

Me and Reading Recovery

Tim
Shanahan



Teacher question:

Would you do an article about your thoughts on the recent report about Reading Recovery?

The first time I heard of Reading Recovery (RR) was in 1987. The editor of the *Journal of Reading Behavior* asked me to review Marie Clay's book, [The Early Detection of Reading Disabilities](#). I knew of the book – even had a copy – but was only aware of the innovative assessment that it presented.

I hesitated to take on the task since the book was already in its third edition and had attracted a reasonable number of reviews already. That was the point, she told me. The instruction proposed in the book had not been reviewed and nor had the research included in its appendix. I'd be the first independent scholar to take a careful look. She thought that would be timely since some professors at Ohio State were then trying to bring the program to US schools.

I conducted the review, attending more to the research claims than the instruction itself, though I noted that the activities were aimed at teaching “directionality of print, locating procedures, spatial layouts of pages, story writing, oral reading, correspondence of spoken and written words, and letter names” and included procedures for “teaching children to read fluently, for helping them to develop self-monitoring and self-correcting strategies during reading” ([Shanahan, 1987](#)).

Notice anything missing? I either didn't or chalked up any omissions to the fact that the program targeted kids who were still not reading well after a full year of teaching. Clay, I assumed, believed that at that point such kids in New Zealand would be decoding and would need lots of re-reading and sentence writing. In any event, I voiced no complaints about the teaching plan, but deemed the studies so poorly designed that one couldn't determine the value of the program on their basis. The flaws in Clay's data misleadingly made the program appear more successful than it had been.

Despite the thoughtful insights in my little review, in the ensuing years, RR became a very big thing in US education. More and more schools adopted it, more and more big-name

reading authorities endorsed it, and more and more data accumulated as to its effectiveness. I wasn't particularly curious – lack of adequate research doesn't mean something doesn't work and I'd been ignored before.

During the mid-1990s, I was approached by one of the Regional Education Labs here in the US. Several governors were considering funding RR in their states and wanted to know what the research said. I was selected for this role because of that earlier review, but my negative take made them wonder if I wasn't too negative about RR. They asked if I would conduct the review with Rebecca Barr who they saw as more of an RR-advocate at that point. Becky and I differed in our views of RR then (not by the end of the process) but we had confidence in each other's integrity, so we agreed.

By then, Ohio State had generated a lot of data, and a handful of independent studies had accumulated too. We wrote the report and proceeded to try to publish a version in *Reading Research Quarterly*. That manuscript went through substantial review and the editors even obtained other prepublication studies for us to consider. That extended report was eventually published, and it even won an award.

We concluded that much of the RR literature was seriously biased ([Shanahan & Barr, 1995](#)). As with the original collection of studies, there were design flaws that systematically made RR appear more effective than it was. Much of the evidence had to be set aside.

However, there were a couple of studies that met acceptable standards (including a particularly well-reported independent randomised trial) and those well-done studies concurred as to its effectiveness.

We also examined some studies that supplemented RR in one way or another: one added explicit phonics instruction ([Iversen & Tunmer, 1993](#)); and the other included parent involvement ([Yukish & Fraas, 1988](#)). In both cases, enrichment

improved efficiency. Students accomplished the program goals with much less instruction.

We also reported the first cost analysis of RR. Program charges varied due to local differences in teacher salaries, but overall, enrolling a student in RR basically doubled the cost of their education for a school year. If a district budgeted \$10,000 per child for a year of schooling, then RR added another \$10,000 for each child enrolled, making it a very expensive intervention.

I mentioned those well-done evaluation studies. One was particularly notable, a study conducted in Australia ([Center, Wheldall, Freeman, Outhred, & McNaught, 1995](#)). This study quickly became the lens through which I viewed RR from then on. It was a randomised control trial with standardised assessment – and with none of the tricks, flaws and biases evident in so many of the other studies. Yola Center and her colleagues found RR to be effective (including in improving students’ phoneme awareness and phonological recoding). This is also why the What Works Clearinghouse has determined that RR works: by focusing only on those studies that were rigorously designed and implemented.

There is more to looking at these kinds of data than identifying statistically significant differences between groups. In this case, the RR learning advantage was not particularly stark.

A full 35% of the RR kids were not discontinued. Despite 12 weeks (60 lessons) or more of RR, they failed to accomplish sufficient learning. With such a high failure rate, it should be clear that RR was not the magic bullet cure being so heavily promoted. If your school

managed to treat 16 RR students (a number rarely reached), only 10 of those students would be expected to succeed. But it gets worse.

How about the control group? How did they do? Those kids got none of the expensive RR intervention, but 31% of them managed to do well in reading anyway. There are many possible reasons why that might be ... maturation, regular classroom instruction, parent efforts ... one of the most intriguing explanations is that the RR screening procedures couldn’t distinguish youngsters with a learning problem from those a bit behind because of limited opportunity to learn (once they got some reading instruction – any reading instruction – they caught up). That latter possibility may not have been likely with the original New Zealand version of the program since RR came only after a year of reading instruction, but the US version jumped right in at the beginning of Grade 1, even when there was little or no Kindergarten reading tuition.

In any event, of those 10 RR kids who did well, five of them likely would have anyway even without RR given the success of the control group.

