Reading Recovery: A Sector-Wide Analysis

SUMMARY

This Learning Curve briefly describes the results of an evaluation examining the impact of Reading Recovery (RR) on students' outcomes in NSW government schools. The results showed some evidence that RR has a modest short-term effect on reading skills among the lowest performing students. However, RR does not appear to be an effective intervention for students that begin Year 1 with more proficient literacy skills. In the longer-term, there was no evidence of any positive effects of RR on students' reading performance in Year 3.

Background

Reading Recovery (RR) is one of the most widely used interventions to improve outcomes for NSW students who are struggling to read in Year 1.

In 2012 RR was offered in over half of NSW government primary schools. RR was developed in New Zealand in the 1970s by Dame Marie Clay as an intensive intervention that aims to lift literacy skills among students performing in the bottom 20 per cent of Year 1 (Department of Education and Communities 2015; What Works Clearinghouse 2008). RR is provided on a one-to-one basis over 12-20 weeks with the intention of raising students' performance to the average level of their Year 1 peers, thereby allowing them to benefit from classroom instruction and successfully progress through the early years of school (Department of Education and Communities 2015; May et al. 2013; 2015).

Students are identified as eligible for RR using a standardised teacher-administered assessment (Clay 2002; Reading Recovery Council of North America 2015)[1]. Students 'successfully discontinue' RR when they have achieved the average reading level for their grade. Students who do not reach this level after 12-20 weeks are referred for further specialist support or for long-term literacy support. Students who do not complete their lessons within a calendar year may have their lessons carried over to the next year. Students may also stop participating in RR if they transfer schools and are not able to continue with RR lessons at their new school.

For students who participated in RR in NSW government primary schools in 2012[2], the majority were 'successfully discontinued' (80.2%). Approximately 15 per cent were referred for further specialist or long-term literacy support, while the remaining students either transferred out of the program or had their lessons carried over to the following year.

---

1 This includes six literacy tasks: Letter Identification, Word Test, Concepts About Print, Writing Vocabulary, Hearing and Recording Sounds in Words and Text Reading (see http://readingrecovery.org/reading-recovery/teaching-children/observation-survey).
2 This includes all students who participated in RR in 2012 and had matched data records from the Best Start assessment in 2011, the Literacy Continuum assessments at the end of Kindergarten (2011) and at the end of Year 1 (2012), and at Year 3 NAPLAN (n = 7,573).

Evidence for the Effectiveness of RR

A considerable amount of research has been conducted world-wide examining the impact of RR on student outcomes. Unfortunately, the evidence derived from most studies has been based on relatively weak research methodologies, which calls into question the reliability of the findings (D'Agostino & Murphy 2004; May et al. 2015; What Works Clearinghouse 2008; 2013). However, among those studies that could be considered rigorous sources of evidence (all Randomised Controlled Trials, RCTs), findings suggest that RR is an effective intervention for increasing student literacy levels (May et al. 2015; What Works Clearinghouse 2013). These positive outcomes are consistent with a recent report commissioned by the NSW Ministerial Advisory Group on Literacy and Numeracy, which concluded that RR was one of the few interventions available in NSW with a reasonably strong evidence base for its efficacy (Australian Council for Educational Research 2013).

While this evidence provides reason to be cautiously optimistic about the effectiveness of RR, previous studies are not without their limitations. For example, each of the studies included in the What Works Clearinghouse review had relatively small sample sizes (fewer than 100 students). Program effects from small, non-representative samples are not necessarily applicable to whole school populations. Furthermore, even though RCTs are the best way of estimating true and unbiased treatment effects, one limitation is that the control groups either receive no additional intervention or a constrained alternative. Both of these options are unlikely to represent how educators respond to low-achieving students in the absence of RR.

RR has also been subject to extensive criticism, particularly in New Zealand, where it has formed a key part of the national literacy strategy for over 25 years (Chapman & Tunmer 2011; Greaney 2011; Tunmer et al. 2013). Tunmer et al. (2013) point out that RR has failed to lift the literacy skills of the poorest performing students, evidenced by the fact that there has been no closing of the achievement gap between low and high performing readers in the 25 years the program has been operating. In support of their argument, Tunmer et al. observe that students at higher starting points are most likely to successfully complete (discontinue) the program and that gains among these higher performing students are not sustained in the longer term (Chapman & Tunmer 2011; Moats 2007; Tunmer & Chapman 2003; Tunmer et al. 2013).

