Commentary

Why Fast and Effortless Reading Is Indispensable for Comprehension

HELEN ABADZI AND TRACY CENTANNI

Dowd and Bartlett (2019) tested the reading competencies of second graders attending Save the Children programs in 11 countries. They used sixty-word passages and comprehension questions without time limits and analyzed data of students reading more than 10 words per minute. They found wide disparities in speed among those who showed high comprehension rates between-countries and within-country. They interpreted the findings as diminishing the role of speed in early-grade assessment and instruction. However, the study has major issues on three fronts: (*a*) use of English-language studies for transparent orthographies, (*b*) disregard of current neurocognitive research on memory functions, and (*c*) sample selectivity. These raise grave concerns regarding the internal and external validity of the conclusions. The authors' dedication is admirable, but their findings do not justify any policy recommendations. To the contrary, given the importance of speed for comprehension, their recommendations may have the perverse effect of disadvantaging the very students that Save the Children is supporting. Researchers and journals have the ethical obligation to publish studies that reflect contemporary reading research.

Background

In May 2019, an article was published in *Comparative Education Review* that questioned the need for reading speed (Dowd and Bartlett 2019). In earlier articles the authors and a few others showed that students from various languages and scripts attain comprehension at different speed levels (Dowd 2011; Graham and van Ginkel 2014; Bartlett et al. 2015; Jukes and Ringole 2016). Dowd and Bartlett have interpreted these disparities as an indication that speed is unreliable as a reading benchmark. They call on governments and donors to de-emphasize this indicator.

The argument reflects a common-sense dilemma confronting educators: Why should schools demand speed? Thoughtful readers are often slow; they carefully consider a text rather than read like parrots. But neuroscientists have looked carefully at what goes on in the brain when we read. Their findings show that common sense may prolong reading instruction and exacerbate difficulties.

Comparative Education Review, vol. 64, no. 2.

0010-4086/2020/6402-0007$10.00

Some Highlights of Reading Neuroscience

Reading originates as a perceptual learning function. Humans have the ability to discriminate between individual shapes and then create larger and meaningful shapes from them. With repeated exposures the shapes are grouped and processed in the brain simultaneously. The ability to quickly recognize visual objects based on their meaning is critical for the survival of many species and likely underlies humans' specialized ability to learn to read print. For example, with practice, medical staff learn to quickly recognize blurry shadows on a scan as tumor tissue. Early in reading instruction, letters are also processed as objects. With practice and instruction, the brain learns to associate these visual objects with language and print, facilitating fluency (Gori and Facoetti 2014). Practice is indispensable for grouping of shapes into letters and words (see Abadzi [2017] for a review). This fundamental and rather mechanical reading stage is rarely discussed by reading theorists.

The letter shapes are processed in certain neuronal pathways that are similar across languages and scripts (Perfetti et al. 2013; Krafnick et al. 2016; Yang et al. 2017). The pathways originate from the visual cortex and move forward, linking sounds and subsequently linguistic processes (Gori and Facoetti 2014). Neuroimaging studies suggest the first 170 milliseconds of the process are visual, and the location of this signal can be used to determine how automated the letter processing is. Linguistic information and comprehension are added about half a second later (Czigler 2010; Eberhard-Moscicka et al. 2016.). The evidence points to a hierarchical, cascaded, interactive model of word recognition (Dufau et al. 2015). Readers get almost instant feedback about sounds and meaning through recurrent loops. Thus, reading involves closely timed sequences, where performance at each stage must be optimized to give reliable and timely input to the next. The visual areas and the face recognition area are involved in the early stages of reading print, while comprehension lies at the end of the path; readers must undergo specialization in these early areas in order to achieve fluency (Younger et al 2017). The faster the readers traverse the early stages, the more easily they access knowledge networks in long-term memory.

The visual and processing regions of the brain are configured to give decisions in milliseconds because ancient organisms had to react instantly to emergencies. We have therefore inherited a mechanism that processes very few essential items for a few seconds: our working memory.[1] All conscious information goes through it, so speed is indispensable in overcoming its limitations. If humans only sent single letters to the working memory, voluminous reading would be impossible. Slow, letter-by-letter readers cannot lift a message off the

[1] The working memory is a temporary storage of information that contains what is in your mind right now (Alloway and Alloway 2013).

page; by the time they get to the end of a sentence, they forget the beginning. A minimum speed of 45–60 words per minute is necessary to make sense of text.