Effect size comparisons with other instructional efforts suggested that RR was comparable, though it was clearly more costly. RR did about the same as many of the other interventions, but this came at some cost. The RR kids needed more instruction to accomplish these outcomes, more individual instruction, and more instruction from the carefully selected “best teachers”.

We examined the available longitudinal evidence and found that the discontinued students did not tend to keep up with their classmates in second

grade and that the relative significance of their initial gains diminished yearly. A big part of the marketing of RR had been to emphasise its long-range value – RR students were going to be self-sustaining reading improvement machines! They wouldn’t need expensive special education or other kinds of extra instructional supports in coming years. The longitudinal data made us sceptical about RR’s lasting power without continued extra help for these students.

Think of it this way: there are two reasons why young children may struggle with reading – causes inside the head and causes outside the head. The inside-the-head barriers include low IQ, serious sensory deficits, cognitive processing problems, learning disabilities, etc., while the second set encompasses poverty, racism, absenteeism, neglect, poor instruction, etc.

RR successfully increases what children know about reading. But that doesn’t alter their brains, nor does it enrich environments permanently. Catching up with the other kids is nice even if temporary, but there was nothing in the instruction that would be a long-term game-changer for most kids. It shouldn’t be surprising that they begin to fall behind again as soon as the RR support is withdrawn.

That isn’t a unique problem for RR. Few early interventions have long-term benefits. But this is a particularly pointed problem for RR, given its extraordinary expense and its profligate promises.

Again, life went on and I ended up in charge of reading programs in the



Chicago Public Schools (CPS). At that time, CPS incentivised schools to adopt RR. I ended that policy immediately and discouraged (but did not ban) individual schools from continuing the program on their own.

My reasoning was this. An average Chicago elementary school at that time enrolled about 850 students, K–8, 85% of whom were likely to be reading below grade level. How could anyone justify spending almost their entire reading improvement budget on successfully raising the reading levels of 4–5 first grade students? Especially when that meant ignoring the reading needs of 700 other kids who were also below grade level, and often much further behind than those first graders.

That to me was a serious ethical problem more than a pedagogical one.

What instigated this question was a recent report from colleagues at my alma mater, the University of Delaware (May, Blakeney, Shrestha, Mazal, & Kennedy, 2022). They issued the results of a longitudinal study on RR earlier this year.

They found that despite positive outcomes at the end of Grade 1, the RR kids had fallen behind comparison kids by fourth grade – surprising to a lot of people who have relied heavily on that program, and yet consistent with the conclusions we drew 27 years ago.

Essentially, the findings suggest that the kids would have been better served without RR – since the kids so like them outperformed them in the long run. I doubt very much that RR was causing damage. But no matter how one interprets that aspect of the study, it should be clear that RR simply fails to provide long-term learning benefits.

My conclusions

1 We owe a debt of gratitude to Marie

Clay for making early reading interventions a thing. Despite the problems with RR, prior to her efforts it was uncommon for educators to respond to reading needs in Kindergarten and Grade 1.

- 2 Reading Recovery, despite some positive research results, neither is effective enough to justify its exceptional cost, nor are its small benefits long-term enough.
- 3 It should be clear, yet again, that explicit decoding instruction tends to be beneficial for students who haven't yet developed those skills. RR advocates would have been wise to adjust more based on the results of the [Iversen & Tunmer study](#).
- 4 There are no magic beans when it comes to early literacy. The trick is to catch kids up early and then to continue to strive to keep them caught up. Don't spend all your resources on that first step, because you'll need them later, too.
- 5 No matter how many ill-conceived studies there might be on a topic, it doesn't justify ignoring the well-designed ones – even if you don't like their results. Following the science does not mean cherry-picking results that are consistent with your beliefs.

References

- Center, Y., Wheldall, K., Freeman, L., Outhred, L., & McNaught, M. (1995). An experimental evaluation of Reading Recovery. *Reading Research Quarterly, 30*, 240–63.
- Clay, M.M. (1979, 1985). *The early detection of reading difficulties*. Heinemann.
- Iversen, J.A., & Tunmer, W.E. (1993). Phonological processing skills and the

Reading Recovery program. *Journal of Educational Psychology, 85*, 112–125.

- May, H., Blakeney, A., Shrestha, P., Mazal, M., & Kennedy, N. (2022, April 23). *Long-term impacts of Reading Recovery through third and fourth grade: A regression discontinuity study from 2011–12 through 2016–17*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Shanahan, T. (1987). Review of Early Detection of Reading Difficulties. *Journal of Reading Behavior, 19*(1), 117–119.
- Shanahan, T., & Barr, R. (1995). Reading Recovery: An evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly, 30*(4), 958–996.
- Yukish, J.F., & Fraas, J.W. (1988). *Success of Old Order Amish children in a strategy-oriented program for children at risk of failure in reading*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

This article originally appeared on the author's blog, [Shanahan on Literacy](#).

Timothy Shanahan [[@ReadingShanahan](#) on Twitter] is Distinguished Professor Emeritus at the University of Illinois at Chicago and was formerly Director of Reading for the Chicago Public Schools, and president of the International Literacy Association. He is a former first-grade teacher and is a parent and grandparent. His website www.shanahanonliteracy.com is popular with parents and teachers.