



Research into Read Write Inc. Phonics

A report on lessons learned

October 2022

Simon Rutt, Caroline Sharp, Kathryn Hurd, Alison Hale and
Connie Rennie



The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus – Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

For more information about the EEF or this report please contact:

- 📍 Jonathan Kay
Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP
- ☎ 0207 802 1653
- ✉ jonathan.kay@eefoundation.org.uk
- 🌐 www.educationendowmentfoundation.org.uk

Contents

About the evaluator	3
Executive summary	4
Background.....	6
The programme	8
Recruitment activity.....	14
Reasons for cancellation	17
Conclusion and lessons learned.....	18
References	20

About the evaluator

This project was conducted by a team from the National Foundation for Educational Research (NFER): Simon Rutt, Caroline Sharp, Kathryn Hurd, Alison Hale and Connie Rennie.

The lead evaluator was Simon Rutt.

Contact details:

Simon Rutt
National Foundation for Educational Research (NFER)
The Mere
Upton Park
Slough
SL1 2DQ

Tel: 01753 574123

Email: s.rutt@nfer.ac.uk

Executive summary

The project

Read Write Inc. Phonics is a literacy programme for children in Reception to Year 4 (age 4–9), aiming to teach children to read and write through a systematic approach. The programme begins with 20-minute daily lessons in term 1 of Reception, building up to an hour a day for children in Year 1 and above. The programme includes systematic synthetic phonics, reading and applying phonics to decodable books, and skills for writing, including handwriting and compositional writing. The programme is delivered by teachers and teaching assistants. In this study, schools were to receive two initial training days and two follow-up training days for the staff member appointed to lead on the programme (the reading leader), two training days for whole-school staff, and 16 in-school development days where a trainer from Ruth Miskin Training (RMT) would visit the school to provide support.

The delivery of the programme began in 2018 as part of a wider Department for Education (DfE) programme that funded its use in schools through the Teaching and Leadership Innovation Fund (TLIF). TLIF is a fund designed to support high-quality professional development for teachers and school leaders in identified priority areas (PAs) and opportunity areas (OAs)¹. The evaluation was designed to evaluate the impact of *Read Write Inc. Phonics* in TLIF schools. It was designed as a quasi-experimental study, where a matched comparison group of similar schools that were not using *Read Write Inc. Phonics* was to be recruited. 72 schools were recruited by RMT to receive the intervention. Comparison schools were to be recruited during the final year of programme delivery. The study was originally designed to assess the impact of *Read Write Inc. Phonics* on pupil attainment in the Phonics Screening Check (PSC) and KS1 reading assessment. A complementary Implementation and Process Evaluation (IPE) was also going to be undertaken, including surveys, interviews and observations of programme delivery.

Study disruption and cancellation

The evaluation began in September 2018 and was expected to finish in July 2020. However, in March 2020 delivery of the *Read Write Inc. Phonics* programme in schools was disrupted by Covid-19. Due to the cancellation of national testing in summer 2020, the evaluators were not able to assess the impact of *Read Write Inc. Phonics* using that year's PSC and KS1 results as planned. The first attempt to restart the evaluation was made in September 2020; however, the trial had to be put on hold again following a further wave of partial school closures and the cancellation of 2021 national assessments. A second attempt to restart was made in September 2021. Delivery of the programme resumed in schools and the evaluators began working to recruit a matched comparison group of schools with characteristics similar to those that were receiving the intervention. However, it was ultimately not possible to recruit a large enough group of schools that had similar characteristics to the schools in the treatment group. Treatment schools had been recruited to this study based on eligibility criteria for TLIF funding, which included being located in an OA or PA, having an Ofsted rating of 'Requires Improvement' or 'Inadequate', or having literacy data that was cause for concern. The comparison group needed to be comparable to the treatment group in terms of overall numbers of schools, proportions meeting each of the TLIF criteria, and the distribution of other school characteristics such as prior attainment and Free School Meal (FSM) eligibility rates. The size of the pool of potential comparison schools in England was smaller by 2022 than it had been when this study was designed in 2018, making recruitment even more challenging than anticipated. Without a well-matched comparison group, the evaluators would be unable to make a meaningful assessment of the impact of the programme, so the decision was made to cancel this evaluation in April 2022.

