Response to O'Sullivan, 2026

Introduction

In *Rethinking early reading intervention policy in Ireland*, O'Sullivan (2026) argues that Reading Recovery (RR) lacks a robust evidence base, is theoretically misaligned with contemporary reading science, and should be replaced by structured literacy alternatives. These claims do not withstand scrutiny when set against the evidence. RR is supported by a substantial international research base demonstrating consistent, highly positive short-term outcomes for the lowest-attaining readers in the early years of schooling, including evidence from randomised and quasi-experimental studies, international meta-analyses, large-scale national evaluations, and UK longitudinal research showing sustained educational advantages into adolescence. O'Sullivan helpfully comments that Ireland has not yet commissioned a national, independent evaluation of RR and that such a study would assist in the evaluation of DEIS strategies (Gillece & Clerkin, 2025); however, the lack of such a study represents an evaluation gap rather than evidence of programme ineffectiveness. Irish outcome data have been systematically collected and reported locally by Reading Recovery Europe for over two decades and, when considered alongside the international and UK evidence, provide a credible foundation for policy decision making. We must differentiate between the current absence of national evaluation and claims of failure or harm. An evaluation gap does not justify abandoning RR, now operating at scale in Ireland. It does not support claims that RR is at odds with effective literacy instruction. Rather, it supports the continued use of RR as a targeted intervention for the most vulnerable learners, alongside the commissioning of a rigorous, independent Irish evaluation. Suggestions to replace an intervention with a long record of effectiveness, scalability, and accountability with alternatives that currently lack comparable longitudinal evidence, contextual relevance, or evaluative rigour are not supported by evidence.

This paper first addresses directly the claims regarding the limitations of RR then responds to the inaccuracies, misrepresentations, and issues of logic in the article.

Inaccuracy #1: Lack of high-quality evidence

Multiple high-quality, peer-reviewed studies show strong immediate impacts

There is substantial evidence (for example, Pinnell (1989); Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S. and Seltzer, M (1994); Neal & Kelly (1999); Ashdown & Simic (2000); D'Agostino & Murphy (2004); Rodgers, Gómez-Bellengé, Wang & Schulz (2005); Schwartz (2005); Burroughs-Lange & Douetil (2006); North American Trainers Group Research Committee (2006); Burroughs-Lange & Douetil (2007); What Works Clearinghouse (2008); Department for Education (2011); Schwartz, Schmitt &Lose (2012); Consortium for Policy Research in Education (2013, 2014, 2016); May, Gray, Sirinides, Goldsworthy, Armijo, Sam, Gillespie & Tognatta (2015); D'Agostino & Harmey (2016); Hurry & Fridkin (2018); Hurry & Fridkin & Holliman (2021) evaluating effectiveness with the lowest-attaining pupils in a wide range of educational contexts.

Extensive evidence from international meta-analyses

An international Meta-Analysis reviewed 203 studies and found **a** large average effect size of 0.59**,** placing RR "in the top 10% of early literacy programs reviewed by the What Works Clearinghouse" (D'Agostino & Harmey, 2016).

Numerous high quality research studies directly counter claims that RR lacks strong evidence or that its evidence base is weak.

Inaccuracy #2: Long-term effects are negative or absent

Long-term UK evidence shows positive impacts up to age 16

The long-term follow-up reported by UCL and KPMG (Hurry, Fridkin and Holliman, 2021) found that children who received RR had **better GCSE outcomes at age 16** than matched peers who did not.
This is a direct *contradiction* of the U.S. May et al. (2023) findings. It also shows that long-term positive results exist in national-level data.

The Literacy Council of North America rebuttal (2022) of the study by May et al. (2023) stresses:

- 75% attrition, making long-term inferences unreliable.
- Non-equivalent groups (controls were not lowest readers).
- Failure to account for classroom literacy conditions years later.

This response also cites multiple U.S. longitudinal studies with positive long-term outcomes.

The validity of negative long-term findings is not a settled conclusion; it is vigorously contested and challenged using evidence from contexts other than the US.

