Economics from the Top Down

new ideas in economics and the social sciences

Why Kids Are Getting Worse at Reading: The Case Against Whole-Language Teaching

Blair Fix

March 28, 2026

CC BY

It's not even an inability to critically think.
It's an inability to read sentences.

— Jessica Hooten Wilson

When it comes to reading, there is something of a moral panic afoot. In the United States, high-school reading scores are tanking . . . and everyone seems to know why.

Raised on a diet of internet slop, today's kids think that 'reading' means scanning the captions of a Tik Tok video. For them, books are becoming incomprehensible relics of a waning literate age. In short, literacy is being murdered, and the killer is the smartphone in every child's hand.

Or is it?

In this essay, I explore a different story about why kids are becoming less literate. It's not a story about kids getting dumbed down by an addictive new technology. It's a story about how adults decided to *not* teach kids how to use a piece of very old technology.

Let me set the stage with a brief parable.

Several millennia ago, humans invented a clever three-step algorithm for encrypting messages. In step one, the user takes a message and decomposes it into a set of distinct sounds. In step two, the user encrypts these sounds into a set of visual symbols. And in step three, the user preserves these symbols on a physical medium.

The purpose of this encryption technology is to transmit meaning across time and space. When another user encounters the encrypted message, they decrypt it by employing the same algorithm in reverse. First, they parse the symbols and convert them into sounds. Next, they parse these sounds and group them into chunks of meaning. Finally, they interpret the decoded message.

With this ancient technology in mind, here is how our story unfolds. For centuries after its invention, the encryption algorithm was used by a small class of administrators who hoarded its secrets. Then, about three hundred years ago, the decryption keys were gradually released to the wider population. Eventually, all children received mandatory decryption training.

But then a funny thing happened. After several generations of formal training, some users began to feel that the decryption algorithm was 'natural', and that its intricacies need not be taught. These users created a new approach to training whereby the decryption algorithm was learned through exposure. New users were shown encrypted messages with convenient pictures and cues that elucidated the meaning. The idea was that through repeated exposure, new users would master the technology.

Or not.

After several decades of this new learning regime, scientists found that decryption skills were in decline, and that the drop was biggest among the weakest decrypters. Meanwhile, new communication technology had proliferated, leading to confusion about the loss of decryption skill. Was the new technology to blame? Or was it something else? Opinions flared, but conclusive evidence was in short supply.

Exiting our parable, let's take stock. In English, the act of encrypting a message is called *writing*. And the act of decrypting it is called *reading*. Both acts rely on an implicit algorithm for decomposing linguistic meaning into sounds, and encoding these sounds into symbols. Good readers have mas-

tered this algorithm. Bad readers have not. And today, bad readers have not mastered this algorithm largely because they were not taught how it works. Or at least, that is my contention here.

The backstory is that in the 1980s and 1990s, a movement called the 'whole language' approach to reading swept through anglophone schools. According to this new zeitgeist, learning to read was as 'natural' as learning to speak. It was a skill that could be learned largely by exposure. Soon this method of 'vibe-based literacy' came to dominate elementary-school pedagogy, with devastating results. Yes, some students flourished. But a large portion of kids were simply left behind, destined to be perpetually poor readers. Today, we are living with the consequences.

In this essay, I make the case that the modern decline of high-school reading ability is in large part, due to the spread of whole-language methods. I will build my case in three parts.

In Part I, I review how reading scores have declined among US high-school students. I show how this decline is not uniform, but is instead marked by a widening skill gap. Among US high-school students, the best readers have actually gotten better over time, but the worst readers have gotten far worse. Importantly, this reading-score gap remains visible across a variety of student demographics. Which is to say that whatever is causing the widening gap, it's not something that government surveys measure.

In Part II, I take a detour into the history of how writing was invented. My purpose is to illustrate exactly why reading is hard to learn, and why students benefit when the requisite skills are explicitly taught. Then I discuss the whole-language movement to *not* teach these skills — a movement based on a misguided view of what it means to read.

In Part III, I build the case that whole-language instruction is the main cause of the widening reading-score gap. I survey many lines of evidence, perhaps the most important of which is that the reading-score gap can be reversed . . . by abandoning whole-language methods and teaching structured literacy.

Of course, my goal here is not to exonerate smartphones and other screen devices from wasting kids' time. Instead, my point is that if we don't properly teach kids *how* to read, they have little chance of actually doing it.

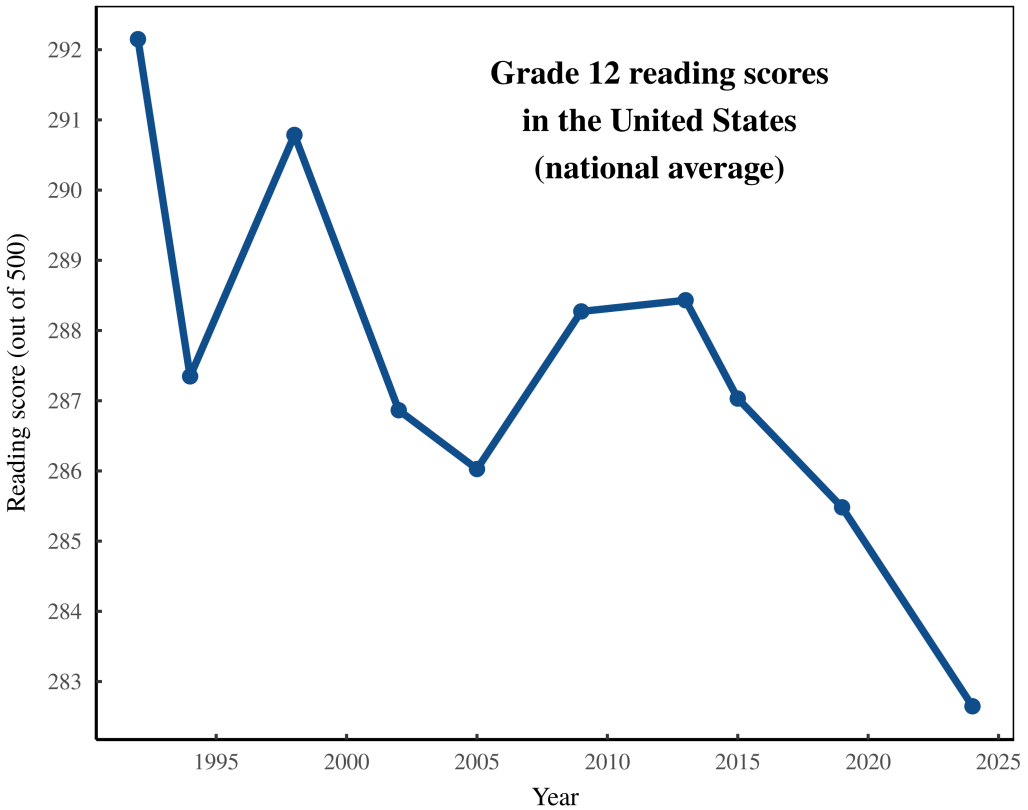

Figure 1: The decline of US high-school reading scores

Since the 1990s, US twelfth-grade students have seen their reading performance drop on the NAEP standardized test. Sources and methods

---

Part I: The decline of high-school reading ability

Our story begins with a much-discussed piece of evidence. Among US high-school students, reading ability is in decline. Figure 1 shows the trend, as captured by the National Assessment of Educational Progress (NAEP).

Looking at this reading-score data, notice that performance drops conspicuously during the 2010s. To many observers, the timing of this drop implicates smartphones, which became popular in the same decade. To be sure, almost no one thinks that smartphones are a panacea for teen literacy (myself included). Still, I see three main problems with the rush to blame declining reading ability solely on phones.

The first problem is that the pandemic probably played a role in worsening reading scores during the latest batch of tests in 2024. (And without this data point, the recent reading-score decline is less conspicuous.) The second

problem is that if we stare more closely at the data, we see that the reading-score decline began in the 1990s, long before smartphones were invented. And the third problem is that the decline in reading scores is not uniform among all students. As we'll soon see, what appears like a universal decline in reading scores is in fact a differential *wedge* — a widening gap between the best and worst high-school readers.

The widening reading-score gap

When faced with a conspicuous trend (like the decline in reading scores shown in Figure 1), a common impulse is to begin searching for the cause. But a wiser approach is to first dissect the trend itself to better understand what's going on.

In the case of high-school reading ability, the decline in the average US score suggests that reading skills have fallen uniformly across all students. However, when we look beneath the average score, a more complicated pattern emerges — a pattern of widening gaps.

Figure 2 illustrates this reading-score wedge. Here, I've taken the trend in high-school reading scores and decomposed the data by reading-score percentile. In this chart, the best readers live at the top, and the worst readers live at the bottom. Looking at the data, what leaps off the page is its *lack* of uniformity. Among the best students, reading scores have actually improved with time. But among the worst students, reading scores have collapsed. It's this low-end decline that's driving the downward trend in average reading scores.

Switching to a snapshot of cumulative change, Figure 3 shows how US high-school reading scores have diverged across percentiles between 1992 and 2024. (The vertical axis shows the change in score as a function of reading-score percentile.) The transformation is quite shocking. For students *above* the 85th percentile, reading scores improved over this period, with the steepest gains at the top. But for students *below* the 85th percentile, reading scores worsened with time, with the steepest losses at the bottom. In short, over the last three decades, US high-school reading scores have been marked by a widening gap between the best and the worst readers.

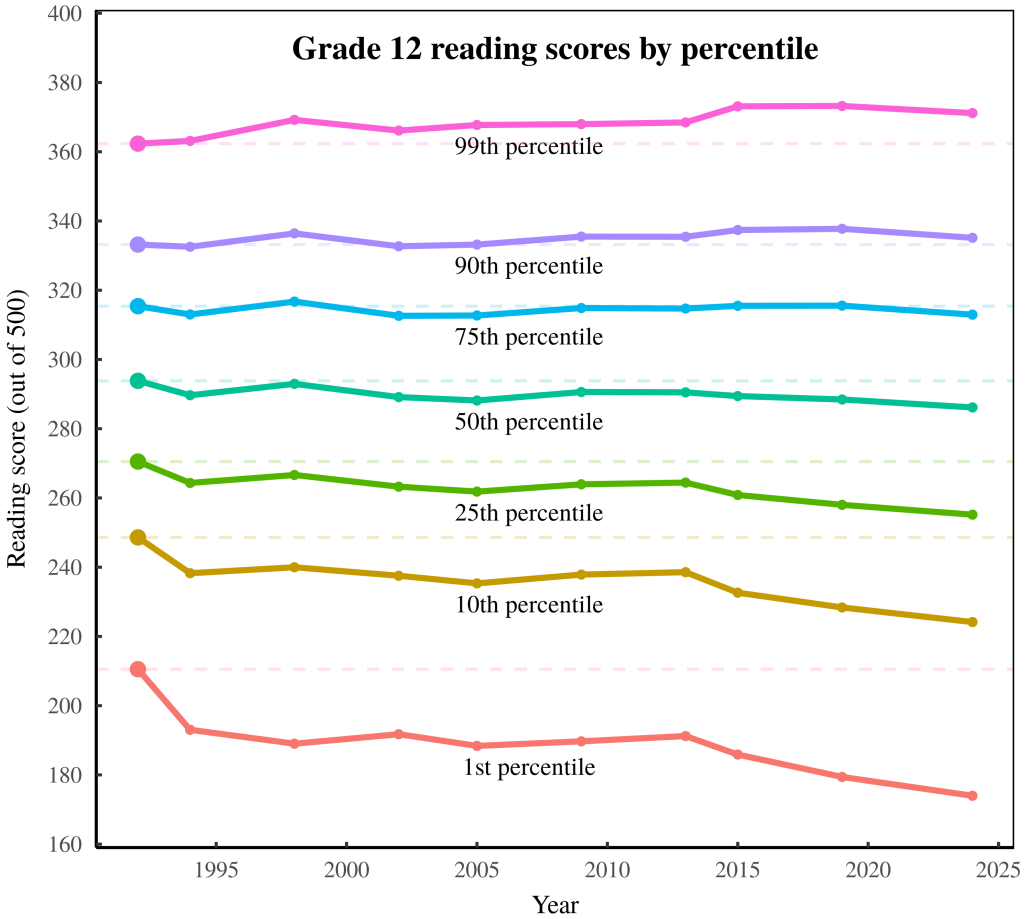

Figure 2: An widening reading-score gap among US high-school students

Over the last three decades, the best US high-school students have gotten slightly better at reading, while the worst students have gotten far worse. This chart shows the widening achievement gap among selected reading-score percentiles. Sources and methods

---

The widening reading-score gap persists across a variety groups

Looking ahead, I'm going to argue that the cause of the widening reading-score gap is something that's not captured by the federal survey data (collected by the NAEP). To make the case, I am going to segment students into a variety of demographic groups based on characteristics like student sex, parent education, school absences, TV use, and reading habits. Student sex aside, these are characteristics that relate strongly to reading ability. And yet without exception, I find that the reading-score gap — the widening gap between the best and the worst readers — persists *within* these groups.

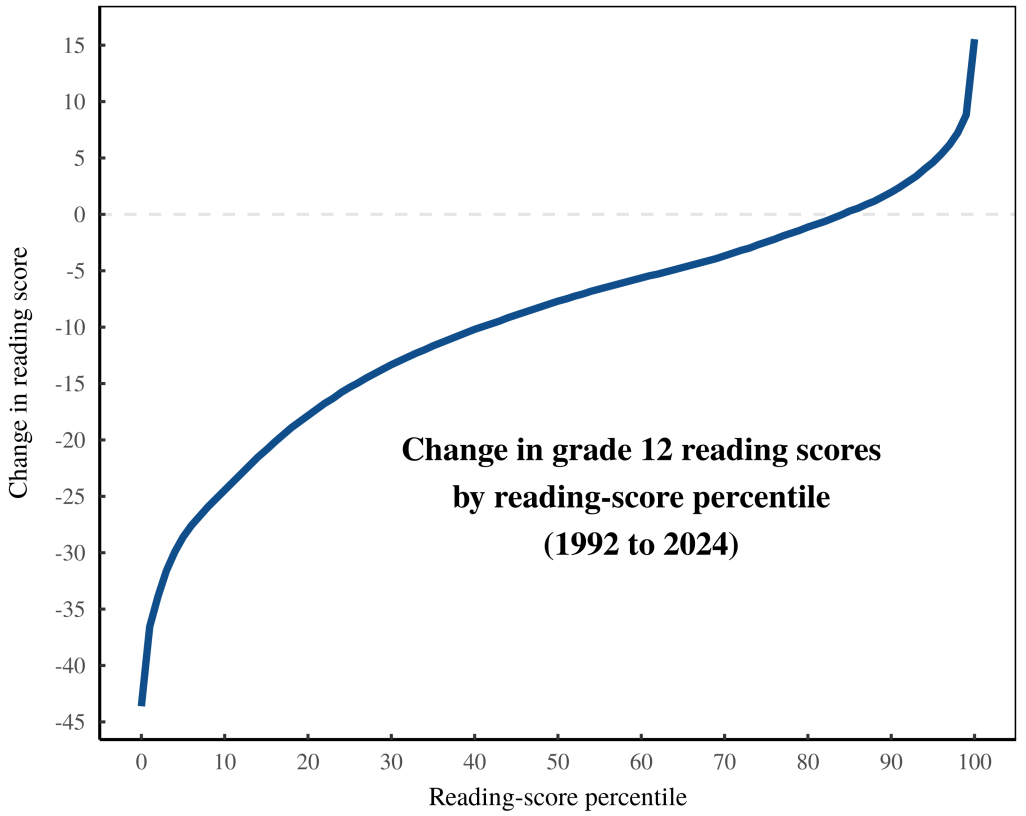

Figure 3: Cumulative change in grade 12 reading scores by percentile, 1992 to 2024

The vertical axis measures the cumulative change in twelfth-grade reading scores as a function of reading score percentile (horizontal axis) over the last three decades. Sources and methods

---

The widening reading-score gap by student sex

Looking at various demographic categories, let's start with student sex. In recent years, there's been much worry about the failure of boys. The reading score data supports this worry, but with some caveats.