While these are important considerations, it is perhaps no surprise that students at higher starting points are more likely to complete the program; they simply require less improvement before they are successfully discontinued. The effectiveness of RR for students at the lowest starting points can only be ascertained by comparing growth among students at similar starting points who do not receive RR. The sustainability of any benefits associated with RR, on the other hand, is a critically important issue that has not been rigorously examined to date.

Objectives of the Current Study

While the balance of the evidence suggests that RR is an effective intervention for raising student literacy levels, most evaluations of RR have been conducted outside Australia. This raises the question of whether RR is equally effective in the NSW context. Second, the extant evidence has not resolved the issues raised by critics regarding the effectiveness of RR for low performing students. Research that carefully accounts for student baseline achievement is needed to assess whether RR is differentially effective for students at low versus high starting points. Finally, the long-term sustainability of the results achieved by RR has not been rigorously addressed in the existing literature.

The primary aims of the current study were to examine the impact of RR on students' literacy outcomes and whether any benefits associated with participating in RR are sustained over the longer term. This evaluation was conducted at the sector-level (state-wide across NSW government schools) and focussed on identifying the impact of RR on student outcomes compared to similar students who attended a school that did not offer RR. An important objective of the current study was to determine whether there was a relationship between students' literacy skills prior to Year 1 and the effectiveness of RR.

The two key research questions addressed in this evaluation were:

1. What is the impact of RR on students' literacy progress at the end of Year 1?

2. Are any benefits of RR sustained to Year 3?

Method

Design

Literacy outcomes for students who participated in RR were compared to a group of students with similar starting reading levels and socio-demographic characteristics who did not participate in RR[3].

Data Sources

Three sources of student-level data were used in this evaluation.

1. Reading Recovery Data: The intervention group was comprised of all students participating in RR in 2012 who were deemed to have completed the program and exited with either a status of 'successfully discontinued' or 'referred' for long-term literacy support[4].

2. Best Start and Literacy Continuum K-10 Data: The NSW Literacy Continuum K-10 contains eight evidence-based literacy aspects that are regarded as critical to literacy success: Reading Texts, Comprehension, Vocabulary Knowledge, Aspects of Writing, Aspects of Speaking, Phonics, Phonemic Awareness and Concepts About Print. As student literacy skills develop within each aspect, they are expected to move progressively across Continuum clusters[5]. Assessment on the Continuum at the beginning of Kindergarten (i.e. Best Start assessment) has been mandatory in all government primary schools since 2010[6]. Following the Best Start assessment, use of the Continuum is not mandatory and is used at the discretion of schools and teachers. In the current study, Literacy Continuum data were sourced for all students who had been tracked against the Continuum at 3 time-points: the beginning of Kindergarten (i.e. the Best Start assessment), at Term 4 of Kindergarten (T4K), and Term 4 of Year 1.

3. NAPLAN Data: To examine the longer-term effects of participating in RR, scaled scores on the 2014 NAPLAN Year 3 Reading assessment were analysed for RR and non-RR students.

---

3 To reduce the risk of selection bias these students were also in schools that did not offer RR.
4 The sample of students who were transferred or carried over to the following year was too small to analyse separately.
5 For further information, see: http://www.curriculumsupport.education.nsw.gov.au/literacy/.
6 This assessment aims to help teachers identify the skills that each student brings to school and includes 7 of the 8 Continuum aspects (Vocabulary Knowledge is excluded).

Impact of RR at the end of Year 1

To determine whether there were differences in reading outcomes at the end of Year 1, students' progress on seven aspects of the Literacy Continuum was compared for students who participated in RR versus those who did not[7]. However, this Learning Curve only reports results for three Continuum aspects: Reading Texts, Comprehension and Phonics. The results for the remaining aspects (detailed in the full report available at http://www.cese.nsw.gov.au) are broadly similar to those for Comprehension and Phonics and are not reported here in the interests of brevity. Separate statistical analyses were carried out for each Continuum aspect.

