British Educational Research Journal
Vol. 48, No. 1, February 2022, pp. 5–21
DOI: 10.1002/berj.3752

Reading intervention at age 6: Long-term effects of Reading Recovery in the UK on qualifications and support at age 16

Jane Hurry*, Lisa Fridkin and Andrew J. Holliman
University College London, UK

One argument for early intervention for reading difficulties is that it can sustainably improve children's reading competence trajectory (the 'inoculation model'), but there are virtually no studies on sustained effects at the end of compulsory schooling. This study reports on a 10-year follow-up of a widely used early literacy intervention, Reading Recovery. UK schools adopting Reading Recovery enrol selected teachers for a year's training, after which they provide one-to-one tutoring and typically act as literacy advisors. In a quasi-experimental, intention to treat, design, 293 6-year-olds with reading difficulties in 42 London schools were assigned to Reading Recovery (RR), standard provision in Reading Recovery schools (RRS) or standard provision in comparison schools (CS). Children were traced at ages 14 (204) and 16 (271) and data collected from the National Pupil Database. At age 14 and 16, significantly fewer RR than CS pupils were officially identified as having special educational needs, a potential consequence of reading difficulties. Using multi-level modelling and controlling for baseline reading and Free School Meal status (an indicator of poverty), at age 16 the RR group significantly outperformed the CS group on academic qualifications (GCSEs) (d = 0.52). However, the RRS group also performed significantly better than the CS group (d = 0.37), consistent with the fact that standard provision for weaker readers in RR schools differed from that provided in CS. Thus, these results support the long-term effects of early intervention but raise questions about the importance of whole-school effects and systemic intervention.

Keywords: reading intervention; longitudinal; social disadvantage; reading difficulty

Introduction

A key task of schooling is to ensure that children become confident readers and writers, prepared for the demands of adult life. Early reading and writing problems frequently persist into late teens and adulthood (Blachman *et al.*, 2014), and adults with poor literacy or qualifications earn less and are more susceptible to unemployment and ill-health (Desjardins *et al.*, 2013). As a response to addressing the long-term consequences of inadequate reading skills, it is widely considered that children with reading difficulties should be offered early intervention, and this is supported by the evidence of its short-term effectiveness (National Reading Panel, 2000; Torgesen, 2000). Specific to Reading Recovery, the focus of this paper and an intensive one-to-

*Corresponding author. Department of Psychology and Human Development, Institute of Education, University College London, 25 Woburn Square, London WC1H 0AA, UK. Email: j.hurry@ucl.ac.uk

one literacy programme for poor readers in their second year of schooling, evidence from previous studies and meta-analyses (Hurry & Sylva, 2007; D'Agostino & Harmey, 2016) demonstrate positive programme effects in the shorter term. However, there is a paucity of research on the durability of the gains made during early reading interventions, either for Reading Recovery or other methods. There are a range of reasons for making a slow start in reading, including biological, developmental and environmental factors. Some of these will inevitably continue to exert their influence as the child matures, some will resolve themselves; however, the potential of early intervention is that it will reduce some of the negative cycles of interaction and provide important foundational skills. The present study reports on a 10-year follow-up of Reading Recovery, examining its longer-term influence on pupils' national qualifications and special educational needs provision.

Proposed mechanisms for early intervention

The first reason why early intervention is promoted as preventative of later literacy problems is that it addresses the otherwise widening gap between poor readers and their peers as they move through school (Chall, 1983; Stanovich, 1986). Poor readers read less than their peers (Adams, 1990), which holds back their reading, language development, general knowledge and IQ (Stanovich, 1986). The second reason is that early intervention provides foundational skills on which later literacy is dependent. The broad underlying dimensions of reading have been usefully captured in the Simple View of Reading (Gough, 1996), which proposes that reading comprehension is a combination of word reading efficiency and language comprehension. Both theory and extensive evidence concur that phonological processing is fundamental to word reading efficiency (National Reading Panel, 2000; Stuart, 2002; Suggate, 2016). 'Phase' models of reading development propose that understanding the alphabetic principle is the critical early hurdle for the child, underpinning word reading efficiency (Frith, 1985; Stanovich, 1986; Byrne, 1998; Ehri, 2005) but by no means the end of development, where fluent reading and comprehension become central. This is supported by reviews of intervention studies (National Reading Panel, 2000; Suggate, 2016), reporting phonological interventions to be more effective for beginning readers and comprehension to be more effective for older readers. However, effectiveness is rarely measured much beyond the end of the intervention period, so long-term implications remain largely unknown.

Issues for sustained impact of early intervention

Relevant to sustained impact, there is a debate about whether early interventions should focus principally on phonological skills or whether comprehension should also be addressed. The searchlights model, being implemented in English schools during the period of this study (National Literacy Strategy; DfEE, 1998), recommended that the teaching of reading should have a broad base from the beginning of a child's education, to include decoding, comprehension, grammatical understanding and a more general experience of different books and texts. Reading Recovery addresses both phonological skills and comprehension. It was developed by Marie Clay, based on her research on children learning to read in their first year of school (Clay, 2005; Watson & Askew, 2009). She conceptualised reading as involving a rapid processing of a range of available information (Clay, 1991, 2013), consistent with the Simple View of Reading but less 'simple', including alphabetic decoding, language, orthographic, semantic and syntactic information, with comprehension and fluency as target outcomes. Additional to supporting word reading efficiency, she proposed that the early years are also critical in the development of metacognitive skills such as activating background knowledge and self-monitoring for comprehension (Bodman & Smith, 2013).

The early developmental stages of literacy acquisition are important in determining later success. However, as children mature, new literacy skills are required as vocabulary and texts become more sophisticated, giving rise to late-emerging difficulties (Chall, 1983). Indeed, children with late-emerging difficulties show a different pattern of cognitive difficulties than children with early-emerging difficulties, particularly relating to comprehension and morphological skills (Leach *et al.*, 2003; Cain *et al.*, 2004; Catts *et al.*, 2012; Hirsch, 2003; Moir *et al.*, 2020), illustrating developmental factors at work.