Additional challenges considered in the decision to cancel the evaluation included the level of exposure to the *Read Write Inc. Phonics* programme within the treatment schools. The original intention was for *Read Write Inc. Phonics* to be delivered in schools for a two-year period, yet the final design resulted in the Year 1 pupils of interest only receiving one full year of uninterrupted classroom learning. The previous Reception year was interrupted by a period of school remote learning between January and March 2021. This would have resulted in a gap in their learning at an important time in the development of reading skills. The gap in phonics data for 2020 and 2021 also introduced problems in the analysis, including issues with the underlying assumptions required for the main analysis. The combination of these

¹ Priority areas are areas that fall into Category 5 or 6 of the DfE's Achieving Excellence Areas classification, including Opportunity Areas, which are local districts identified as facing the greatest challenges to social mobility.

factors led to the decision to cancel the evaluation, as to continue would have placed too much burden on schools and teachers at a challenging time, without resulting in a robust evaluation.

Conclusions and lessons learned

The arrival of the Covid-19 pandemic in 2020 severely impacted on the running of this evaluation. The consequences of partial school closures and the cancellation of statutory testing in the summers of 2020 and 2021 resulted in multiple redesigns to cope with amended primary outcomes and the number of schools involved within the treatment and comparison groups. A lesson learned from this evaluation is that there could be value in involving an advisory group or other such third party in reviewing evaluations that have been significantly disrupted or changed since project inception. Review points for a stop/go decision could be planned and overseen by an advisory group, to support the challenging process of determining at what point an evaluation has been disrupted to such an extent that it should be cancelled.

A further lesson learned relates to the challenges of recruiting comparison schools to take part in evaluations. When signing up to a project, schools consider the requirements and benefits for them as a school or teacher in participating. When recruiting schools to comparison groups, schools are likely to feel that the benefits of participating are more limited than in studies where there is a possibility of them being allocated to the treatment group to receive a promising intervention. For this project, a thank you payment for comparison schools was introduced, to help incentivise schools to sign up for the study. This helped to increase sign-up rates. However, the Covid-19 pandemic has also put additional pressures on schools and in certain areas it has impacted on the willingness of schools to sign up to research projects.

Where possible, using secondary data sets such as the National Pupil Database (NPD) could help to reduce the need to recruit comparison schools taking part in quasi-experimental studies and reduce burden on schools. The NPD is a rich and well-used resource for educational research. However, the use of secondary data sets also has disadvantages. In the case of this study, disadvantages of using the NPD included that it would not have been possible to accurately identify and exclude from the study schools that were using *Read Write Inc. Phonics* resources, to assess whether schools were delivering similar phonics instruction, or to establish how Covid-19 had impacted on literacy teaching in comparison schools. Use of the NPD also would not have detracted from the known issues in the delivery of the programme during a significant period of disruption to pupil learning, meaning there would have still been major questions of how to interpret the results to provide meaningful insights to schools.

Background

Overview of the report and its context

The project to assess the impact of RMT: *Read Write Inc. Phonics* was originally commissioned in 2018 and was part of a wider DfE programme that funded its use in schools through the TLIF. TLIF is a fund designed to support high-quality professional development for teachers and school leaders in the identified PAs² and OAs.³ Three organisations were involved in this project; EEF (the commissioner), NFER (the evaluator) and RMT (the provider). Table 1 provides a timeline from the inception of the project to project closure. The original timeline can be found in the study plan.

The research was originally designed to look at attainment outcomes from the summer of 2020, using outcome data collected directly from schools. The design was a quasi-experimental design (QED) with a complementary implementation and process evaluation (IPE).⁴ The design included the PSC and the KS1 reading assessment as primary and secondary outcomes. Whilst teacher assessments are available on the NPD, the evaluation team decided to use the scaled score as the KS1 outcome.⁵ This data is only available directly from schools, so a recruited comparison group was included within the evaluation design, rather than relying on the NPD. The matched comparison group of schools would follow the TLIF eligibility criteria (i.e., schools located in priority areas or Opportunity Areas, a high proportion of schools with OFSTED ratings of 'Requires Improvement' or 'Inadequate', or schools where reading attainment was a cause for concern). Further details of the matching process can be found in the *Impact evaluation* section of this report. With the first Covid-19 lockdown occurring in March 2020, the project was put on hold due to the cancellation of all statutory testing. The project was redesigned with the evaluation delayed by 12 months. Due to continued periods of lockdown and the cancellation of statutory testing in 2021, the evaluation was again put on hold. The original TLIF funding expired, so additional funding for the continuation of the programme, and the evaluation, were instead provided directly by the Education Endowment Foundation (EEF).

Discussions between NFER, EEF and RMT about the future of the evaluation, continued throughout 2021. In Spring 2021, the evaluation team raised their concerns about the potential difficulties in recruiting a comparison group of schools in the Spring of 2022, for a data collection in the Summer of 2022. An alternative design and data collection method involving the NPD had been proposed in the case that not enough schools agreed to be part of the comparison group. This alternative design is discussed later in the report (see *Lessons learned* section).