Each analysis accounted for the impact of relevant student- and school-level factors known to be associated with academic performance. These included: gender, Aboriginal status, country of birth, Non-English Speaking Background, English as a Second Language (ESL) Phase, parent qualification, parent occupation group, parent's spoken language at home, school participation in the Priority Schools Funding Program, school location, school socio-economic status, and school region. To account for important differences in RR and non-RR students' reading skills prior to the availability of RR in Year 1, the cluster (level)[8] achieved on the relevant aspect (Reading Texts, Comprehension or Phonics) at Best Start and at the end of Kindergarten were also included in each aspect analysis.

Impact of RR on Year 3 Reading Outcomes

To examine the longer-term effect of RR, students' scaled scores on the Year 3 NAPLAN Reading assessment were compared for students who participated in RR versus students who had similar prior achievement levels in schools that did not offer RR[9]. This analysis included the same student- and school-level factors described above to account for other factors that are known to be associated with academic achievement.

Results

Impact of RR at the end of Year 1

The impact of RR on students' literacy progress at the end of Year 1 varied depending on how well students were assessed to be reading at the end of Kindergarten. Results for Reading Texts (summarised in Figure 1) are presented as odds ratios which represent the likelihood of achieving a higher Reading Texts level at the end of Year 1 for RR students relative to non-RR students. As shown in Figure 1, while RR was effective at improving reading outcomes for students at the lowest starting reading levels, it was not effective for students at higher starting points relative to non-RR students. Students who participated in RR and were assessed at Level 1 or below in Reading Texts at the end of Kindergarten had odds of progressing to a higher Reading Texts level at the end of Year 1 that were 2.7 times higher than their non-RR counterparts. The magnitude of this effect decreased to 1.5, and was not statistically significant for students who started at Level 2 in Reading Texts at the end of Kindergarten. Above Level 2, results showed that non-RR students had significantly higher odds of progressing on Reading Texts than RR students at the end of Year 1, even after accounting for important contextual variables.

Figure 1. Estimated effect of RR on the likelihood of achieving a higher Reading Texts level at the end of Year 1

Less likely ← (Odds ratio is less than 1.0) **→ More likely** (Odds ratio is greater than 1.0)
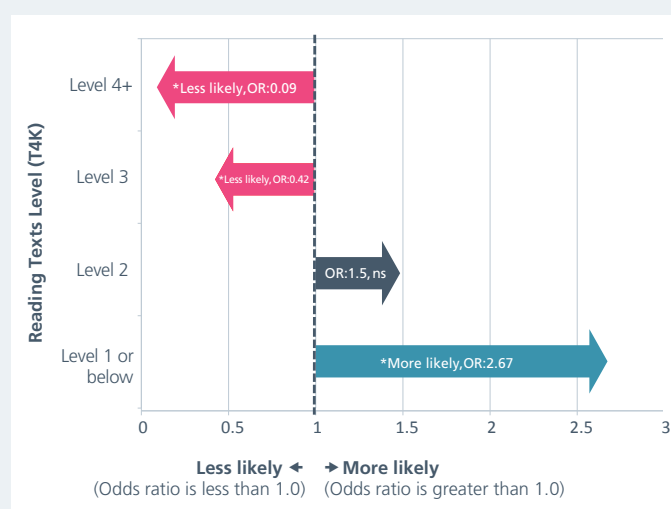
Note. Odds ratios represent the odds of achieving a higher level on Reading Texts at the end of Year 1 for RR students relative to non-RR students. Results coloured in **cyan** show that RR students had higher odds of progressing on Reading Texts relative to non-RR students; results coloured in **pink** show that non-RR students had higher odds of progressing on Reading Texts relative to RR students; results coloured in **navy** show that RR and non-RR students were equally likely to progress on Reading Texts.

In relation to progress on Comprehension at the end of Year 1, results presented in Figure 2 show no positive effects for RR. For students who were assessed at the lowest levels (Level 1 or below) on Comprehension at the end of Kindergarten, there were no significant differences in Comprehension progress at the end of Year 1 for RR and non-RR students. However, for students at higher starting points (Level 2 and above), all odds ratios presented in Figure 2 are less than 1, indicating that non-RR students had significantly higher odds of progressing on Comprehension than RR students at the end of Year 1.