Other factors also exert their influence on developing children, such as genetic factors relating to cognitive and linguistic abilities, their behaviour and environmental factors relating to home and school. The question remains, therefore, whether early gains can be maintained to some degree in the face of these ongoing influences on literacy development.

Long-term follow-up of the outcomes of early literacy interventions

There is little empirical evidence of the long-term effects of early literacy intervention to inform this debate. Apart from the UK studies by Hurry & Sylva (2007) and the follow-ups of the current evaluation (Hurry, 2012; Holliman & Hurry, 2013) when children were aged 9 and 11 years, the longest follow-up of Reading Recovery children has been in the USA to Grade 3 (UK Year 4) (D'Agostino & Harmey, 2016; D'Agostino *et al.*, 2017). Effects of Reading Recovery appear to be reliably sustained to Grade 2 or 3 (Hurry & Sylva, 2007; Holliman & Hurry, 2013; D'Agostino *et al.*, 2017) but are more uncertain beyond this point. Hurry & Sylva (2007) only found significant effects on reading at age 10-11 years for the bottom half of their sample, who were complete non-readers at 6 years old (Cohen's d = 0.54 and 0.59). Similarly, for other early reading interventions, the longer-term benefits have only rarely been researched. Suggate (2016) conducted a meta-analysis of studies of the long-term effects of 71 reading interventions (N = 8,161 at post-test), but the average time from immediate post-test to follow-up was only 11 months. Suggate reported an average of small but useful immediate post-test effects (Hedges d_w = 0.37) that decreased at follow-up (Hedges d_w = 0.22). Interventions which addressed comprehension (Reading Recovery fell into this category) had greater sustained effects than those that were solely focused at the word and phoneme level. One final study by Blachman *et al.*, (2014) reports moderate effect on children's word reading (Cohen's d = 0.53 on Woodcock Basic Skills Cluster, Cohen's d = 0.62 on Woodcock Word Identification) a decade after children had received an intensive and reasonably broad-based reading intervention for early reading difficulties. The intensity, duration and focus of the intervention were not too dissimilar to Reading Recovery: the intervention emphasised the phonologic and orthographic connections in words, focusing on accurate decoding and word recognition, but also worked on fluency, spelling and reading of both phonetically controlled text as well as books that were not phonetically controlled reading. Blachman and colleagues speculated that the lack of long-term effects on comprehension reflected the word-level focus of the intervention. It is of interest, therefore, to see if the broader range of elements addressed in Reading Recovery, including an emphasis on meaning and comprehension, might lead to a broader range of sustained benefits for children's reading.

The current study

The purpose of this study is to provide information on the long-term effects of the early intervention Reading Recovery in a UK context. Children's outcomes from the same study have previously been reported at the end of Year 1 (6-year-olds; Burroughs-Lange & Douëtil, 2007); Year 2 (7-year-olds; Tanner *et al.*, 2010); Year 4 (9-year-olds; Holliman & Hurry, 2013) and the end of elementary school (11-year-olds; Hurry, 2012). Children receiving Reading Recovery had made significantly greater reading progress than a comparison group drawn from non-Reading Recovery schools at all four of these follow-ups (effect size immediately post-intervention, word reading Cohen's $d = 0.52$; 3 years later, reading Cohen's $d = 0.53$ and writing Cohen's $d = 0.46$, Holliman & Hurry, 2013; effect size 5 years post-intervention, reading Cohen's $d = 0.39$ and writing Cohen's $d = 0.33$, Hurry, 2012). They had not made significantly greater progress at ages 9 and 11 than the weaker readers in their own schools not being offered Reading Recovery (RR school comparison group), raising the question of whether Reading Recovery has whole-school effects (Holliman & Hurry, 2013), particularly because this RR school comparison group also significantly outperformed the comparison group in non-RR schools by age 11 (Cohen's $d = 0.24$).

Two questions were explored in this 10-year follow-up study to assess the long-term effectiveness of the Reading Recovery programme:

1. Did children who received Reading Recovery perform better on national qualifications at age 16 than similar children not receiving Reading Recovery? National qualifications speak to broad academic implications of early reading intervention.
2. Were children who received Reading Recovery less likely to be identified with special educational needs (SEN) at age 14 and 16 than similar children not receiving Reading Recovery? Making a slow start in reading does not equate to having a special educational need but is a risk factor for more pervasive needs later in school and a consequence with resource implications.

Method

Design

In a quasi-experimental design, in 2005, 5-year-olds who were the weakest readers in their schools were assessed at baseline in 42 London schools. Half of the schools offered Reading Recovery and half of the matched comparison schools did not.

Within Reading Recovery schools, the weakest two-thirds of pupils were then given Reading Recovery (the RR group), the other third were not (the Reading Recovery Schools group, RRS). Children in comparison schools (the CS group) received a range of standard provision. In the current article, pupils were followed up through a national database (National Pupil Database, NPD) at age 14 and 16 (end of Years 9 and 11). Group differences on GCSE score (national qualification at 16 years) were tested for statistical significance using multi-level hierarchical regression (child and baseline school levels), controlling for baseline literacy and demographics. Two grouping strategies were used in the analyses reported here. The first, a conventional approach within Reading Recovery research, compared CS, RR and RRS groups (referred to below as the 'Full sample' grouping). Since the selection process to RR or RRS groups relied to some extent on teacher judgement, a second grouping strategy was also tested using a more transparent, objective selection process to mitigate selection bias: the four lowest scoring children in each school on the baseline literacy measure were selected from the full sample (referred to below as 'Bottom4' grouping) and those in Reading Recovery schools (RR_bottom4), not all of whom were given RR, were compared with those in comparison schools (CS_bottom4). Group differences in SEN status at age 14 and 16 were also tested.

