton, C.A., Fletcher, J.M., Taylor, W.P., Barth, A.E., & Vaughn, S. (2014). An experimental evaluation of guided reading and explicit interventions for primary-grade students at-risk for reading difficulties*. Journal of Research on Educational Effectiveness*,*7*(3), 268-293.

“Over the past decade in the United States, levelled texts, or ‘little books’ with finely graduated levelling of text difficulty, have regained their status as a literacy staple for beginning readers. Despite their resurgence, questions remain regarding the reliability and quality of these books. In this study, we conducted a detailed analysis of 20 sample texts at levels 5, 10, 15 and 20, from one series of levelled books published in the United States. We found that the levelling system used in this series is not a particularly reliable indicator of text difficulty and that the quality of these books varies considerably between and within levels. We suggest that close attention to text levels could be detrimental in the reader–text matching process. We further recommend that in order to judge and select quality books that answer the needs of their students, teachers become more aware of the various factors influencing text comprehensibility and quality.” (p. 43)

Picher, B., & Fang, Z. (2007). Can we trust levelled texts? An examination of their reliability and quality from a linguistic perspective. *Literacy, 41*(1), 43-51.

“Hoffman, Roser, and Salas (2001) found that teachers using the Fountas and Pennell leveling structure can reliably level text. However, when texts leveled in that manner assess student reading performance, they produce highly unreliable results. Running record scores that are acquired from a single-leveled text reading would not necessarily represent a student’s true reading level. … Making absolute decisions with a running record requires the teacher to average student scores on at least three passages with at least two raters. Our results indicate that the most limiting factor in rendering students’ running record scores reliable is the number of passages used. That finding supports the contention of Ross (2004) that passage might exhibit a sizable source of error variance when scoring running records. Using a single score obtained from reading a single passage to portray that student’s universe score would be highly questionable.” (p. 123) … Teachers should recognize that traditional text-leveling procedures do not fully account for all factors that affect the difficulty of a text. Even when controlling for text level, type, and structure, there are still naturally occurring topical variations between texts that will render one more difficult than another.” (p.125)

Fawson, P.C., Ludlow, B.C., Reutzel, D.R., Sudweeks, R., & Smith, J.A. (2006). Examining the reliability of running records: Attaining generalizable result. *The Journal of Educational Research, 100*(2), 113-126.

“This article has raised a number of concerns about the use of running records. These concerns include: (1) a lack of clarity in the guidelines about whether running records are appropriate for beginning and fluent readers, (2) problems with the comparability of running records taken on different texts, and the lack of assessment of comprehension, (3) the absence of evidence to support the use of self-corrections as an indicator of effective reading, and (4) erroneous interpretation of the meaning of oral reading errors.” (p. 16-17).  … “The problems associated with the analysis of errors in running records mean that this practice cannot be recommended. … It is time to reappraise the widespread use of running records as the main assessment of children’s reading in the first years at school. There is a need for other assessments that link more directly to what it is that children require for fluent reading. Teachers need to have access to carefully constructed tests of reading accuracy and comprehension, as well as measures of grapheme-phoneme correspondences, phonological awareness, and word recognition. The use of such tests will provide teachers with crucial information about what children currently know and what they still need to learn to become successful independent readers.” (p.18) … “ … following these [running records] guidelines provides a misleading picture of a child’s reading. A child who has difficulty decoding some words may still be assessed as making effective use of visual information if there is any visual connection between the child’s errors and the correct words” (p. 248).

Blaiklock, K. (2004). A critique of running records of children’s oral reading. *New Zealand Journal of Educational Studies, 39,* 241-252. Retrieved from [http://unitec.researchbank.ac.nz/bitstream/handle/10652/1428/Blaiklock%2C%202004.%20A%20critique%20of%20running%20records%20of%20children%27s%20oral%20reading.pdf?sequence=3&isAllowed=y](https://protect-au.mimecast.com/s/yYyvCBNqgBCBPw71Uz7KBgI?domain=unitec.researchbank.ac.nz).

[http://www.shanahanonliteracy.com/2011/08/rejecting-instructional-level-theory.html#uds-search-results](https://protect-au.mimecast.com/s/EQPGCJyBZ6t5p6qRh86yyn?domain=shanahanonliteracy.com)

[**Shanahan on Literacy**](https://protect-au.mimecast.com/s/GlTBCK1DOrCm8X2lFX6qMe?domain=shanahanonliteracy.com/)

Information for teachers and parents on teaching and assessing reading, writing, and literacy.

**Sunday, August 21, 2011**

**Rejecting Instructional Level Theory**

A third bit of evidence in the complex text issue has to do with the strength of evidence on the other side of the ledger. In my two previous posts, I have indicated why the common core is embracing the idea of teaching reading with much more complex texts. But what about the evidence that counters this approach?  
  
