INTERNATIONAL LITERACY ASSOCIATION READING RESEARCH QUARTERLY

What Constitutes a Science of Reading Instruction?

Timothy Shanahan
University of Illinois at Chicago, USA

ABSTRACT

Recently, the term *science of reading* has been used in public debate to promote policies and instructional practices based on research on the basic cognitive mechanisms of reading, the neural processes involved in reading, computational models of learning to read, and the like. According to those views, such data provide convincing evidence that explicit decoding instruction (e.g., phonological awareness, phonics) should be beneficial to reading success. Nevertheless, there has been pushback against such policies, the use of the term *science of reading* by "phonics-centric people", and their lack of instructional knowledge and experience. In this article, although the author supports pedagogical decision making on the basis of a confluence of evidence from a variety of sources, he cautions against instructional overgeneralizations based on various kinds of basic research without an adequate consideration of instructional experiments. The author provides several examples of the premature translation of basic research findings into wide-scale pedagogical application.

Science of reading is a term that has been used for more than 200 years. Throughout this history, it has been used most frequently to refer to the pronunciation and decoding of words on the basis of basic research. In this article, I situate the term historically and recommend caution in formulating agendas of instructional practice or policy primarily or solely on the basis of basic research. As such, I make claims about the nature of the empirical evidence that should comprise any science of reading instruction.

The purpose of research in any field of study, including reading education, is to increase our knowledge comprehensively about reading and its development. Yet, because reading is widely accepted as a public good, reading research also has a more specific responsibility. It should help determine how best to increase the literacy franchise, to raise levels of literacy performance, to ensure equal access, and to make the delivery of reading instruction more certain and more efficient. That first goal, the comprehensive expansion of knowledge, is not my focus here, but the second one, figuring out how best to teach reading and to use these results to inform instructional practice and public policy, is what I will explore.

In current use of the term *science of reading*, authors often try to make pedagogical and policy claims mainly on the basis of basic research done in the cognitive sciences and neurosciences, particularly with regard to beginning reading (e.g., Seidenberg, 2017). As such, the term is a bit of a misnomer because those using it today tend to reason directly

INTERNATIONAL LITERACY ASSOCIATION

Reading Research Quarterly, 55(S1)
pp. S235–S247 | doi:10.1002/rrq.349
© 2020 International Literacy Association.

from basic research to the prescription of instruction; the conversation seems to be less about a science of reading than a science of reading instruction. Unfortunately, this approach ignores why we use research in this endeavor at all; we seek to reduce uncertainty in the solving of practical problems. We, as a society, are uncertain about how best to educate our children. We know that large percentages of students leave school with literacy levels too low to allow them full participation in the benefits of society (Shanahan & Shanahan, 2008). We know that socioeconomic status, racial identity, and language heritage too often determine how successful students will be in gaining reading ability in our schools (National Center for Education Statistics, 2013). In other words, how well reading is taught in schools is important, as it can have long-lasting impacts on the well-being of individuals. The purpose of bringing research to this enterprise is to try to increase the likelihood of success; we want to promote practices and policies that will have the greatest possibility of ensuring equity and excellence in reading. We seek to reduce our uncertainty and to decrease the size and consequences of the inferences we must make to make that work.

For example, let's say that a study reveals that readers are more likely to remember the information higher in the information hierarchy of a text than the information lower in the structure (Meyer, 1975). Such research provides useful information about how our minds work, how we gain information, and how we learn. It suggests that if we want to create better readers, then it would be wise to promote a more thorough or effective processing of subordinate information. Perhaps knowing that individuals are less likely to process subordinate information reduces our uncertainty as to how to proceed, but past experience suggests that it does not do so sufficiently to justify major pedagogical or policy investments. Sometimes basic research, including basic research in education that depends on descriptive, correlational, or qualitative data, identifies phenomena that may not be addressable pedagogically.

If we want to know how best to promote this better reading, then we need studies aimed at determining whether we can teach such abilities. That means developing curricula and instructional regimes and then evaluating their effectiveness. With regard to improving readers' recall of subordinate information, that has been researched, and indeed, this aspect of behavior can be improved, removing a significant inference from the causal chain (A. Adams, Carnine, & Gersten, 1982), and this success should increase our certainty with regard to future possibilities of success. If this instructional study were then replicated repeatedly with diverse groups of students, that should increase our certainty even more in the value of this approach to reading improvement (Hebert, Bohaty, Nelson, & Brown, 2016; Pyle et al., 2017; Wijekumar, Meyer, & Lei, 2017). Also, as will be explained, even that level of certainty may be inadequate, both practically and ethically, for providing sufficient and responsible guidance for education policy and teaching.

Although we have much to learn from basic studies of reading processes and neurological functions, it is important not to overrely on such evidence in determining how best to teach reading. High-quality research reduces uncertainty, and a confluence of high-quality empirical data from multiple sources should go far in increasing our confidence that certain policies and practices will be effective and beneficial. In this article, I argue not for ignoring the fruits of basic research but for adjudicating matters of pedagogy chiefly on the basis of experimental studies of instructional practice itself. If our goal is to determine how best to teach reading, then we must rely on data that evaluate the effectiveness of teaching, rather than depending solely or even mainly on studies of reading processes or of other noninstructional phenomena, which are then applied to teaching through analogy or logical deduction or from premature conclusions drawn from empirical investigations that do no more than describe or correlate. The role of basic research in shaping instruction quite appropriately lies either in identifying pedagogical innovations that can be evaluated through studies of instruction or in providing evidence that further buttresses or explains the results of such experimental pedagogical study.

Basic research by its very nature cannot reduce our uncertainty sufficiently to justify its use to determine public policy any more than the genetic mapping of COVID-19 should be used to promote a particular regimen of wide-scale inoculation. In that sense, it is necessary but not sufficient. Basic research may point in helpful directions or warn us off from likely failures, but to be really certain of the effectiveness and safety of a vaccine, it is necessary to try it out on scale with varied populations. Identifying agents that might work effectively against a virus is one thing, and figuring out how such agents can actually be used effectively is something else altogether. Both of these kinds of research are essential in both medicine and education, but the key point here is that practice and policy, ultimately, must depend on evidence showing the practical effectiveness and safety of the approaches taken with patients or students, not on the basic research that may have generated the original insights that led to the development of those practices that worked.

The impact of the term *science of reading* as used today is as much bound up in its tone as in its meaning; it now often seems to be used as a rhetorical cudgel to challenge those not adhering to some particular conception of it, hence the arguments over who even has the right to use the term (Calkins, 2020; National Education Policy Center & Education Deans for Justice and Equity, 2020). One's comfort with today's science of reading seems to depend on which instructional approaches one advocates and what one is willing to accept as determinative evidence. As such,

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

in this article, I delve into the nature of the kind of evidence that should be the basis of a science of reading instruction.

This consideration of what should be included in a science of reading instruction surfaces many philosophical and practical issues, not all of which can be addressed here: Which research questions are worthwhile from a "science of reading instruction" perspective? Should we promote basic or applied science? Should pedagogy be governed by standards of instruction or professional autonomy? What are the nature and qualities of research most likely to contribute to a science of reading (e.g., types of studies, methodological rigor, criteria for amounts and types of evidence)?

What Is the Science of Reading?