Sample

Twenty-one elementary schools providing Reading Recovery in 2005–6 were well matched with 21 similar comparison schools in terms of: uptake of Free School Meals (an indicator of socioeconomic status); number of children with English as an additional language; school size; and attainment of Year 1 children in September 2005 (see Burroughs-Lange & Douëtil, 2007 for further details). The bottom 12% of 5-year-old (Year 1) readers in each school were assessed (approximately seven children). In Reading Recovery schools, consistent with Reading Recovery standard practice, the teacher and teacher leader selected roughly the four lowest scorers on the Observation Survey of Early Literacy Achievement (see Measures) in their school to receive Reading Recovery, but were also informed by children's performance in the previous year. The remaining children in the Reading Recovery schools comprised the comparison group within Reading Recovery schools (the RRS group) and the children in comparison schools (the CS group). At follow-up at age 14 and 16, the sample sizes had reduced. From baseline to age 14 there was an overall attrition of 16% and from baseline to age 16 of 8%. The children who were untraced did not differ significantly from those traced, either on demographic factors or literacy levels. Table 1 shows the sample over the three time points by grouping (Full sample and Bottom4).

Measures

Baseline. At baseline (September 2005), children were assessed on the Observation Survey of Early Literacy Achievement (OSELA; Clay, 2013; D'Agostino *et al.*, 2018) and the Word Recognition And Phonic Skills test (WRAPS; Moseley, 2003). OSELA is the core assessment in the Reading Recovery programme and comprises six components to cover key aspects of early literacy (Table 2). D'Agostino *et al.*, (2018)

Table 1. Sample by grouping strategies at baseline and follow-ups at age 14 and 16

	Baseline	Age 14	Age 16
Full sample			
CS	148	128	138
RR	91	76	84
RS	54	42	49
Total	293	246	271
Bottom four scoring children in each school on baseline OSELA			
CS_bottom4	84		79
RR_bottom4	84		75
Total bottom 4	168		154

Notes: CS = comparison school; RR = Reading Recovery; RRS = Reading Recovery school, not RR; CS_bottom4 = comparison school, bottom four scoring children in each school on baseline OSELA; RR_bottom4 = Reading Recovery school, bottom four scoring children in each school on baseline OSELA.

recommend using an overall score for OSELA, summing all the sub-tests.[1] Accordingly, an overall score was calculated by deriving z scores for each sub-test, summing these scores and deriving a z score of that sum (mean = 0, standard deviation = 1). The resulting OSELA z score was normally distributed with skewness = 0.858 and kurtosis = 0.531.

All tests were administered individually. Testers were Reading Recovery teachers previously trained in OSELA procedures. Background data were collected on each child at baseline, on: uptake of Free School Meals (FSM; no FSM = 0, FSM = 1); English as an additional language (EAL; non-EAL = 0, EAL = 1); gender; age; and whether or not children were identified as having special educational needs (SEN).

Table 2. Baseline measures

The Observation Survey of Early Literacy Achievement (OSELA)	
BAS II Word Reading Test (Elliott *et al*., 1996)	Assesses ability to read single words of increasing difficulty.
Letter Identification	Assesses ability to recognise letters (name, sound or a word beginning with the letter).
Concepts about Print	Assesses knowledge of the conventions of print.
Writing Vocabulary	Assesses spelling vocabulary. Children are asked to write as many words as they can within 10 minutes.
Hearing and Recording Sounds in Words	Assesses phonetic skills in writing. Children are asked to write a dictated sentence with marks for phonetic accuracy.
Book Level	Assesses text reading, establishing which of a series of texts, graded from 1 (very simple caption books) to 26 children can read with 90% accuracy.

Word Recognition And Phonic Skills test (WRAPS)
50-item test of word recognition. Words are read out and children select the correct word from a number of options.

During the intervention year, information was collected on additional support provided to RRS and CS children.

10-Year follow-up

Educational qualifications. English pupils take their first national qualifications at age 16 (end of Year 11), the General Certificate of Secondary Education (GCSE), and data were collected from the NPD for our sample at the end of their Year 11. A suite of GCSEs is accepted as the record of achievement at age 16, similar to a leaving certificate or baccalaureate qualification. GCSE scores provide a continuous measure of qualification, the sum of each subject result [highest score for each qualification = 9 (A*), lowest = 0] (DfE, 2017). Results were also categorised by number of GCSEs at A* to G for each subject. This was valuable because at the time of this study, entrance to higher-level education required five A* to C passes, including Maths and English. Both the GCSE scores and grade A* to G data were collected from the NPD.

Special Educational Needs and Disabilities (SEND) status. When pupils were aged 14 (end of Year 9), their SEN status was also collected from the NPD. English schools at this time followed a SEND code of practice (involving roughly 20% of children overall), as follows: a status of School Action was recorded when a child was not making progress and action was needed to meet learning or behaviour difficulties; School Action Plus was used where a child had not made adequate progress under School Action (external advice being sought); finally, for the most persistent and serious problems, the child would be assessed by experts and receive a Statement of SEND (roughly 2% of children in English schools). These data were again collected from the NPD when pupils were aged 16 (end of Year 11), at which time the practices and legislation had changed somewhat: School Action and School Action Plus were grouped together and a Statement of SEND and the newly introduced Education Health and Care Plans (EHCP; equivalent to a Statement of SEND) were grouped together.

Intervention

Reading Recovery. The intervention received by the children in this study followed standard UK practice. Following assessment, in UK schools the weakest 6-year-old readers (typically the bottom 5–10%) were selected for Reading Recovery. Children received daily, one-to-one, half-hour lessons with a Reading Recovery teacher for between 12 and 20 weeks, until either they were reading at the average level for their class (‘discontinued’) or they were ‘referred’ back to school in need of ongoing support. The lessons addressed phonics, language, reading fluency, writing and comprehension, responsive to the specific needs of the individual child based on daily assessment (Douëtil *et al.*, 2013).

Training and fidelity of instruction was managed within the Reading Recovery programme in a ‘three-tiered system’ (Schmitt *et al.*, 2005), with national leaders based in universities running Master’s degrees and continuing professional development for teacher leaders, who worked in local regions to provide professional development for school-based teachers. Teacher leaders undertook 1 year of in-depth training.