As Figure 4 demonstrates, over the last three decades, both sexes have seen their worst high-school readers get worse and their best readers get better. However, this widening reading-score gap is more pronounced among males then among females. So the message is not that boys in general are getting worse at reading. The message is that something is driving a wedge between the best and worst readers, and this wedge is thicker among boys than among girls.

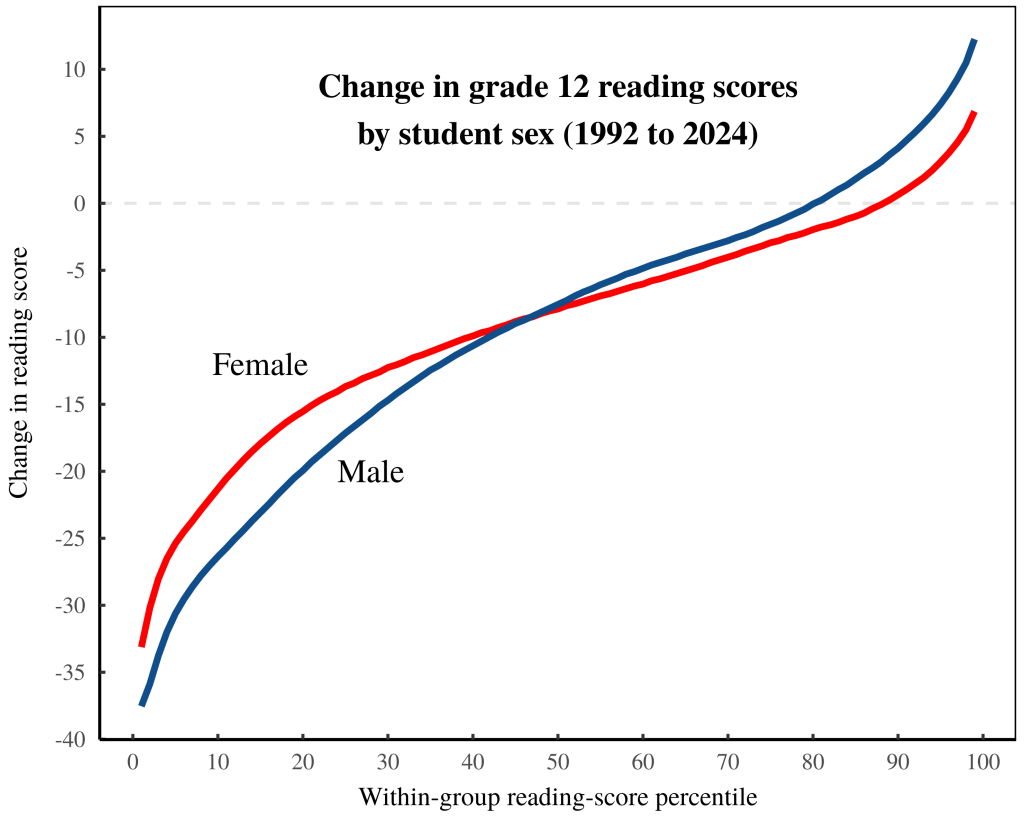

Figure 4: The widening reading-score gap occurs among both sexes

This chart shows the change in twelfth-grade reading scores between 1992 and 2024, isolated by student sex. The horizontal axis shows reading-score percentile within each sex. The vertical axis shows the corresponding change in test score over the last three decades. Note that the widening reading-score gap exists within both sexes, but is more severe among males. Sources and methods

---

The widening reading-score gap by parent education

Let's move on to demographic factors that are well known to affect student success. Parent education is a big one. As Figure 5 shows, students with more educated parents tend to be better readers, likely because educated parents care more about their kids' schooling, and they have more time and money to invest in learning. That said, when we switch to measuring the *change* in reading scores over the last three decades, a different picture emerges — one in which parent education is largely irrelevant.

Figure 6 shows the pattern. Here, I've grouped high-school kids by their parents' education, and then measured the change in reading scores within each group between 1992 and 2024. The picture that emerges is one of

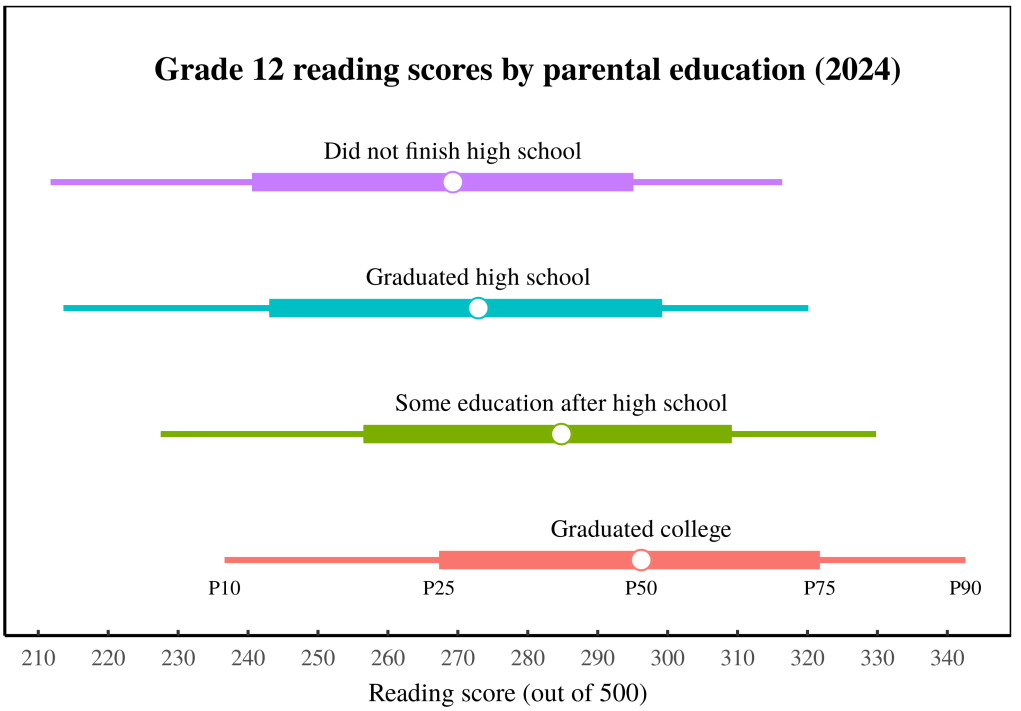

Figure 5: High-school reading scores by parental education in 2024

This chart shows the spread in twelfth-grade reading scores as a function of parental education (the highest level of education attained by either parent). As expected, students with more educated parents tend to be better readers. Sources and methods

demographic intransigence. Regardless of their parents' education, the best high-school readers have gotten better and the worst readers have gotten worse.

The widening reading-score gap by school absenteeism

Continuing our demographic journey, kids who miss more school tend to be worse students, likely because school works best if you actually go. So it's unsurprising that high-school students with more absences tend to be worse readers, as Figure 7 shows. Yet student absences are *not* what's driving the widening gap between the best and worst readers.

Figure 8 shows the evidence. Here, I've grouped high-school students by their school absences, and then measured the change in reading score between 2002 and 2024. Again, we see a pattern of intransigence. Regardless of school absence rates, the best high-school readers have gotten better, and the worst readers have gotten worse.

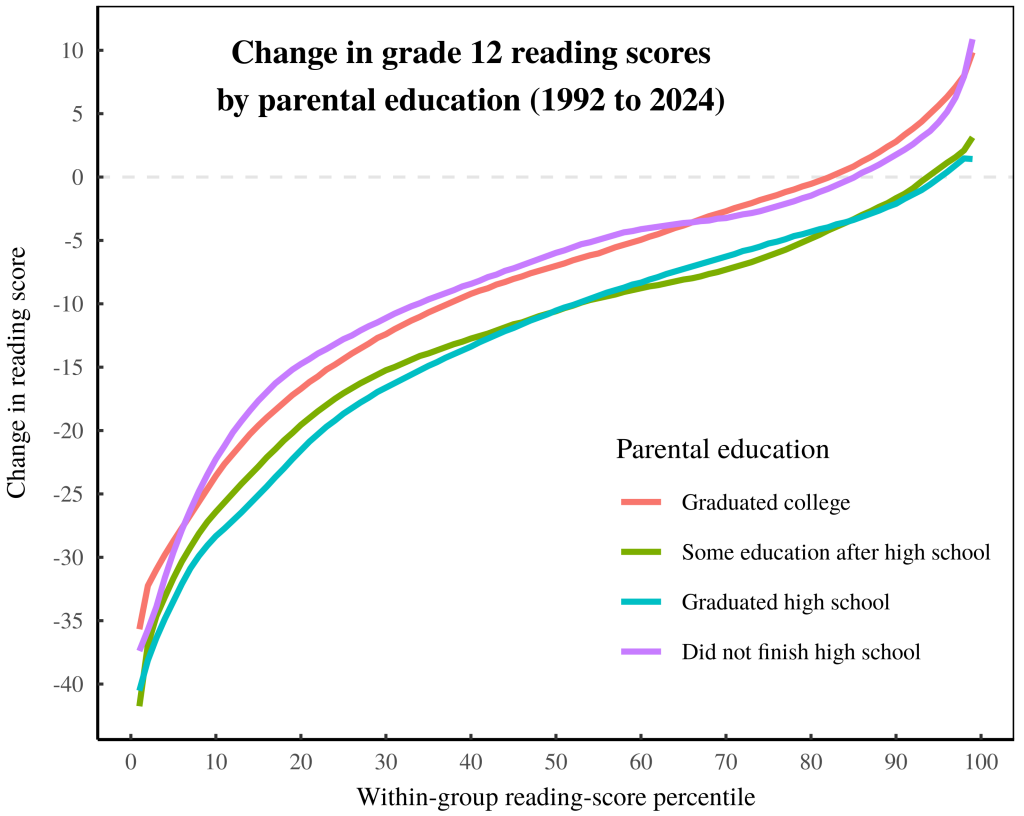

Figure 6: The widening reading-score gap occurs within all parental education groups

This chart shows the change in twelfth-grade reading scores between 1992 and 2024, isolated by parental education. The horizontal axis shows reading-score percentile within each parental education group. The vertical axis shows the corresponding change in test score over the last three decades. Sources and methods

---

The widening reading-score gap by TV use

Now to some demographic characteristics that reflect students' intellectual behavior. First up is TV use. As we might expect, students who spend more time watching TV tend to be worse readers. Figure 9 shows the disparity in 1998.

So how do TV habits relate to the *change* in reading score over time? Here, the problem is that the NAEP only surveyed students' TV habits during the 1990s, so we don't know much about the long-term trend. That said, the pattern during the 1990s shows similar signs of demographic intransigence.

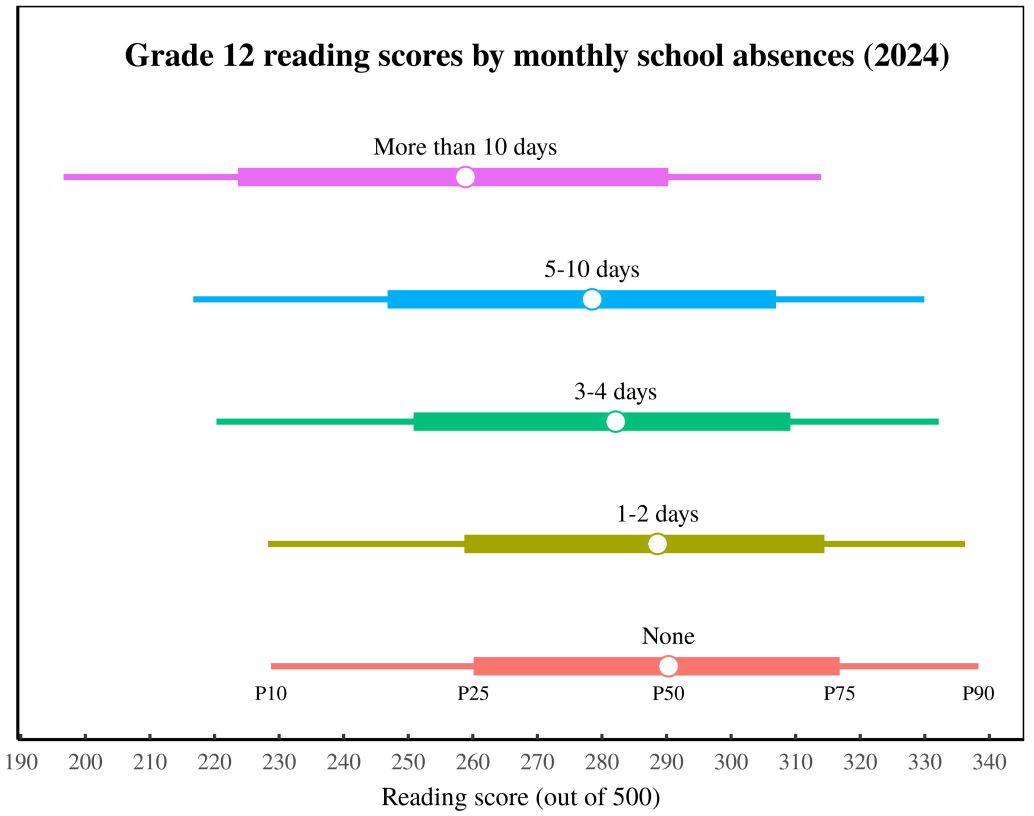

Figure 7: High-school reading scores by monthly school absences in 2024

This chart shows the spread in twelfth-grade reading scores as a function of the students' monthly school absences. As expected, students with fewer absences tend to be better readers. Sources and methods

---

Figure 10 illustrates. Here, I've grouped high-school students by their TV habits, and then measured the change in reading scores between 1992 and 1998. Again, we find a widening gap between the best and worst readers, and we find that the gap is largely unaffected by kids' TV habits.

The widening reading-score gap by pleasure-reading habits

Now to what is perhaps the most obvious trait of a good reader . . . choosing to actually read. As Figure 11 shows, kids who spend more time reading for pleasure also tend to be better readers. Shocking!