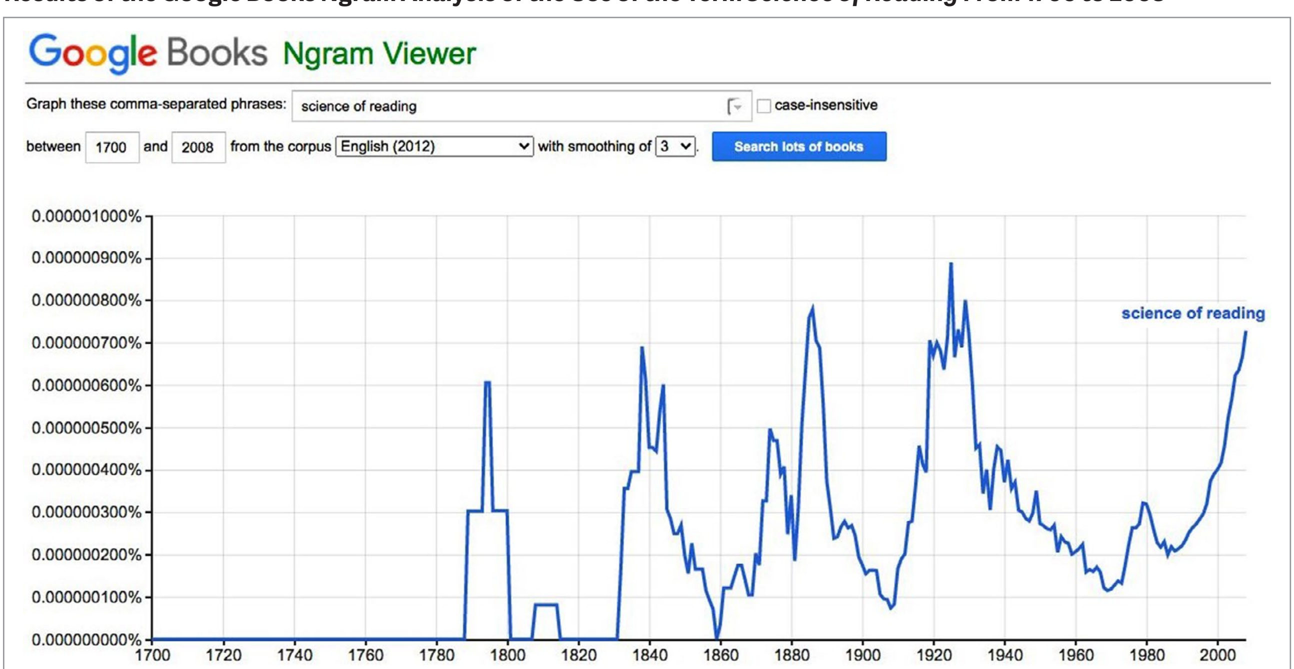
Science of reading is an old term. I conducted a Google Books Ngram analysis of its use in a corpus of 8 million books, limited to those written in English. This collection of books, fiction and nonfiction, represents 6% of all books ever published, drawn from more than 40 university libraries. Ngram searched these books for all appearances of the three-word string "science of reading." Problems with Ngram analyses have been identified (e.g., scanning problems, overinclusion of scientific texts), but those should not be an issue here because our purpose is simply to see how long the term has been in use, to gain a sense of its popularity, and to identify its more common meanings, rather than any exhaustive or nuanced examination of usage.

The term was used first to refer to text reading during the 18th century, coinciding with the birth of linguistics, the scientific study of language. The original agenda of scientific linguistics was the determination of proper pronunciations of ancient languages (Allan, 2010). Thus, during the early 19th century, the term was used to refer to how one should read the Koran or the Bible, particularly with regard to the pronunciation of the words (e.g., Stewart, 1809).

Science of reading was first used pedagogically during the 1830s. Consistent with its original meaning in linguistics, it was used in education to discuss teaching students to sound out words properly (Experience, 1836; Pitman, 1843). Use of the term has waxed and waned over two centuries, with occasional torrents of use in the 1840s, 1880s, and 1920s (see Figure 1). Although that historically early reading research was not limited to issues of decoding (e.g., Huey, 1908/1968), the term *science of reading* usually has been reserved for decoding, often with an emphasis on noninstructional research (including studies of eye movements and linguistic analyses of the English spelling system).

It is worth noting that efforts to apply research to reading instruction have increased markedly since the 1950s and that these efforts rarely employed the term *science of reading*. This increased emphasis seems to be beyond argument, but it can be documented in many ways: increases in amount of reading research (e.g., the 72 volumes of *Annual Summary of Investigations Relating to Reading* published 1925–1998: Roser & Weintraub, 2009), the appearance of specialized journals of reading research (e.g., *Reading Re-*

FIGURE 1
Results of the Google Books Ngram Analysis of the Use of the Term *Science of Reading* From 1700 to 2008



19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

search Quarterly, *Journal of Reading Behavior/Journal of Literacy Research*, *Scientific Studies of Reading*), increases in national funding for reading research (e.g., Center for the Study of Reading, National Reading Research Center), institutional efforts to interpret reading research for practitioners and policymakers (e.g., M.J. Adams, 1991; Anderson, Hiebert, Scott, & Wilkinson, 1985; August & Shanahan, 2006; National Early Literacy Panel [NELP], 2008; National Reading Panel [NRP], 2000; RAND Reading Study Group, 2002; Snow, Burns, & Griffin, 1998), and government policies that require practical concordance with reading research: Reading Excellence Act (U.S. Department of Education, 2000), No Child Left Behind (U.S. Department of Education, 2002b), Reading First (U.S. Department of Education, 2002c), Early Reading First (U.S. Department of Education, 2002a), Striving Readers (U.S. Department of Education, 2005), Striving Readers Comprehensive Literacy Program (U.S. Department of Education, 2015b), and the Every Student Succeeds Act (U.S. Department of Education, 2015a). These efforts to increase the application of research in reading have included a heavy emphasis on decoding but usually have gone well beyond that, considering domain knowledge, vocabulary, reading comprehension, metacognition, oral language, and a plethora of other instructional issues (e.g., comprehension strategies, oral reading fluency, read-alouds to students, text readability).

Now, the term *science of reading* has returned to vogue. This resurgence has included its appearance in scholarly publications (e.g., Snowling & Hulme, 2005). However, the surge of attention to the science of reading is attributable to its use by the media and in the public debate. Analysis of the term in indexes of newspapers and other popular publications (via Nexis Uni) over a three-decade period reveals an upsurge in the use of the term since 2018. This increase is specifically due to coverage of policy initiatives of the International Dyslexia Association, the publication of Seidenberg's (2017) book *Language at the Speed of Sight: How We Read, Why So Many Can't, and What Can Be Done About It*, and an exploration of issues raised in that book by Hanford (2018), a radio journalist for American Public Media.

Seidenberg (2013, 2017) did not explicitly define *science of reading* in his book or in an earlier article, but it is easy enough to piece together what counts and what does not. His insights about instruction are based almost entirely on conclusions drawn from basic scientific research concerning the mechanisms of skilled reading, neural circuits, computer simulations of learning, and the like. Instructional studies are barely mentioned, and when they are, it is not to support the pedagogy that he promotes. Hanford (2018) focused public attention on Seidenberg's book and used the term much as he did to refer to a wide range of empirical studies drawn from cognitive science and neural science showing the importance of phonological processing in proficient reading. Although I agree with Seidenberg's and Hanford's conclusion—the science of reading actually supports the teaching of explicit phonics—I do so not because I agree that basic research has provided the greatest certainty of that but because basic research findings have concurred with the preponderance of evidence drawn from the direct study of that instructional question.

Given the well-documented contentiousness of the reading field (Stanovich, 2000; Taylor, 1998), an embrace of a term like *science of reading* by some will only arouse opposition among others, as has been the case here (Calkins, 2020; National Education Policy Center & Education Deans for Justice and Equity, 2020). Those opposed to the current use of this term argue that it is used too narrowly (Hanford, 2018, mentions *decoding* and related terms 86 times and all other aspects of reading only once) and that the instructional practices it promotes are overemphasized and often inappropriate. The critics also take an ad hominem swing. They question the value of ideas communicated by journalists and noneducators who do not know classroom instruction or who are not scholars of reading and literacy. Unfortunately, these opponents of current uses of the term are no more likely to rely on appropriate instructional research evidence to support their pedagogical claims.