Figure 2. Estimated effect of RR on the likelihood of achieving a higher Comprehension level at the end of Year 1

Less likely ← (Odds ratio is less than 1.0) **→ More likely** (Odds ratio is greater than 1.0)
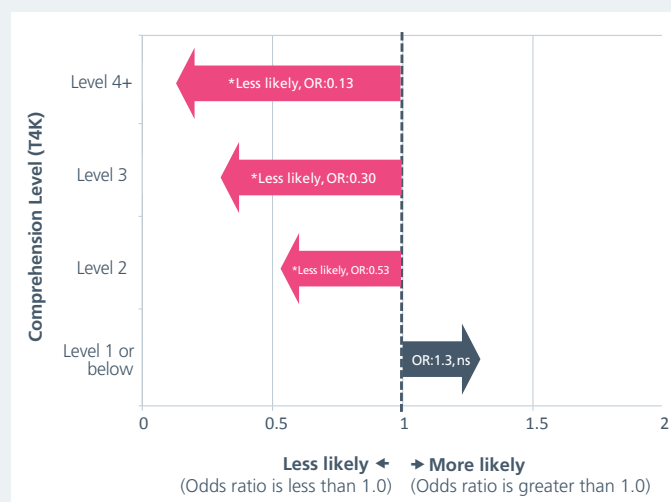
ote. Odds ratios represent the odds of achieving a higher level on Comprehension at e end of Year 1 for RR students relative to non-RR students. Results coloured in **pink** ow that non-RR students had higher odds of progressing on Comprehension lative to RR students; results coloured in **navy** show that RR and non-RR students were equally likely to progress on Comprehension.

---

7   Students were included in each analysis if they had complete demographic information and valid scores (i.e. not missing, blank or unknown) on the relevant aspect at Best Start and at Term 4 Kindergarten and Term 4 of Year 1. Since assessment on the Continuum is not mandatory after Best Start, a number of RR and non-RR students had missing records at the end of Kindergarten and at the end of Year 1 and could not be included in each analysis. See full report (http://www.cese.nsw.gov.au) for details.

8   In all subsequent analyses, Continuum clusters are treated as levels, with Level 0 equal to Cluster 1; Level 1 equal to Cluster 2 etc.

9   Students were included in the analysis if they had complete demographic information, a valid (i.e. non-missing) NAPLAN score on the Reading domain in Year 3, as well as valid scores on the Literacy Continuum aspects for Reading Texts and Comprehension in Best Start and at Term 4 Kindergarten.

In relation to progress on Phonics at the end of Year 1, results presented in Figure 3 also show no positive effects for RR. For students who were assessed at the lowest levels (Level 2 or below) on Phonics at the end of Kindergarten, there were no significant differences in Phonics progress at the end of Year 1 for RR and non-RR students. However, for students at higher starting points (Level 3 and above), all odds ratios presented in Figure 3 are less than 1, indicating that non-RR students had significantly higher odds of progressing on Phonics than RR students at the end of Year 1.

Figure 3. Estimated effect of RR on the likelihood of achieving a higher Phonics level at the end of Year 1
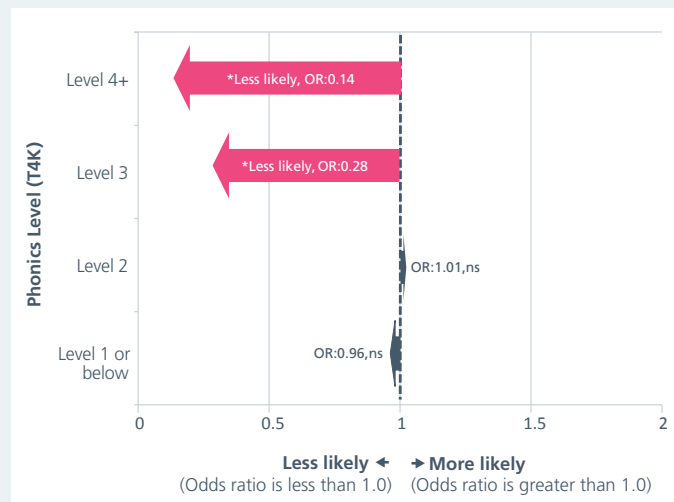

Note. Odds ratios represent the odds of achieving a higher level on Phonics at the end of Year 1 for RR students relative to non-RR students. Results coloured in **pink** show that non-RR students had higher odds of progressing on Phonics relative to RR students; results coloured in **navy** show that RR and non-RR students were equally likely to progress on Phonics.