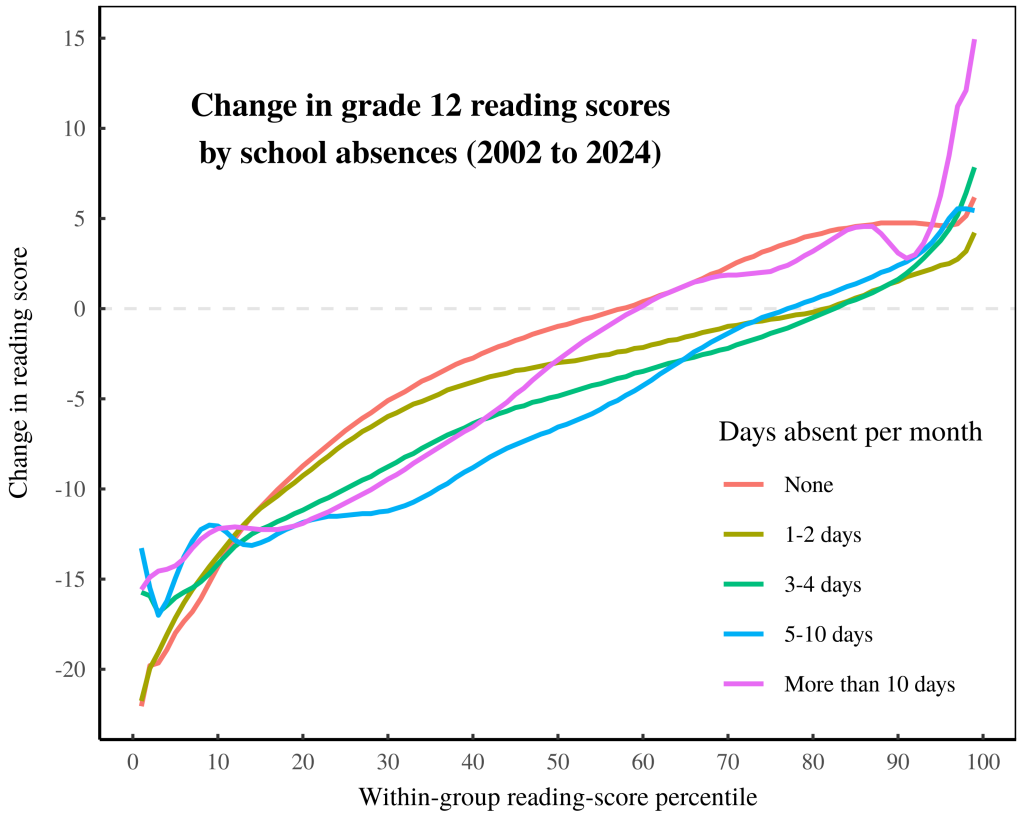

Figure 8: The widening reading-score gap occurs within all school absenteeism groups

This chart shows the change in twelfth-grade reading scores between 2002 and 2024, isolated by students' monthly school absences. The horizontal axis shows reading-score percentile within each absenteeism group. The vertical axis shows the corresponding change in test score over the last two decades. Sources and methods

---

Sarcasm aside, it's now well established that Americans are spending less time reading during their spare time. And since doing something is the best way to maintain a skill, it follows that if kids are reading less, this lack of practice might make them worse at reading.

Unfortunately, the available evidence shoots down this otherwise plausible theory for why high-school reading scores have declined. When we group kids by their reading habits and measure the change in score over time, we find a widening reading-score gap *within* all groups.

Figure 12 shows the pattern. Now the caveat here is that this data is limited to the 1990s. (For some reason, the NAEP stopped surveying students about their reading habits at the very moment when time spent reading started to

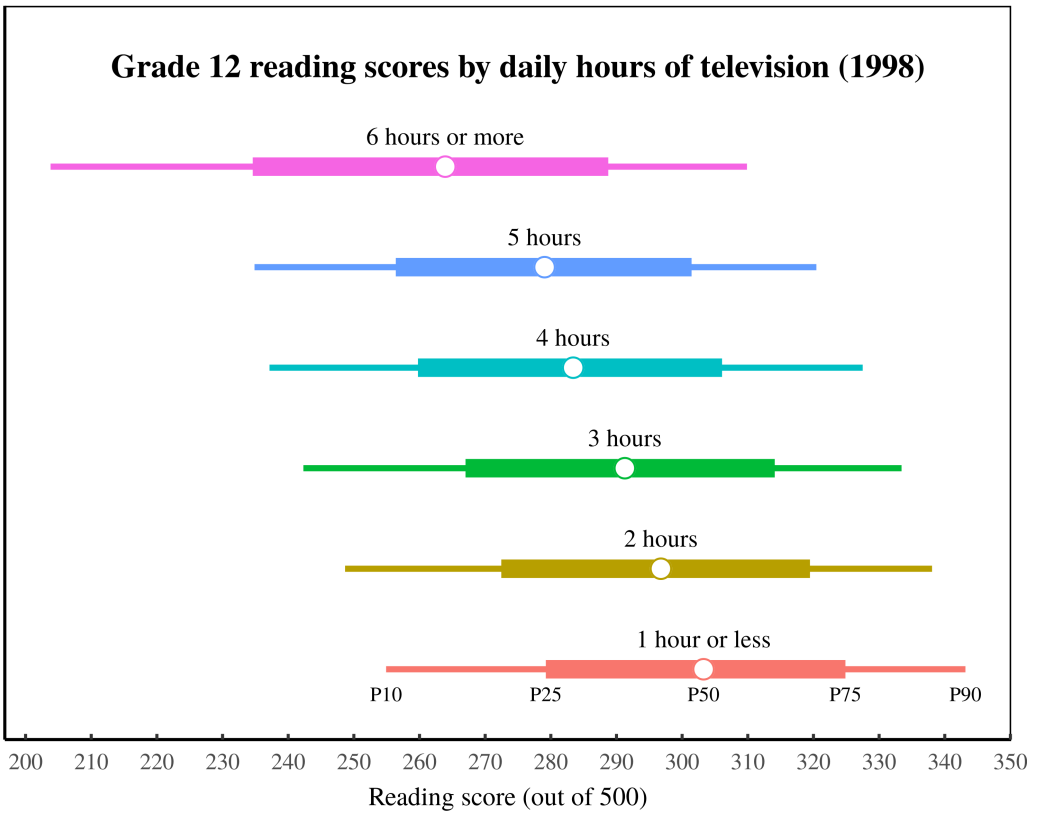

Figure 9: High-school reading scores by student TV use in 1998

This chart shows the spread in twelfth-grade reading scores as a function of students' TV use on school days. As expected, students who watch less TV tend to be better readers. Sources and methods

---

drop precipitously.) At any rate, we can conclude that whatever is driving a wedge between the best and worst readers, it's likely not their pleasure reading habits.

What's driving the reading-score gap?

This concludes my tour of the US read-score data. Let me summarize the main results. Over the last three decades, high-school reading scores have declined, but in a non-uniform way. Over this period, the best high-school readers got better, while the worst readers got far worse. And perhaps most importantly, this widening reading-score gap persists across a variety of demographic groups — groups which are well known to affect school outcomes. In short, something is driving a wedge between the best and worst readers, and this driver is not captured by the federal survey data.

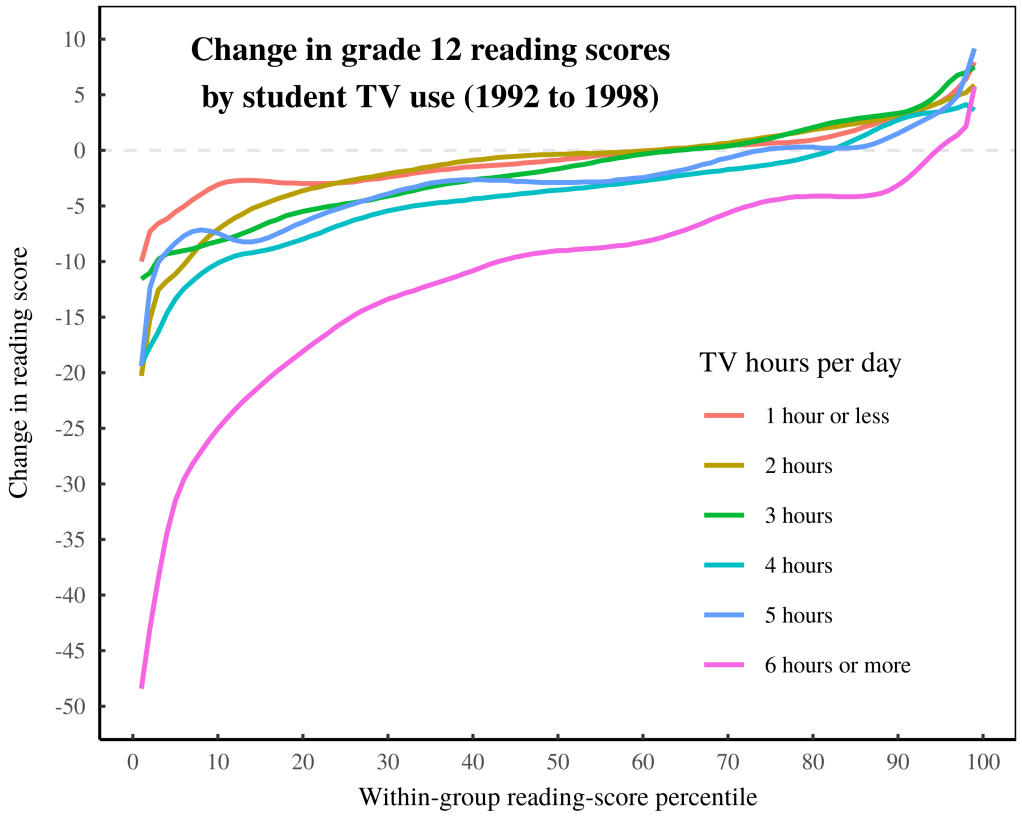

Figure 10: The widening reading-score gap occurs within all TV-use groups

This chart shows the change in twelfth-grade reading scores between 1992 and 1998, isolated by students' TV habits. The horizontal axis shows reading-score percentile within each TV-use group. The vertical axis shows the corresponding change in test score over the six-year period. Sources and methods

---

Now, it might seem that this lack of survey evidence leaves a world of possible causes. But then again, much is known about how and why kids fail at reading. Indeed, a large body of research shows that when kids struggle to read, it's usually because they were not taught the low-level skills that are required. And here, the cruel irony is that for the last three decades, many anglophone educators thought it wise to *not* teach these skills. That's because they were guided by the whole-language approach to teaching, which preached that reading could be learned largely through exposure.

We'll get to the spread of whole-language teaching in Part III. But first, we need some prior knowledge. We need to understand how reading works, why it is hard to learn, and why it's disastrous to not teach kids how to operate the encryption algorithm.

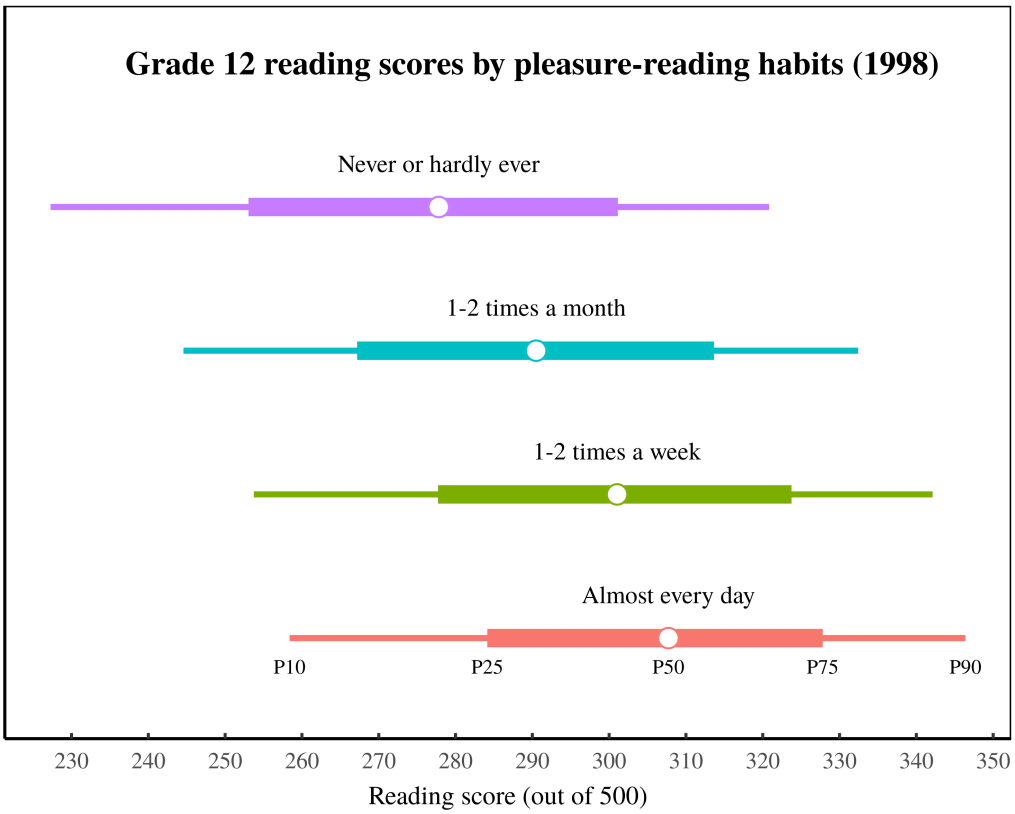

Figure 11: High-school reading scores by student pleasure-reading habits in 1998

This chart shows the spread in twelfth-grade reading scores as a function of students' pleasure-reading habits. As expected, students who read more tend to be better readers. Sources and methods

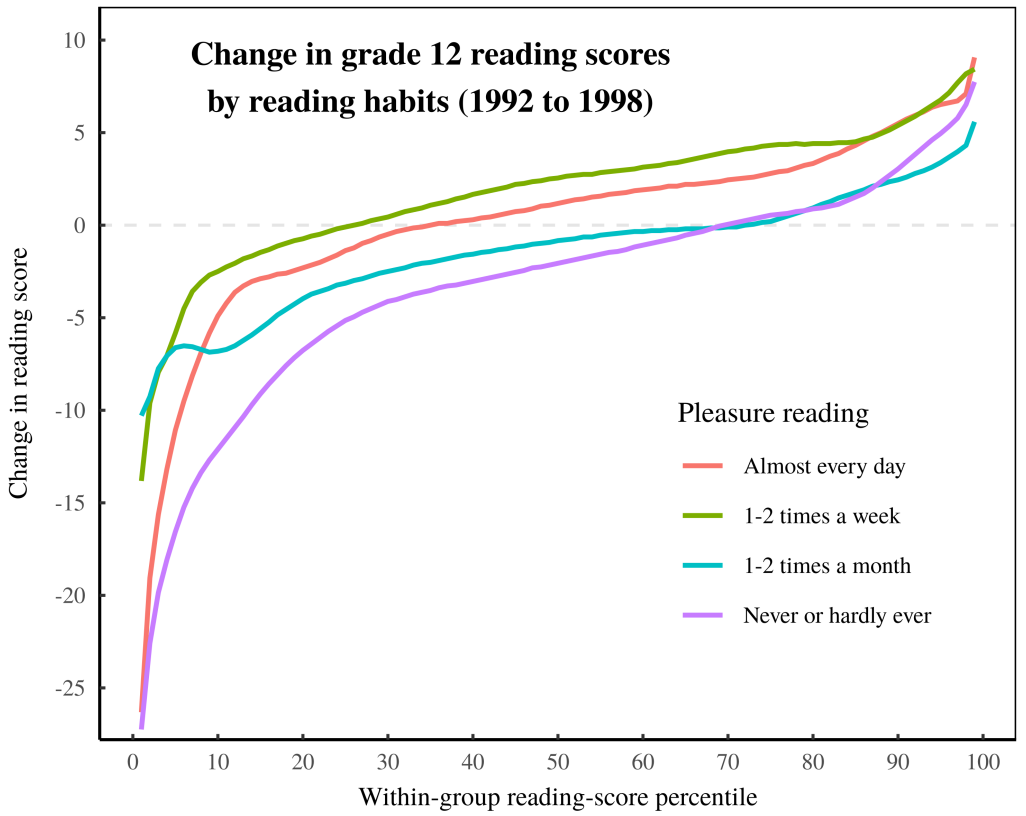

Figure 12: The widening reading-score gap occurs within all reading-habit groups

This chart shows the change in twelfth-grade reading scores between 1992 and 1998, isolated by students' pleasure-reading habits. The horizontal axis shows reading-score percentile within each reading-habit group. The vertical axis shows the corresponding change in test score over the six-year period. Sources and methods

Part II: Why reading is hard to learn

In literate cultures, writing is so ubiquitous that few people realize that it is a form of technology, and a miraculous one at that. Writing is an ingenious technique for encrypting spoken language into visual symbols — symbols that can then be decrypted by another person.