Many years ago, when I was a primary grade teacher, I was struggling to teach reading. I knew I was supposed to have groups for different levels of kids, but in those days information about how to make those grouping decisions was not imparted to mere undergraduates. I knew I was supposed to figure out which books would provide the optimal learning experience, but I had no technology to do this.  
  
So, I enrolled in a master’s degree program and started studying to be a reading specialist. During that training I learned how to administer informal reading inventories (IRI) and cloze tests and what the criteria were for independent, instructional, and frustration levels. Consequently, I tested all my students, and matched books to IRI levels using the publisher’s readability levels. I had no doubt that it improved my teaching and students’ learning.  
  
I maintained my interest in this issue when I went off for my doctorate. I worked with Jack Pikulski. Jack had written about informal reading inventories (he’d studied with Johnson and Kress), and as a clinical psychologist he was interested in the validity of these measures. He even sent a bunch of grad students to an elementary school to test a bunch of kids, but nothing ever came of that study. Nevertheless, I learned a lot from Jack about that issue.  
  
He had (has) a great clinical sense and he was skeptical of my faith in the value of those instructional level results. He recognized that informal reading inventories were far from perfect instruments and that at best they had general accuracy. They might be able to specify a wide range of materials for a student (say from grade 2 to 4), but that they couldn’t do better than that. (Further complicating things were the readability estimates. These had about the same level of accuracy.)  
  
For Jack, the combination of two such rough guestimates was very iffy stuff. I liked the certainty of it though and clung to that for a while (until my own clinical sense grew more sophisticated).  
Early in my scholarly career, I tracked down the source of the idea of independent, instructional, and frustration levels. It came from Emmett Betts’ textbook. He attributed the scheme to a study conducted by one of his doctoral students. I tracked down that dissertation and to my dismay it was evident that they had just made up those designations without any empirical evidence, something I wrote about 30 years ago!  
  
Since then, readability measures have improved quite a bit, but our technologies for setting reading levels have not. Studies by William Powell in the 1960s, 70s, and 80s showed that the data that we were using did not result in an identification of optimum levels of student learning. He suggested more liberal placement criteria, particularly for younger students. More liberal criteria would mean that instead of accepting 95% word reading accuracy as Betts had suggested, Powell identified 85% as the better predictor of learning—which would mean putting kids in relatively more difficult books.  
  
Consequently, I have sought studies that would support the original contention that we could facilitate student learning by placing kids in the right levels of text. Of course, guided reading and leveled books are so widely used it would make sense that there would be lots of evidence as to their efficacy.  
  
Except that there is not. I keep looking and I keep finding studies that suggest that kids can learn from text written at very different levels (like the studies cited below by Morgan and O’Connor).  
  
How can that be? Well, basically we have put way too much confidence in an unproven theory. The model of learning underlying that theory is too simplistic. Learning to read is an interaction between a learner, a text, and a teacher. Instructional level theory posits that the text difficulty level relative to the student reading level is the important factor in learning. But that ignores the guidance, support, and scaffolding provided by the teacher.  
  
If the teacher is doing little to support the students’ transactions with text then I suspect more learning will accrue with somewhat easier texts. However, if reasonable levels of instructional support are available then students are likely to thrive when working with harder texts.  
  
The problem with guided reading and similar schemes is that they are focused on helping kids to learn with minimal amounts of teaching (something Pinnell and Fountas have stated explicitly in at least some editions of their textbooks). But that switches the criterion. Instead of trying to get kids to optimum levels, that is the levels that would allow them to learn most, they have striven to get kids to levels where they will likely learn best with minimal teacher support.  
  
The common core standards push back against the notion that students learn best when they receive the least teaching. The standards people want to know what it takes for kids to learn most, even if the teacher has to be deeply involved. For them, challenging text is the right ground to maximize learning… but the only way that will work is if kids are getting substantial teaching support in the context of that hard text.  
  
P.S. Although Lexiles have greatly improved readability assessment (shrinking standard errors of measurement and improving the amount of comprehension variance that can be explained by text difficulty), and yet we are in no better shape than before since there are no studies indicating that if you teach students at particular Lexile levels more learning will accrue. (I suspect that if future studies go down this road, they will still find that the answer to that issue is variable; it will depend on the amount and quality of instructional support).”  
  
Betts, E. A. (1946). *Foundations of reading instruction.* New York: American Book Company.  
  
Morgan, A., Wilcox, B. R., & Eldredge, J. L. (2000). Effect of difficulty levels on second-grade delayed readers using dyad reading. *Journal of Educational Research, 94,* 113–119.  
  