With schools returning to a near-normal within-school delivery from September 2021, the research was restarted from September 2021, with just one intended outcome measure identified as the PSC for Year 1 pupils in the Summer of 2022. KS1 Reading Assessments were omitted from the study plan due to the data collection burden imposed upon schools. The evaluation team planned to go ahead with recruitment, and the alternative NPD design would be discussed as a back up in the case that recruitment was not possible. In December 2021, it became evident that it would be unlikely for the evaluation to recruit a fully comparable sample of comparison schools. The pool of available comparison schools was becoming smaller due to some schools starting to use RMT programmes, and with general recruitment rates to evaluations post-Covid being reduced. This meant that to get a large enough, and similar enough comparison group, the evaluation team would have needed to have an extremely high recruitment rate, from a diminishing population of schools. The changing nature of the design and the number of treatment schools remaining within the evaluation meant that to maintain power, a comparison group of approximately 70 schools would need to be recruited. This was based on 65 schools receiving the treatment. As discussed later, the number of treatment schools that had actively provided the initial data was potentially as low as 58. A treatment group of this size would require approximately 79 comparison schools. The evaluation team recommended at this point to proceed to the alternative design, which would rely purely on NPD data without any active involvement from comparison schools, and that some schools could still be recruited to inform the IPE work. This was discussed at length between NFER, EEF and RMT, and a decision was made to continue with the original design involving recruitment of a comparison group of schools that would provide outcome data due to concerns about limitations of the alternative design (discussed later in this report).

² Blackpool, Northumberland, Sheffield & Doncaster, Stoke-on-Trent, South Central

³ <https://www.gov.uk/government/publications/social-mobility-and-opportunity-areas>

⁴ The published study plan can be found here:

https://d2tic4wvo1iusb.cloudfront.net/documents/projects/RMT_TLIF_-_Study_Plan_2019.04.09_Updated.pdf?v=1630925453

⁵ The collection of scaled scores was considered more robust to the evaluation as they provide greater differentiation between pupil performance than that available through the use of Teacher Assessments.

Following extensive recruitment activity in early 2022, 93 schools agreed to be part of the comparison group for this evaluation. Although the number of schools met our sample size calculations, the characteristics of the schools were not similar enough to our treatment schools (see Table 5 in *Reasons for cancellation*). Following discussions between EEF, RMT and NFER, the decision was made not to proceed with the study using the matched design, due to the poor quality of the matches between treatment and comparison schools. It was also decided that we would not attempt to switch to an alternative design using administrative data due to concerns about limitations of this option. The reasons for this are explained in greater detail later in this report, but cover issues around the lack of a complementary IPE, not being able to identify unofficial use of RMT products in comparison schools, and the wider disruption caused by the Covid-19 pandemic.

Table 1. Project timeline

Date	Activity
February 2018	All treatment schools recruited by RMT.
March 2018	First cohort of treatment schools commence programme.
September 2018	Evaluation commissioned by EEF.
	Project set-up meeting between EEF, NFER and RMT
September 2018	Second cohort of treatment schools commence programme. First cohort continue with programme.
December 2018	Baseline survey for teachers in the second cohort of treatment schools.
September 2019	Second year of treatment starts.
February 2020	Recruitment of schools for comparison group starts.
March 2020	Schools closed due to Covid-19 pandemic. Recruitment activities stopped.
March 2020	DFE announces cancellation of statutory testing in primary schools for 2020.
April 2020	Project put on hold.
April–August 2020	Ongoing discussions between EEF, NFER and RMT about restart of project.
September 2020	Project restarts with amended target pupils but still with phonics and KS1 reading from 2021 as main outcomes. Pupils of interest were now those that started Reception in September 2019 and pupils who started Year 1 in September 2019.
December 2020	DFE announces cancellation of KS1 testing in primary schools for 2021.
January 2021	Schools closed due to Covid-19 pandemic.
January 2021	DFE announces cancellation of Phonics Screening Check (PSC) for 2021.
January–August 2021	Discussions between EEF, NFER and RMT as to whether evaluation can continue.
March 2021	Planned re-opening of schools.
September 2021	Project restarts with Year 1 in September 2021 (the main pupils of interest for the evaluation), using PSC as primary outcome. To assist the recruitment of comparison schools and to reduce the burden of data collection for all schools, KS1 reading results are dropped from evaluation.
December 2021	NFER undertake matching of treatment schools from pool of available comparison schools.
February 2022	NFER informed EEF that responding sample of comparison schools will not be as comparable as evaluation requires. Continuation would therefore be a risk.
April 2022	Following discussions between EEF, NFER and RMT, the project was cancelled, and all schools informed.