Inaccuracy #3: Methodological problems in RR research / implementation fidelity issues

Evidence that some studies used over-reliantly in O'Sullivan (2026) may have serious methodological flaws

A comprehensive research review found that five of six rigorously reviewed quantitative studies reported positive results**,** and that some studies used to critique RR have major methodological flaws. McGee (2006) shows that studies critical of RR often suffer from design issues, whereas most high-quality studies find positive effects. When analyses were re-examined using appropriate methods, outcomes for pupils are more positive than critics report**.**

Evidence of high fidelity and strong outcomes in large-scale national implementations

Evaluations of the Every Child A Reader approach in the UK (Burrough-Lange & Douëtil, 2006; Department for Education, 2011) and subsequent follow-ups show national RR roll-outs with strong fidelity and sustained outcomes.

A large number of peer reviewed research studies (listed above) directly undermine the claim that the evidence for the effectiveness of RR is weak because of poor research quality.

Inaccuracy #4: Exclusion of lowest-performing readers

Evidence shows RR specifically **targets** *the lowest-achieving pupils and supports them to succeed*

Multiple large-scale investigations consistently show RR is *most effective* for children at the very bottom of the attainment distribution (D'Agostino & Harmey, 2016). Every child taught in RR, even if only for a day is reported in local and national reporting

RR has well-evidenced immediate, highly positive outcomes for the lowest-achieving children around the age of 6.

Inaccuracy #5: RR is misaligned with required reading skills (especially phonics)

Evidence shows RR improves decoding and phonics performance.

A University College, London study (Harmey & Anders, 2024) present strong evidence that pupils' phonic knowledge was improved during their RR programme. In the U.K., primary schools emphasise phonics in early literacy instruction, so pupils identified for RR had already had systematic phonics instruction for their first year in school yet were not decoding well. It appears that RR, with its one-on-one instruction, frequent diagnosis, and linking reading with writing, found a way to crack the phonics code for most students. This evidence demonstrates that:

- Some pupils do not learn to decode in classroom early literacy programmes.
- RR improves phonic skills as evidenced by performance on the National Assessment, the Phonics Screening Check.
- Students who complete RR outperform peers who receive RR later or not yet at all.

This strongly counters the claim that RR does not effectively support phonics or decoding.

Inaccuracy #6: RR effects fade out quickly

Evidence of lasting gains in UK longitudinal study

The ten-year follow up (Hurry, Fridkin & Holliman, 2021; Hurry & Fridkin, 2018) demonstrate sustained outcomes to GCSE level at age 16. It is worth noting that RR is the only early literacy intervention being held to account for results many years later, acknowledged by May when he said that "RR is the only intervention to step up and submit itself to rigorous design." (AERA Conference, 2022).

Shanahan (2022) made a demarcation between what we can expect of early literacy intervention and performance, saying *"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."* Initial acceleration of rate of progress at one point (demonstrated by May et al 2015) followed by normal developmental variation is being used to misleadingly critique RR. Evidence to support a conclusion that RR effects fade out quickly would need to include:

- long-term negative outcomes directly attributable to the intervention
- consistent patterns across multiple high-quality independent studies in different geographical regions and contexts.

Phonics study (Harmey & Anders, 2024) shows increasing gains over time

The longer the interval between completing RR and the Year-1 Phonics Screening Check, the stronger children performed, suggesting gains accumulate rather than fade.

There is simply not the evidence to conclude that any long-term dip reported by May et al. (2023) can be attributed to inclusion in RR rather than to the quality of subsequent literacy instruction, or the realities of children's lives as they move through school.

Inaccuracy #7: RR is not cost-effective or evidence-based

Unbiased and balanced evidence endorse RR as effective

The Springer meta-analysis notes RR was previously judged effective by:

- What Works Clearinghouse (WWC, 2013)
- Evidence for ESSA (Johns Hopkins University)

These gold-standard, external evaluators defend RR, contradicting the critique and call for removing RR from the policy landscape in Ireland.

Refuting that RR is not effective in disadvantaged contexts (such as DEIS schools)

International evidence in socio-economic disadvantaged contexts

In the UK, research (Burroughs-Lange & Douetil, 2006, 2007) demonstrated that RR was effective in poor urban settings. Harmey & Anders (2024) emphasise that RR is particularly valuable in contexts with socioeconomic disadvantage, where early literacy vulnerabilities are more pronounced.

A comparison study implemented in Scotland (Gourlay & Harmey, 2020) found positive results when RR was newly introduced across multiple local authorities seeking to close attainment gaps.