Sustainability to Year 3

While the magnitude of the impact of RR on students' Year 3 NAPLAN Reading scores varied depending on how well students were assessed to be reading at the end of Kindergarten, results showed that RR students achieved lower scores than non-RR students irrespective of their starting points. As shown in Figure 4, students who participated in RR achieved significantly lower scores on the Year 3 NAPLAN Reading assessment at all starting levels, with much larger differences observed for students at higher starting points. For students at the lowest levels at the end of Kindergarten (Level 1 or below), the average difference in scaled NAPLAN Reading scores was 25.2 points in favour of non-RR students. For students at the highest levels at the end of Kindergarten (Level 4 or above), the difference in Year 3 NAPLAN Reading scores increased to 86.7 points in favour of non-RR students.

Figure 4. Estimated effect of RR on scaled NAPLAN Reading scores in Year 3 (2014)
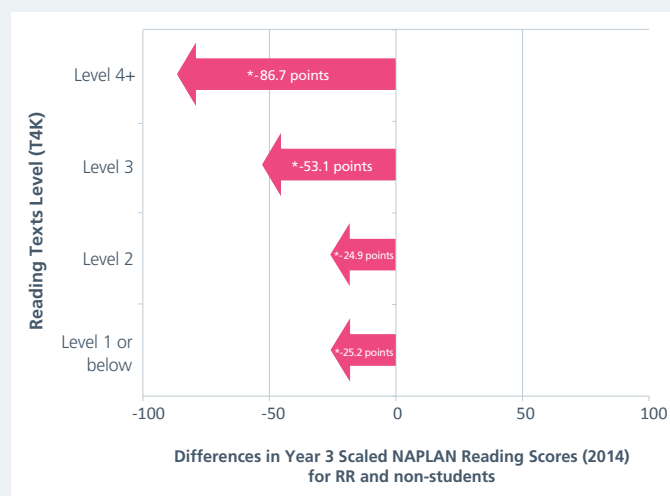
Note. Results coloured in **pink** show that RR students achieved lower NAPLAN Reading scores compared to non-RR students.

Discussion

The results from this retrospective evaluation provide some evidence that RR is effective at improving short-term reading outcomes at the end of Year 1. However, this was only evident on the Continuum aspect corresponding directly to Reading Texts and was only observed for students identified as the lowest performing readers at the end of Kindergarten. For Comprehension and Phonics, the results for the lowest performing readers were equivalent for RR and non-RR students. Students achieving higher reading levels at the end of Kindergarten showed less growth on the Continuum if they participated in RR compared to similar students in non-RR schools. These findings suggest that RR is an effective intervention for improving short-term reading outcomes among the poorest performing readers, which is the primary intention of the intervention.

In relation to the longer-term intervention effects, results showed no evidence that RR has any positive effects on students' NAPLAN Reading performance in Year 3. Irrespective of students' starting reading ability, students who participated in RR achieved significantly lower reading scores in NAPLAN in Year 3 compared to their non-RR counterparts.

The lack of sustained benefit associated with RR should not necessarily be interpreted as a program failure. The duration of the program is only 12-20 weeks so it is equally possible that RR students do not receive the level of support they need to sustain any short-term RR effects beyond Year 1. There is a relatively large time lag between student participation in RR in Year 1 and NAPLAN testing in Term 2 of Year 3. Students, especially those identified as at-risk of falling behind, are often exposed to a range of initiatives throughout the early years of school. While the current study provides the first rigorous sector-level analysis of the relative longer-term effect of RR in NSW, the impact of RR on students' Year 3 outcomes cannot be isolated from the potential effects of any exposure to other initiatives in the intervening period between Year 1 and Year 3.