The programme

Read Write Inc. Phonics aims to support the development of pupil literacy regardless of socio-economic status, special need or language status, by providing a whole-school approach to teaching phonics and early reading. Reading is a more important driver of social mobility than socio-economic status (OECD, 2002). Children from deprived backgrounds read better and enjoy reading more when they have excellent teachers (Read On Get On, 2016). Leaders who build a school culture of supportive professional development and teachers who have a love of reading have the biggest impact on children's literacy outcomes.

This study involved the evaluation of a professional development and teaching programme which is based on the *Read Write Inc.* Phonics programme. The programme was developed by Ruth Miskin and is delivered by RMT. Oxford University Press (OUP) publish the programme. The *Read Write Inc.* programme itself is on the list of validated phonics programmes that the Department of Education display on their website. The list of programs can be found here: <https://www.gov.uk/government/publications/choosing-a-phonics-teaching-programme/list-of-phonics-teaching-programmes>.

Theory of change

The theory of change (ToC) is a way of presenting how a specific initiative and its planned activities are linked and will lead to its intended short- and long-term impacts. Below is the ToC for this evaluation.

The study plan identifies the full list of activities treatment schools would have received from RMT. Between March 2018 and March 2020 these were delivered as normal. Activities included tailored training to the reading leader, whole-school training, as well as a number of development days for the leadership team and teachers. Supply cover was also provided during this time, to release the reading leader for specific training. As a consequence of the pandemic, there were no funded activities for the academic year 2020/21 but additional funding for 2021/22 was secured from EEF that allowed access to online training, an hour-long CPD session with the reading leader, three development days and three remote progress meetings. The third year of funding did not provide supply cover.

Assumptions

- Evidence suggests that reading is a more important driver of social mobility than socio-economic status.
- Children from deprived backgrounds read better and enjoy reading more when they have excellent teachers.
- Leaders who build a school culture of supportive professional development and teachers who have a love of reading have the biggest impact on children's literacy outcomes.
- RMT supports schools teaching the *Read Write Inc.* Phonics programme.
- Schools develop a cohesive approach to teaching phonics and early reading and teach children to decode as the primary strategy for word reading.
- Word reading is embedded and practised in closely matched storybooks, enabling children to develop reading fluency which in turn aids comprehension.

Strategies and activities

What is the approach?

- Training for the reading leader before whole-school training
- Two whole-school staff training days, with an SLT meeting at the end of the first day
- 16 support visits to each school, including a leadership implementation day
- Two headteacher and reading leader Literacy Leaders training days
- Half-termly assessment of children using sound and word reading, allowing the reading leader to regroup as necessary
- Daily individual tutoring for the slowest progress 20% of children in Reception, Year 1 and Year 2.

Resources include:

- Resources include Teacher Handbooks, resources for teaching phonic lessons, decodable storybooks and writing books for children
- Resources are available for schools to purchase from Oxford University Press.

Short-term outcomes (1–2 years)

Pupil impact:

Primary outcome:

Improved phonics attainment, measured by the PSC.

Secondary outcome:

Improved reading attainment, measured by KS1 reading outcomes. This outcome is not being used in the main analysis.

School-level impacts:

- Teacher confidence in teaching phonics and early reading improves
- Reading leader has enhanced ability to lead literacy in school, including practices such as cycle of practice, coaching and feedback.

Target groups

- Schools: Primary
- Regions: PAs (as set by DfE)
- Pupils: All pupils in Reception, Year 1 and Year 2

Longer term outcomes (2–5 years)

- Improved attainment in reading, at KS2
- Improved attainment in writing, at KS2

Neither of the longer term outcomes were part of this evaluation.

Impact evaluation

Quasi-experimental design (QED) method

At the conception of the evaluation, in-depth discussions took place between EEF, NFER and RMT as to the most appropriate design. It was decided the design should be able to identify the long-term performance in reading skills within treatment schools prior to the introduction of the programme. The evaluation therefore planned to use a difference-in-differences (diff-in-diff) design with a matched schools approach, which aimed to assess the impact of the programme on pupils' reading ability and phonic knowledge. This design relied on the parallel trends assumption, which assumes that, in the absence of the intervention, the treatment and comparison groups will follow the same trend (Liu et al., 2010). Comparison schools were to be chosen by selecting eligible schools based on observable characteristics, using a propensity score matching (PSM) approach. Details on how this was carried out are provided below.