Fortunately, another brain area helps us attain fluency and comprehension. It is the formation and specialization of the visual word form area (VWFA). A certain part of this region preferentially responds to print stimuli over other nonprint objects (Dehaene and Cohen 2011). Many years of practice are required to integrate this region into the reading network and achieve fluent reading (Cantlon et al. 2011; Centanni et al. 2017; Younger et al. 2017). The degree of specialization in this region is correlated with reading ability in young children, signifying the importance of the VWFA for reading (Centanni et al. 2018). When this system does not develop the necessary speed, reading acquisition fails as in dyslexia (van der Mark et al. 2009). Largely due to a lack of practice, many children in low-income countries also show halting reading and poor comprehension.

Fluency matters a lot. Innumerable behavioral studies highlight speed as a prerequisite for comprehension (Laberge and Samuels 1974; Daneman and Carpenter 1980; Pikulski and Chard 2005). Small losses in word recognition lead to substantial comprehension reduction, because the mind searches for meaningful solutions. Oral reading fluency measures correlate also with the comprehension of more complex texts (correlation 0.91; Fuchs et al. 1988). The relationship between speed and comprehension is stronger in consistently spelled languages and in the lower levels of education (RTI 2010).

Research therefore shows that speed is critical early in instruction and that comprehension is a by-product. When people become fluent and automatic readers, they spend little time on visual detection. The brain constantly integrates the print in the current visual window with the information that has already been consumed.

Reading comprehension is a multidimensional concept that also involves speed. It is often defined as the interaction between written words and knowledge triggered outside the text (Rayner et al. 2001). It involves the construction of a coherent mental representation of the text in readers' memory, from which questions of various types are answered (van den Broek et al. 2005; McNamara and Magliano 2009; Kendeou et al. 2012). Reading should activate the same brain regions involved in speech comprehension (Devlin 2010). A large, randomized control trial in the United States found that word reading skills and oral language skills both contributed to comprehension, with each contributing unique variance (Language and Reading Research Consortium 2018).

The comprehension needs differ somewhat between languages with complex orthographies, such as English and Khmer, and those with consistent (or transparent) orthographies, such as Spanish and Hindi. In English, students must know many of the words read in order to pronounce them, while most other languages do not have this requirement. Predictably English reading is harder to automatize, so the vast majority of reading studies have been conducted in English. Some findings are applicable to other languages, but many are not.

Memory functions and reading neuroscience are rarely taught in colleges of education. Therefore, graduates do not clearly understand why speed is a necessary prerequisite to complex thinking about text. Inevitably, reading specialists are influenced by English-language methods and may apply them to consistent orthographies that are used by the vast majority of languages. Without a neuroscience or memory background, they may form theories of reading that contravene research findings.

Comments on the Dowd and Bartlett Research Design

The Dowd and Bartlett (2019) article offers little information about the neuroscientific aspects of reading. Its research review relied heavily on older behavioral studies, English-based studies, and the US National Reading panel of 2000. The researchers measured the speed and comprehension of second graders in 11 countries who attended Save the Children programs. The organization focuses on underprivileged children, and its reading programs train teachers and promote family literacy. The study used data from 11 countries collected around 2013–14. Many children spoke languages other than the language of the test. Staff in each country constructed 60-word passages, along with 5–10 questions that aimed to measure literal and inferential comprehension. Students who answered 75–80 percent of the questions were considered "readers with comprehension." The researchers found that students in Malawi attained this level reading on average 30 correct words per minute, while the Vietnamese needed about 95 words per minute. This disparity was presented as evidence that speed is an unreliable indicator.

However, the study had technical problems that make it impossible to arrive at this conclusion or perhaps at any conclusion. The most important are:

> *Reliance on outlier populations.* The sample consisted of 6,250 second graders; 3,118 (or 49.8 percent) read 10 or more words per minute; 2,138 (or 34 percent) were considered readers with comprehension. The latter read at an average of about 51 words per minute and answered 75–80 percent of questions correctly. At a time when many students are still learning to read, this performance level is surprising. For example, 33 percent of the second graders in poor communities of Asiut, Egypt, read about 35 words per minute and correctly answered inferential questions in standard Arabic. Their native language is listed as standard Arabic, which is not normally spoken anywhere. If assessments were accurate, these second graders are outliers. They may have higher intelligence, longer working memory, a higher socioeconomic status, a better vocabulary, or educated parents who gave them lots of practice (Clarke et al. 2010). Comprehension depends

on these variables, so findings cannot be generalized to second graders everywhere.