School-based teachers received 1 year's training in their local context, and continuing support from their teacher leader. Data were gathered annually to monitor implementation at local and national levels (Amott *et al*., 2013).

Comparison children in Reading Recovery and comparison schools. An earlier study reported on additional teaching offered to the comparison children during the intervention year (Year 1; Burroughs-Lange, 2006). It is reported here to give detail on what comparison children received instead of Reading Recovery, and for the insight it provides of differences between Reading Recovery and comparison schools. Additional support for RRS typically involved phonics plus comprehension; the most common additional support was 'Supported Reading' based on Reading Recovery principles. CS children were more likely to receive a phonics-only intervention; Ruth Miskin Literacy. However, for around half of RRS children (53%) and around one-third of CS children (63%), no additional support was reported (Burroughs-Lange & Douëtil, 2007).

RR_bottom4. At follow-up at age 16, the RR_bottom4 group was made up of 53 children who received Reading Recovery in the intervention year and 22 children who did not (grouped as RRS in the analysis of the Full sample). Of this 22, three received Reading Recovery in the following year.

Data analysis

Only pupils with data available at age 16 were included in analyses. In the analyses for both the Full sample and the Bottom4, group differences on GCSE score were tested for statistical significance using multi-level hierarchical regression in Stata 15, at child and baseline school levels, and controlling for baseline variables to account for individual differences, as follows: OSELA, WRAPS, Free School Meals and English as an additional language. For the Full sample, groups were entered as two dummy variables with CS as the control = 0 and RR and RRS alternately = 1. For the Bottom4 grouping, CS_bottom4 = 0, RR_bottom4 = 1. Because of significant difference in uptake of Free School Meals between the key comparison groups, RR and CS (see Table 3), multi-level models were also run separately for children with and without Free School Meals to reduce threats to internal validity. Group differences on nominal and ordinal level variables were tested using chi-square tests with adjusted residuals.

Results

Baseline

The children were initially selected from economically disadvantaged areas with just over half taking Free School Meals (i.e. poor, national average = 16%) and 48% with English as an additional language (Table 3). The majority had weak word reading skills at baseline, 50% failing to read even one word on the BAS II Word Reading and 84% either not reading or only able to read the most basic level books (Level 1 of 26).

Table 3. Baseline characteristics of children for whom data was available both at 6 and 16 years of age, by group

	% Free School Meals	% English additional language	% Boys	Age mean (SD)	OSELA score mean (SD)	WRAPS score mean (SD)
Full sample						
CS (N = 138)	61.6%	53.6%	66.4%	5y 8m (2.4m)	−0.12 (0.83)	10.99 (6.06)
RR (N = 84)	42.9%[a]	47.6%	61.4%	5y 8m (2.8m)	0.02 (0.79)	11.17 (6.32)
RRS (N = 49)	51.0%	32.7%[b]	47.9%[c]	5y 8m (3.4m)	0.22 (1.58)	12.27 (10.02)
Total (N = 293)	53.9%	48.0%	61.5%	5y 8m (2.7m)	−0.01 (1.04)	11.27 (7.00)
Bottom four scoring children in each school on baseline OSELA						
CS_bottom4 (N = 79)	65%	48%	68%	5y 8m (2.5m)	−0.37 (0.65)	9.24 (5.71)
RR_bottom4 (N = 75)	41%[d]	56%	60%	5y 8m (3.3m)	−0.30 (0.86)	9.97 (7.11)
Total (N = 154)	53%	52%	64%	5y 8m (2.9m)	−0.34 (0.75)	9.60 (6.42)

CS = comparison school; RR = Reading Recovery; RRS = Reading Recovery school, not RR; CS_bottom4 = comparison school, bottom four scoring children in each school on baseline OSELA; RR_bottom4 = Reading Recovery school, bottom four scoring children in each school on baseline OSELA; Age = age on 1 September 2005.

[a]CS vs. RR, $\chi^2(1, N = 222) = 7.39, p = 0.007$.

[b]CS vs. RRS, $\chi^2(1, N = 187) = 6.37, p = 0.012$.

[c]CS vs. RRS, $\chi^2(1, N = 185) = 5.11, p = 0.024$.

[d]CS_bottom4 vs. RR_bottom4, $\chi^2(1, N = 154) = 8.33, p = 0.004$.

They did have some skills in place as measured by OSELA, such as some letter knowledge and concepts about print.

There were no significant group differences in literacy at baseline between any of the groups, although the RRS children scored slightly higher, as was to be expected on the basis of the selection criteria. However, a significantly lower proportion of RR and RR_bottom4 children took Free School Meals compared to CS and CS_bottom4 children. Comparing CS and RRS, a significantly lower proportion of RRS children had English as an additional language and a significantly lower proportion were boys (Table 3).

Follow-up on academic progress at age 16 (end Year 11)

There were moderately sized differences in GCSE scores at age 16 between intervention groups and by EAL status (Table 4). Pupils who had been given Reading Recovery or attended an RR school at age 5–6 years had higher GCSE scores at age 16 than those attending comparison schools. Children for whom English was an additional language at age 5 scored higher on GCSEs than their peers. Although children receiving Free School Meals scored slightly lower than their peers, the effect size was small.

Differences between RR, RRS and CS on GCSE scores were tested for statistical significance using multi-level hierarchical regression in Stata 15, with child and school levels and controlling for baseline: OSELA, WRAPS, Free School Meals and

Table 4. Mean, SD and ES for intervention groups and demographic variables at 16 years of age

	Full sample grouping			Bottom4 grouping		FSM status		EAL status	
	CS	RR	RRS	CS_bottom4	RR_bottom4	No FSM	FSM	Non-EAL	EAL
Mean	32.0	42.5	40.2	30.1	37.6	38.0	34.8	31.9	42.0
SD	19.8	19.3	25.9	19.8	22.7	22.3	20.4	22.5	19.3
ES	RR vs. CS = 0.52 RRS vs. CS = 0.37			0.36		0.15		0.48	

Notes: SD = standard deviation; ES = Cohen's *d*; CS = comparison school; RR = Reading Recovery; RRS = Reading Recovery school, not RR; CS_bottom4 = comparison school, bottom four scoring children in each school on baseline OSELA; RR_bottom4 = Reading Recovery school, bottom four scoring children in each school on baseline OSELA.