O’Connor, R. E., Swanson, H. L., & Geraghty, C. (2010). Improvement in reading rate under independent and difficult text levels: Influences on word and comprehension skills.*Journal of Educational Psychology, 102,* 1–19.  
  
Pinnell, G. S., & Fountas, I. C. (1996). *Guided reading: Good first teaching for all children.*Portsmouth, NH: Heinemann.  
  
Powell, W. R. (1968). *Reappraising the criteria for interpreting informal inventories.* Washington, DC: ERIC 5194164.  
  
Shanahan, T. (1983). The informal reading inventory and the instructional level: The study that never took place. In L. Gentile, M. L. Kamil, & J. Blanchard (Eds.),*Reading research revisited,* (pp. 577–580). Columbus, OH: Merrill.

“Another panelist—Timothy Shanahan, an emeritus professor at the University of Illinois and the author or editor of over 200 publications on literacy—went on to debunk [a popular approach](https://protect-au.mimecast.com/s/19oNCL7Eg9f4NGRAixQf1y?domain=rand.org) that goes hand in hand with teaching comprehension skills: To help students practice their “skills,” teachers give them texts at their supposed individual reading levels rather than the level of the grade they’re in.

According to Shanahan, [no evidence](https://protect-au.mimecast.com/s/SqeACMwGj8CWxwqgF7tBiV?domain=shanahanonliteracy.com) backs up that practice. In fact, Shanahan said, recent research [indicates](https://protect-au.mimecast.com/s/r_f2CNLJxkiXVO0ASP1thC?domain=ernweb.com/) that students actually learn more from reading texts [that are considered too difficult for them](https://protect-au.mimecast.com/s/zNW3COMK7Yc457pJilx2Z-?domain=shanahanonliteracy.com)—in other words, those with more than a handful of words and concepts a student doesn't understand. What struggling students need is guidance from a teacher in how to make sense of texts designed for kids at their respective grade levels—the kinds of texts those kids may otherwise see only on standardized tests, when they have to grapple with them on their own.

That view was endorsed by Marilyn Jager Adams, a cognitive and developmental psychologist who is a visiting scholar at Brown University. “Giving children easier texts when they’re weaker readers,” she said during the panel discussion, “serves to deny them the very language and information they need to catch up and move on.”

Wexler, N. (2018). *Why American students haven't gotten better at reading in 20 years*. Retrieved from [https://www.theatlantic.com/education/archive/2018/04/-american-students-reading/557915/?utm\_source=twb](https://protect-au.mimecast.com/s/vpA3CP7L1NfMvjKOspgXpK?domain=theatlantic.com)

[https://edexcellence.net/articles/leveled-reading-the-making-of-a-literacy-myth](https://protect-au.mimecast.com/s/b-ewCQnM1WfrB2k8u62KTC?domain=edexcellence.net)

Leveled reading: The making of a literacy myth

Robert Pondiscio September 24, 2014

“Among opponents of the Common Core, one of the more popular targets of vitriol is the standards’ focus on improving literacy by introducing higher levels of textual complexity into the instructional mix. The move to challenge students with more knotty, grade-level reading material represents a shift away from decades of general adherence to so-called “instructional level theory,” which encourages children to read texts pitched at or slightly above the student’s individual reading level. New York public school principal Carol Burris, an outspoken standards critic and defender of leveled reading, recently published an anti-Common Core missive on the Washington Post’s Answer Sheet blog that was fairly typical of the form. Where, she wondered, “is the research to support: close reading, increased Lexile levels, the use of informational texts, and other questionable practices in the primary grades?”

The blog post, which has already been intelligently critiqued by Ann Whalen at Education Post, expanded on remarks delivered by Burris earlier this month at an Intelligence Squared U.S. debate with Fordham president Michael Petrilli and former assistant secretary of education Carmel Martin. There, too, she demanded evidence of literacy improvements arising from the use of complex texts.

A fair request and one that warrants a thorough response. But first, for the benefit of readers who are neither teachers nor literacy specialists, a quick explainer on how these two theories of reading work: In leveled reading, a teacher listens as her student reads a piece of text at a given reading level. If the child makes two-to-five mistakes per one hundred words, that is considered her “instructional” level. Zero or one mistakes means the book is too easy; six or more mistakes and that level is deemed her “frustration” level. Children are then offered lots of books at their “just right” level on the theory that if they read extensively and independently, language growth and reading proficiency will follow, setting the child on a slow and steady climb through higher reading levels. It sounds logical, and, as we will see, there are definite benefits to getting kids to read a lot independently.

By marked contrast, Common Core asks teachers to think carefully about what children read and choose grade-level texts that use sophisticated language or make significant knowledge demands of the reader (teachers should also be prepared, of course, to offer students support as they grapple with challenging books). Instead of asking, “Can the child read this?” the question might be, “Is this worth reading?”