The current study was not designed to identify why the short-term benefits of RR were not sustained to Year 3. This can only be done by closely inspecting the way RR is being delivered in NSW government schools and how students are supported beyond their participation in the program. One explanation that is asserted strongly by RR critics is that RR does not provide sufficient tuition in phonics and phonemic awareness to effectively remediate literacy performance among struggling readers (Center et al. 1995; Chapman & Tunmer 2011; Greaney 2011; Moats 2007; Reynolds & Wheldall 2007; Tunmer & Chapman 2003; Tunmer et al. 2013). It is well-accepted in the scientific literature that systematic early instruction in phonics is critical to the development of early reading skills, particularly for struggling students, and should form an integral component of a balanced approach to literacy development (Education Endowment Foundation 2015; Ehri et al. 2001; Hattie 2009; U.S. Department of Health and Human Services 2000; Rowe 2005). Unfortunately, information is not available at the sector-level on what types of remedial literacy interventions are offered in non-RR schools. It may be the case that non-RR students received interventions that comprised more comprehensive instruction in phonics and were therefore more effective in remediating student literacy skills through to Year 3. This is an important question that should be the focus of future research.

Another possible factor that may limit the effectiveness of RR relates to teacher quality and intervention fidelity, neither of which could be assessed in the current analysis. A key assumption of RR is that high quality instruction is essential to accelerating the literacy skills of struggling students (May et al. 2013; 2015). While the RR guidelines in NSW stipulate that all RR teachers are required to undergo training for one year in the implementation of RR and undertake ongoing professional development (Department of Education and Communities 2015), the possibility remains that the actual quality and consistency of implementation may vary across schools. Future research should include an in-depth analysis of intervention fidelity to better understand whether, and to what extent, these factors may influence student outcomes.

Finally, an important concern identified in the current study was that there were a number of students participating in RR who do not appear to fit the profile of struggling readers. For example, analysis of Literacy Continuum data showed that approximately 30 per cent of students who go on to participate in RR in Year 1 are judged to be meeting expected learning outcomes on Reading Texts at the end of Kindergarten. While the Continuum benchmarks may require some further refinement and could under-estimate students' needs for further literacy support, these findings could also point to some issues with the implementation of RR in NSW. Indeed, it is clear that at a sector-level, there are currently some relatively high-performing students participating in RR. While these students may be lower performing readers in a relatively high-achieving school or class context, they may not be appropriate candidates to participate in an intervention such as RR that was developed to target the lowest performing students in the state. It is possible that these relatively high-achieving students would be better-placed receiving high quality classroom-based instruction.

Limitations

While the current study represents a rigorous retrospective analysis assessing the sector-level impact of RR in NSW, there are two key limitations that must be acknowledged.

1. The Literacy Continuum was not developed as a robust assessment tool to measure student progress. Apart from Best Start, judgements about students' progress are subjectively made by teachers by extrapolating the outcomes from school-based assessments and classroom observation to align with the Continuum clusters. This raises a number of potential issues in relation to the reliability of these judgements, the extent to which school-based assessments align with the Continuum, and the ability of teachers to appropriately map school-based assessments to Continuum benchmarks. The assumption made in the current analysis was that the variability in consistency of teacher judgement did not differ systematically across RR and non-RR schools.

2. While the current analysis accounted for all available and measurable demographic and prior achievement indicators, there were some variables that may be causally related to student outcomes that could not be measured and accounted for in determining the impact of RR. For example, it was not possible to measure any within-school factors (e.g. teacher quality) that may have impacted student performance. Furthermore, it was not possible to account for all important student-level differences. The most notable omitted variable was an indicator of students' learning disabilities, which is likely to have a significant impact on student achievement. The over-riding assumption in the current analysis was that any important omitted variables were equally distributed across RR and non-RR schools.

Summary and Implications

The results from the current analysis provide some evidence that RR is an effective short-term intervention for remediating Reading Text skills among the lowest performing students. However, RR does not appear to be an effective intervention for students who begin Year 1 with more proficient literacy skills, at least compared to other interventions or initiatives that are available in non-RR schools. The implication of these findings is that, currently, the most cost-effective method of implementing RR in NSW may be to target only the students performing at the lowest levels at the end of Kindergarten (at a sector- not a school-level) or to restrict RR to schools that are identified as having a high number of students who are not meeting performance benchmarks in Kindergarten or early Year 1. The limitations of this analysis also highlight the strong need to collect better information on the teaching practices and interventions being offered in non-RR schools and to develop valid and reliable measures of students' literacy progress throughout the early years of school.

References

Australian Council for Educational Research 2013, *Literacy and numeracy interventions in the early years of schooling: A literature review,* Report to the Ministerial Advisory Group on Literacy and Numeracy.