English as an additional language (Table 5). In Model 1, the effect of EAL on GCSE scores was highly significant when baseline literacy and FSM were accounted for, with EAL pupils doing better. FSM status was not a significant predictor. In Model 2, both RR and RRS groups had significantly higher GCSE scores than CS. As a further sensitivity analysis, to address the difference between groups on uptake of Free School Meals, the regression analyses were repeated separately by FSM status. RR had significantly higher GCSE scores than CS for both analyses, but the effect was slightly larger for those not taking Free School Meals at baseline (**no FSM**, $N = 101$, $B = 11.75$, SE = 4.0, $p < 0.003$, RR $M = 43.9$, SD = 18.4; CS $M = 33.4$, SD = 20.1, Cohen's $d = 0.53$; **FSM**, $N = 121$, $B = 7.57$, SE = 3.7, $p < 0.040$, RR $M = 40.6$, SD = 20.6; CS $M = 31.2$, SD = 19.8, Cohen's $d = 0.46$). Similarly, RRS children scored significantly higher than CS children for both analyses (**no FSM**, $N = 77$, $B = 10.52$,

Table 5. Factors at baseline (6 years of age) predicting GCSE new points score at 16 years, by groupings CS, RR, RRS: multi-level hierarchical regression

	Model 1					Model 2				
Fixed effects	*B*	SE	*p*	95% Conf. int.		*B*	SE	*p*	95% Conf. int.	
OSELA	4.44	1.72	0.010	1.06	7.81	4.02	1.65	0.015	0.76	7.27
WRAPS	0.24	0.23	0.298	−0.22	0.71	0.27	0.23	0.233	−0.17	0.72
FSM	−2.04	2.43	0.401	−6.81	2.72	−1.69	2.41	0.484	−6.41	3.03
EAL	12.29	2.46	0.001	7.47	17.11	12.60	2.42	0.001	7.85	17.35
RR						10.22	3.28	0.002	3.79	16.64
RRS						8.92	3.76	0.018	1.56	16.29

Random effects	Variance	SE	95% Conf. int.		Variance	SE	95% Conf. int.	
School-level variance	61.8	25.9	27.2	140.7	40.8	20.8	15.0	110.8
Child-level variance	332.4	31.2	276.6	399.5	331.0	30.9	275.6	397.5

Notes: RR = dummy variable RR 1, CS & RRS 0; RRS = dummy variable RRS 1, CS & RR 0; FSM = Free School Meals; EAL = English as an additional language.

SE = 5.2, $p < 0.042$, RRS M = 41.6, SD = 31.0; CS M = 33.4, SD = 20.1, Cohen's d = 0.34; **FSM**, N = 110, B = 8.94, SE = 4.2, $p < 0.034$, RRS M = 38.8, SD = 20.3; CS M = 31.2, SD = 19.8, Cohen's d = 0.37).

The multi-level analyses were repeated for the Bottom4 (Table 6). The was once again evident in Model 1. In Model 2, children from Reading Recovery schools had significantly higher GCSE scores than children from comparison schools.

Group differences were also evident in the attainment necessary for educational progression (5+ GCSEs at A* to C, including English and Maths), with a statistically significant chi-square for the overall results shown in Table 7 [$\chi^2(6, N = 220) = 19.13, p < 0.004$]. Where the adjusted residual in any cell is greater than 2, this indicates that the cell is significantly larger than expected ($\alpha < 0.05$) or less than −2 significantly smaller than expected. RR pupils were significantly more likely to achieve five or more GCSEs at A* to C including English and Maths than CS pupils and significantly less likely to achieve five GCSEs at A* to G (insufficient to advance to higher-level study in England) than CS pupils. The same pattern of results was evident for both children taking Free School Meals and those not taking Free School Meals, though more marked for those not taking Free School Meals.

Special educational needs status at age 14 and 16

At baseline, of the 271 children for whom data was available at age 16, 8% of children had been identified with some level of SEN (School Action, School Action plus or a Statement of SEN); CS = 8 (6%), RR = 9 (11%), RRS = 5 (10%). Between-group differences were not statistically significant.

When the children were aged 14 (end Year 9), there were significantly fewer RR pupils with a SEN status than CS pupils (Table 8) and significantly fewer RR pupils with a statement of SEND/EHCP than CS pupils. When the children were aged 16

Table 6. Factors at baseline (6 years of age) predicting differences in GCSE new points score at 16 years, by groupings CG_bottom4, RR_bottom4: multi-level hierarchical regression

	Model 1					Model 2				
Fixed effects	*B*	SE	*p*	95% Conf. int.		*B*	SE	*p*	95% Conf. int.	
OSELA	5.45	2.41	0.024	0.73	10.17	4.86	2.33	0.037	0.29	9.42
WRAPS	0.51	0.28	0.070	−0.04	1.06	0.51	0.27	0.061	−0.02	1.04
FSM	−0.55	3.11	0.860	−6.63	5.54	0.76	3.11	0.808	−5.34	6.85
EAL	14.62	3.14	0.001	8.46	20.78	14.88	3.08	0.001	8.83	20.92
RRvCS						7.50	3.18	0.018	1.26	13.74

Random effects	Variance	SE	95% Conf. int.		Variance	SE	95% Conf. int.	
School-level variance	16.3	27.4	0.60	439.0	6.7	24.5	0.1	8,915.0
Child-level variance	340.6	45.8	261.7	443.4	337.3	45.0	259.6	438.2

Notes: EAL = English as an additional language; FSM = Free School Meals; RRvCS = RR_bottom4 vs. CS_bottom4.