RR is often implemented in areas of socio-economic disadvantage. In 2024-25 across Europe, 58.65% of pupils with completed programmes were from areas of disadvantage (in countries where this is recorded) (Reading Recovery Europe, 2025).

There is evidence indicating that RR implementations are effective across DEIS-like contexts.

In addition, to refuting the claims made by O'Sullivan (2026), it is also necessary to draw attention to some concerns regarding the trustworthiness of the critique. This next section presents some serious flaws in the approach taken to scholarship.

Flawed Logic

Research Gaps and Inaccuracies

The critique leans heavily on orthographic mapping theory to argue misalignment, yet RR pupils do not all present with persistent word-reading difficulties. Reading is a meaning-making process requiring metacognition, prediction, summarisation, and comprehension monitoring. RR aligns with the Primary Language Curriculum of Ireland. This does not appear to have been taken into consideration.

Misrepresentation of Research

O'Sullivan draws extensively on May et al. (2023) to claim RR causes negative long-term outcomes. This study suffers major methodological flaws: massive attrition, nonequivalent groups, untracked instruction quality, and was rejected by the What Works Clearinghouse. Short-term acceleration followed by normal variation cannot justify claims of harm. Multiple high-quality studies using appropriate methodologies researching across multiple contexts would be required to build a critique of this nature.

Selective Use of Evidence

The critique misrepresents RR by excluding prompts such as 'Does that look right?' which direct

children to visual/phonological cues. It inaccurately labels RR as 'whole language,' which has been repeatedly debunked. Slavin's research is misquoted; he did not compare synthetic phonics and RR. The critique would seem to be built on selective use of research information.

Lack of Practitioner and Contextual Evidence

The article offers no practitioner perspectives, despite implementation science emphasising context, fidelity, and teacher adaptation. Irish settings may integrate systematic phonics more consistently than those in other countries and therefore offer a quite different context than the study cited to discredit long term gains for RR. This does not seem to have been considered.

Issues of Logic

Generalisation: Conclusions extend beyond available evidence.
Confirmation bias: Only negative or fading studies are cited.
False cause: National literacy gaps are incorrectly attributed to RR.
False dichotomy: Framing the debate as RR vs structured literacy ignores system complexity.
Unsupported analogy: New Zealand's policy decisions do not transfer automatically to Ireland.
Argument from silence: Absence of long-term Irish data does not imply ineffectiveness.
Equivocation: Misalignment with theory does not imply practical inefficacy.

Therefore, there are flaws in the logic behind the construction of the argument.

Missing Evidence

The critique claims that there is no Irish outcome data; there has been evidence of effectiveness for the last 25 years. The researcher may be unaware of it and not looked this evidence up before writing the critique, but that does not mean it does not exist and is already used by policy makers to make support the continuation of RR in Ireland.

The alternatives suggested have no long-term comparative data, no randomized studies in any contexts and have not been in existence to make any claims about what they may offer long term. The author suggests that a switch is made from an intervention with long term evidence of scalability and effectiveness in a variety of national contexts, to classroom programmes (not interventions for the lowest attaining literacy learners) with at best little evidence of any kind.

Implications for Ireland

Short-term gains are consistent internationally, that is not under attack in the article. However, it may be that long-term outcomes depend heavily on subsequent instruction, a possible explanation for varied and contested evidence for outcomes many years after the initial intervention. Ireland's coherent phonics-informed curriculum and more stable educational system differentiate it from the fragmented US context. UK longitudinal studies—more comparable to Ireland—show positive outcomes and offer more relevant information for the Irish context.

RR fits securely within Ireland's balanced Primary Language Curriculum. RR aligns with whole-school literacy practices, SET support, consistent phonics teaching, and a literacy-rich environment and offers a personalised approach to tailor instructional emphases to the needs of each individual child.

Conclusion

The critique advanced by O'Sullivan (2026) does not provide a secure basis for recommending policy change in relation to RR in Ireland. Its conclusions rest on selective use of evidence, contested interpretations of a small number of studies, and an over-extension of findings from non-comparable contexts. In particular, the absence of a national, independent Irish evaluation is incorrectly treated as evidence of ineffectiveness rather than recognised as a gap that warrants further investigation.