Center, Y, Wheldall, K, Freeman, L, Outhred, L, & Mcnaught, M 1995, 'An evaluation of Reading Recovery', *Reading Research Quarterly*, vol. 30, pp. 240–263.

Chapman, JW & Tunmer, WE 2011, 'Reading Recovery: Does it work?', *Perspectives on Language and Literacy*, Fall, pp. 21-24.

Clay, M 2002, *An Observation Survey of Early Literacy Achievement.* 2nd edn, Heinemann, Portsmouth, N.H.

D'Agostino, JF & Murphy, J 2004, 'A meta-analysis of Reading Recovery in United States Schools', *Educational Evaluation and Policy Analysis*, vol. 26, no. 1, pp 23-38.

Department of Education and Communities 2015, *Reading Recovery Guidelines for 2015 and beyond*, Department of Education and Communities, viewed 27 March 2015, http://www.curriculumsupport.education.nsw.gov.au/earlyyears/reading_recovery/docs/rr_guidelin_15.pdf

Education Endowment Foundation 2015, *Evidence and Data: Phonics*, viewed 27 March 2015 http://educationendowmentfoundation.org.uk/toolkit/toolkit-a-z/phonics/

Ehri, LC, Nunes, SR, Stahl, SA & Willows, DM 2001, 'Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's Meta-Analysis', *Review of Educational Research*, vol. 71, no. 3, pp. 393-447.

Greaney, K 2011, 'The multiple cues or "searchlights" word reading theory: Implications for Reading Recovery', *Perspectives on Language and Literacy*, Fall, pp. 15-19.

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

May, H, Gray, A, Gillespie, JN, Sirinides, P, Sam, C, Goldsworthy, H, Armijo, M & Tognatta, N 2013, *Evaluation of the i3 scale-up of Reading Recovery: Year one report, 2011-2012,* Consortium for Policy Research in Education, viewed 27 March 2015 http://cpre.org/sites/default/files/researchreport/1488_readingrecoveryreport.pdf

May, H, Gray, A, Sirindes, P, Goldsworthy, H, Armijo, M, Sam, C, Gillespie, JN & Tognatta, N 2015, 'Year one results from the Randomized Evaluation of the i3 scale-up of Reading Recovery', *American Educational Research Journal,* advance online publication.

Moats, L 2007, *Whole-language high jinks: How to tell when "scientifically-based instruction" isn't.* Thomas Fordham Foundation.

Reading Recovery Council of North America 2015, *Clay's Observation Survey,* viewed 5 November 2015 http://readingrecovery.org/reading-recovery/teaching-children/observation-survey.

Reynolds, M & Wheldall, K 2007, 'Reading Recovery 20 years down the track: Looking forward, looking back' *International Journal of Disability, Development and Education,* vol. 54, no. 2, pp. 199-223.

Rowe, K 2005, Teaching Reading literature review: *A review of the evidence-based research literature on approaches to the teaching of literacy, particularly those that are effective in assisting students with reading difficulties.* A report of the Committee for the National Inquiry into the Teaching of Literacy, Australian Government Department of Education, Science and Training, Canberra, ACT.

Tunmer, WE & Chapman, JW 2003, 'The Reading Recovery approach to preventative early intervention: As good as it gets?', *Reading Psychology,* vol. 24, pp. 337-360.

Tunmer, WE, Chapman, JW, Greaney, KT, Prochnow, JE. & Arrow, AW 2013, 'Why the New Zealand National Literacy Strategy has failed and what can be done about it: Evidence from the Progress in International Reading Literacy Study (PIRLS) 2011 and Reading Recovery monitoring reports' *Australian Journal of Learning Difficulties,* vol. 18, no. 2, pp. 139-180.

U.S. Department of Health and Human Services 2000, *Report of the National Reading Panel: Teaching children to read – findings and determinations of the National Reading Panel by Topic Areas,* viewed 27 March 2015, http://www.nichd.nih.gov/publications/pubs/nrp/pages/findings.aspx.

What Works Clearinghouse 2008, *Intervention report: Reading Recovery,* United States Department of Education, Institute of Education Sciences.

What Works Clearinghouse 2013, *What Works Clearinghouse Intervention Report: Reading Recovery.* United States Department of Education, Institute of Education Sciences.