Self-selection into the program. The researchers could have examined why one-third of the sample read with comprehension at speeds that are rare in other low-income environments (Gove 2013). The Save the Children programs surely deserve credit for the students' good performance, but teachers and families self-select into them.[2] Overall, it is hard to generalize from self-selected high-performing samples to the rest of the world.

Unexplored comprehension inconsistencies. The authors give little information about the 50 percent of the second graders who read 0–9 words per minute, but they mention a low correlation between speed and comprehension. Early Grade Reading Assessments (EGRA) exist for most of the sample countries, and their data show otherwise. For example, reports on Nicaragua show a sizable correlation between speed and comprehension (RTI 2010; second graders read 46.5 words per minute and understood 57 percent of questions on average). Poor comprehension by native speakers of a transparent orthography could raise questions of poor test administration and measurement bias. In principle, findings could be replicated using later samples and administrations, but the authors only reported data from 2013–14.

Comprehension confounded with language knowledge. According to the authors, it was not possible to assess vocabulary of non-native speakers. But the study does not even state in which languages students were tested. For example, were the Ethiopian students in Tigre tested in Tigrinya, Amharic, or something else? This is unknown. High reading speed accompanied by limited comprehension could reflect such confusion. Furthermore, the authors defined comprehension in ways that vary from prevailing research (e.g., Seigneuric et al. 2000; Cain et al. 2004; Cutting and Scarborough 2006; Keenan et al. 2008). No comprehension studies were used to justify their approach.

Variable timing of the tests. It is unknown if all 11 countries measured reading simultaneously or at different times. It is also unknown how much of the second grade had elapsed when students were tested. In this early reading stage, a few months matter a great deal and could account for some of the variability in test scores.

Variability and confusion of literal and inferential questions. For comparability, reading tests usually ask literal questions in order to assess what students

[2] See https://www.savethechildren.org/us/what-we-do/global-programs/education/literacy-boost.

can retain in working memory after a minute of reading. The Dowd and Bartlett study measured comprehension through a mix of literal and inferential questions. However, these have different accuracy and reaction times (see, e.g., RTI 2014a and 2014b). Furthermore, different countries in their sample used a different mix, thus magnifying confusion.

Psychometric weaknesses. To compare comprehension across 11 countries, test items ought to have roughly similar difficulty. However, no discrimination or difficulty indices were reported. In short reading tests, few questions can be asked, so test developers often skip pretesting or item statistics, but some questions are often easier than in others. This constitutes an internal validity violation that compromises the results.

Working memory confounding. Students were allowed to re-read the text before replying, thus renewing the contents of their working memory. This additional confounding event makes it impossible to compare comprehension results with those of other tests. One further issue concerns the "serial effects" of working memory. Slow readers may remember the beginning and the end of a message but miss the middle (Gupta et al. 2005). Allowing unlimited reads to answer questions therefore eliminates the diagnostic value of this memory feature.

Neglect of within-groups variance. The researchers did not report most means or standard deviations in the article, but a visual inspection of the scores suggests large variations in reading speed within every country. However, the discussion focused on between-country (or between-language) differences. This may signal experimenter bias toward a favored hypothesis.

Speed measures as raw words per minute. Words per minute usually reflect the number of utterances as between blank spaces, but writing traditions set spaces differently. For example, Malawi has Bantu languages that are agglutinative and words appear longer, while Vietnam has tonal languages with multiple short words. Students in these two languages happened to comprehend with the highest and lowest number of words per minute. Humans do have a sense of words (McBride et al. 2012), so methods exist to overcome this linguistic variation. For example, for agglutinative languages components could be counted. Also, the number of words in longer texts of various languages could have been counted and ratios could have been produced and used (Abadzi 2013). However, such options were not considered. This is one more reason why the speed differences may be a measurement artifact.