A Science of Reading Instruction

Historically, basic science has been distinguished from applied science on the basis of its apparent distance from practical problems: Basic science aims at answering fundamental questions, whereas applied science tries to solve practical ones. This distinction is about neither who does the science (neurologists, psychologists, ethnographers vs. educators) nor their motivation (even those conducting the most basic forms of science hope that their insights will have some practical value). In this article, *basic science* refers to any investigation that is not directly aimed at evaluating whether a particular approach to instruction works. That includes those brain science and computer simulation studies that are the basis of public debate in reading education today. They are basic because they tell us something about how the reading brain works and how we may learn to recognize words, but these studies provide no direct test of any instruction, because no instruction is provided in them. This definition of *basic science* includes descriptive studies that can reveal who is doing well in reading and who struggles with status quo reading instruction. It also includes various correlational studies and complex statistical modeling that exposes the relations that various cognitive abilities may have with reading, or ethnographic observations in classrooms that describe patterns of interaction that might be useful.

In 2002, the National Research Council (NRC) issued a prestigious report in book form called *Scientific Research*

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

in Education. The purpose of the report was both to promote the use of research as the basis of educational policy and to attempt to heal the rift between quantitative and qualitative researchers. The NRC said that "the distinction between basic and applied science [in education] has outlived its usefulness" (p. 20). That may suggest to some that the point being made here is inappropriate or outdated, which is not the case. The NRC argument was not that all research is equal when it comes to educational policy but that all research well implemented was scientific, a position I heartily endorse. "What makes research scientific is not the motive for carrying it out, but the manner in which it is carried out" (p. 20). NRC was saying that basic research was not any more scientific than applied research and that both should be accepted as being integral to the scientific endeavor. What is being argued here is not whether basic research is scientific or even useful (indeed, it is both), but whether it is proper, either intellectually or ethically, to prescribe pedagogical practice and policy on its basis alone.

No matter how scientific basic research may be, ultimately any science of instruction will have to depend on applied studies of teaching, that is, those studies that require smaller inferences to application. This is not a rehash of the ad hominem concerns of the critics of current sources on science of reading but a matter of practical epistemology. No matter how sound the studies of neural processing, perception, and memory, we must recognize the possibility that they, at least in some cases, could be irrelevant, inconsequential, or misleading with regard to teaching.

A famous example is the first U.S. study of reading, or more properly of word perception (Cattell, 1886). Cattell (1886) found that readers recognized words more quickly than letters, which was interpreted to mean that we read words as wholes rather than decoding them. The pedagogical interpretation of this was that word memorization rather than decoding should be taught, and this study was cited well into the 20th century as evidence of that. What Cattell's study revealed was accurate and reliable—people actually recognize letters within words faster than isolated letters—but the interpretation of this finding and its application to teaching was neither accurate nor reliable. Studies quite consistently have found decoding instruction to be advantageous (M.J. Adams, 1990; Chall, 1996; NELP, 2008; NRP, 2000).

Another example of pedagogical conclusions from noninstructional studies can be found in analyses of oral reading errors. Goodman (1967) analyzed errors and found that they revealed the use of orthographic/phonemic, syntactic, and/or semantic information. A reader might, for instance, read the word *automobile* as *car* (relying on the semantic and syntactic context or cues). Goodman concluded that this was how readers read, figuring words out as much from meaning and context as from letters and sounds. This is what is meant by the three-cueing systems. Never mind that later studies only found such responses when readers erred, not when reading words proficiently and that this dependence on semantic and syntactic information was prominent with poor readers, not with good ones (Stanovich, 1980).

Goodman (1967) was not the first to recommend this kind of guessing on the basis of minimum visual information, nor was he the first to do so without any instructional evidence showing that it conferred a learning advantage. Earlier, Bond and Bond (1943), in their popular reading textbook, made the same recommendation, in their case on the basis of Gestalt psychology. According to Gestalt psychology, perceptions arise from patterns or gestalts rather than from an analysis of constituent parts. Thus, Bond and Bond concluded that readers should not devote much energy to figuring out words letter by letter, which is a fair generalization from the basic psychological studies but one not particularly in much agreement with direct studies of reading instruction.

Again, in Goodman's (1967) case, his empiricism was sound. Readers, when distracted or struggling, try to compensate for this failure by inferring words that might make sense in context. However, no one has shown that teaching students to compensate in this way improves reading achievement, and other basic research has weakened the original claim because proficient readers look at pretty much every letter during reading, and where they look is not affected by semantics or syntax (Rayner, Binder, Ashby, & Pollatsek, 2001). (Although no research has shown that learning benefits from teaching cueing systems, there likewise is no evidence showing such teaching to be harmful; Hanford, 2019.)

There has been much basic physiological and psychological study into what distinguishes struggling readers from those who learn literacy more easily. These studies have led to myriad theories of how best to address the needs of struggling readers. An early research-based theory was the claim that reading disability was linked to left-handedness (Kushner, 2017). Fortunately, one is no longer likely to see educational prescriptions based on this long-standing scientific correlation. A century ago, it was widely embraced by psychologists, physicians, and educators, which led to retraining practices in which children were forced to use their right hands through binding or corporal punishment. Needless to say, these practices were never found to improve reading ability (e.g., Fagan, 1931).

On the basis of various comparisons, correlational studies, and theories of brain–behavior relations, scientists have categorized students into disability subtypes, with the idea that each subtype requires special instruction addressing particular deficits (Johnson & Myklebust, 1967; Kephart, 1960; Kirk & Kirk, 1971). These deficiencies were not specific to reading but referred to

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

limitations in underlying cognitive or neural processes, including perceptual deficits and problems with hemispheric differentiation (Holmes, 1965; Kavale & Forness, 1985; Werner & Strauss, 1939). For example, there were several decades of provocative research into visual perception and visuomotor processing and their roles in learning to read (e.g., Robinson, Mozzi, Wittick, & Rosenbloom, 1960; Snyder, & Freud, 1967). Frostig's (Frostig & Horne, 1964) theories of perceptual-motor coordination and their implications for reading instruction emerged from such research. This research also led to the widespread use of perceptual measures in classrooms and clinics, including the Bender Visual-Motor Gestalt Test, Frostig's Developmental Test of Visual Perception, and the Metropolitan Reading Readiness Test, all aimed at trying to predict who might have trouble in learning to read or to identify what perceptual training was required. Likewise, at least for a couple of generations, young students practiced picture and shape discrimination in kindergarten and first grade in the name of reading readiness because of the misapplication of such psychological data (Durkin, 1980).

Some of the earliest reading studies tracked eye movements during reading (Javal, as cited in Huey, 1908/1968). Although the original methodology for this research was somewhat medieval (and required the administration of cocaine), the basic findings continue to be confirmed in more recent technologically advanced eye motion studies (Rayner et al., 2015). These data reveal much about reading, but attempts to translate them to instruction have been problematic. Because studies have shown that poorer readers have less efficient eye movement, this has led to optometric training of the eyes to read better (Heath, Cook, & O'Dell, 1976; Keogh & Pelland, 1985); such interventions have come and gone since the 1930s but are still with us today (Apperson, 1940; Murphy, 2017). These days, there are even computerized systems that supposedly transform one into a faster, more efficient reader by banishing eye movements altogether, as studies have shown that eye movements require too much of a reader's time (Spritz, n.d.).

Another line of theory and research drawn from clinical data and studies of perception can be found in Orton's (1925) theories of hemispheric dominance that encouraged tracing and other multisensory rehearsals of words, with the idea that this would result in more apt memory storage in the appropriate neural hemisphere. Unfortunately, this neurological theory, although quite influential, lacked a basis as empirically accurate as the earlier examples. Orton concluded that the hallmark of dyslexia was the visual reversal of words and letters, which he attributed to inappropriate dominance of the neural hemispheres. More rigorous studies have concluded that reversals play no special role in dyslexia (Fischer, Liberman, & Shankweiler, 1978), and to this day, no convincing studies have shown positive learning effects from tracing letters or words, yet those practices hang on.