Leveled reading is intuitive and smartly packaged (who wants kids to read “frustration level” books?), but its evidence base is remarkably thin. There is much stronger research support for teaching reading with complex texts.

What’s the source of the blind faith that Burris and others have in leveled reading instruction? “In the decades before Common Core, an enormous amount of the instruction in American elementary and middle schools has been with leveled text,” says David Liben, a veteran teacher and Senior Content Specialist at Student Achievement Partners. “The generally poor performance of our children on international comparisons speaks volumes about its effectiveness. To become proficient, students need to have the opportunity to read, with necessary support, rich complex text. But they also need to read—especially if they are behind—a huge volume and range of text types just as called for in the standards.” Students could read many of these less complex texts independently. “Instruction with complex text at all times is not what is called for, even by Common Core advocates,” Liben takes care to note.

Burris and others, however, offer a reflexive defense of leveled instruction. At the Intelligence Squared event, she claimed that “We know from years of developmental reading research that kids do best when they read independently with leveled readers.” Such surety is belied by a surprising lack of rigorous evidence. Literacy blogger Timothy Shanahan, a Distinguished Professor Emeritus of urban education at the University of Illinois at Chicago, recently detailed his discovery of the inauspicious origins of instructional level theory as a young scholar.

Made famous in Emmett Betts’s influential, now-little-remembered 1946 textbook Foundations of Reading Instruction, leveled reading theory actually emerged from a more obscure study conducted by one of Betts’s doctoral students. “I tracked down that dissertation and to my dismay it was evident that they had just made up those designations without any empirical evidence,” Shanahan wrote. When the study—which had in effect never been conducted—was “replicated,” it yielded wildly different results. In other words, there was no study, and later research failed to show the benefits of leveling. “Basically we have put way too much confidence in an unproven theory,” Shanahan concluded.

Experts have spent much of the last four decades unraveling elements of Betts’s thesis, as Douglas Fisher and Nancy Frey recently demonstrated in The Reading Teacher, a popular journal. The authors, who work closely with the International Reading Association (IRA), were longtime advocates of leveled reading. Reexamining the published research in light of the new standards, however, they found that the use of leveled text beyond the very first years of primary school yielded no achievement gains in students. The belief that young readers should only be taught from texts that they understood to a level of 95 percent or higher—a stringent notion of comprehension first envisioned by Betts—has been found to be erroneous. Researchers William R. Powell and C.G. Dunkeld, as early as 1971, said that the 95 percent–cutoff was too high; and, more recently, academic Juliet Halladay condemned it as “somewhat arbitrary.”

Even more striking to Fisher and Frey was the abundance of support for the use of more difficult reading material: “Surprisingly, we did find studies suggesting that students learn more when taught with texts that were above their instructional level.” One such prominent study, though unheralded in their review, was that of the Science IDEAS model put forward by researchers Michael Vitale and Nancy Romance. The program, which replaces eight weeks of English Language Arts lessons with a regimen of complex science instruction for a group of third- to fifth-graders, was shown to not only enhance scientific aptitude among the group, but also accelerate reading comprehension through the use of complex science texts.

Another trial, organized by specialists at Brigham Young University, divided a swath of struggling students into three groups of “paired readers,” each furnished with texts of a set difficulty level. Paired reading, a method by which two pupils read aloud together, has proven broadly successful in generating literacy gains among children; indeed, all three groups improved through the use of the paired system. But the greatest advance was made by the group using text that was two years above its instructional level. Burris has dismissed paired reading and the study as “idiosyncratic”; her meaning here is obscure, but she might have more simply described it as a proven, effective, and inexpensive way of helping children learn to read.

In addition to these studies, Shanahan, in the IRA journal Reading Today Online, lists twenty studies showing the efficacy of instruction with more complex text. Thus we have a significant and growing body of research providing support for this initiative.

To be emphatically clear, none of this is data should be taken to advocate for a total phasing-out of texts students can read independently, many of which would be at lower levels of complexity. “Nowhere in the Common Core standards,” Liben concurs, “or in the work of these experts is it recommended that we abandon this practice. This is why the Core standards call for all students to read ‘widely and deeply.’ Not doing so would make it impossible to grow the vocabulary and knowledge essential to success.”

Russ Walsh, a teacher and curriculum director, making the case for leveled instruction in another Answer Sheet post, finally concedes that the best approach “is to balance our instruction between independent level, on-level, and frustration level texts.” On this we agree. But before Common Core, such balance was far less likely, too often denying our most needy students the opportunity to read, enjoy, and benefit from a full range of rich texts. As Alfred Tatum noted in the Fisher article cited above, “Levelled texts lead to levelled lives.”