The primary analysis was to investigate the impact of the RMT: *Read Write Inc. Phonics* programme, on the overall PSC score. Secondary analysis would have converted the raw score into a threshold indicator identifying whether a pupil had reached the expected standard. The diff-in-diff approach requires data prior to the commencement of the programme; in this case the relevant phonics data was to be requested from the NPD. For all schools, whether in the treatment or comparison group, the analysable data sets would have contained historical phonics check data between 2015 and 2019, and the planned 2022 data collection (years 2020 and 2021 were cancelled due to Covid-19). Models would have aimed to identify any noticeable change in performance in 2022, given the trend in performance between 2015 and 2019. A two-year gap in assessment data is slightly problematic within the design as the evaluation is unable to determine the trajectory in these years. The assumption would have been that treatment and comparison schools continued on the same trajectory as seen between 2015 and 2019. This may not have been the case and so may have resulted in an over- or underestimation of programme effects, had the evaluation gone ahead.

The original design included the secondary analysis of KS1 reading outcomes for Year 2 pupils in the same academic year. Following discussions between EEF, NFER and RMT, it was decided to drop the KS1 element from the evaluation as it was felt that this would significantly reduce the burden being placed on schools and therefore help in the recruitment of comparison schools.

The design of this evaluation was heavily reliant on the process of matching the schools undertaking the *Read Write Inc. Phonics* programme with eligible comparison schools. This process is described in the *Matching* section below.

Treatment schools

The original design for this evaluation had four cohorts of primary school pupils (see Table 2): Two groups of 36 schools, each with two cohorts (Year 1 and Year 2 pupils) resulting in 72 schools recruited by RMT in 2018. The school population was maintained primary schools in England with KS1 pupils. Originally, pupils to be tested were from Years 1 and 2.

TLIF eligibility criteria, as specified by the DFE were used to recruit the intervention sample. These criteria included:

- At least 70% of schools had to be from a PA⁶ or OA.⁷
- Schools had to have an Ofsted rating of 'Requires Improvement' or 'Inadequate'.
- Ofsted-rated 'Good' schools with literacy data that was causing concern were considered on a case by case basis if in an OA or PA.
- Ofsted-rated 'Outstanding' schools were not considered.
- Schools could not have had in-school training or support from RMT in the previous two years.

Of the 72 schools recruited by RMT prior to the evaluation being commissioned, a majority (62%) of the intervention sample were Ofsted-rated 'Requires Improvement' or 'Inadequate' and from an OA or PA. The other 38% represented two other combinations:

- 1 rated 'Requires Improvement' or 'Inadequate', and outside of the OAs and PAs
- 2 Ofsted rated 'Good' but had literacy data that was causing concern⁸ and were in an OA or PA.

⁶ Blackpool, Northumberland, Sheffield & Doncaster, Stoke-on-Trent, South Central

⁷ <https://www.gov.uk/government/publications/social-mobility-and-opportunity-areas>

⁸ Causing concern indicates schools that were below the national average for KS2 reading.

At the commencement of the project, it had been agreed with DfE that if any school dropped out after training had occurred they would not be replaced. Since the start of the project in 2018, some attrition did occur. Three of the original schools dropped out of the study prior to commencing in September 2018 and were replaced by RMT as agreed with DfE. Since September 2018, and prior to December 2019, two further schools stopped receiving the programme, one due to a school closure and the other did not wish to continue. Between September 2020 and February 2022, a further five schools dropped out of the evaluation, having not wished to continue into another year of evaluation or provide the relevant data. Matching was therefore conducted on the remaining 65 treatment schools. However, after matching, a further three schools dropped out. This resulted in 62 schools receiving the RMT: *Read Write Inc.* Phonics programme and being included in the final revised design of the evaluation.

Pupils to be selected for analysis were on roll and in their Year 1 class by 30th September 2021. Late arrivals were to be excluded from analysis.

The following table identifies the cohorts of schools and pupils that were identified to be treatment schools and how this changed during the life of the project.

Table 2. Cohorts of schools recruited to the intervention

Cohort	Start of intervention delivery	Year group at start	End of intervention delivery	Year group at end	Outcome	N
1	Spring 2018	Reception (mid)	Spring 20	Year 2	KS1 reading	72
2	September 2018	Reception (beginning)	Spring 20	Year 1	PSC	
3	September 2018	Reception (beginning)	Summer 20	Year 1	PSC	
4	September 2018	Year 1 (beginning)	Summer 20	Year 2	KS1 reading	
Design amended due to cancellation of 2020 statutory assessments (Nov 2020)						
5	September 2019	Reception (beginning)	Summer 2021	Year 1	PSC	67
6	September 2019	Year 1 (beginning)	Summer 2021	Year 2	KS1 reading	67
Design amended due to cancellation of 2021 statutory assessments (Nov 2021)						
7	September 2021	Year 1 (beginning)	Summer 2022	Year 1	PSC	65

It can be seen from Table 2 that at each stage of restart a number of schools from the original treatment group were lost to the evaluation, despite reminders from NFER as to the importance of remaining within the evaluation and the efforts of RMT. It is also worth noting that in the final months of the evaluation between January 2022 and March 2022, a further three schools failed to re-sign MOUs and four had not at that point provided the necessary data, although NFER and RMT were still confident the schools who had not provided data would do so.