Yet another theory of instruction to be derived from basic psychological study is the idea of learning styles. Although this idea is most associated with Carbo (e.g., 1983), who elaborated on it greatly, parts of it have had a very long empirical life (Henmon, 1912). The basic premise is that individuals learn in different ways and that teachers must identify the type of learner a student is so instruction can be modified appropriately. This is not a theory focused on differences in reading ability; thus, prescribing phonics instruction to poor decoders or repeated reading to the disfluent would not capture its essence. No, the idea here is that students vary in their ability to learn through different modalities; there are visual and auditory learners, and teachers must teach to the strongest modality. This is another interesting theory that has not panned out (Rogowsky, Calhoun, & Tallal, 2020).

There are many other examples of how basic research, even when empirically accurate, may mislead instruction. Educators, for example, used to delay reading instruction rather than provide early interventions (Morphett & Washburne, 1931). Psychological studies led to the discouragement of oral reading instruction in U.S. classrooms for approximately 50 years (Singer, 1981), although research has found learning benefits from some oral reading approaches (NRP, 2000). Laboratory studies focused on the teaching of artificial characters have found no benefit to letter name teaching, which led to the recommendation that we should not teach the ABCs to beginning readers (Samuels, 1970), yet classroom studies with the actual alphabet have revealed learning benefits (Byrne, Fielding-Barnsley, & Ashley, 2000). There are more examples of this kind of overgeneralization, misinterpretation, or premature application of basic research, but these should be sufficient to make my point.

The examples provided here should not be misunderstood. Their point was not to expose basic research as having no value, nor to suggest that it is usually wrong. Basic research can provide insightful hypotheses that may eventually lead to the development of valuable pedagogical developments and can help us better understand why a pedagogical approach may be effective. The purpose of these examples is to illustrate the dangers of attempting to move directly from basic research findings to the formulation of public policy or to the widespread adoption of particular instructional practices without direct, rigorous, and repeated evaluations of the ability of those insights to improve instructional practice. In each of these examples, later studies either called into question the accuracy of the original empiricism or convincingly exposed the inadequacies of the pedagogical inferences drawn from sound data (Berninger, Lester, Sohlberg, & Mateer, 1991; Luchow & Shepherd, 1981; Pryzwansky, 1972; Williams, 1969).

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The cognitive and neuropsychological studies that Seidenberg (2017) examined on how readers read words are an important part of a science of reading instruction but not just or even mainly because they are high-quality studies. These studies are valuable because they have contributed to our understanding of reading instruction through their consistency with the findings of extensive instructional experiments that have demonstrated consistently and overwhelmingly that explicit and systematic teaching of decoding is beneficial (NELP, 2008; NRP, 2000). Without those instructional studies, Seidenberg's results could offer no more than provocative hypotheses that would still need to be evaluated through instructional studies. With those instructional studies, Seidenberg's results provide insightful explanations of why systematic phonics instruction is so advantageous—a real contribution but not the one being touted publicly in the debates over the science of reading.

Making predictions about what kinds of instruction will be effective on the basis of basic research is a fraught enterprise. When the predictions are incorrect, they encourage poor pedagogy. When they are sound, their value can only be determined by their consistency with the findings of instructional studies. As such, the predictions reinforce what we learn from instructional studies, strengthening our trust in those pedagogical findings through their consonance with the predictions. Again, this does not denigrate the value of basic research for identifying potential pedagogical innovations or insightful explanations that could lead to even greater future innovation. Yet, no matter how good the ideas of basic research, they must be tried out instructionally and shown to be beneficial in improving reading ability or its dispersion in some way before they should be recommended to educators and policymakers.

A functional magnetic resonance imaging study has shown that phonics instruction alters particular areas of the brain where reading is known to take place (Temple et al., 2003), but remember that this kind of study cannot tell us how to teach reading. This particular study could not possibly do so because it did not provide an evaluation of any kind of teaching. The instruction in this study had already been shown to be effective; the students who received it outperformed those who did not, which was why their brains were observed: The instructional group had already shown that their brains were changed (learning had happened), and the study was aimed at revealing why that instruction may have been so successful. This study, and others like it, may enlighten us as to why phonics may be more effective than word memorization by divulging the neural correlates of the learning that we already knew had taken place. Maybe it will encourage future improvements to instruction, but for now, the main reason for teaching phonics explicitly, thoroughly, and well in the primary grades is because in approximately 100 instructional experiments (NELP, 2008; NRP, 2000), such teaching has consistently given students a clear learning advantage.

Of course, cognitive science has come up with more than the centrality of decoding to successful reading. Cognitive science has implicated other processes in reading as well. For instance, there is an extensive body of evidence suggesting the importance of theory of mind in reading comprehension (e.g., Y.-S.G. Kim, 2020). This concept refers to the ability to understand the mental states or perspectives of others. The readers' theory of mind is what allows them to draw inferences about intents, goals, and emotions, and it has been shown to be correlated with comprehension abilities (both listening and reading). Then, there is also rapid automatic naming, a measure of the speed at which individuals can process cognitive information; studies have consistently shown this to be one of the most robust correlates of reading ability (Araújo, Reis, Petersson, & Faísca, 2015). Also, a great deal of research has demonstrated the importance of working memory and other executive functions in reading (Castles, Rastle, & Nation, 2018; Ober, Brooks, Homer, & Rindskopf, 2020). Working memory and its limitations clearly play a crucial role in reading comprehension.

As much research as there has been on each of these aspects of reading, and as persuasive as that research's findings may be because of the strength and consistency of the correlations, the role these variables play in complex statistical models, and the elegant theories that connect them to reading, those variables still cannot be a part of a science of reading instruction. This distinction is made because no one has yet found that they can successfully teach these variables in ways that improve reading. Unlike with decoding, for which there is a rich and extensive collection of such studies, these insights from cognitive science are, at this point, disconnected from the instructional enterprise. They are provocative but not proven, and as such should not be recommended to teachers.

Let's consider a counterexample: the earlier noted research on text structure. There is some basic science underlying this work but not a great deal (Meyer, 1975). None of the publicly prominent "science of reading" advocates even mention the value of teaching students about discourse structure. Nevertheless, there is an extensive and rigorous body of instructional research showing that the teaching of text structure promotes higher levels of reading comprehension with a diverse group of students under a wide array of circumstances (Hebert et al., 2016; Pyle et al., 2017; Wijekumar et al., 2017). It might be worthwhile for cognitive psychologists and neuroscientists to try to provide deeper understanding of how we process text structure in our minds and brains, but until they do, based on the preponderance of available instructional evidence, it seems obvious that teaching text structure should be encouraged, even given the dearth of cognitive and neural evidence buttressing these instructional studies. By contrast, we should

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

not teach theory of mind, rapid naming, or working memory despite their strong cognitive and neural foundations, because there is not yet evidence that we know how to teach them effectively. Teachers benefit from knowing both what works and why it works, but it is the what, ultimately, that leads to improved reading ability, at least most immediately. Education is necessarily an applied science, not a basic one.

As noted earlier, some oppose this renewed call for a reliance on a science of reading, but the basis of their opposition seems to be mainly to the specific instructional practices being promoted rather than to the nature of the research that is the basis of this promotion. Their concern is not with the epistemological premise of drawing instructional conclusions from basic research (they do this, too, just focusing on different assemblages of other nonconclusive evidence). From this, it appears that if this same cognitive science and neuroscientific data were used to support their own cherished, although often untested, instructional theories, they would gladly accept this basic scientific support (even if journalists were the ones making this information public).

In the end, the only way to know if any instructional approach is effective is to try it out in classrooms and to measure its impact on student learning, the same approach adopted by various research review panels (August & Shanahan, 2006; NELP, 2008; NRP, 2000), research institutions (e.g., Institute of Education Sciences, National Institute of Child Health and Human Development, What Works Clearinghouse), and independent scholars (e.g., Castles et al., 2018; Graham & Hebert, 2011). Of course, the better designed such educational experiments are and the more often they are independently replicated, the more likely that the instructional approach under study can be made to work in other classrooms.