Discussion

The above issues go a long way toward explaining the comprehension differences described in the Dowd and Bartlett (2019) article. Multiple variables are confounded in multiple ways, making interpretation impossible. Rather than invalidating the use of speed, the study highlighted the consequences of ignoring memory functions. In particular, the study downplayed the critical relationship between working memory and comprehension. The authors briefly acknowledged the role of this concept, but they ignored it later.

Dowd and Bartlett (2019) sensibly state that the need for speed deserves to be interrogated. Indeed it makes little difference if students read 200 or 250 words per minute; at that speed, the priority is to comprehend nuances effortlessly and for that, proficient readers may read difficult text slowly. But in the early stages, speed is what gets a text to pass through working memory in one piece and be processed by comprehension areas.

The recommendation to deemphasize speed has bigger consequences. Neural networks need practice to specialize and transmit information in milliseconds. Real life demands volumes of reading! Pages and pages await, and textbook assignments involve hundreds or thousands of words. They need to flow effortlessly through the reading paths. Struggling readers do not have infinite time. If they read a text slowly or inaccurately, their working memory becomes clogged with letters, and they fail to recall important details. Mental effort is unpleasant, and slow readers may quickly stop (Mizuno et al. 2011).

We applaud the researchers' commitment to educating the poor. And given the anglocentric tendencies of many educators, we understand the syllogisms that set this study in motion. But to give valid policy advice, it is crucial to use state-of-the-art knowledge. Otherwise, results and recommendations may have the perverse effect—if followed—of actually disadvantaging the very students the donors attempt to support.

Given the learning crisis worldwide, we must all be accountable for our policy advice. Nowadays information travels instantly and widely, so questionable findings may be impossible to rein in. Journals thus have the ethical obligation to be proactive. It is crucial to demand experimental rigor, request evidence of data fidelity, and use reviewers who are well versed in the relevant research. To fulfill the targets of the Sustainable Development Goals by 2030, the choices are limited. Governments and donors must be guided by accurate publications reflecting the latest scientific advances.

References

Abadzi, H. 2013. "Developing Cross-Language Metrics for Reading Fluency Measurement: Some Issues and Options." Global Partnership for Education Working Paper Series on Learning no. 6, World Bank, Washington, DC.

Abadzi, H. 2017. “Turning a Molehill into a Mountain? Why Reading Programs Are Failing the Poor Worldwide.” *Prospects* 177. https://doi.org/10.1007/s11125-017-9394-9.

Alloway, T. P., and R. Alloway. 2013. *The Working Memory Advantage.* New York: Simon & Shuster.

Bartlett, L., A. J. Dowd, and C. Jonason. 2015. “Problematizing Early-Grade Reading: Should the Post-2015 Agenda Treasure What Is Measured?” *International Journal of Educational Development* 40:308–14.

Cain, K., J. Oakhill, and P. Bryant. 2004. “Children’s Reading Comprehension Ability: Concurrent Prediction by Working Memory, Verbal Ability, and Component Skills.” *Journal of Educational Psychology* 96 (1): 31.

Cantlon, J. F., P. Pinel, S. Dehaene, and K. A. Pelphrey. 2011. “Cortical Representations of Symbols, Objects, and Faces Are Pruned Back during Early Childhood.” *Cerebral Cortex* 21 (1): 191–99.http://dx.doi.org/10.1093/cercor/bhq078.

Centanni, T. M., L. W. King, M. D. Eddy, S. Whitfield-Gabrieli, and J. D. Gabrieli. 2017. “Development of Sensitivity versus Specificity for Print in the Visual Word Form Area.” *Brain and Language* 170:62–70.

Centanni, T. M., E. S. Norton, A. Park, S. D. Beach, K. Halverson, O. Ozernov-Palchik, and J. D. Gabrieli. 2018. “Early Development of Letter Specialization in Left Fusiform Is Associated with Better Word Reading and Smaller Fusiform Face Area.” *Developmental Science* 21 (5): e12658.

Clarke, P. J., M. J. Snowling, E. Truelove, and C. Hulme. 2010. “Ameliorating Children’s Reading-Comprehension Difficulties: A Randomized Controlled Trial.” *Psychological Science* 21 (8): 1106–16.