Matching

The first matching exercise to create a comparison group of schools was undertaken in December 2019 and was based on the 72 original intervention schools. After a pause to the project due to Covid-19, the matching process was rerun in late 2021. As described previously, during the course of the evaluation, and prior to the 2021 matching process, seven schools dropped out of the evaluation. Matching for the redesigned evaluation was therefore conducted on the remaining 65 treatment schools. In 2021, RMT also provided a full list of schools that had taken part in any RMT programme within the last five years. This list was used to exclude those schools from the pool available for the comparison sample, as agreed between EEF, NFER and RMT. The comparison population consisted of all maintained primary schools in England with KS1 pupils who met the original TLIF criteria. It was a specific request from RMT that the recruitment process was undertaken as late as possible, due to an increased uptake of the programme and they wanted to allow schools the maximum opportunity to sign up to their programme. A requirement of being within the comparison group was that they would not sign up.

In total, 779 maintained primary schools in England were eligible to form the comparison population; that being schools that had not taken part in RMT programmes in the last five years and fulfilled TLIF criteria. However, within this population, some were more similar in characteristics to the intervention sample than others. At the start of the project this potential pool of comparison schools was over 1,000, but due to various changes in how schools met the eligibility criteria and with more primary schools engaging with RMT, this pool was much reduced. For the purpose of calculating sample size, and likelihood of being able to recruit a similar comparison sample, in December 2021, PSM matching was

carried out between the intervention sample and the pool of available comparison schools. The results of the PSM were used to inform the recruitment process, again in Spring 2022.

Four variables were used for the PSM: KS2 attainment, region, percentage of pupils eligible for FSM, and school size. These variables were selected to be part of the matching process due to their association with attainment outcomes. Ideally, we would have used prior attainment as measured by phonics data; however, this was not publically available. Only KS2 performance data is available on publicly available data sets, which restricted us slightly given the age of the treatment pupils and KS2 test takers. FSM eligibility is a measure regularly used as an indicator of deprivation. Whilst attainment and deprivation are quite regularly included within matching processes, the other two categories were included following discussions between EEF, NFER and RMT. The variables and their categories were as follows:

KS2: Quintiles of Reading performance in 2019

Region: North, Midlands, South

FSM: Percentage of pupils eligible for FSM (continuous variable)

School size: Number of pupils (continuous variable).

The RMT and comparison groups of schools were then split into three subsamples, according to the TLIF selection criteria:

- 1 in an OA/PA and Ofsted-rated 'Requires Improvement' or 'Inadequate' (78 potential comparison schools)
- 2 not in an OA/PA and Ofsted-rated 'Requires improvement' or 'Inadequate' (132 potential comparison schools)
- 3 not in an OA/PA and Ofsted-rated 'Good' but KS2 reading rated below the national average (569 potential comparison schools).

Matching was carried out on 65 RMT schools, in R using the 'Matchit' package for each subgroup separately. The method was nearest match, with Region and KS2 prior attainment quintiles being exact matches, and a matching ratio of 10.

Schools varied in how similar they were to the RMT sample characteristics. Table 3 below shows the number of RMT schools by the TLIF selection criteria, the results of the matches by the TLIF categories, and the matching variables used.

Table 3: Matching the comparison population to the treatment group

	RMT schools and their possible comparison schools					
	RMT schools	Similar schools				Non-similar Schools
		Matching variables				
TLIF selection criteria		KS2 attainment+ region+ FSM+ school size	KS2 attainment+ region+ FSM	KS2 attainment+ region	KS2 attainment	0 vars
OA/PA and Ofsted-rated 'Requires Improvement' or 'Inadequate'	39	72	2	0	4	0
Not OA/PA and Ofsted-rated 'Requires Improvement' or 'Inadequate'	6	26	1	0	30	75
Not OA/PA and Ofsted-rated 'Good' and KS2 reading below national average	20	96	77	48	50	298
Matching category N		194	80	48	84	373
Total N	65		406			373

In the first subsample category (OA/PA and Ofsted-rated 'Requires Improvement' or 'Inadequate'), 72 of the 78 possible comparison schools matched on all four variables used. A further two schools matched on KS2 attainment, region and FSM percentage, but did not match on school size. Four further schools matched on KS2 attainment only. In the second category (Not OA/PA and Ofsted-rated 'Requires Improvement' or 'Inadequate'), 26 schools matched on all four variables, one school matched on all but school size, and 30 schools matched on KS2 quintile only and 75 of the 132 schools did not match to the RMT sample on any of the variables. In the last category (Not OA/PA and Ofsted-rated 'Good' and KS2 reading below national average), 96 schools matched on all four variables, 77 schools matched on three variables, 48 on two variables, 50 on one variable, and 298 schools did not match on any variables.