As some of the critics of the science of reading have pointed out, a great deal of relevant research has gone well beyond phonics instruction (Calkins, 2020). Instructional studies have identified the importance of explicit teaching of phonological awareness (NELP, 2008; NRP, 2000), oral reading fluency (Kuhn & Stahl, 2003; NRP, 2000), reading comprehension strategies (NRP, 2000; Shanahan et al., 2010), text structure (Hebert et al., 2016; Pyle et al., 2017), vocabulary (NRP, 2000; Stahl & Fairbanks, 1986), the use of complex text (Shanahan, 2019), the impact of writing on reading (Graham & Hebert, 2011), and several other curricular and instructional approaches and interventions (e.g., Hattie, 2009). Consequently, those too must be considered part of any acceptable definition of a science of reading instruction.

The Pieces That Do Not Fit

Yet, what of those instructional practices, supported by pedagogical research but inconsistent with basic research findings? Are these practices like the fabled hummingbird that supposedly, according to aeronautical research, could not fly? The idea is to use discrepancies between basic research and successful instructional approaches to raise questions about how to make an approach more successful.

An example of this is Reading Recovery, a program of beginning reading instruction aimed at struggling readers. This instructional scheme has often been challenged by critics (e.g., Baker et al., 2002; Greaney, 2001, 2011; Wood, 1994) unhappy because of the inconsistency of that program with what is known about effective decoding instruction. Despite this inconsistency, qualitative syntheses (e.g., Shanahan & Barr, 1995), meta-analyses (e.g., D'Agostino & Harmey, 2016), and specific high-quality studies (What Works Clearinghouse, 2008) have all concluded that Reading Recovery improves reading. The research on Reading Recovery has often been flawed by seriously misleading research practices, and these limitations have been well documented (Chapman & Tunmer, 1991, 2016; Shanahan, 1987; Shanahan & Barr, 1995; What Works Clearinghouse, 2008). However, even when those questionable studies are dismissed, Reading Recovery has been found to provide learning benefits, albeit at what to many is a prohibitively high cost.

Reading Recovery appears to be a hummingbird! Although we know the program works, we cannot be certain about why it does. Often, instructional approaches are complex, including a variety of content and instructional activities. In this, Reading Recovery is not an exception. In its daily 30-minute lessons, students read little books according to their reading levels, engage in the rereading of books they have read previously, receive feedback on their oral reading errors based on the three-cueing systems, write sentences, and engage in other learning activities under the one-to-one supervision of a highly trained teacher. Unfortunately, studies have indicated that the entire constellation of content and activity can be advantageous, but have not identified which features of the program are active ingredients. Is it possible, for instance, that any of those components could be omitted without lessening the positive results?

This kind of component analysis has not been common in education, but when it has been used, such as with vocabulary instruction, it has been very informative (McKeown, Beck, Omanson, & Pople, 1985) both in helping to tailor better instruction and, as with basic research, by providing clues to what makes this instruction work. Although we lack component analyses of Reading Recovery, some relevant investigations have provided clues. For example, the one-to-one teaching approach may neither explain the program's success nor be necessary for its successful implementation (Iversen, Tunmer, & Chapman, 2005). Investigations have shown that Reading Recovery has no impact on phonological awareness (Chapman, Tunmer, & Prochnow, 2001), and in a study in which Reading Recovery was

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

supplemented with explicit phonics teaching, learning accelerated significantly (Iversen & Tunmer, 1993).

It seems clear from the neurosciences (D'Mello & Gabrieli, 2018) that in terms of brain function, we all read in the same way, no matter how we were taught. It is not clear, however, what readers learn that enables this universal process. We know about the coordination of phonological processing and visual processing, and we know that teaching a broad array of sound–symbol relations and spelling patterns enhances reading achievement, but we do not know what is coded into memory. Do nascent readers learn rules, patterns, or the words themselves? *Rules* refers to a system of constraints that limit structural possibilities, such as *i* after *e* except after *c*. Some psychologists think that reading requires the mastery of such rules (Pinker, 1999). Another possibility is that we learn to recognize patterns (Venezky, 1995). *Patterns* here refers to the allowable sequences of letters that relate spelling to sound. According to this idea, learners come to recognize legal sequences of letters (i.e., those that occur with regularity). Finally, there is the possibility that we learn words more directly, as templates perhaps, which enables us to recognize similar shapes and sequences in the future (Pinker, 1984), making reading a remarkable memory task.

However, what is going on when instruction does so little to explicitly familiarize students with the relations between orthographic patterns and phonology, such as in Reading Recovery? Somehow, students who are being taught in this way are still ending up reading much as the kids who receive explicit decoding instruction. The same could be said of approaches to reading that only teach words (Barr, 1974). As already noted, such approaches do not do as well as explicit decoding instruction in improving reading, yet how do students learn from them at all? According to basic research studies, they should not work; that they do should be a matter of more than intellectual curiosity.

What physicists and engineers knew about aerodynamics was not consistent with the flight behaviors they could observe in hummingbirds (Ransford, 2008). That led them to a great deal of study of hummingbirds, expanding what we now know about flying and hovering. Instead of assuming that the basic knowledge of aerodynamics was complete and correct, researchers decided that it was worth probing those instances where practice was not in accord with empirically grounded theory.

Another example of discrepancies between successful instruction and basic science concerns the role of text complexity in reading instruction. Teachers have long been told to teach students to read at their so-called instructional levels and that, by doing so, students would make optimum progress (Betts, 1946). Accordingly, a large number of instructional programs that successfully teach many students to read use this approach (e.g., Ransford-Kaldon et al., 2010). Yet, there has been a great accumulation of basic cognitive data showing that under at least some circumstances, engaging students in more complex texts leads to greater learning (McDaniel & Einstein, 2005). Instructional studies are starting to bear this out as well (Shanahan, 2020).

These examples demonstrate the value that basic research in cognitive science and the neurosciences brings to reading. Instructional experiments may indicate that an instructional approach is beneficial. When such results appear to contradict basic cognitive research, it is worthwhile to pursue the discrepancy empirically, as this may lead to further refinements of successful practices or to the development of more effective alternatives. Those new ideas themselves, then, must eventually be evaluated through the use of teaching experiments.

As the NRC (2002) concluded, a guiding principle that should underlie all scientific investigation is the need to employ research methods that permit direct investigation of a question: "[Research] methods can only be judged in terms of their appropriateness and effectiveness in addressing a particular research question" (p. 3). In other words, if we are trying to find out the most effective way to teach reading or the educational practices most likely to provide the fairest distribution of reading ability, then we need to directly investigate those questions with appropriate methodology. Given the nature of the kinds of practical questions of reading pedagogy under discussion here (e.g., what is it that schools or teachers can do that will result in…?), direct investigation would require that research rigorously evaluate the effectiveness of the recommended pedagogy.

The NRC (2002) went a step further, too:

> Particular research designs and methods are suited for specific kinds of investigations and questions, but can rarely illuminate all the questions and issues in a line of inquiry. Therefore, very different methodological approaches must often be used in various parts of a series of related studies. (p. 4)

The NRC's immediate purpose was to highlight the value of investigations that combine quantitative and qualitative methods. Nevertheless, this point also supports the kinds of combinations of basic and applied research promoted here. Basic research is not appropriate for answering the practical pedagogical questions raised by public policy exigencies (despite recent "science of reading" claims), but basic research will continue to play an important role when used in combination with applied research.

Issues of Research Quality

If such teaching experiments are necessarily the epistemological heart of the science of reading (because they

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

answer the questions most directly and with methodology appropriate to the questions), then all of the quality criteria for conducting and reporting such studies must be adhered to. Teachers and groups of students in the various conditions have to be truly comparable (through randomization or relevant pretesting), and the conditions being compared must be equivalent, too (e.g., all groups receiving equal amounts of teaching, fidelity checks). The measures used to evaluate learning must be reliable and valid, and replication matters, too (J.S. Kim, 2019). If the findings of multiple independent studies do not concur, then there can be no certainty. These multiple studies should be meta-analyzed properly and consistently. Finally, the insights drawn from implementation science must be honored (Kelly & Perkins, 2012; Wasik & Snell, 2019); there is a great distance between cognitive studies produced in a laboratory and instructional studies implemented in classrooms, but there is an equally wide chasm between instructional effectiveness studies and successful large-scale implementations.