Table 7. GCSE grade, by group

		5+ A* to C GCSE inc. Eng. & Maths	5+ A* to C GCSE	5+ A* to G	GCSE below G	No passes
CS	*N* (%)	31 (22.8%)	16 (11.8%)	64 (47.1%)	15 (11.0%)	10 (7.4%)
N = 136	Adj. res.	−4.3	0.9	2.6	0.7	1.3
RR	*N* (%)	41 (49.4%)	10 (12.0%)	24 (28.9%)	6 (7.2%)	2 (2.4%)
N = 83	Adj. res.	3.3	0.7	−2.4	−0.9	−1.5
RRS	*N* (%)	22 (44.9%)	1 (2.0%)	18 (36.7%)	5 (10.2%)	3 (6.1%)
N = 49	Adj. res.	1.6	−2.1	−0.4	0.1	0.2

Notes: Adj. res. = adjusted residuals; CS = comparison schools; GCSE inc. Eng. & Maths = General Certificate of Secondary Education, including English and Maths; RR = Reading Recovery; RRS = Reading Recovery school, not RR.

(end Year 11), there were significantly fewer RR pupils with a statement of SEND/EHCP than CS pupils, but no other significant differences.

Discussion

We report here on the effects of an early reading intervention, Reading Recovery, 10 years after intervention, when children had taken their GCSEs, the first high-stakes national qualifications with value for employment (Wolf, 2011) and the first major branching point for a young person's educational career (Gayle *et al.*, 2016). For the level of qualification traditionally required for progression to further education in England (five or more GCSEs at grade A* to C including English and Maths), 49% of 16-year-olds assigned to Reading Recovery at age 5 achieved this level and 45% of their classmates who were also weak readers at baseline but were not assigned to Reading Recovery. This compared to 24% in comparison schools, and against a national average of 54% in 2016 (DfE, 2017). Thus, RR pupils in this study, in the bottom 10% of readers in their schools at age 5, were performing only 5% below the national average at age 16 and were twice as likely to reach this important qualification threshold as CS pupils. This significant difference between RR and CS pupils in their GCSE results (Cohen's $d = 0.52$) was confirmed when their scores were compared in a multi-level regression analysis. This suggests that early reading intervention can have a long-lasting and meaningful impact on readers who make a slow start. However, the fact that the comparison children in Reading Recovery schools also outperformed the CS group (Cohen's $d = 0.37$) complicates the interpretation of these results. Two possible explanations for the good performance of RRS seem plausible. The first is that Reading Recovery may have an effect on the performance of all low-performing young readers. The second is that RR schools in the study were different at baseline to comparison schools, in ways other than the presence of Reading Recovery.

Considering first a whole-school effect of Reading Recovery, this beneficial effect for all weaker readers in RR schools is consistent with the findings from at least two other British samples (Hurry & Sylva, 2007; Tanner *et al.*, 2010). The intervention

Table 8. SEN status at age 14 and 16

		No SEN	School Action	School Action+	Statement
Age 14					
CS	*N* (%)	62 (48.4%)	31 (24.2%)	22 (17.2%)	13 (10.2%)
N = 128	Adj. res.	−2.0	1.6	−0.1	1.5
RR	*N* (%)	48 (63.2%)	11 (14.5%)	16 (21.1%)	0 (0%)
N = 76	Adj. res.	1.8[a]	−1.5	1.3	−3.0[b]
RRS	*N* (%)	24 (57.1%)	8 (19.0%)	4 (9.5%)	6 (14.3%)
N = 42	Adj. res.	0.4	−0.2	−1.5	1.7
Total (*N* = 246)		134 (54.5%)	50 (20.3%)	43 (17.5%)	19 (7.7%)

		No SEN	School Action, School Action+	Statement or EHCP
Age 16				
CS	*N* (%)	93 (67.4%)	33 (23.9%)	12 (8.7%)
N = 138	Adj. res.	−0.3	0.6	1.7
RR	*N* (%)	60 (71.4%)	24 (28.6%)	0 (0%)
N = 84	Adj. res.	0.7	0.8	−2.9[c]
RRS	*N* (%)	32 (65.3%)	12 (24.5%)	5 (10.2%)
N = 49	Adj. res.	−0.5	−0.2	1.3
Total (*N* = 271)		185 (69.4%)	69 (24.3%)	17 (6.3%)

CS = comparison school; EHCP = Education Health and Care Plan; RR = Reading Recovery; RRS = Reading Recovery school, not RR; School Action+ = School Action Plus; SEN = special educational needs; Statement = Statement of SEN.
[a]CS vs. RR, $\chi^2(1, N = 204) = 12.55, p < 0.006$.
[b]CS vs. RR, $\chi^2(1, N = 204) = 8.24, p < 0.004$.
[c]CS vs. RR, $\chi^2(1, N = 222) = 8.90, p < 0.006$.

has a range of potential impacts across the school: weaker readers are systematically identified, tested and monitored; Reading Recovery teachers are extensively trained in supporting weaker readers and in the UK are expected to play a leadership role in this area in their schools; Reading Recovery introduces whole-school practices, such as the levelling of reading books and the use of plastic letters in supporting decoding and spelling; classroom teachers are freed up to focus on the other weaker readers in their class; class average reading levels are brought up by virtue of the intervention. All these impacts have been evidenced in one way or another (Hurry & Sylva, 2007; Tanner *et al.*, 2010).

Addressing the second potential explanation, the quasi-experimental design has removed the classic method of achieving equivalence, randomisation. The schools and children were similar at baseline in terms of literacy and percentage of children with EAL, but the CS children were more socially disadvantaged (though RR and CS schools were similar on percentage FSM; Burroughs-Lange & Douëtil, 2007). This difference was dealt with in two ways in the analyses. Firstly, regression analyses controlled for FSM status and secondly, a further sensitivity regression analysis was conducted, splitting the sample into those taking FSM and those not, thus ensuring that experimental comparisons were comparing like with like in this regard. RR and RRS pupils in both groups had significantly better GCSE outcomes than CS pupils.