O'Sullivan (2026) has not made claims for addressing the needs of the struggling learners in her article. Her argument suggests a belief that a structured program meets the needs of all learners equally. The struggling literacy learner is developing differently, and irrespective of the classroom program needs special attention and often an alternative approach. RR addresses this need, and for 40 years, annual evaluation research of RR implementations internationally confirm positive results.

A substantial international evidence base—including randomised evaluations, meta-analyses, large-scale national implementations, and long-term UK longitudinal studies—demonstrates that RR produces strong short-term gains for the lowest-attaining readers and, in some contexts, sustained educational benefits into adolescence. Claims that RR is theoretically misaligned, fades rapidly, excludes the poorest readers, or disadvantages pupils over time are not supported when the full body of high-quality evidence is considered.

Importantly, the Irish context differs materially from that of the United States, where much of the contested long-term evidence originates. Ireland's coherent Primary Language Curriculum, systematic phonics instruction, and more stable educational structures align more closely with the UK contexts in which positive long-term outcomes have been documented. Within this framework, RR functions as a targeted, short-term intervention that complements, rather than competes with, whole-school literacy provision.

Policy decisions affecting the most vulnerable learners require caution, proportionality, and strong evidential foundations. Replacing an intervention with decades of international evaluation and demonstrated scalability with alternatives lacking comparable long-term or contextual evidence represents a high-risk strategy. A more responsible course is to support a rigorous, independent national evaluation of RR in Ireland while continuing to strengthen system-wide literacy coherence and instructional quality.

Reading Recovery Europe welcomes robust, balanced, and contextually grounded scrutiny. What is required is not premature displacement, but careful evaluation—ensuring that policy

decisions are guided by comprehensive evidence and remain firmly focused on safeguarding effective support for children who struggle most to learn to read and write.

Call to Action

Reading Recovery Europe urges policymakers, school leaders, researchers, and training bodies to:

• Support a national, independent evaluation of RR in Ireland.
• Ensure any policy changes are grounded in rigorous, context-specific evidence.
• Strengthen whole-school literacy alignment, ensuring consistent phonics instruction before and after RR.
• Invest in high-quality teacher professional learning and implementation fidelity monitoring.
• Avoid premature replacement of RR with programmes lacking established Irish evidence.
• Commit to safeguarding the most vulnerable learners by basing decisions on comprehensive, unbiased, and rigorous research.

Together, we must ensure that every child struggling to gain initial literacy has access to the targeted, evidence-informed support they need to thrive as readers and writers , accelerate their learning, and return to their classrooms prepared to continue progressing in literacy without ongoing special support.

Reading Recovery Europe, April 2026

References:

- Ashdown, J. and Simic, O. (2000) 'Is early literacy intervention effective for English language learners? Evidence from Reading Recovery'. *Literacy Teaching and Learning*, 5 (1), 27-42.
- Burroughs-Lange, S. and Douëtil, J. (2006) Evaluation of Reading Recovery in London Schools Every Child A Reader 2005-2006.pdf *University of London: Institute of Education*.
- Burroughs-Lange, S. and Douëtil, J (2007) 'Literacy progress of young children from poor urban settings: A Reading Recovery comparison study'. *Literacy Teaching and Learning*, 12 (1), pp 19-46.
- Consortium for Policy Research in Education (2013), '*Evaluation of the i3 Scale-up of Reading Recovery: Year One Report, 2011-12*', RR-76.
- What Works Clearinghouse Intervention Report: Reading Recovery. (July 2013). United States Department of Education, Institute of Education Sciences.
- Consortium for Policy Research in Education (CPRE). (2013, 2014, 2016). Evaluation of the i3 Scale-up of Reading Recovery reports.
- D'Agostino, J.V., & Harmey, S.J. (2016), An International Meta- Analysis of Reading Recovery, Journal of Education for Students Placed at Risk (JESPAR), Vol 21, NO.1, 29-46.