Furthermore, using the figures in Table 3, it is possible to compare the percentages of RMT schools to the comparison schools in the different TLIF categories. It can be seen that over 50% of the RMT sample fit into the first TLIF category. However, the majority of the comparison schools fit into the bottom TLIF category. As such, it can be seen that the matched comparison sample does not have the same distribution of TLIF criteria characteristics. The differences in percentages are discussed in further detail in the *Reasons for cancellation* section below.

For the purposes of recruitment, the sample was split into two waves in order to prioritise the closest matches. Schools that matched on between one and four variables were categorised as similar schools. Schools that did not match on all variables were categorised as non-similar schools. In total, 406 potential comparison schools made up the similar group of schools, and 373 potential comparison schools made up the non-similar group.

Recruitment activity

This section describes the recruitment activity from inception of the evaluation in 2018, up until the cancellation of the project, just after the matching process in March 2022.

Intervention sample recruitment 2018

Intervention schools were recruited to the project in 2018, prior to when the evaluation was originally commissioned. Four cohorts of primary school pupils, as described above in Table 2, were recruited to the evaluation by RMT.

RMT provided the names and email addresses of the headteachers and reading leaders for each of the intervention schools to NFER in September 2018 and this information was used for all contact between NFER and these schools.⁹ This included the signing of Memorandums of Understanding (MoUs) and general contact to keep them informed of changes to the evaluation's design.

Comparison school recruitment February 2020

As the project was originally designed to collect attainment outcomes from summer 2020, comparison group recruitment was planned from February 2020 onwards. NFER contacted 167 schools in February 2020 inviting them to be part of the study and asking them to sign the MoU. The MoU outlined the roles and responsibilities of NFER and the comparison schools who wished to take part in the study; participation would include:

- completing a school survey
- sharing pupil data and KS1 raw scores for the PSC for pupils in Year 1 and KS1 reading raw score data for pupils in Year 2 in June 2020
- schools that completed these activities would receive an incentive payment of £200.

Following the initial invitation email, up to three reminder emails were sent to schools in February and March. Twenty-five schools agreed to take part in the comparison group and 29 schools asked to be withdrawn from the sample.

When the schools closed due to the Covid-19 pandemic in March 2020 the project was put on hold and schools were contacted to inform them of the project status. During the summer term 2020, NFER contacted the 25 schools who had agreed to take part as comparison schools to let them know that the project had been delayed.

Towards the end of 2020, work restarted on the project and a new sample of comparison schools was drawn. It was planned that recruitment of these schools would begin in January 2021; however, due to the schools' closures and the cancellation of statutory assessments in 2021, it was agreed that the project would be delayed for another year.

Intervention school communications

From February 2020 until project closure, NFER had regular contact with intervention schools informing them of changes to the design and updating them on the data collection expectations. NFER employed its normal contact protocols and at each change of design and project re-start schools were contacted and requested to sign a new MoU. Where there were requests for data, as in February 2022 for pupil data, a reminder strategy was used to maximise response. School attrition occurred throughout the project with a few schools dropping out during project restarts. At the point of final project closure, 58 schools had provided the requested pupil data.

Comparison sample activities 2022

A total of 779 schools were identified as potential comparison schools. This sample included 406 matched schools that could be matched using between one and four of the matching variables, and 373 schools who could not be matched (please see Table 4).

Discussions took place about what would be a suitable incentive or thank you payment for comparison schools. Due to the increasing concern as to whether schools would want to sign up to the evaluation this incentive was discussed as a way of increasing that response rate. It was agreed that all participating comparison schools would be offered an

⁹ This process complied with the GDPR guidance available at the time.

incentive payment of £500 upon receipt of their summer 2022 Year 1 PSC pupil scores and with the completion of a school-level survey.

Initial invitations were sent to the full comparison sample on 28 February 2022, and schools were contacted by email to invite them to sign up to the study by completing an online MoU. A paper version of this invite letter was also dispatched on 2 March. To encourage sign up from the matched sample, these schools were phoned from the 7–10 March. On the 10 March an additional email reminder was also sent to all schools.

Table 4. Sign up figures by 16 March 2022:

Sample	Number of schools	Number of schools signed MoU	% sign up
Matched sample	406	56	12%
Unmatched sample	373	37	10%

By 16 March, 93 schools had completed the MOU. Whilst the recruitment rate was as anticipated given the new challenges facing schools, the percentage of similar and non-similar schools was not ideal. Please see the *Reasons for cancellation* section for a fuller discussion of this point. At this stage, due to the imbalance between the treatment and comparison schools, further recruitment activities were halted following an EEF project review meeting.