The fact that children were selected for Reading Recovery, to an extent based on teacher judgement, introduced another possible threat to interpreting any differences between treatment and control groups. To address this threat, the lowest-scoring children on the objective literacy assessment at baseline in Reading Recovery and comparison schools were compared. Only around three-quarters of these weakest readers in the Reading Recovery schools had received RR. These weakest readers attending Reading Recovery schools in Year 1 significantly outperformed children in comparison schools on GCSEs at age 16 (Cohen's $d = 0.37$). On balance, this suggests that the Reading Recovery effect is real in this study but that the more conservative estimate of effect size might be sensible and that the effect of Reading Recovery operates both through one-to-one intervention and also at a wider school level.

These long-term outcomes are suggestive in two different ways, one practical, relating to cost and embedded systems, and one theoretical, relating to the nature of reading development. In terms of cost, whilst Reading Recovery has generally been evaluated as an effective intervention, it is also costly, requiring daily individual tuition for 12–16 weeks with a highly trained teacher and a further two organisational layers to maintain quality and training. The finding that the RR group outperformed CS pupils on GCSEs has positive implications for the economy, which arguably offset these costs. GCSE performance is a strong determinant of future success (Gayle, *et al.*, 2016), both in education (Payne, 2003; Babb, 2005) and employment (Jones *et al.*, 2003; Babb, 2005; Murray, 2011). Being identified as having special educational needs also has cost implications, particularly where a student has a Statement of SEND (Gross, 2006). Whilst a slow start in reading has a range of possible explanations and only a small percentage of the children in this study were identified with SEN at age 5, it raises the risk of later difficulties. Around one-third of CS pupils were identified as having SEN at 16, compared with a national average of around 10%, though social disadvantage (receiving FSM) more than doubles the risk of being identified with SEN (DfE, 2018) and our sample was relatively socially disadvantaged. Nonetheless, RR pupils were significantly less likely to be identified as having SEN at age 14 and 16 than CS pupils. The good performance of the RRS group raises the question as to whether these good results could be achieved by having Reading Recovery teachers and systems in place but no one-to-one tuition, a much cheaper option. Indeed, embedded systems may be a necessary element of effectiveness of one-to-one intervention. We do not have an answer to this question, but it seems an important one. Of course, the effectiveness of Reading Recovery for a larger group of pupils than those directly receiving the intervention makes the intervention more cost effective.

In terms of the nature of reading development, there are reasons to anticipate that early literacy intervention will prevent later reading difficulties by addressing foundation skills and narrowing the gap between poor and strong readers in subsequent reading, but there are counter factors which challenge this hypothesis, the continuing effects of biological, developmental and environmental factors. The significantly better progress of pupils for whom English was an additional language in this study illustrates one such factor. It seems likely that those children at age 5 were being held back in their reading by their English language skills, but that as they progressed through school and their English language caught up with their peers, so did their reading. The other children in the study will have had a range of explanations for their slow start with reading. That Reading Recovery in this study has moderate measurable effects 10 years on supports the case for early intervention and its role in reading development. The effect reported here compares with the substantial 10-year follow-up effects reported by Blachman *et al.*, (2014) (Cohen's $d = 0.53$ for Basic Skills cluster and Cohen's $d = 0.62$ for Word Identification), following a similarly intensive early intervention for children with reading difficulties. The outcome measure of Reading Recovery is broader than that reported by Blachman *et al.*, (2014), extending beyond an impact on word-level measures. Suggate (2016) reports that interventions which address comprehension, a category in which he includes Reading Recovery, have greater long-term effects than phonological only interventions, but the research base is currently too small to be confident about the significance of type of intervention for long-term effect. However, the long-term effectiveness of the Reading Recovery programme on a broad measure of academic attainment, reported here, provides support for Clay's proposition of the complexity of reading behaviour, even in the early stages of reading development, and the value of addressing a broad range of skills which include not only phonics but also metacognition and comprehension monitoring.

Limitations

There are three main limitations in this study. Firstly, the assignment to Reading Recovery and comparison conditions was not random, and though there were no significant differences in the literacy levels of the children in the two groups, the comparison group was significantly more economically disadvantaged than the Reading Recovery group. Secondly, the sample size is smallish (271 children). Thirdly, the children in this study were economically disadvantaged and a relatively high proportion spoke English as an additional language, and this may have implications for generalisability to more affluent monocultural communities.

NOTE

[1] The individual sub-tests of OSELA tend to be skewed in different directions over the course of intervention, with some having floor effects at baseline (e.g. Book Level and BAS word reading) and others having ceiling effects at post-test. This reflects the rapid development of children's reading skills over this period. The use of an overall score addresses this problem.

References

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

Amott, P., Hindmarsh, V. & Morris, H. (2013) Creatively responding to the imperative of scaling up Every Child a Reader, chapter 5, in: S. Burroughs-Lange & A. Ince (Eds) *Reading Recovery and Every Child a Reader: History, policy, and practice* (London, IOE Press), 120–149.

Babb, P. (2005) *A summary of focus on social inequalities* (London, Office for National Statistics).

Blachman, B.A., Schatschneider, C., Fletcher, J.M., Murray, M.S., Munger, K.A. & Vaughn, M.G. (2014) Intensive reading remediation in Grade 2 or 3: Are there effects a decade later?, *Journal of Educational Psychology*, 106, 46–57.

Bodman, S. & Smith, J. (2013) The theoretical and pedagogical base of Reading Recovery, chapter 3, in: S. Burroughs-Lange & A. Ince (Eds) *Reading Recovery and Every Child a Reader: History, policy, and practice* (London, IOE Press), 63–91.

Burroughs-Lange, S. (2006) *Evaluation of Reading Recovery in London schools: Every Child a Reader 2005–06*. Available online at: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.603.4006&rep=rep1&type=pdf (accessed 28 September 2020).

Burroughs-Lange, S. & Douëtil, J. (2007) Literacy progress of young children from poor urban settings: A Reading Recovery comparison study, *Literacy Teaching and Learning*, 12, 19–46.

Byrne, B. (1998) *The foundation of literacy: The child's acquisition of the alphabetic principle* (Hove, Psychology Press).

Cain, K., Oakhill, J. & Bryant, P.E. (2004) Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills, *Journal of Educational Psychology*, 96, 31–42.