- D'Agostino, J.V., and Murphy, J.A. (2004) 'A meta-analysis of Reading Recovery in United States schools'. *Educational Evaluation and Policy Analysis*, 26 (1), 23-38.
- Department for Education (2011), '*Evaluation of Every Child a Reader (ECaR)',* DFE-RR114.
- Evidence for ESSA. (n.d.). Reading Recovery. John Hopkins University. https://www.evidenceforessa.org/program/reading-recovery/. Accessed March 2026.
- Harmey, S. J. & Anders, J. (2024) The link between completing Reading Recovery and performance on a phonics screening check, *Journal of Education for Students Placed at Risk* (JESPAR), 29:4, 311-331, DOI: 10.1080/10824669.2023.2179057.
- Hurry, J. and Fridkin, L. (2018), The impact of Reading Recovery ten years after intervention', UCL Institute of Education.
- Hurry, J., Fridkin, L. and Holliman, A. J. (2021). Reading intervention at age 6: Long-term effects of Reading Recovery in the UK on qualifications and support at age 16. British Educational Research Journal Vol 48, Issue 1.
- Gilleece, L., & Clerkin, A. (2025). Towards more robust evaluation of policies and programmes in education: identifying challenges in evaluating DEIS and Reading Recovery. Irish Educational Studies, 44(3), 445–473. https://doi.org/10.1080/03323315.2024.2334704
- Gourlay, G., & Harmey, S. (2020). A Reading Recovery Comparison Study: Supporting a New Implementation in Scotland. *Journal of Reading Recovery, 20* (1), 59 – 70.
- Literacy Council of North America (2022) *Revisiting the Recently Reported Longitudinal Study of Reading Recovery* Available at: https://myliteracycouncil.org/revisiting-the-recently-reported-longitudinal-study-of-reading-recovery/ Accessed on 29th March 2026
- May, H., Gray A., Sirinides P., Goldsworthy H., Armijo M., Sam C., Gillespie J.N., Tognatta N. (2015), *'Year One Results From the Multisite Randomized Evaluation of the i3 Scale-Up of Reading Recovery',* American Educational Research Journal, 52, (3).
- May, Henry, Aly Blakeney, Pragya Shrestha, Mia Mazal, and Nicole Kennedy (2023) Long-Term Impacts of Reading Recovery through Third and Fourth Grade: A Regression Discontinuity Study. *Journal of Research on Educational Effectiveness* 1:1–26. Available at: https://doi.org/10.17605/osf.io/jc9bm. Accessed 29th March 2026.
- McGee, Lea M. (2006) Research on Reading Recovery: What Is the Impact on Early Literacy Research? *Literacy Teaching and Learning*, v10 n2 p1-50, 2006.
- Neal, J.C. and Kelly, P.R. (1999) 'The success of Reading Recovery for English language learners and descubriendo la lectura for bilingual students in California'. *Literacy Teaching and Learning*, 4 (2), 81-108.
- North American Trainers Group Research Committee (2006) '*Six Reading Recovery studies: Meeting the criteria for scientifically based research'*. Reading Recovery Council of North America: Columbus, OH, USA.
- O'Sullivan, J. (2026). Rethinking early reading intervention policy in Ireland: evaluating the case for moving beyond Reading Recovery? *Irish Educational Studies*, 1–20. https://doi.org/10.1080/03323315.2026.2644149
- Pinnell, G. S. (1989). Reading Recovery: Helping At-Risk Children Learn to Read. *The Elementary School Journal*, 90, 161-183.

- Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S. and Seltzer, M (1994) 'Comparing instructional models for the literacy education of high-risk first graders'. *Reading Research Quarterly*, 29 (1), 8-39.
- Reading Recovery Europe (2025) Available at: https://readingrecoveryeurope.org/wp-content/uploads/2025/09/europe_report_for_reading_recovery_2023-24.pdf Accessed March 2026.
- Rodgers, E., Gómez-Bellengé, F., Wang, C. and Schulz, M. (2005) Predicting the literacy achievement of struggling readers: Does intervening early make a difference? Paper presented at the annual meeting of the *American Educational Research Association*, Montreal, Quebec, Canada.
- Schwartz, R. M. (2005) 'Literacy learning of at-risk first-grade students in the Reading Recovery early intervention'. *Journal of Educational Psychology*, 97 (2), 257-267.
- Schwartz, R.M. Schmitt, M.C. and Lose, M.K. (2012), *'Effects of teacher-student ratio in response to intervention approaches'* The Elementary School Journal, 112 (4), 547-567.
- What Works Clearinghouse (2008) 'Intervention report: Reading Recovery' U.S. Department of Education, Institute of Education Sciences.