Clearly, the idea that certain data will be more apt for answering particular kinds of questions challenges those who claim that it is better to treat all research paradigms as equal. Placing instructional experiments above basic cognitive research, or above instructional studies that are descriptive, qualitative, or correlational, may seem to some to be unfair or antidemocratic (Cunningham, 2001; Pearson, 2004). The issue, however, is not one of fairness but rather of what kinds of questions can be answered satisfactorily by particular empirical methods (NRC, 2002; NRP, 2000). If one claims that an approach to teaching confers learning advantages, that is a causal claim. The trustworthiness of a causal claim about instruction will differ depending on whether the empirical evidence was drawn directly from a systematic evaluation of its use or from the measurement of some distant, underlying phenomena that then must be linked back to instruction logically rather than empirically (as is often done with basic research). Likewise, our ability to depend on an instructional claim should be enhanced when the approach has been applied and found to improve learning. Certainly, such evidence should be more persuasive than qualitative or quantitative correlations or descriptions. At best, these can suggest the possibility that something may be effective. Correlations may point us in the right direction, but given the importance of literacy and the inequality of its distribution in society, it would be best to adopt practices with a high certainty of effectiveness above those that provide no more than provocative possibilities.

Current proponents of the science of reading are correct that there is a substantial body of high-quality cognitive and neuroscientific evidence, and it is evident that instruction consistent with that research has not been emphasized much in teacher education recently (Cohen, Mather, Schneider, & White, 2017; Joshi et al., 2009). Yet, these arguments have characterized this problem too narrowly, ignoring most issues of reading instruction beyond decoding and beginning reading. These arguments have underestimated the challenges inherent in applying any research findings on scale and may have overestimated the likely payoffs from such applications. Consider the National Assessment of Educational Progress gains obtained during the past phonics awakenings in the 1990s and early 2000s, which were positive for sure yet modest considering the current rhetoric (National Center for Education Statistics, 2013).

Basic research has an important role to play in reading science but can never be the final determinant of practice or policy; that should always depend on studies that directly evaluate the effectiveness of a practice or policy. Reports such as those by the NRP (2000) and the NELP (2008) are the most promising foundation for practice and policy determinations because the panels directly evaluated the effectiveness of instruction. Likewise, the research agendas of the Institute of Education Sciences and the National Institute of Child Health and Human Development are promising both because they encourage high-quality basic research studies and because they require that these lines of investigation eventually result in experimental evaluations of practical applications. Such programs of research should allow us to take pedagogical action with greater certainty and with a higher possibility of success.

REFERENCES

Adams, A., Carnine, D., & Gersten, R. (1982). Instructional strategies for studying content area texts in the intermediate grades. *Reading Research Quarterly*, *18*(1), 27–55. https://doi.org/10.2307/747537

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Allan, K. (2010). *The Western classical tradition in linguistics* (2nd ed.). Sheffield, UK: Equinox.

Anderson, R.C., Hiebert, E.H., Scott, J.A., & Wilkinson, I.A.G. (1985). *Becoming a nation of readers: The report of the Commission on Reading*. Washington, DC: National Institute of Education, U.S. Department of Education.

Apperson, S.V. (1940). The effectiveness of orthoptic training as a means of remedial instruction of reading. *The Journal of Experimental Education*, *9*(2), 160–166. https://doi.org/10.1080/00220973.1940.11010201

Araújo, S., Reis, A., Petersson, K.M., & Faísca, L. (2015). Rapid automatized naming and reading performance: A meta-analysis. *Journal of Educational Psychology*, *107*(3), 868–883. https://doi.org/10.1037/edu0000006

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Erlbaum.

Baker, S., Berninger, V.W., Bruck, M., Chapman, J., Eden, G., Elbaum, B., … Wolf, M. (2002). *Experts say Reading Recovery is not effective, leaves too many children behind: An open letter from reading researchers*. Retrieved from https://www.wrightslaw.com/info/read.rr.ltr.experts.htm

Barr, R. (1974). The effect of instruction on pupil reading strategies. *Reading Research Quarterly*, *10*(4), 555–582.

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Berninger, V.W., Lester, K., Sohlberg, M.M., & Mateer, C. (1991). Interventions based on the multiple connections model of reading for developmental dyslexia and acquired deep dyslexia. *Archives of Clinical Neuropscyhology*, *6*(4), 375–391.

Betts, E. (1946). *Foundations of reading instruction: With emphasis on differentiated guidance*. Chicago, IL: American Book.

Bond, G.L., & Bond, E. (1943). *Teaching the child to read*. New York, NY: Macmillan.

Byrne, B., Fielding-Barnsley, R., & Ashley, L. (2000). Effects of preschool phoneme identity training after six years: Outcome level distinguished from rate of response. *Journal of Educational Psychology*, *92*(4), 659–667. https://doi.org/10.1037/0022-0663.92.4.659

Calkins, L. (2020). *No one gets to own the term "the science of reading"*. New York, NY: Teachers College Reading and Writing Project, Columbia University. Retrieved from https://readingandwritingproject.org/news/no-one-gets-to-own-the-term-the-science-of-reading

Carbo, M. (1983). Research in reading and learning style: Implications for exceptional children. *Exceptional Children*, *49*(6), 486–494. https://doi.org/10.1177/001440298304900601

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Cattell, J.M. (1886). The time taken up by cerebral operations. *Mind*, *11*(44), 524–538. https://doi.org/10.1093/mind/os-XI.44.524

Chall, J.S. (1996). *Learning to read: The great debate* (3rd ed.). Fort Worth, TX: Harcourt Brace.

Chapman, J.W., & Tunmer, W.E. (1991). Recovering Reading Recovery. *Australia and New Zealand Journal of Developmental Disabilities*, *17*(1), 59–71. https://doi.org/10.1080/07263869100034271

Chapman, J.W., & Tunmer, W.E. (2016). Is Reading Recovery an effective intervention for students with reading difficulties? A critique of the i3 scale-up study. *Reading Psychology*, *37*(7), 1025–1042. https://doi.org/10.1080/02702711.2016.1157538

Chapman, J.W., Tunmer, W.E., & Prochnow, J.E. (2001). Does success in the Reading Recovery program depend on developing proficiency in phonological-processing skills? A longitudinal study in a whole language instructional context. *Scientific Studies of Reading*, *5*(2), 141–176. https://doi.org/10.1207/S1532799Xssr0502_2

Cohen, R.A., Mather, N., Schneider, D.A., & White, J.M. (2017). A comparison of schools: Teacher knowledge of explicit code-based reading instruction. *Reading and Writing*, *30*(4), 653–690. https://doi.org/10.1007/s11145-016-9694-0

Cunningham, J.W. (2001). The National Reading Panel report [Review]. *Reading Research Quarterly*, *36*(3), 326–335. https://doi.org/10.1598/RRQ.36.3.5

D'Agostino, J.V., & Harmey, S.J. (2016). An international meta-analysis of Reading Recovery. *Journal of Education for Students Placed at Risk*, *21*(1), 29–46. https://doi.org/10.1080/10824669.2015.1112746

D'Mello, A.M., & Gabrieli, J.D.E. (2018). Cognitive neuroscience of dyslexia. *Language, Speech, and Hearing Services in Schools*, *49*(4), 798–809. https://doi.org/10.1044/2018_LSHSS-DYSLC-18-0020

Durkin, D. (1980). *Teaching young children to read*. Boston, MA: Allyn & Bacon.

Experience. (1836). On teaching reading. In W.C. Woodbridge (Ed.), *American annals of education and instruction, for the year 1836* (Vol. 6, pp. 65–69). Boston, MA: John Allen.