Cutting, L. E., and H. S. Scarborough. 2006. “Prediction of Reading Comprehension: Relative Contributions of Word Recognition, Language Proficiency, and Other Cognitive Skills Can Depend on How Comprehension Is Measured.” *Scientific Studies of Reading* 10 (3): 277–99. https://doi.org/10.1207/s1532799xssr1003_5.

Czigler, I. N. 2010. “Representation of Regularities in Visual Stimulation.” In *Unconscious Memory Representations in Perception,* ed. I. Czigler and I. Winkler. Amsterdam: Benjamins.

Daneman, M., and P. A. Carpenter. 1980. “Individual Differences in Working Memory and Reading.” *Journal of Verbal Learning and Verbal Behavior* 19:450–66.

Dehaene, S., and L. Cohen. 2011. “The Unique Role of the Visual Word Form Area in Reading.” *Trends in Cognitive Sciences* 15 (6): 254–62.

Devlin, K. 2010. “The Mathematical Brain.” In *Mind, Brain, and Education,* ed. D. De Souza. Bloomington, IN: Solution Tree.

Dowd, A. J. 2011. “Fluency, Accuracy and Comprehension across the Globe: Unpacking the Utility of 60 Words per Minute.” Presentation to the Comparative and International Education Society, Montreal, Canada.

Dowd, A. J., and L. Bartlett. 2019. “The Need for Speed: Interrogating the Dominance of Oral Reading Fluency in International Reading Efforts.” *Comparative Education Review* 63 (2): 189–212.

Dufau, S., J. Grainger, K. Midgley, and P. J. Holcomb. 2015. “A Thousand Words Are Worth a Picture: Snapshots of Printed-Word Processing in an Event-Related Potential Megastudy.” *Psychological Science* 26 (12): 1887–97.

Eberhard-Moscicka, A., L. Jost, L. Fehlbaum, S. E. Pfenninger, and U. Maurer. 2016. “Temporal Dynamics of Early Visual Word Processing—Early versus Late

N1 Sensitivity in Children and Adults." *Neuropsychologia* 91 (October): 509–18. http://dx.doi.org/10.1016/j.neuropsychologia.2016.09.014.

Fuchs, L. S., D. Fuchs, and L. Maxwell. 1988. "The Validity of Informal Reading Comprehension Measures." *Remedial and Special Education* 9 (2): 20–28. https://doi.org/10.1177/074193258800900206.

Gori, S., and A. Facoetti. 2014. "Perceptual Learning as a Possible New Approach for Remediation and Prevention of Developmental Dyslexia." *Vision Research* 99 (June) 78–87. dx.doi.org/10.1016/j.visres.2013.11.011.

Gove, A., S. Habib, B. Piper, and W. Ralaingita. 2013. "Classroom-Up Policy Change: Early Reading and Math Assessments at Work." *Research in Comparative and International Education* 8 (3): 373–86.https://doi.org/10.2304/rcie.2013.8.3.373.

Graham, B. E., and A. J. van Ginkel. 2014. "Assessing Early Grade Reading: The Value and Limits of 'Words per Minute.'" *Language, Culture and Curriculum* 27 (3): 244–59.

Gupta, P., J. Lipinski, B. Abbs, and P-H. Lin. 2005. "Serial Position Effects in Nonword Repetition." *Journal of Memory and Language* 53:141–62.

Keenan, J. M., R. S. Betjemann, and R. K. Olson. 2008. "Reading Comprehension Tests Vary in the Skills They Assess: Differential Dependence on Decoding and Oral Comprehension." *Scientific Studies of Reading* 12 (3): 281–300. https://doi.org/10.1080/10888430802132279.

Kendeou, P., T. C. Papadopoulos, and G. Spanoudis. 2012. "Processing Demands of Reading Comprehension Tests in Young Readers." *Learning and Instruction* 22 (5): 354–67. https://doi.org/10.1207/s1532799xssr1003_5.

Krafnick, A., L. H. Tan, D. L. Flowers, et al. 2016. "Chinese Character and English Word Processing in Children's Ventral Occipitotemporal Cortex: fMRI Evidence for Script Invariance." *NeuroImage* 133 (2): 302–12

LaBerge, D., S. J. Samuels. 1974. "Toward a Theory of Automatic Information Processing in Reading." *Cognitive Psychology* 6:293–323.