Catts, H.W., Compton, D., Tomblin, J.B. & Bridges, M.S. (2012) Prevalence and nature of late-emerging poor readers, *Journal of Educational Research*, 104, 166–181.

Chall, J.S. (1983) *Stages of reading development* (New York, McGraw-Hill).

Clay, M.M. (1991) *Becoming literate: The construction of inner control* (Auckland, Heinemann).

Clay, M. (2005) *Literacy lessons designed for individuals. Part two: Teaching procedures* (Portsmouth, NH, Heinemann).

Clay, M. (2013) *An observation survey of early literacy achievement* (London, Heinemann).

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

D'Agostino, J.V., Lose, M.K. & Kelly, R.H. (2017) Examining the sustained effects of Reading Recovery, *Journal of Education for Students Placed at Risk*, 22(2), 116–127.

D'Agostino, J., Rodgers, E. & Mauck, S. (2018) Addressing inadequacies of the observation survey of early literacy achievement, *Reading Research Quarterly*, 53, 56–69.

Desjardins, R., Thorn, W., Schleicher, A., Quintini, G., Pellizzari, M., Kis, V. *et al* (2013) *OECD skills outlook 2013: First results from the survey of adult skills* (Paris, OECD Publishing).

DfE (2017) *Revised GCSE and equivalent results in England, 2015 to 2016*. SFR03/2017 (London, Department for Education).

DfE (2018) *Special educational needs in England: January 2018* (London, Department for Education).

DfEE (1998) *The national literacy strategy: Framework for teaching* (London, Department for Education and Employment).

Douëtil, J., Hobsbaum, A. & Maidment, P. (2013) Reading Recovery, and early literacy intervention, chapter 1, in: S. Burroughs-Lange & A. Ince (Eds) *Reading Recovery and Every Child a Reader: History, policy, and practice* (London, IOE Press), 6–27.

Ehri, L.C. (2005) Learning to read words: Theory, findings and issues, *Scientific Studies of Reading*, 9, 167–188.

Elliott, P., Smith, P. & McCulloch, K. (1996) *British Abilities Scales* (Windsor, NFER-Nelson).

Frith, U. (1985) Beneath the surface of developmental dyslexia, in: K. Patterson, M. Coltheart & J. Marshall (Eds) *Surface dyslexia* (London, LEA), 301–330.

Gayle, V., Murray, S. & Connelly, R. (2016) Young people and school General Certificate of Secondary Education attainment: Looking for the 'missing middle', *British Journal of Sociology of Education*, 37, 350–370.

Gough, P. (1996) How children learn to read and why they fail, *Annals of Dyslexia*, 46, 3–20.

Gross, J. (2006) *The long-term costs of literacy difficulties* (London, KPMG Foundation).

Hirsch, E.D. Jr (2003) Reading comprehension requires knowledge—of words and the world: Scientific insights into the fourth-grade slump and the nation's stagnant comprehension scores, *American Educator*, 27, 10–29.

Holliman, A.J. & Hurry, J. (2013) The effects of Reading Recovery on children's literacy progress and special educational needs status: A three-year follow-up study, *Educational Psychology*, 33, 719–733.

Hurry, J. (2012) The impact of Reading Recovery five years after intervention. Available online at: https://lemosandcrane.co.uk/resources/The%20impact%20of%20reading%20recovery%20five%20years%20after%20intervention.pdf (accessed 28 September 2020).

Hurry, J. & Sylva, K. (2007) Long-term outcomes of early reading intervention, *Journal of Reading Research*, 30, 227–248.

Jones, J., Joyce, M. & Thomas, J. (2003) Non-employment and labour availability, *Bank of England Quarterly Bulletin*, Autumn, 291–303.

Leach, J. M., Scarborough, H. S. & Rescorla, L. (2003) Late-emerging reading disabilities, *Journal of Educational Psychology*, 95(2), 211–224.

Moir, T., Boyle, J. & Woolfson, L.M. (2020) Developing higher-order reading skills in mainstream primary schools: A metacognitive and self-regulatory approach, *British Educational Research Journal*, 46(2), 399–420.

Moseley, D. (2003) *Word recognition and phonic skills* (2nd edn) (Abingdon, Hodder & Stoughton).

Murray, S.J. (2011) Growing up in the 1990s: Tracks and trajectories of the 'Rising 16's': A longitudinal analysis using the British Household Panel Survey. Ph.D. thesis, University of Stirling.

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

Payne, J. (2003) *Vocational pathways at age 16–19*. Research Report No. 501 (London, Department for Education and Skills).

Schmitt, M., Askew, B., Fountas, I., Lyons, C. & Pinnell, G. (2005) *Changing futures: The influence of Reading Recovery in the United States* (Worthington, OH, Reading Recovery Council of North America).

Stanovich, K.E. (1986) Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy, *Reading Research Quarterly*, 21(4), 360–406.

Stuart, M. (2002) Using the dual-route cascade model as a framework for considering reading development, in: R. Stainthorp & P. Tomlinson (Eds) *Learning and teaching reading* (Leicester, The British Psychological Society), 45–60.

Suggate, S.P. (2016) A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions, *Journal of Learning Disabilities*, 49, 77–96.

Tanner, E., Brown, A., Day, N., Kotecha, M., Low, N. & Morrell, G. et al (2010) Evaluation of Every Child a Reader (ECaR). Department for Education Research Report DFE-RR14. Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/182631/DFE-RR114.pdf (accessed 28 September 2020).

Torgesen, J.K. (2000) Individual differences in response to early interventions in reading: The lingering problem of treatment resisters, *Learning Disabilities Research & Practice*, 15, 55–64.

Watson, B. & Askew, B. (2009) *Boundless horizons: Marie Clay's search for the possible in children's literacy* (Rosedale, Pearson Education).

Wolf, A. (2011) Review of vocational education – The Wolf Report. Available online at: www.gov.uk/government/publications/review-of-vocational-education-the-wolf-report (accessed 28 September 2020).