Fagan, L.B. (1931). A case study of dextral training of a left-handed boy and its effect on speech, reading and writing. *The Psychological Clinic*, *19*, 291–293.

Fischer, F.W., Liberman, I.Y., & Shankweiler, D. (1978). Reading reversals and developmental dyslexia: A further study. *Cortex*, *14*(4), 496–510. https://doi.org/10.1016/S0010-9452(78)80025-2

Frostig, M., & Horne, D. (1964). *The Frostig program for the development of visual perception: Teacher's guide*. Chicago, IL: Follett.

Goodman, K.S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, *6*(4), 126–135. https://doi.org/10.1080/19388076709556976

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, *81*(4), 710–744. https://doi.org/10.17763/haer.81.4.t2k0m13756113566

Greaney, K. (2001). An investigation of teacher preferences for word identification strategies. *Australian Journal of Language and Literacy*, *24*(1), 21–30.

Greaney, K.T. (2011). The multiple cues or "searchlights" word reading theory: Implications for Reading Recovery®. *Perspectives on Language and Literacy*, *37*(4), 15–19.

Hanford, E. (2018, September 10). Hard words: Why aren't kids being taught to read? *APM Reports*. Retrieved from https://www.apmreports.org/episode/2018/09/10/hard-words-why-american-kids-arent-being-taught-to-read

Hanford, E. (2019, August 22). At a loss for words: How a flawed idea is teaching millions of kids to be poor readers. *APM Reports*. Retrieved from https://www.apmreports.org/episode/2019/08/22/whats-wrong-how-schools-teach-reading

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Heath, E.J., Cook, P., & O'Dell, N. (1976). Eye exercises and reading efficiency. *Academic Therapy*, *11*(4), 435–445. https://doi.org/10.1177/105345127601100408

Hebert, M., Bohaty, J.J., Nelson, J.R., & Brown, J. (2016). The effects of text structure instruction on expository text comprehension: A meta-analysis. *Journal of Educational Psychology*, *108*(5), 609–629. https://doi.org/10.1037/edu0000082

Henmon, V.A.C. (1912). The relation between mode of presentation and retention. *Psychological Review*, *19*(2), 79–96. https://doi.org/10.1037/h0072813

Holmes, J.A. (1965). Basic assumptions underlying the substrata-factor theory. *Reading Research Quarterly*, *1*(1), 5–28. https://doi.org/10.2307/746971

Huey, E.B. (1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press. (Original work published 1908)

Iversen, S., & Tunmer, W.E. (1993). Phonological processing skill and the Reading Recovery Program. *Journal of Educational Psychology*, *85*(1), 112–126. https://doi.org/10.1037/0022-0663.85.1.112

Iversen, S., Tunmer, W.E., & Chapman, J.W. (2005). The effects of varying group size on the Reading Recovery approach to preventive early intervention. *Journal of Learning Disabilities*, *38*(5), 456–472. https://doi.org/10.1177/00222194050380050801

Johnson, D., & Myklebust, H.R. (1967). *Learning disabilities: Educational principles and practices*. New York, NY: Grune & Stratton.

Joshi, R.M., Binks, E., Graham, L., Ocker-Dean, E., Smith, D.L., & Boulware-Gooden, R. (2009). Do textbooks used in university reading education courses conform to the instructional recommendations of the National Reading Panel? *Journal of Learning Disabilities*, *42*(5), 458–463. https://doi.org/10.1177/0022219409338739

Kavale, K.A., & Forness, S.R. (1985). Learning disability and the history of science: Paradigm or paradox? *Remedial and Special Education*, *6*(4), 12–24. https://doi.org/10.1177/074193258500600404

Kelly, B., & Perkins, D.F. (Eds.). (2012). *Handbook of implementation science for psychology in education*. New York, NY: Cambridge University Press.

Keogh, B.K., & Pelland, M. (1985). Vision training revisited. *Journal of Learning Disabilities*, *18*(4), 228–236. https://doi.org/10.1177/002221948501800410

Kephart, N.C. (1960). *The slow learner in the classroom*. Columbus, OH: Merrill.

Kim, J.S. (2019). Making every study count: Learning from replication failure to improve intervention research. *Educational Researcher*, *48*(9), 599–607. https://doi.org/10.3102/0013189X19891428

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Kim, Y.-S.G. (2020). Theory of mind mediates the relations of language and domain-general cognitions to discourse comprehension. *Journal of Experimental Child Psychology*, *194*, article 104813. https://doi.org/10.1016/j.jecp.2020.104813

Kirk, S.A., & Kirk, W.D. (1971). *Psycholinguistic learning disabilities: Diagnosis and remediation*. Chicago: University of Illinois Press.

Kuhn, M.R., & Stahl, S.A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, *95*(1), 3–21. https://doi.org/10.1037/0022-0663.95.1.3

Kushner, H.I. (2017). *On the other hand: Left hand, right brain, mental disorder, and history*. Baltimore, MD: Johns Hopkins University Press.

Luchow, J.P., & Shepherd, M.J. (1981). Effects of multisensory training in perceptual learning. *Learning Disability Quarterly*, *4*(1), 38–43. https://doi.org/10.2307/1510711

McDaniel, M.A., & Einstein, G.O. (2005). Material appropriate difficulty: A framework for determining when difficulty is desirable for improving learning. In A.F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 73–85). Washington, DC: American Psychological Association.

McKeown, M.G., Beck, I.L., Omanson, R.C., & Pople, M.T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly*, *20*(5), 522–535. https://doi.org/10.2307/747940

Meyer, B.J.F. (1975). *The organization of prose and its effects on memory*. Amsterdam, Netherlands: North-Holland.

Morphett, M.V., & Washburne, C. (1931). When should children begin to read? *The Elementary School Journal*, *31*(7), 496–503. https://doi.org/10.1086/456609

Murphy, R. (with Heiting, G.). (2017, April). Are learning-related vision issues holding your child back? *All About Vision*. Retrieved from https://www.allaboutvision.com/parents/learning.htm

National Center for Education Statistics. (2013). *The Nation's Report Card: Trends in academic progress 2012* (NCES 2013-456). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.

National Education Policy Center & Education Deans for Justice and Equity. (2020). *Policy statement on the "science of reading"*. Boulder: National Education Policy Center, University of Colorado Boulder. Retrieved from https://nepc.colorado.edu/publication/fyi-reading-wars

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Washington, DC: National Institute of Child Health and Human Development.

National Research Council. (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Ober, T.M., Brooks, P.J., Homer, B.D., & Rindskopf, D. (2020). Executive functions and decoding in children and adolescents: A meta-analytic investigation. *Educational Psychology Review*. Advance online publication. https://doi.org/10.1007/s10648-020-09526-0

Orton, S.T. (1925). "Word-blindness" in school children. *Archives of Neurology and Psychiatry*, *14*(5), 581–615. https://doi.org/10.1001/archneurpsyc.1925.02200170002001

Pearson, P.D. (2004). The reading wars. *Educational Policy*, *18*(1), 216–252. https://doi.org/10.1177/0895904803260041

Pinker, S. (1984). Visual cognition: An introduction. *Cognition*, *18*(1–3), 1–63. https://doi.org/10.1016/0010-0277(84)90021-0

Pinker, S. (1999). *Words and rules: The ingredients of language*. New York, NY: Basic.

Pitman, I. (1843). Report of the proceedings at the Birmingham Phonographic Festival, held in the assembly room of Dee's Royal Hotel, July 18, 1843. *The Phonotypic Journal*, *2*(20), 109–120.