Language and Reading Research Consortium and Y. D. Chiu. 2018. "The Simple View of Reading across Development: Prediction of Grade 3 Reading Comprehension from Prekindergarten Skills." *Remedial and Special Education* 39 (5): 289–303.

Makel, M. C., and J. A. Plucker. 2014. "Facts Are More Important Than Novelty: Replication in the Education Sciences." *Educational Researcher* 43 (6): 304–16. https://doi.org/10.3102/0013189X14545513.

McBride-Chang, C., H. C. Chen, B. Kasisopa, D. Burnham, R. Reilly, and P. Leppänen. 2012. "What and Where Is the Word?" *Behavioral and Brain Sciences* 35 (5): 295–96. https://doi.org/10.1017/S0140525X1200009X.

McNamara, D. S., and J. P. Magliano. 2009. "Towards a Comprehensive Model of Comprehension." In *The Psychology of Learning and Motivation*, ed. B. Ross. New York: Academic.

Mizuno, K., M. Tanaka, K. Yamaguti, O. Kajimoto, H. Kuratsune, and Y. Watanabe. 2011. "Mental Fatigue Caused by Prolonged Cognitive Load Associated with Sympathetic Hyperactivity." *Behavioral and Brain Functions* 7 (1): 17.

Perfetti, C., F. Cao, and J. Booth. 2013. "Specialization and Universals in the Development of Reading Skill: How Chinese Research Informs a Universal Science of Reading." *Scientific Studies of Reading* 17 (1): 5–21.

Pikulski, J. J., and D. J. Chard. 2005. "Fluency: Bridge between Decoding and Reading Comprehension." *The Reading Teacher* 58:510–19. http://www.eduplace.com/state/author/pik_chard_fluency.pdf.

Rayner, K., B. Foorman, C. Perfetti, D. Pesetsky, and M. Seidenberg. 2001. "How Psychological Science Informs the Teaching of Reading." *Psychological Science in the Public Interest* 2 (2): 31–74.

Ringole, A., and M. Jukes. 2016. "Data-Driven Approaches to Identifying Reading Fluency Benchmarks." Paper presented at the Comparative and International Education Society Conference, Vancouver, British Columbia.

RTI (Research Triangle Institute). 2010. "Promoting Quality Education in Nicaragua." Final report (February), prepared for Bureau for Economic Growth, Agriculture and Trade (EGAT/ED), United States Agency for International Development, Washington, DC.

RTI (Research Triangle Institute). 2014a. "Egypt Grade 3 Early Grade Reading 2nd National Assessment." USAID EdData IITask Order No. 27. Prepared for Education and Training USAID/EGYPTUnit.

RTI (Research Triangle Institute). 2014b. "Indonesia: The National Early Grade Reading Assessment (EGRA) and Snapshot of School Management Effectiveness (SSME) Survey Report of Findings." Eddata II Technical and Managerial Assistance, Task Number 23, Contract Number AID-497-BC-13-00009 Strategic Objective 3.

Seigneuric, A., M. F. Ehrlich, J. V. Oakhill, and N. M. Yuill. 2000. "Working Memory Resources and Children's Reading Comprehension." *Reading and Writing* 13 (1–2): 81–103.

van den Broek, P., D. N. Rapp, and P. Kendeou. 2005. "Integrating Memory-Based and Constructionist Processes in Accounts of Reading Comprehension." *Discourse Processes* 39:299–316.

van der Mark, S., K. Bucher, U. Maurer, E. Schulz, S. Brem, J. Buckelmüller, and D. Brandeis. 2009. "Children with Dyslexia Lack Multiple Specializations along the Visual Word-Form (VWF) System." *NeuroImage* 47 (4): 1940–49.

Yang, Y., J. Wang, C. Bailer, V. Cherkassky, and M. A. Just. 2017. "Commonality of Neural Representations of Sentences across Languages: Predicting Brain Activation during Portuguese Sentence Comprehension Using an English-Based Model of Brain Function." *NeuroImage* 146 (February): 658–66. https://doi.org/10.1016/j.neuroimage.2016.10.029

Younger, J. W., E. Tucker-Drob, and J. R. Booth. 2017. "Longitudinal Changes in Reading Network Connectivity Related to Skill Improvement." *NeuroImage* 158:90–98.