The advantages of writing over speaking are profound. A spoken sentence is gone the moment it is uttered. A written sentence can be preserved indefinitely. A spoken sentence requires a living speaker. A written sentence can be read long after the writer is dead, and in places the writer never imagined. In short, reading and writing come with many advantages over spoken language. But they also have one big downside, which is that compared to speaking, learning to read and write is far more difficult.

There's a reason for this difficulty — a reason that every grade-school teacher should have plastered on their wall. As the linguist Mark Liberman observes, reading is hard to learn "*for the same reasons that writing was hard to invent*". The lesson here is that each time a child struggles to read, they are replicating, in shortened form, humanity's lengthy struggle to invent the written word.

To get a sense for this formative struggle, realize that humans developed spoken language as early as 135,000 years ago. Yet formal writing dates back a mere 5,000 years. Now, I say 'formal' writing, because long before humans encoded words into abstract symbols, we drew pictures that conveyed meaning. For example, more than 40,000 years ago, humans in present-day Spain drew a picture of a bull that is instantly recognizable to anyone today. It is with such pictures that writing begins.

For the first proto-writers, the most obvious way to encode meaning was to draw an image of the thing they had in mind. To elicit the idea of a bull, they drew a bull. To elicit the idea of a snake, they drew a snake. And so on. Now, the intriguing part is that even today, drawing pictures remains the most obvious way to encode meaning.

For example, if you show a toddler a picture of a snake, they'll happily shout 'snake'. Indeed, using this image-based method, a toddler will happily decrypt the meaning of pictographs that are thousands of years old. (See Figure 13 for some ancient Egyptian examples.[1]) But note what the toddler will *not* do.

[1] I've used the word 'pictograph' to refer to Egyptian symbols that are meant to be interpreted literally. In contrast, Egyptian 'hieroglyphs' use the rebus principle to encode more complex forms of thought.

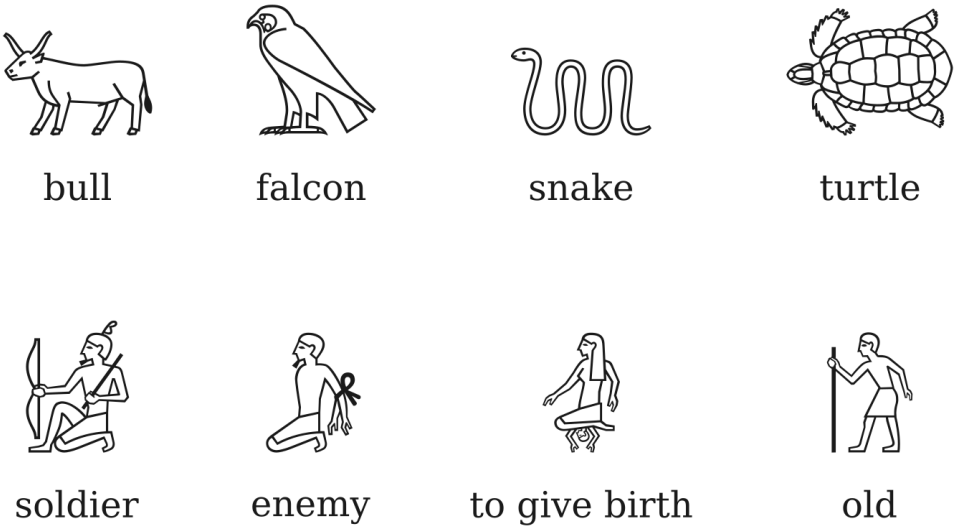

Figure 13: Pictures are the most obvious path to writing

Here is a collection of Egyptian pictographs that are easily decoded by a modern observer. (Fun fact: these symbols are available as Unicode characters.) Note that I use the word 'pictograph' to refer to the literal meaning of each symbol. Egyptian 'hieroglyphs' are more complicated because they use the rebus principle to attach linguistic sounds to symbols.

When shown the word 'snake', the toddler will *not* shout 'snake'. And they will not do so for the same reason that humans found it difficult to move from pictographs to more complex forms of writing. The path forward was not obvious.

For early writers, the hurdle was to somehow use pictures to encode messages that went beyond one-word statements like 'bull' or 'snake'. For example, how would you use pictures to write this sentence?

> The farmer went to the market, sold his bull for $1000, invested the money in a Trump memecoin, and then went bankrupt.

Intriguingly, Egyptian hieroglyphs illustrate a key downside of the rebus principle, which is that the symbols can only be deciphered if the reader has knowledge of the spoken language being encoded. For example, to the French speaker, the image of a bee and a leaf (Figure 14) would *not* elicit the concept of 'belief' ('croyance' in French). It is because of this oral dependence that the Egyptian hieroglyphs spent thousands of years as unintelligible nonsense. When the ancient Egyptian dialect died out, no one had a clue what the surviving symbols meant. The key to decoding these messages came only in 1799, with the discovery of the Rosetta stone — a tablet that contained identical messages encoded in Greek, hieroglyphic Egyptian, and demotic Egyptian.

Figure 14: The rebus principle — using symbols to encode linguistic sounds

Linguists believe that the use of the rebus principle is a crucial step for the development of writing. With the rebus principle, one interprets images in terms of the linguistic sounds they represent (rather than the literal objects being displayed). In English, the symbols of a bee and a leaf can combine to represent the word 'belief'.

Given the immense temporal gap between simple picture drawing and formal writing, we can infer that ancient humans struggled greatly with this problem. But they eventually discovered a good solution, which was to use pictures in a more abstract way. Instead of using pictures to represent real-world things, they used pictures to represent *linguistic sounds*.

This method is called the 'rebus principle', and it works as follows. To encode a message that has no simple associated image, we first decompose the message into sounds. Then we map these sounds onto a set of simple images. For example, if we combine the image of a bee and a leaf (as shown in Figure 14), we can use the rebus principle to encode the English word 'belief'. Clever!

Now, to the modern (literate) observer, the rebus principle seems both intuitive and obvious. After all, it's the principle that underpins how we read and write. (More on that in a moment.) And yet, for the first proto-writers, the rebus principle was unintuitive. Why?

Well, the problem seems to be that for humans, speaking is so effortless that we have little knowledge of how we do it. Yes, we speak by making sounds . . . but it is the *meaning* of these sounds that we say and hear. Indeed, we have almost no conscious knowledge of the sounds themselves. Or put another way, our brains naturally clump meaning into spoken units called 'words'. But our brains find it quite foreign to consciously decompose these words into meaningless sounds. Yet the task for the first writers was to do just that — to reverse engineer what our brains do subconsciously. By all accounts, early writers found this task difficult. And to this day, children struggle to read largely because they struggle to decompose the sounds within words.

At any rate, once the rebus principle was discovered, it allowed writing systems to evolve in two different ways. First, the symbols themselves tended to become more abstract with time. The effect of this abstraction was to simplify the act of writing, but to make learning how to read more difficult.

Second, the rebus principle allowed writers to explore the phonetic level at which to encode their language. Here, we might think that alphabetic writing would be the most obvious path forward. But it was not. The alphabetic system was the *least* obvious approach to writing.

To the oral speaker, the natural unit of meaning is the individual word. Hence the most obvious way of writing was to equate symbols with whole words. This 'logographic' system works perfectly well, and is the basis for Chinese writing.

The next step down the sound ladder is to use symbols to encode individual syllables. This 'syllabic' approach is the basis for Japanese writing.

Finally, the lowest step down the sound ladder is to use symbols (an alphabet) to encode individual 'phonemes', which are the smallest unit of linguistic sound. The advantage of this 'alphabetic' approach is that it is by far the most efficient encoding method, requiring the user to memorize the least number of symbols. [2] However, the disadvantage, as Mark Liberman observes,

[2]Western scholars have long argued that the alphabet is the most efficient way to encode language, since it requires memorizing the least number of symbols. (Some scholars have gone so far as to claim that the alphabet was essential for rationality.) Of course, arguments

is that alphabetic writing places a "special burden" on the reader, since it requires that they gain conscious access to the lowest unit of linguistic sound — a unit that is "relatively inaccessible to introspective scrutiny".

It is likely because of this "special burden" that whereas logographic and syllabic forms of writing evolved multiple times in different places, alphabetic writing seems to have evolved only once. It arose out of the Semitic tradition, which developed a consonant-only system of writing. When the ancient Greeks later borrowed these symbols, they adapted some of the letters to notate vowels. All modern alphabets seem to have sprung from this singular lineage.

The point is that when children struggle to read alphabetic writing, they are in good company. Alphabetic writing is the least obvious way to store linguistic meaning. And it is also the writing method that is most susceptible to *corruption*.

For example, in modern English, there are nine different ways to encode the long 'a' sound (as in halo, aid, ate, say, they, rein, great, eight, and straight). Who invented this absurd system? No one did. These different notations evolved from older phonetic principles that have since been corrupted, as pronunciation changed but spelling did not. Thus, English is littered with silent letters that were once pronounced (as in 'knight'), homonyms that once had distinct pronunciations (as in 'meat' and "meet"), and a myriad of ways to spell the same sound.[3]

about the superiority of alphabetic writing are no doubt driven in part by ethnocentrism. But in light of the way we now use computers, I think the 'efficiency' argument is probably correct.

To enter non-alphabetic characters on a computer, the solution is invariably to use the Latin alphabet, at least in part, to phonetically type the desired non-alphabetic symbol. (See, for example, the pinyin system for typing Chinese symbols.) Hence, it seems unlikely that we could have modern computers without the alphabet.

[3]Much of the phonetic corruption in modern English results from an event called the 'Great Vowel Shift'. During this shift, which began in the 1400s, vowel sounds that were previously distinct slowly evolved into the same sound. Since the spelling of these sounds didn't change, the rules of English phonetics became more confusing. Similar complications happen when languages merge. The new hybrid is left with phonetic principles from both parent tongues. English is itself a mutt of a language — part Germanic, part Latin, and part French.

No one would design such a convoluted system by choice. And yet, it is the system that English children must learn. We fool ourselves if we think it is easy or 'natural'.[4] And we doubly fool ourselves if we think that children can deduce the principles behind English phonetic encoding without explicit instruction.

Hiding the decryption key

To summarize our foray into the history of writing, reading is difficult to learn because it requires developing a set of low-level skills that are not intuitive or natural to the oral speaker.

Now, the tyranny is that for a good reader, these low-level skills have become so automatic that they are subconscious. Hence, a good reader might think it reasonable to teach a child to read without giving explicit instruction on the principles involved. This form of teaching is a strikingly bad idea; and yet it is the method that dominates anglophone schooling.

More on that in a moment. But first, let me convince you (a good reader) that you have a set of low-level decoding skills that are largely subconscious. Go ahead and read the following words aloud:

guttorply melochection intifittle swooflia

As you read, you no doubt realized that these are not English words. They are *pseudowords* that are assembled using the principles of English phonetics. (You can generate more pseudowords here.)

The purpose of these pseudowords is to illustrate that good readers have internalized the low-level principles behind English phonetic encoding. Good readers understand how letters combine to represent sounds, and how these

[4] In a fascinating 2021 paper, Xavier Marjou used a neural network to estimate how difficult various languages are to read and write. He fed the neural network fixed-length samples of various languages, and then asked it predict pronunciation from spelling (i.e. read) and predict spelling from pronunciation (i.e. write). In his test, English performed horribly. The bot had a write accuracy of 36% and read accuracy of just 31%. In contrast, languages like Arabic and Finnish had read/write accuracies well above 80%. (Intriguingly, Marjou found that Chinese writing was more lopsided. The bot could read Chinese with ease, but failed miserably at writing it.) In short, English may be the lingua franca, but only because it was the native tongue of two successive global empires (British and then American). English is actually a dreadful language in which to become literate.

sounds can be combined into words (real or not). Of course, the corollary is that bad readers lack these low-level skills. Which is why, when faced with pseudowords, bad readers fall flat on their face.

With this failure in mind, here is the trick to bad reading instruction: ask students to read without teaching them the decryption key.

To describe a lesson in bad reading instruction, it seems only fair to use myself as an example. When my daughter started school, she was enrolled in French immersion, which meant that she'd learn to read French at school and English at home. Her home-school teacher would be me ... Blair Fix, PhD & VGR (very good reader).

Now, at the time, I'd bought into the whole-language approach to reading instruction.[5] Forget phonics, I thought. We'll just start reading together, and my daughter will naturally pick things up. But she didn't. So with the benefit of hindsight, let me illustrate how things went wrong.

Our first step was to read simple picture books like the one shown in Figure 15. These books are called 'levelled readers', and they're meant to be co-read, with prompts from the pictures. First, I'd read the book. Then my daughter would 'read' it back to me. Things seemed to be going well . . . or so I thought.

[5]How I came to (briefly) believe in whole-language teaching is a somewhat embarrassing case study in how bad science spreads. When my daughter was around the age of four, the topic of reading instruction was on my mind. As luck would have it, one day when I was substitute teaching at Leaside High School in Toronto, I was sitting in the staffroom and came upon a dusty book on reading instruction. Now, I have no recollection of the name of the book or its author. But it could have plausibly been *Reading Without Nonsense*, written by the psycholinguist Frank Smith, one of the main architects of whole-language theory.

At any rate, the book made two arguments that, at the time, I found convincing. First, the author noted that when children are faced with an unfamiliar word, their first instinct is to guess at the word from the context. Kids' *last* instinct is to actually sound out the unknown word. Second, the author noted that English phonics are a complicated mess, with very few rules that are exception free. Hence teaching this mess, he argued, is a waste of time. It's better to focus on building kids' bank of memorized words.

What's embarrassing is that I found these arguments convincing, even though the author provided zero evidence for the efficacy of his proposed approach. Let's run through the problems.

First up is the argument that kids avoid decoding words. This observation is true, but is now recognized to be an avoidance tactic. Decoding words is difficult, and so kids tend to avoid the task, especially if they've not been taught the principles to actually read the word in question.

Second, it is true that compared to languages like Arabic and Finnish, English is a phonological mess. But what scientists now recognize is that the complexity of English orthography means that *more* (not less) time must be spent giving kids explicit and systematic phonics instruction.

Third, the idea that memorizing words is easier than memorizing phonetics is pure silliness based on borderline innumeracy. A quick look at structured literacy programs shows that they teach at most about 100 phonological principles. In contrast, the English language contains at least 460,000 words. So if you're going to defend the wrote memorization of words, you're defending a task that is 5,000 times more expansive.

Back to my brief acceptance of these arguments. I found the whole-language approach convincing for understandable (but mistaken) reasons. The arguments sounded plausible . . . from the standpoint of a literate adult who had long since become unaware of his low-level decoding skills. Of course, when I tried to put whole-language methods in action to teach my daughter, they failed badly. It was only in the face of this failure that I bothered to do some actually research, and learned the folly of my beliefs. (I discovered the science of reading by googling the word 'dyslexia'. That led me to Sally Shaywitz's excellent book *Overcoming Dyslexia*, which lays out the reasons why struggling readers need explicit phonics instruction.)