Pryzwansky, W.B. (1972). Effects of perceptual-motor training and manuscript writing on reading readiness skills in kindergarten. *Journal of Educational Psychology*, *63*(2), 110–115. https://doi.org/10.1037/h0032648

Pyle, N., Vasquez, A.C., Lignugaris/Kraft, B., Gillam, S.L., Reutzel, D.R., Olszewski, A., … Pyle, D. (2017). Effects of expository text structure interventions on comprehension: A meta-analysis. *Reading Research Quarterly*, *52*(4), 469–501. https://doi.org/10.1002/rrq.179

Ransford, M. (2008, March 4). Animals that hover. *Popular Science*. Retrieved from https://www.popsci.com/scitech/article/2008-03/animals-hover/

Ransford-Kaldon, C.R., Flynt, E.S., Ross, C.L., Franceschini, L., Zoblotsky, T., Huang, Y., & Gallagher, B. (2010). *Implementation of effective intervention: An empirical study to evaluate the efficacy of Fountas & Pinnell's Leveled Literacy Intervention system (LLI)*. Memphis, TN: Center for Research in Educational Policy, University of Memphis.

Rayner, K., Abbott, M.J., Schotter, E.R., Belanger, N.N., Higgins, E.C., Leinenger, M., … Plummer, P. (2015). *Keith Rayner eye movements in reading data collection*. San Diego: University of California, San Diego Library Digital Collections. https://doi.org/10.6075/J0JW8BSV

Rayner, K., Binder, K.S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research*, *41*(7), 943–954. https://doi.org/10.1016/S0042-6989(00)00310-2

RAND Reading Study Group. (2002). *Reading for understanding: Toward a R&D program in reading comprehension*. Santa Monica, CA: RAND.

Robinson, H.M., Mozzi, L., Wittick, M.L., & Rosenbloom, A.A. (1960). Children's perceptual achievement forms: A three year study. *American Journal of Optometry and Archives of American Academy of Optometry*, *37*(5), 223–237. https://doi.org/10.1097/00006324-196005000-00001

Rogowsky, B.A., Calhoun, B.M., & Tallal, P. (2020). Providing instruction based on students' learning style preferences does not improve learning. *Frontiers in Psychology*, *11*, article 164. https://doi.org/10.3389/fpsyg.2020.00164

Roser, N.L., & Weintraub, S. (2009). From Gray to Google: Learning within a profession. In J.V. Hoffman & Y.M. Goodman (Eds.), *Changing literacy for changing times: An historical perspective on the future of reading research, public policy, and classroom practices* (pp. 17–33). New York, NY: Routledge.

Samuels, S.J. (1970). *An experimental program for teaching letter names of the alphabet*. Washington, DC: Office of Education, U.S. Department of Health, Education, and Welfare.

Seidenberg, M.S. (2013). The science of reading and its educational implications. *Language Learning and Development*, *9*(4), 331–360. https://doi.org/10.1080/15475441.2013.812017

Seidenberg, M. (2017). *Language at the speed of sight: How we read, why so many can't, and what can be done about it*. New York, NY: Basic.

Shanahan, T. (1987). *The Early Detection of Reading Difficulties* (3rd ed.), by Marie Clay [Book review]. *Journal of Reading Behavior*, *19*(1), 117–119.

Shanahan, T. (2019). Why children should be taught to read with more challenging texts. *Perspectives on Language and Literacy*, *45*(4), 17–19, 22–23.

Shanahan, T. (2020). Limiting children to books they can already read. *American Educator*, *44*(2), 13–17, 39.

Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly*, *30*(4), 958–996. https://doi.org/10.2307/748206

Shanahan, T., Callison, K., Carriere, C., Duke, N.K., Pearson, P.D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: A practice guide* (NCEE 2010-4038). Washington, DC: National Center for Education Evalua-

19362722, 2020, S1, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.349 by Columbia University Libraries, Wiley Online Library on [07/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

tion and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, *78*(1), 40–59. https://doi.org/10.17763/haer.78.1.v62444321p602101

Singer, H. (1981). Teaching the acquisition phase of reading development: An historical perspective. In O.J.L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 291–311). Hillsdale, NJ: Erlbaum.

Snow, C.E., Burns, M.S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Snowling, M.J., & Hulme, C. (2005). *The science of reading: A handbook*. Malden, MA: Blackwell.

Snyder, R.T., & Freud, S.L. (1967). Reading readiness and its relation to maturational unreadiness as measured by the spiral aftereffect and other visual-perceptual techniques. *Perceptual and Motor Skills*, *25*(3), 841–854. https://doi.org/10.2466/pms.1967.25.3.841

Spritz. (n.d.). *Reading reimagined*. Retrieved from https://spritz.com/

Stahl, S.A., & Fairbanks, M.M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, *56*(1), 72–110. https://doi.org/10.3102/00346543056001072

Stanovich, K.E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*(1), 32–71. https://doi.org/10.2307/747348

Stanovich, K.E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York, NY: Guilford.

Stewart, C. (1809). *A descriptive catalogue of the Oriental library of the late Tippoo sultan of Mysore*. Cambridge, MA: University Press.

Taylor, D. (1998). *The spin doctors of science: The political campaign to change America's mind about how children learn to read*. Urbana, IL: National Council of Teachers of English.

Temple, E., Deutsch, G.K., Poldrack, R.A., Miller, S.L., Tallal, P., Merzenich, M.M., & Gabrieli, J.D.E. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(5), 2860–2865. https://doi.org/10.1073/pnas.0030098100

U.S. Department of Education. (2000). *Reading Excellence Act*. Retrieved from https://www2.ed.gov/pubs/promisinginitiatives/rea.html

U.S. Department of Education. (2002a). *Early Reading First*. Retrieved from https://www2.ed.gov/programs/earlyreading/index.html

U.S. Department of Education. (2002b). *No Child Left Behind: Elementary and Secondary Education Act (ESEA)*. Retrieved from https://www2.ed.gov/nclb/landing.jhtml

U.S. Department of Education. (2002c). *Reading First*. Retrieved from https://www2.ed.gov/programs/readingfirst/legislation.html

U.S. Department of Education. (2005). *Striving Readers*. Retrieved from https://www2.ed.gov/programs/strivingreaders/resources.html

U.S. Department of Education. (2015a). *Every Student Succeeds Act (ESSA)*. Retrieved from https://www.ed.gov/ESSA

U.S. Department of Education. (2015b). *Striving Readers Comprehensive Literacy Program*. Retrieved from https://www2.ed.gov/programs/strivingreaders-literacy/index.html

Venezky, R.L. (1995). From orthography to psychology to reading. In V.W. Berninger (Ed.), *The varieties of orthographic knowledge* (pp. 23–46). New York, NY: Springer.

Wasik, B.A., & Snell, E.K. (2019). Synthesis of preschool dosage: How quantity, quality, and content impact child outcomes. In A.J. Reynolds (Ed.), *Sustaining early childhood learning gains: Program, school, and family influences* (pp. 31–51). New York, NY: Cambridge University Press.

Werner, H., & Strauss, A. (1939). Types of visuomotor activity in their relation to low and high performance ages. *Proceedings of the American Association on Mental Deficiency*, *44*(1), 163–168.

What Works Clearinghouse. (2008, December). *Reading Recovery*® [WWC intervention report]. Washington, DC: Institute of Education Sciences. Retrieved from https://readingrecovery.org/wp-content/uploads/2016/12/wwc_reading_recovery_report_08.pdf

Wijekumar, K.K., Meyer, B.J.F., & Lei, P. (2017). Web-based text structure strategy instruction improves seventh graders' content area reading comprehension. *Journal of Educational Psychology*, *109*(6), 741–760. https://doi.org/10.1037/edu0000168

Williams, J.P. (1969). Training kindergarten children to discriminate letter-like forms. *American Educational Research Journal*, *6*(4), 501–514. https://doi.org/10.3102/00028312006004501

Wood, K. (1994). Reading Recovery: What is it? How effective is it? *Educational Psychology in Practice*, *10*(1), 3–13. https://doi.org/10.1080/0266736940100101

Submitted April 20, 2020
Final revision received July 3, 2020
Accepted July 16, 2020

TIMOTHY SHANAHAN is a Distinguished Professor Emeritus at the University of Illinois at Chicago, USA; email shanahan@uic.edu. His research focuses on how to improve reading achievement, disciplinary literacy, and reading–writing relations.