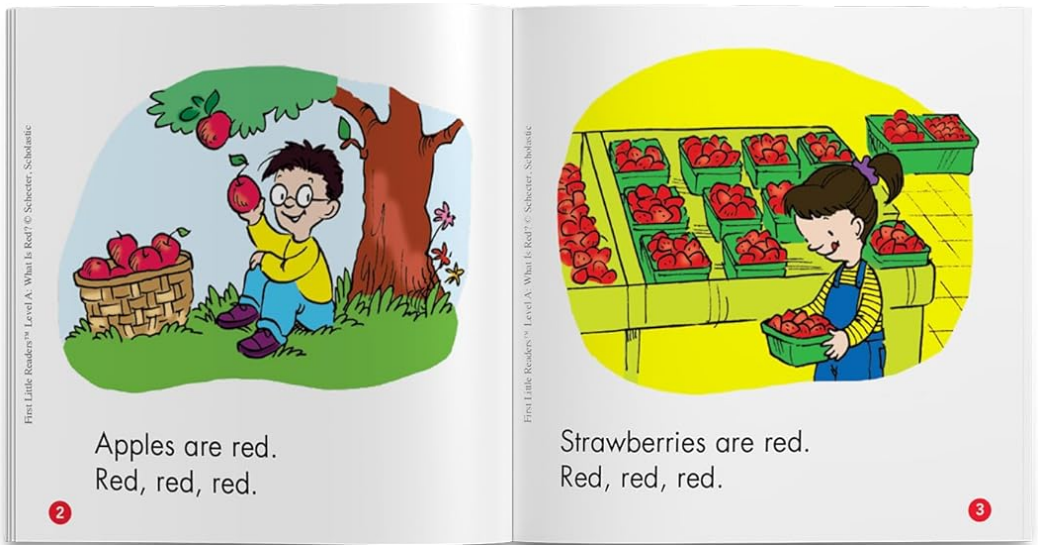

Figure 15: Teaching children how to avoid reading

This is an excerpt from the book *First Little Readers – What Is Red?* It is typical of the levelled-reader format. Note the repetitive sentence structure, which can be easily memorized and recited based on cues from the pictures. Note also the difficulty of the actual decoding task. In this supposed 'Level A' reader, kids are asked to decode the words 'apples' and 'strawberries' — words that are phonetically complicated. In short, this book practically begs children to *not* read the text, and to instead recite the message from pictographic cues.

What I realize now is that these levelled readers actually teach children how to *avoid* reading. Again, the history of writing is instructive as to why. To develop alphabet writing, people had to realize that linguistic meaning (i.e. words) could be broken down into sounds, and that these sounds could be encoded into symbols:

$$\text{meaning} \rightarrow \text{sounds} \rightarrow \text{symbols}$$

Learning to read therefore involves the reverse operation. The learner must take symbols, map them onto sounds, and decode the meaning:

$$\text{symbols} \rightarrow \text{sounds} \rightarrow \text{meaning}$$

Looking at levelled readers, these books make no attempt to teach the reading decryption algorithm. Instead, they take a repetitive sentence structure and map it onto a set of richly illustrated images. Of course, kids love this

format, because it provides a far more obvious path for encoding meaning. Indeed, it is the same path taken by the first writers, who took meaning and mapped it onto pictures:

$$\text{meaning} \rightarrow \text{pictures}$$

When we ask a child to 'read' an illustrated (levelled) reader, they naturally take the easiest path by decoding the message from the picture:

$$\text{pictures} \rightarrow \text{meaning}$$

Now, the problem here is that asking a child to decode meaning from pictures actively misleads them about how reading actually works. Historically, the path from pictographs to alphabetic writing was long and torturous. Indeed, the path was so difficult that it happened only once. Hence, it is completely unreasonable to request that a child use pictographs as a tool for deducing alphabetic principles. Many children will not make this conceptual leap.

Instead, when children are shown pictographs with associated text, what tends to happen is that they memorize a few of the simpler looking words, and then they recite (or simply guess) the rest of the words from the context (which includes the picture and the repetitive sentence structure). Unsurprisingly, researchers have observed this guessing behavior countless times. The strategy occurs because actually reading the words on the page is the hardest and least intuitive option for decoding meaning. Guessing the message from the pictures and the context is far easier. And so that's what kids do.

With word guessing in mind, here is where things get weird. When whole-language theorists observed children's guessing tactic, they interpreted it as a *learning strategy*. Indeed, they enshrined it in a pedagogical approach called the 'cueing method', in which children are taught to read text by looking first to the picture, next to the syntactic context, and last to the actual word.

In short, the whole-language method took a reading avoidance tactic (guessing words from context) and transformed it into a 'strategy' for learning how to read. As you might guess, the effect of this strategy is to leave the worst readers behind.

Teaching the decryption key

The irony of whole-language pedagogy is that when it swept through anglophone schools in the 1980s and 1990s, cognitive scientists were codifying the best practices in reading instruction. Suffice it to say that these best practices are *not* what whole-language theory preaches.

In a nutshell, cognitive scientists discovered that kids struggle to read for the same reason that humans struggled to develop writing. For the oral speaker, it is unnatural to decompose words into sounds. Hence, kids who struggle to read typically lack 'phonemic awareness' — they do not hear the sounds within words. And so the solution is to *teach* this skill. Teach kids that English words are built from a repertoire of several dozen sounds. Teach kids how to segment words into sounds. Teach kids how sounds map onto the alphabet. Yes, this instruction is low-level work. But for young children, it is exciting new knowledge. They love it!

Once kids grasp the basics of how the alphabet represents sounds, they're ready to read three-letter words like 'sit' and 'hat'. Next, they can read these words in simple sentences. Crucially, the practice reading must lack picture cues, because pictures provide a way to avoid the cognitively demanding task of decoding words. (Figure 16 shows a sample reading task from Treasure Hunt Reading, a free structured literacy program.)

Finally, reading ability can be built iteratively by adding new sounds and new ways to encode each sound. (Remember, English is a phonetic mess.) At each step, kids read text that obeys only the phonetic principles that they've been taught. And because kids have been taught the skills for success, something surprising happens. They learn to read!

This evidence-backed approach to reading instruction is called 'structured literacy', and it should be standard practice in every elementary school. But it is not. Instead, the dominant approach (at least until recently) has been to shower kids in text and hope that they deduce the key for decryption. Many kids fail to make this deduction, and they grow up to be life-long dysfunctional readers.

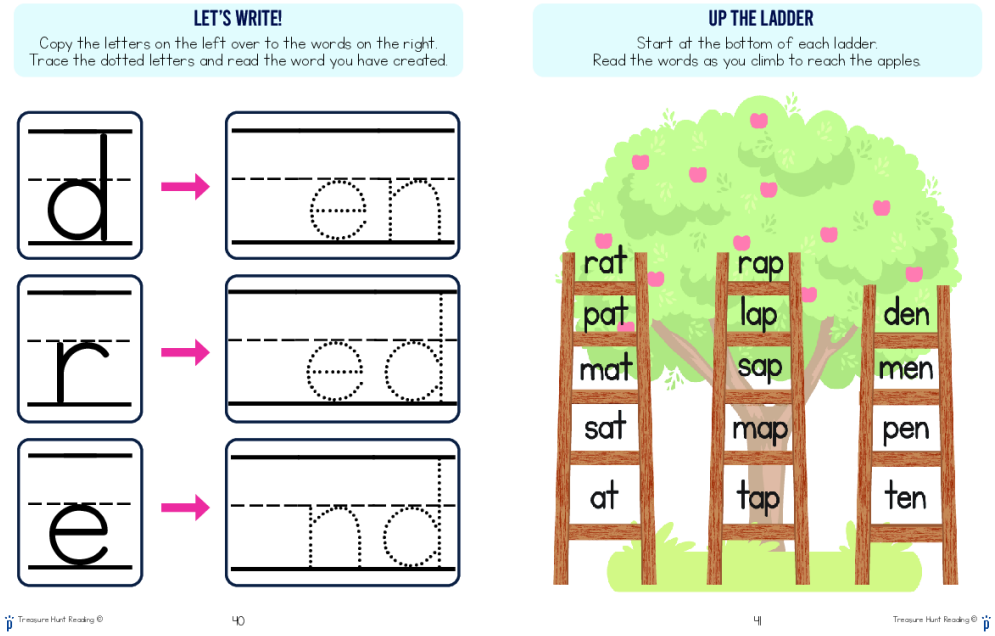

Figure 16: A typical reading task in a structured literacy program

In this example from Treasure Hunt Reading, kids are asked to form and read simple three letter words with a repeating phonic pattern. Notice that this task comes on page 40 of the workbook, after the required consonant and vowel sounds have been taught.

Part III: The case against whole-language instruction

Now that we understand why reading is hard to learn, I want to build the case against whole-language pedagogy. I'm going to argue that whole-language teaching is the main wedge that's widening the gap between the best and worst high-school readers.

My case rests on seven lines of evidence, listed below. At root, my argument is simple. If we do not teach kids the code for decrypting English words, it is the worst readers who suffer, with effects that last a lifetime.

The case against whole-language instruction

1. Whole-language methods are ubiquitous
2. Whole-language instruction harms struggling readers
3. If whole-language instruction is to blame, the timing checks out
4. Early reading ability determines later success
5. A decoding deficit creates a compound reading failure
6. A decoding deficit is a lifelong problem
7. The widening reading-score gap is reversible

1. Whole-language methods are ubiquitous

The first piece of evidence implicating whole-language methods in the widening reading-score gap is the fact that these methods are ubiquitous. For example, a 2019 survey found that 75% of US K-2 teachers taught the 'three cueing' method for deducing words from their context — a method that mistakenly reinforces the evasion tactic that struggling readers use to avoid decoding words.[6]:

Now, the good news is that in the last few years, the cueing method has grown increasingly unpopular, and has even been banned in 15 US states. But the bad news is that for today's high-school students, the damage has already been done.

2. Whole-language instruction harms struggling readers

The second piece of evidence implicating whole-language instruction in the widening reading-score gap is the fact that this method harms struggling readers.

That was the conclusion drawn from a large study of the 'Reading Recovery' program, which is a whole-language intervention designed to bring below-grade-level readers up to speed. In this study, a large group of struggling first-grade readers received intensive one-on-one instruction from a Reading Recovery expert. The students' progress was then tracked over time and compared to a group of struggling readers who did not receive an intervention.

[6]Whole-language methods became popular not just in the US, but across the anglophone world. For example, Timothy Mills reports that a 2013 UK survey found that a stunning "89% of teachers believed that children needed to use a variety of cues to extract meaning from text".

In 2017, the researchers reported that the initial intervention was a success: by the end of their tutoring, the Reading Recovery students were doing substantially better than their peers. Then came the cold water. When the researchers returned to track student progress a few years later, they found a strikingly different pattern. By third and fourth grade, the students who'd participated in the Reading Recovery program were not ahead of their peers . . . they were between a *half to a full grade level behind*.

Although the researchers struggled to explain this negative effect, the problem seems obvious in hindsight. Whole-language methods promote the wrote memorization of words, which is an effective strategy when the vocabulary is small. But as the reading material gets more complex, memorization methods fail, and the students' decoding deficit rears its head. And because whole-language instruction actively dissuades students from learning how to decode words, their decoding deficit grows worse with time.

3. If whole-language instruction is to blame, the timing checks out

The third piece of evidence implicating whole-language instruction in the widening reading-score gap is the fact that the timing checks out. It was in the mid-1980s when whole-language instruction began to spread. And it was in the mid-1990s when high-school reading scores began to drop. If we add a decade delay to allow the first whole-language recipients to advance through school, we'd expect to see the high-school effects of this method appear during the mid-1990s . . . just when high-school reading scores began to drop.

Of course, we don't have rigorous data that tracks the spread of whole-language methods. But we can get a rough sense for the popularity of this approach by tracking the frequency of the phrase 'whole language' within English books. When we do so, we find that the phrase became popular in the 1980s. The top panel in Figure 17 shows the pattern. If we suppose that this word frequency indicates the popularity of whole-language instruction during the early years of schooling, we'd expect to see the high-school effects of these methods appear about a decade later, during the mid-1990s (dashed curve).

Now, by the mid-2000s, the term 'whole language' had become less fashionable, in large part because the reading wars had prompted a backlash against this (ineffective) approach. But instead of being abandoned, whole-language methods were simply rebranded as 'balanced literacy'. And judging by the

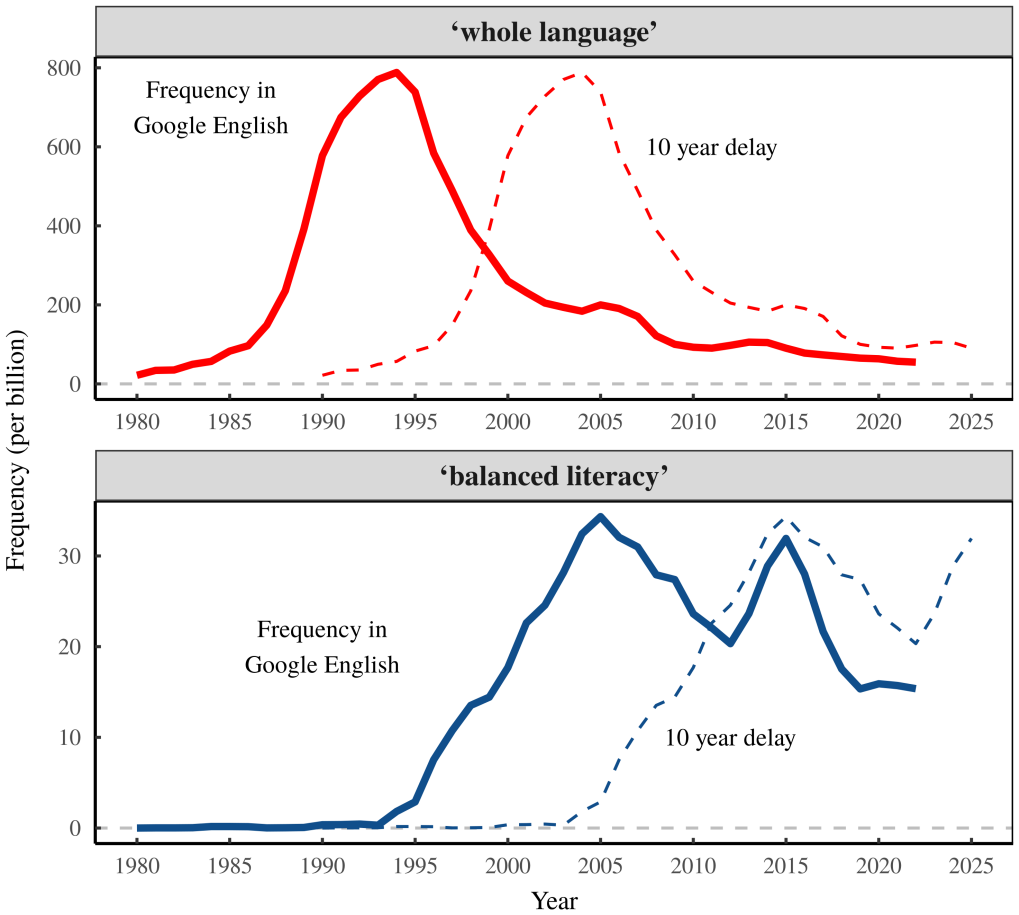

Figure 17: The whole-language impulse

The top panel shows the frequency of the phrase 'whole language' in the Google English corpus (a large sample of English books). The dashed-red line shows this frequency with a ten year delay, used to indicate the potential high-school effects of whole-language education on those who endured it during the early years of schooling. The bottom panel shows the frequency of the phrase 'balanced literacy', which became a common euphemism for whole-language methods (such as the three-cueing strategy). The dashed-blue line shows this frequency with a ten year delay. Sources and methods

---

frequency of this latter phrase (shown in the bottom panel of Figure 17), balanced literacy became prominent during the mid-2000s, with a second wave of popularity during the mid-2010s. If we add a ten-year delay to this pattern, we find that the high-school effects of balanced literacy should appear within the last decade.

In short, if whole-language methods are behind the widening reading-score gap that's appeared over the last three decades, the timing checks out.

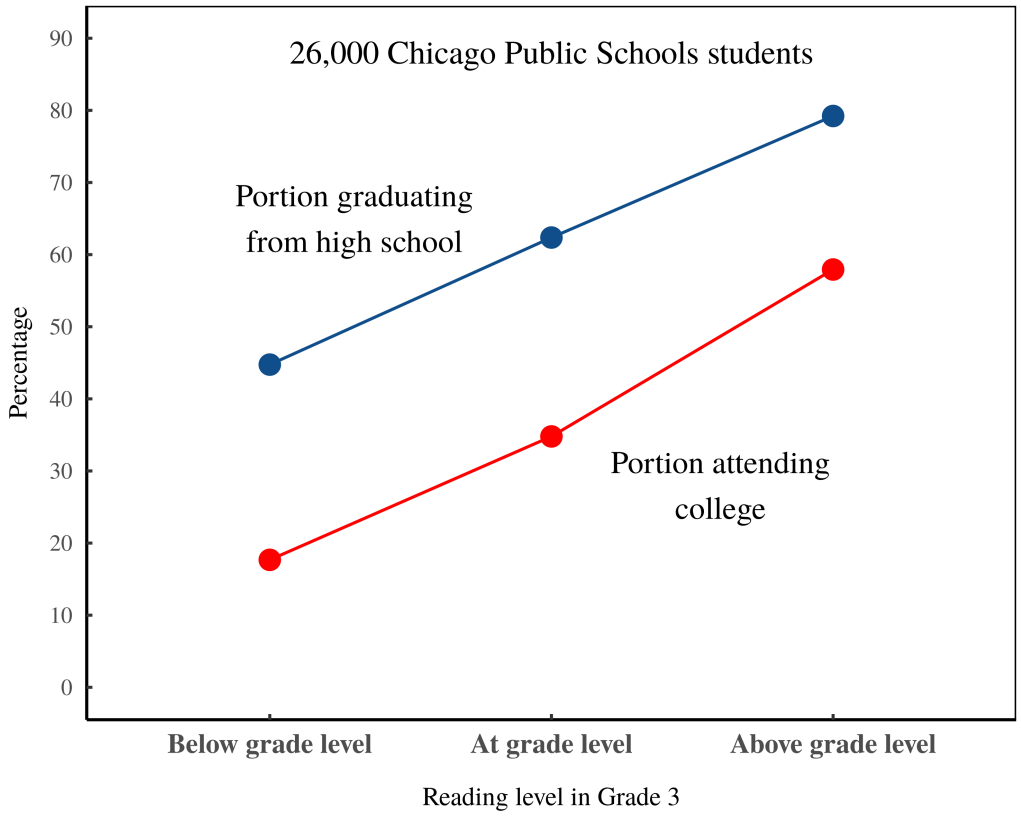

Figure 18: Third-grade reading ability strongly determines success during high school and beyond

In the mid-1990s, Chapin Hall researchers measured the reading ability of a cohort of third-grade students in the Chicago Public Schools system. Then they tracked student progress over time. The researchers found that high-school graduation rates and college attendance were both strongly determined by reading ability in grade three, as shown here. Sources and methods

---

4. Early reading ability determines later success

The fourth piece of evidence implicating whole-language instruction in the widening reading-score gap is the fact that learning during the first few years of school determines success later on.

For example, a 2010 study of 26,000 Chicago elementary school students found that reading levels in Grade 3 strongly predicted high-school graduation rates. Among third-grade students whose reading was below grade level, only 44% would later graduate from high school. But among third-grade students whose reading was above grade level, 79% would later graduate from high school. The portion of kids going to college showed a similarly stark gap, as illustrated in Figure 18.

5. A decoding deficit creates a compound reading failure

The fifth piece of evidence implicating whole-language methods in the widening reading-score gap is the fact that by downplaying decoding skills (and encouraging students to guess at words), the approach likely creates a compound reading failure.

To understand this effect, realize that reading comprehension is the product of two distinct skills. Reading comprehension depends firstly on the *ability to decode words*. And it depends secondly on *oral comprehension*. According to the 'simple view of reading', reading comprehension is the product of these two abilities:

$$\text{reading comprehension} = (\text{decoding skill}) \times (\text{oral comprehension})$$

Now in principle, decoding skill is separate from oral comprehension, which is why someone can be fluent but illiterate. That said, in literate cultures, oral comprehension tends to be a function of decoding ability itself. And that's because writing is the most potent source of knowledge.

Simply by reading, a good decoder can learn about new words and new ideas, causing their oral comprehension to grow with time. In contrast, a poor decoder struggles to unlock the knowledge contained within books, which means that their oral comprehension remains stunted. But here is the kicker. Because reading comprehension depends on the *product* of decoding and oral abilities, poor decoders suffer a compound failure. Their poor decoding hinders their oral comprehension, and so their reading comprehension suffers doubly.

Figure 19 shows an example of this compound failure. The data comes from a longitudinal study of 485 students in Iowa and Illinois, whose school progress was tracked throughout 1990s. The chart illustrates how students' vocabulary depended on their decoding skills, which were measured in grade 4. (Vocabulary is a crude proxy for oral comprehension.) Over time, students' vocabularies tended to increase, but not at the same rate. Instead, the best fourth-grade decoders had the greatest vocabulary growth, while the worst fourth-grad decoders had the least vocabulary growth.

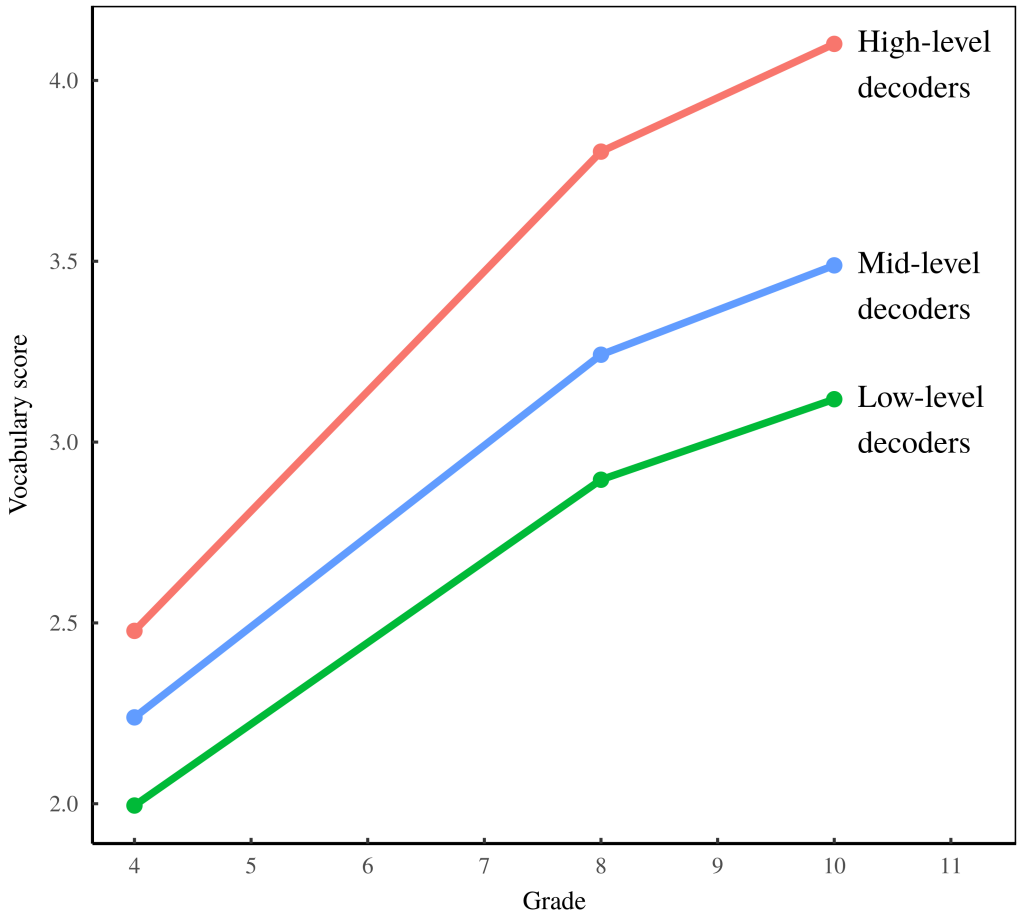

Figure 19: Poor decoding skills lead to a stunted vocabulary

In the 1990s, Dawna Duff and colleagues tracked vocabulary growth among a cohort of American students. This chart shows how vocabularies grew as a function of decoding ability, measured in fourth grade. While all students saw their vocabularies increase with time, the poor decoders saw the least growth. Assuming that these poor decoders retained their decoding deficit throughout school, by high school they would suffer from a compound reading failure, with worse decoding skills *and* a smaller vocabulary. Sources and methods

---

If the poor fourth-grade decoders retained their decoding deficit as they aged (a reasonable assumption), we can surmise that by high-school, they suffered from a compound failure: both their decoding and their oral comprehension lagged behind their peers, leading to significantly worse reading comprehension. In short, to the extent that whole-language instruction hinders the development of decoding skills, we expect that by adulthood, it creates a compound failure of reading comprehension.

6. A decoding deficit is a lifelong problem

The sixth piece of evidence implicating whole-language methods in the widening reading-score gap is the fact that decoding deficits are not solely a childhood problem. They typically remain present well into adulthood.

True, poor decoders often develop coping skills for parsing text — skills like memorizing sight words and guessing unfamiliar words from their context. Unfortunately, these coping strategies do not solve the core decoding deficit, which remains visible to anyone who cares to look.

The way to unearth a decoding deficiency is ask people to read pseudowords — fake words that are constructed from real English phonetics. When tasked with parsing such words, poor readers reveal their core deficit: they have not mastered the principles of English phonetics.

Figure 20 shows a striking example of this deficit. The data comes from research conducted by Molly Minus during the 1990s. As part of her PhD research, Minus measured decoding ability among three groups of adults:

1. Upper-level university students
2. College students taking remedial reading courses
3. Prisoners receiving reading instruction

Minus found stark differences between these groups. When tasked with reading 50 pseudowords, the university students got almost all of them correct (top panel). The college students faired worse (middle panel). And the prisoners? Well, they were abysmal, getting an average of 8 words correct (bottom panel).

The point here is that well into adulthood, the main driver of functional illiteracy is the inability to decode words — an inability that whole-language teaching actively promotes.

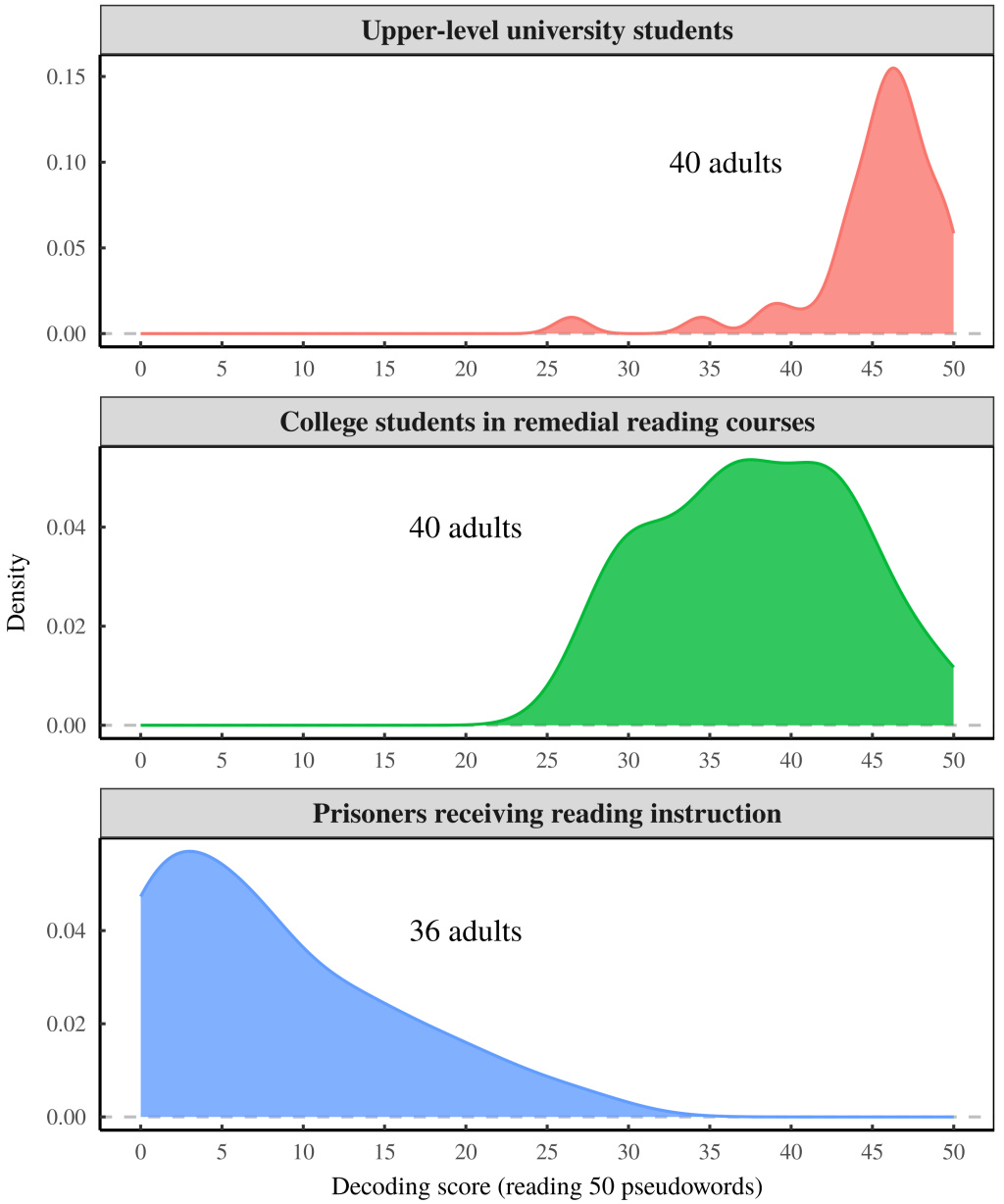

Figure 20: Decoding ability among three groups of adults

In the early 1990s, Molly Minus measured decoding skills among three groups of adults: upper level university students (top panel), college students in remedial reading courses (middle panel), and prisoners receiving reading instruction (bottom panel). This chart shows the distribution of decoding scores within each group. (The task was to read fifty pseudowords.) Minus found that decoding skill varied predictably by group. The university students were excellent decoders, while the prisoner were horrible. Sources and methods

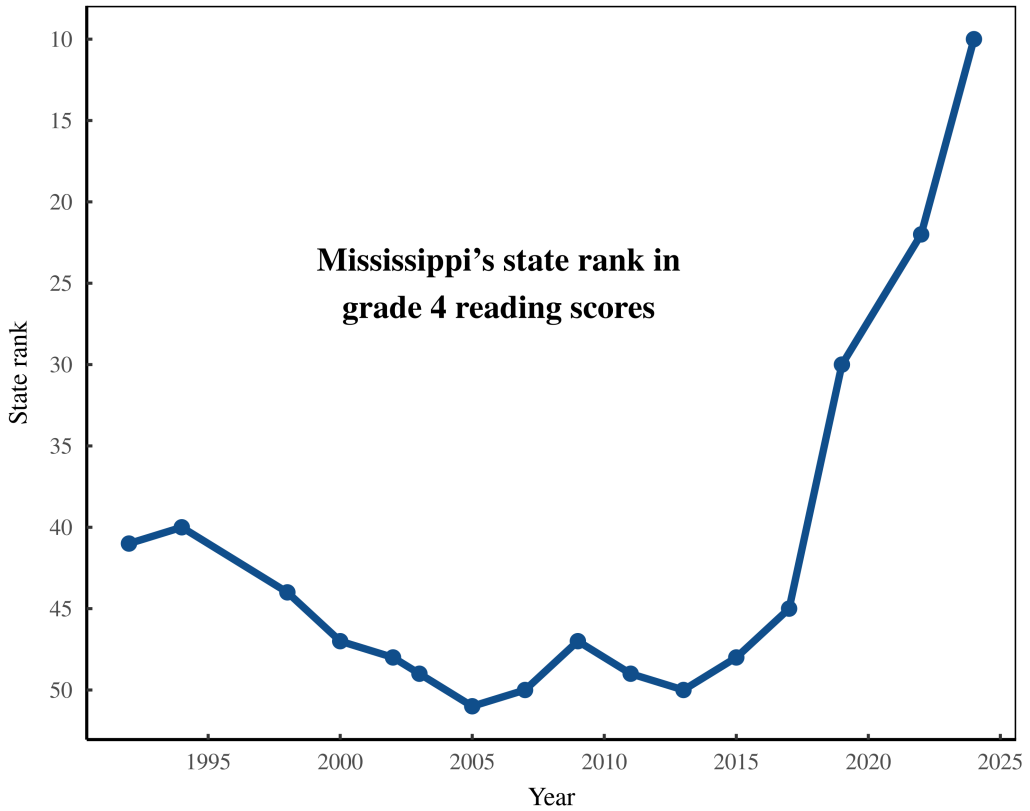

Figure 21: The Mississippi miracle

This chart shows Mississippi's state rank in fourth-grade reading scores. (Note the reverse scale on the vertical axis.) For decades, Mississippi sat at the bottom of the reading-score heap. But after 2013, it clawed its way into the top ten — a remarkable transformation that's often dubbed the 'Mississippi miracle'. Sources and methods

---

7. The widening reading-score gap is reversible

Now to what is perhaps the key piece of evidence implicating whole-language teaching in the widening reading-score gap. It seems that this growing gap is not inevitable, and that it can be reversed by dumping whole-language methods and replacing them with structured literacy. Interestingly, we can thank the state of Mississippi for demonstrating this fact.

Perpetually the poorest US state, Mississippi once produced some of the nation's worst readers. As recently as 2013, Mississippi's fourth-grade reading scores were in last place. Yet by 2024, it had turned things around and was among the top ten states. Figure 21 shows the transformation.

So what happened? Did Mississippi suddenly get rich? Did its children stop watching TV? No and no. What happened is that in 2013, Mississippi overhauled how it taught kids to read.

To make sense of this overhaul, realize that in the 1990s, Mississippi followed other states in embracing the whole-language approach to reading. Curriculum documents from the time illustrate the educational dogma. In 1996, Mississippi's Department of Education declared that among first-grade students, "[r]eading and writing are no longer viewed as isolated tasks to be taught and tested". Then, as if to foreshadow what would follow, the document argues for a "harmony" of strategies as students "*attempt* to understand how reading and writing work" (my emphasis).

It's with this possibility for failure that we should interpret the patterns in Figure 22. In the top panel, I've plotted the change in Mississippian fourth-grade reading scores between 1992 and 1994 (measured as a function of reading-score percentile). Notice the Z-shaped trend: over this period, a widening gap emerged between the best and worst readers. Well that's curious. That's the same type of gap that emerged among US high-school students over the last three decades. (See Figure 3.). Perhaps Mississippi's embrace of whole-language methods is telling us something.

Now let's look at the bottom panel in Figure 22, which plots the change in Mississippian fourth-grade reading scores from 2011 to 2019. Notice that the pattern looks starkly different than the one above. During the 2010s, all fourth-grade Mississippian readers improved. However, the worst readers improved the most, while the best readers improved the least.

What caused this reverse effect? Fortunately, we know exactly what happened. In 2013, Carey Wright became Mississippi's state Superintendent of Education. A proponent of evidence-based instruction, Wright oversaw a massive change in how Mississippi taught reading. Whole-language methods were abandoned and replaced with a focus on systematic phonics and phonemic instruction. The new approach included extensive support for teachers, funding for the early detection of reading problems, and (most controversially) a 'third-grade gate', which required that all third-grade students pass a mandatory reading test before proceeding to the next grade.[7]

[7]Making students repeat a grade is a controversial tactic because there isn't much evidence that it works. And one can see why. If a kid fails to read because they don't understand the connection between sounds and alphabetic symbols, forcing them through another year

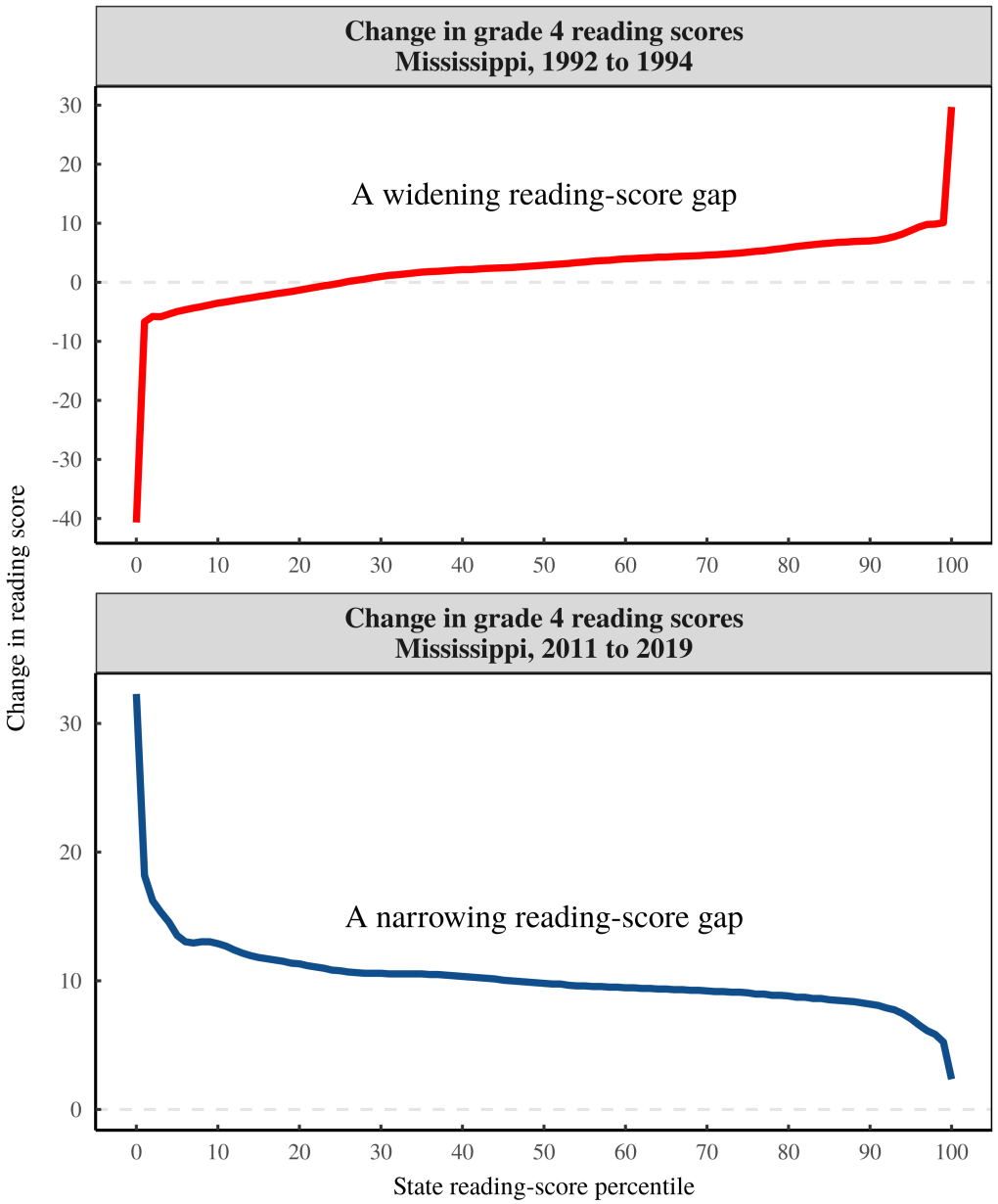

Figure 22: Mississippi's widening and then narrowing reading-score gap

This chart shows the change in Mississippian fourth-grade reading scores as a function of reading-score percentile, captured during two different eras. The top panel shows the reading-score change between 1992 and 1994, an era when the state was rapidly adopting whole-language methods. Over this period, a widening reading-score gap emerged, with the best fourth-grade readers getting better, and the worst readers getting worse. In contrast, the bottom panel shows the reading-score change between 2011 and 2019, an era when the state abandoned whole-language methods and implemented a structured-literacy approach to reading instruction. (I've halted the data in 2019 to avoid any pandemic-related artifacts.) Over this period, the reading-score gap narrowed significantly. While all fourth-grade readers improved, the best readers improved the least, while the worst readers improved the most. Sources and methods

The results of this teaching experiment were nothing short of dramatic. Mississippi's fourth-grade reading scores were vaulted from worst in class to among the top ten. But more interesting, in my view, is the structure of this 'Mississippi miracle' — the fact that it targeted the worst readers and helped them the most. Mississippi's experiment strongly suggests that the widening skill gap between the best and worst high-school readers is not inevitable, and that it's been driven by whole-language pedagogy — a teaching approach that systematically fails the worst readers.

When the method creates the disease

When parents send their kids to school, they no doubt assume that the teacher uses the best methods for instruction, much as a patient assumes that their doctor uses the best forms of medicine. Unfortunately, if the history of science tells us anything, it is that best practices are not guaranteed.

The problem is that once a flawed practice becomes institutionalized, it is difficult for the practitioner to discern that their method is unsound. If you practice only what your mentor preached (and you're surrounded by others who do the same), the failure of your method becomes invisible. Which is why doctors practiced bloodletting for centuries, yet were oblivious to the fact that it tended to kill their patients.

It's within this context that we should interpret the whole-language movement. It promoted a flawed teaching method that, once institutionalized, became invisible to its practitioners. When children *did* learn to read (despite their poor instruction), it was proof that the method worked. And if the child failed, they were just 'slow learners' who'd eventually get the gist of reading. In short, it was the *method* that succeeded, but the individual *child* who failed.

of the same failed instruction won't help them. That said, it is wildly unethical to let children proceed through school if they cannot read. Doing so sets them up for a lifetime of failure. So is there a better alternative than mandatory retention?

Perhaps the most innovative approach comes from the 'Success for All' program, described in Episode 12 of the podcast Sold A Story. In Success for All program, the entire school has a shared reading-instruction block, in which students are grouped by reading ability and not by grade. This ability-based streaming allows teachers to ramp up the intervention for kids who are falling behind. In short, the goal is to make it impossible for kids to fall behind in reading. I think the logic here is impeccable. Decoding skills must be mastery based. If kids do not get the basics of decoding, everything else should be put on hold until they master this skill.

Still, there's the question of how such a flawed practice managed to become a tradition. Did whole-language teaching become popular because it contained a kernel of truth? I think the answer is no. The whole-language approach became popular because it told a story that people wanted to hear. Whole-language theory sold the idea that learning to read is as natural as learning to speak. The message, then, is that the teacher can empower their students largely by *getting out of the way*. So forget lectures. Forget drills. Forget explicit instruction. Just give kids a chance to explore great literature, and they will naturally learn how to read.

To grasp the appeal of this idea, we must understand the context in which it emerged. Prior to the 1960s, anglophone schooling was a dictatorial affair. It was a place where the teacher was a sergeant and the kids were compliant enlistees. It was an environment that did not exactly stimulate high-minded thinking. It was this dictatorial environment that whole-language proponents rejected. They sold (and still sell) their approach as a "democratic" and "humanistic" method that empowers "teachers and students alike". In short, the spread of whole-language teaching had more to do with political ideals than with any evidence that the method was sound.[8]

Now, the irony is that although whole-language methods were billed as a tool for empowering *all* children, what they actually did was to empower the *best* students — the kids who learned to read effortlessly, and who benefited from have time (and resources) to practice their skills. But for the kids who struggled to decode words, the whole-language environment was downright confusing, since they were being asked to 'practice' a task that they found

[8]We can get a sense for the political nature of whole-language theory by reading some of the later writings of whole-language theorist Kenneth S. Goodman. In his 1998 book *In defense of good teaching*, Goodman wastes no time arguing that phonics instruction is a far-right plot:

> For well over half a century the far right has used a campaign for phonics to elect ultraconservative school board members and attack both public education and teacher education.

Goodman is also keen to cast cognitive scientists as partisan hacks who are the enemy of good teachers:

> Their "true science" can produce "research-based programs." The far right can concentrate on keeping pressure on politicians and administrators at all levels, on being ready to pack a hearing room or school board meeting, while the disinformers and "scientists" push the "truth" on teachers and administrators.

One gets the feeling that when Goodman is faced with empirical evidence that contradicts whole-language theory, he responds by vilifying scientists. It's not a good look.

bewildering. It was the equivalent of giving a ten-year old integral calculus and saying "Do the math, kid. I'm empowering you by not teaching you how it works."

Of course, the failure of whole-language methods doesn't mean that wrote dictatorial teaching is the way to go. It is not. The real message is that higher-level abilities do not come from the ether; they depend on the mastery of lower-level skills.

Great jazz musicians improvise effortlessly because they've internalized low-level skills like scales and chord progressions. Great mathematicians produced abstract proofs because they've mastered the low-level skills behind arithmetic and algebra. Great writers produce compelling literature because they've been taught the low-level mechanics of their language. And good readers can parse text accurately and rapidly because they grasp the low-level algorithm of how symbols encode linguistic sounds.

In each case, choosing to *not* teach these low-level skills isn't 'democratic'. It isn't 'humanistic'. It is the definition of *regressive*. Choosing to not teach low-level skills is a recipe for selecting the least gifted students and systematically leaving them behind. It is a recipe for producing the outcomes we now observe … a widening gap between the best readers and the worst readers.

Of course, the evidence against whole-language teaching does not get smartphones and algorithmic slop off the hook for polluting our social environment. But then again, when kids struggle to merely decode words, they have little reason to get off their phones and actually read.

Support this blog

Hi folks, Blair Fix here. I'm a crowdfunded scientist who shares all of my (painstaking) research for free. If you think my work has value, consider becoming a supporter. You'll help me continue to share data-driven science with a world that needs less opinion and more facts.

Become an ETD supporter

Resources

If you are teaching a child to read (or write), here are some helpful resources:

- Treasure Hunt Reading. A free structured literacy program developed by Prenda. It comes with a workbook that's free to print, and a collection of instruction videos that systematically teach phonics principles. It's what I used to teach my daughter to read.
- All About Spelling. Compared to decoding words (reading), encoding them (spelling) is the more difficult task. Struggling readers will also struggle to spell, and they benefit from structured instruction. All About Spelling offers a series of highly structure lesson-books and workbooks that build spelling competency bit by bit, with no gaps. The materials are pricey, but worth it. And the letter tile app is excellent.
- Sold a Story. A documentary series from Emily Hanford about how whole-language methods came to dominate anglophone education. It's a jaw-dropping exposé about why the school system leaves kids behind.

Sources and methods

US reading scores

All US reading-score data comes from the NAEP and can be browsed here: https://www.nationsreportcard.gov/ndecore/landing

Data series are as follows:

- Figure 1: Average scale scores for grade 12 reading, by all students [TOTAL]
- Figure 2 and 3: Distribution percentages for grade 12 reading, by all students [TOTAL]. (I interpolate this data with a spline function. See quantile function notes below.)
- Figure 4: Distribution percentages for grade 12 reading, by sex [GENDER]. (I interpolate this data with a spline function. See quantile function notes below.)
- Figure 5: Percentile scores for grade 12 reading, by parental education level, from 2 questions [PARED]

- Figure 6: Distribution percentages for grade 12 reading, by parental education level, from 2 questions [PARED]. (I interpolate this data with a spline function. See quantile function notes below.)
- Figure 7: Percentile scores for grade 12 reading, by days absent from school in the last month [B018101]
- Figure 8: Distribution percentages for grade 12 reading, by days absent from school in the last month [B018101]. (I interpolate this data with a spline function. See quantile function notes below.)
- Figure 9: Percentile scores for grade 12 reading, by amount of TV or video watched on school day [B001801]
- Figure 10: Distribution percentages for grade 12 reading, by amount of TV or video watched on school day [B001801]. (I interpolate this data with a spline function. See quantile function notes below.)
- Figure 11: Percentile scores for grade 12 reading, by read for fun on your own time [R810901]
- Figure 12: Distribution percentages for grade 12 reading, by read for fun on your own time [R810901]
- Figure 21: Average scale scores for grade 4 reading, by all students [TOTAL] and jurisdiction. (Note that some states have missing data in certain years. To construct a complete interstate ranking across all years, I've filled in missing state data with a linear interpolation.)
- Figure 22: Distribution percentages for grade 4 reading, by all students [TOTAL] and jurisdiction. (I interpolate this data with a spline function. See quantile function notes below.)

Inferring the quantile function for reading scores

The most fine-grain data provided by the NAEP consists of 'distribution percentages' — the percentage of students with reading scores within a given ten-point range. To work with this data, I smooth it using a spline function.

Figure 23 shows my approach. Here, the points show the empirical data from the NAEP. Each point is located at the midpoint of the reading-score bin, and shows the percentage of students within each bin. To infer the smoothed distribution behind this binned data, I interpolate between data points using a spline function (the curve in Figure 23). I then treat this curve as a probability distribution for reading scores.

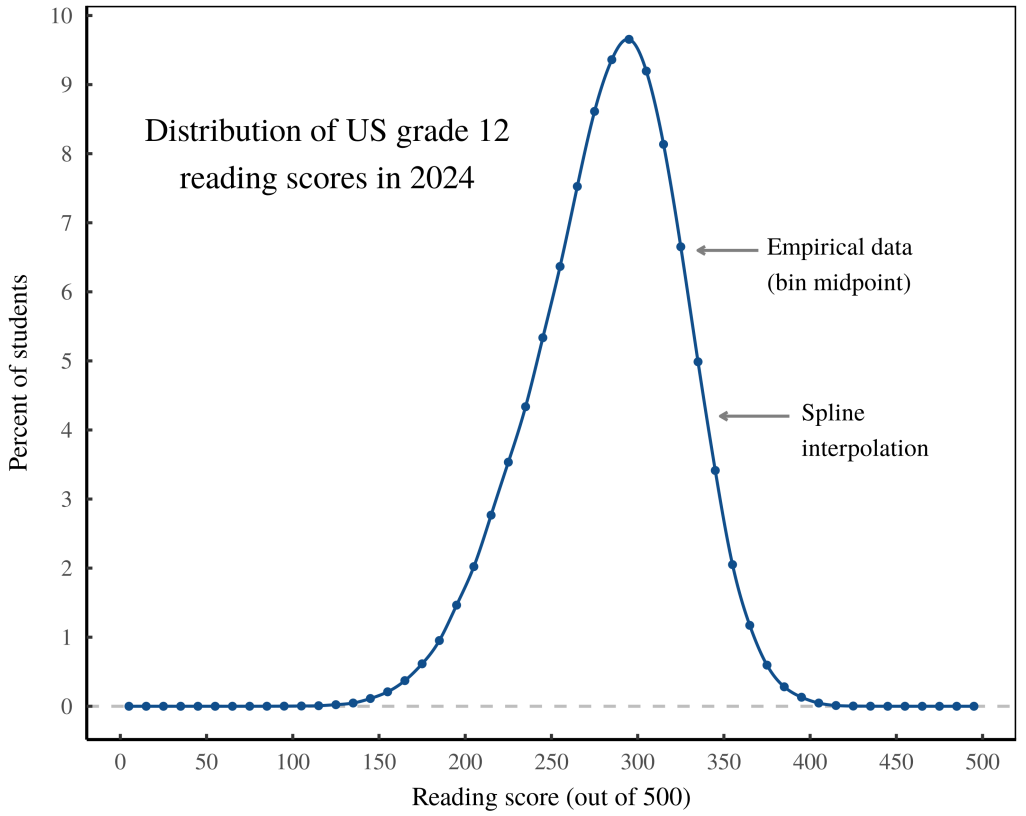

Figure 23: Distribution of US grade 12 reading scores in 2024

This chart illustrates my method for smoothing the 'distribution percentages' data provided by the NAEP. The empirical data consists of binned values reporting the portion of students with reading scores within a given ten-point range. Here, blue points show this empirical data, with the point placed at the midpoint of each reading-score bin. To estimate the complete distribution beneath this binned data, I interpolate between points using a spline function, as illustrated by the curve. I then treat this curve as probability distribution for reading scores.

From this estimated probability distribution, I then infer the quantile function, which consists of reading scores as a function of reading-score percentile. Figure 24 shows the inferred quantile function for all US grade 12 students in 2024. Once I've created the quantile functions for all of the desired data (various years and various demographic groups), I use these functions to estimate changes in percentile score over time.

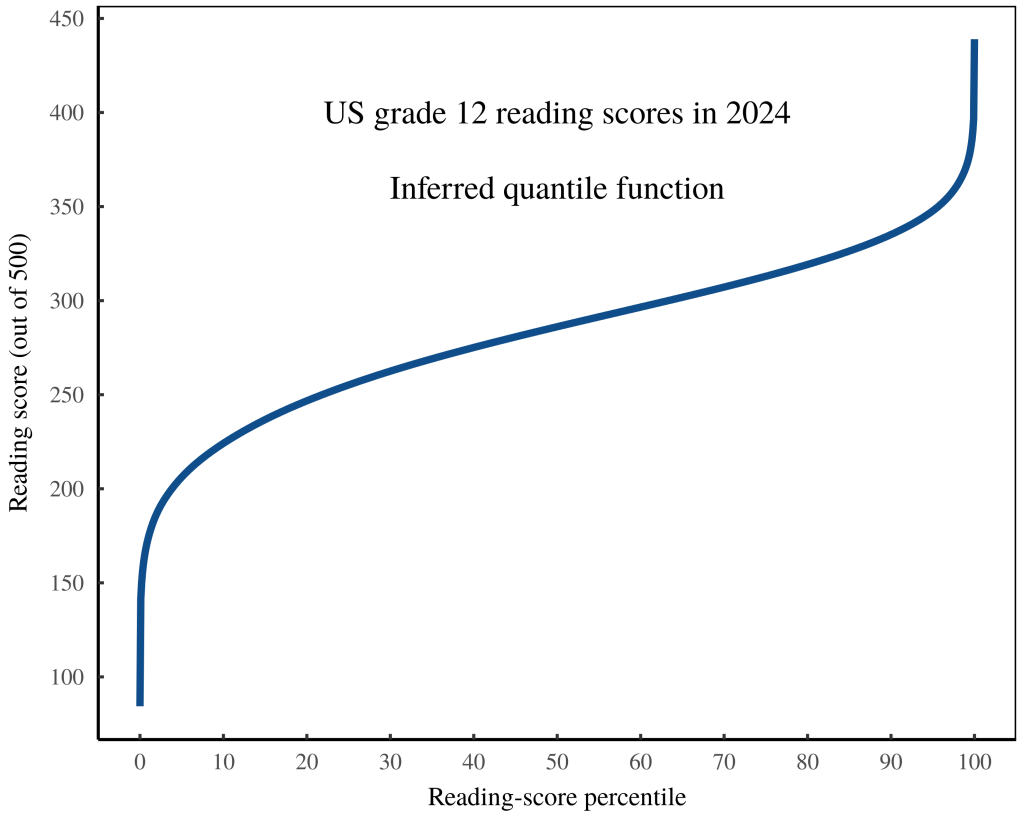

Figure 24: The inferred quantile function for US grade 12 reading scores in 2024

This chart shows inferred (smoothed) values for US grade 12 reading score as a function of read-score percentile. To construct this quantile function, I use the estimated probability distribution from Figure 23. I then use this smoothed quantile function to estimate changes in percentile reading score over time. (These changes are not shown here.)

Word frequency (Figure 17)

Data for the frequency of 'whole language' and 'balanced literacy' is from the Google Ngrams dataset, downloaded with the R package `ngramr`.

High-school graduation rates by third-grade reading ability (Figure 18)

Data comes from Figures 4 and 5 in the report 'Third Grade Reading Level Predictive of Later Life Outcomes', by Lesnick, Goerge, and Smithgall. The study tracks a large cohort of Chicago Public Schools students as they progress through school. (To work with this data, I digitized it using Engauge Digitizer.) Note that Lesnick and colleagues define reading grade level as follows:

- below grade level: reading scores below the national 25th percentile

- at grade level: reading scores above the national 25th percentile and below the national 75th percentile
- above grade level: reading scores above the national 75th percentile

Vocabulary growth as a function of decoding skill (Figure 19)

Data is from Figure 2 in the paper 'The Influence of Reading on Vocabulary Growth: A Case for a Matthew Effect' by Duff, Tomblin, and Catts. The study tracked the progress of 485 children in Iowa and Illinois, beginning in 1993. Note that what I call 'decoding ability', Duff and colleagues call 'reading ability'. But what they actually measure is the fourth-grade ability to decode pseudowords and selected sight words. (To work with this data, I digitized it using Engauge Digitizer.)

Adult decoding ability (Figure 20)

Data is from Figures 6, 7, and 8 in Molly Minus' paper 'The Relationship of Phonemic Awareness to Reading Level and the Effects of Phonemic Awareness Instruction on the Decoding Skills of Adult Disabled Readers'. (I digitized the data with Engauge Digitizer.)

Note that Minus also measured 'phonemic awareness' in her three sample groups. (Phonemic awareness is the ability to identify the sounds within words.) As with decoding ability, she found that phonemic awareness varied starkly between groups, as shown in Figure 25. This evidence reinforces the standard scientific picture of reading. Good readers can parse the sounds within words. Bad readers cannot.

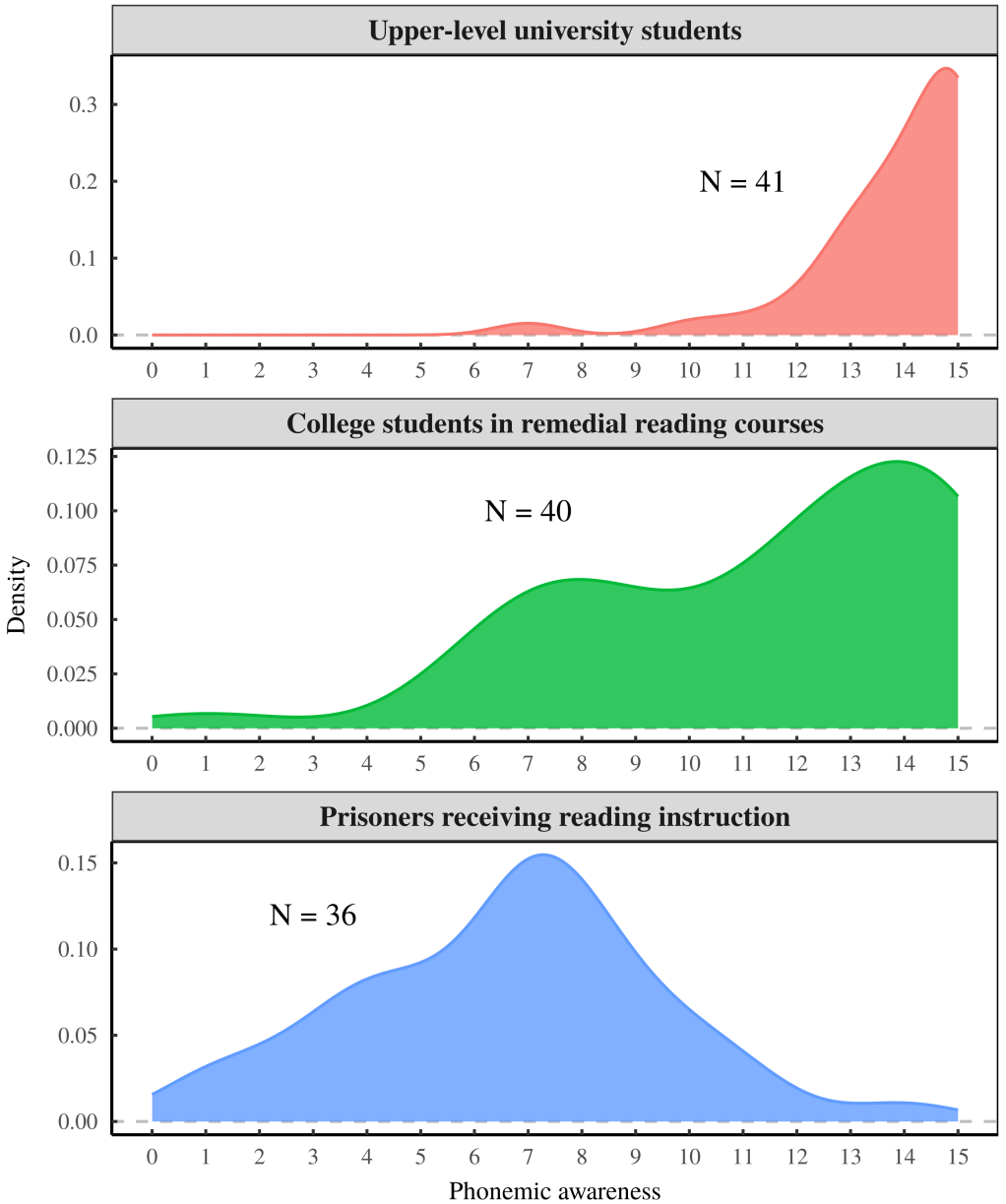

Figure 25: Phonemic awareness among three groups of adults

This chart shows Molly Minus's measurements of phonemic awareness (the ability to decompose the sounds within words) in three samples of adults. Data is from her Figures 10, 11, and 12 in Minus (1992). Note the large disparity between prisoners and university students.

Further reading

Bentz, C., & Dutkiewicz, E. (2026). Humans 40,000 y ago developed a system of conventional signs. *Proceedings of the National Academy of Sciences*, *123*(9), e2520385123.

Bone, J. K., Bu, F., Sonke, J. K., & Fancourt, D. (2025). The decline in reading for pleasure over 20 years of the American time use survey. *Iscience*, *28*(9).

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.

Duff, D., Tomblin, J. B., & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a matthew effect. *Journal of Speech, Language, and Hearing Research*, *58*(3), 853–864.

Lesnick, J., Goerge, R., Smithgall, C., & Gwynne, J. (2010). Reading on grade level in third grade: How is it related to high school performance and college enrollment. *Chicago, IL: Chapin Hall at the University of Chicago*, *1*, 12.

Marjou, X. (2021). OTEANN: Estimating the transparency of orthographies with an artificial neural network. *Proceedings of the Third Workshop on Computational Typology and Multilingual NLP*, 1–9.

May, H., Blakeney, A., Shrestha, P., Mazal, M., & Kennedy, N. (2024). Long-term impacts of reading recovery through 3rd and 4th grade: A regression discontinuity study. *Journal of Research on Educational Effectiveness*, *17*(3), 433–458.

Minus, M. A. E. (1993). *The relationship of phonemic awareness to reading level and the effects of phonemic awareness instruction on the decoding skills of adult disabled readers*. The University of Texas at Austin.

Moats, L. C. (2000). *Whole language lives on: The illusion of "balanced" reading instruction*. Diane Publishing.

Ryan, H., & Goodman, D. (2016). Whole language and the fight for public education in the US. *English in Education*, *50*(1), 60–71.

Shaywitz, S. (2003). *Overcoming dyslexia*. New York.