As a gesture of goodwill, when the project was cancelled, schools in the comparison group were offered a thank you payment of £150 for their time. In order to make this payment, NFER collected the schools' bank details, and some schools did ask why we were offering to pay them the money since the evaluation had been discontinued. Out of the 93 schools who signed up only one school did not provide their bank details.

Baseline data collection for IPE

A number of data collections were planned to support the IPE component of this evaluation. Whilst not all these activities took place, a number did take place in 2018 and 2019.

We conducted telephone interviews with a number of developers/trainers, including the director/s of RMT; two members of the provider team who had been involved in planning, designing and delivering the intervention; and two further trainers, identified by the provider. Interviews explored the intervention characteristics, including expected dosage, CPD and support arrangements, expected classroom implementation and permissible tailoring, and any planned changes to implementation. Interviews took place in the early autumn of 2018 and mainly focused on the following dimensions:

- **Adaptation:** where and why it may be necessary to adapt the training sessions and how this may affect fidelity.
- **Programme support:** how the trainer/developer identifies who needs more support; methods of receiving requests for help; how much this has been a feature of the intervention in previous years.
- **Cost:** the cost of the programme, both from a school's and from the developer's perspectives.

The evaluation team planned to observe a complete set of CPD¹⁰ delivery received by one school, ensuring a complete picture of the scope of training, although the pandemic meant that this was impossible to complete. These observations would have aimed to ensure the evaluation team fully understood the intervention materials and expectations of teachers/schools. Targeted dimensions were as follows.

- **Fidelity:** the extent to which trainers adhere to the set training materials and timings, or shape training to fit the audience (or for other reasons).
- **Quality and responsiveness:** the quality of training provided, including the level of attendee engagement, the clarity and overall strength of the trainer.

As well as the above data collection exercises, two school staff surveys were also planned: baseline data collection from the treatment group and an endpoint survey of both the treatment and comparison group. The baseline survey took

¹⁰ These activities included: two-day training for the reading leader, two one-day whole-school training sessions, additional training days for the reading leader and literacy coordinators, 16 in-school development days with leadership team and teachers as well as being able to access a number of free events through RMT.

place in September 2018 but the endpoint survey, planned for May 2022, was cancelled. The baseline survey was delivered online and covered the following areas:

- Teachers' existing knowledge and confidence in teaching phonics, reading and writing.
- The school's existing strategies for supporting children who are falling behind (including 1:1 tuition).
- Any other literacy programme or CPD the school has been involved with.

Being part of the main TLIF evaluation, the delivery of the programme had already started in accordance with that funding, but was prior to the EEF evaluation being put in place. The first cohort of 36 schools were part of this initial delivery commencing before this evaluation was commissioned. This survey was therefore not thought to be appropriate for these schools. This baseline survey was sent to all 36 schools that were part of the second cohort of school recruitment. When sent to schools, the instruction was that the survey should be completed by every teacher who was involved in the delivery of the programme, including the headteacher and the reading leader. The survey was kept open for a number of weeks and a reminder strategy utilising further emails and telephone calls was carried out. At closure of the survey, 32 of the 36 schools (89%) had returned at least one survey, with a total of 196 responses.

Data collection was considerably disrupted by the Covid-19 pandemic. In total, we completed interviews with two providers and two trainers, together with four observations of training and the baseline survey in treatment schools mentioned above. We do not feel it is appropriate to report the findings from these data collection activities because the information is partial and dated. They were also focused on the requirements of TLIF, which funded schools' access to the programme. This means that the data collected does not reflect the way in which the programme has operated since 2018.

Reasons for cancellation

The main reason for the cancellation of this evaluation was the lack of comparability between the schools signed up to be comparison schools and the schools receiving *Read Write Inc. Phonics* as part of the DFE's TLIF initiative. As previously mentioned, the matching process relied on being able to exclude schools that had purchased the *Read Write Inc. Phonics* programme, including training from RMT. This resulted in a pool of possible comparison schools and Table 3 identifies the breakdown of these schools around the three main eligibility criteria used as part of the TLIF funding. . Due to the need to increase the pool of possible comparison schools, it was not possible to recruit from just the cohort of similar schools, and so the full population of similar and dissimilar schools were contacted. Table 5 below identifies the 93 schools that responded positively to their inclusion as a comparison school and how they compare to the treatment schools. Again, the 93 schools came from the two groups of schools, similar and non-similar.

Table 5: TLIF matching criteria for treatment and comparison schools

	Treatment schools	Comparison schools
1 = OA/PA and Ofsted-rated 'Requires Improvement' or 'Inadequate'	39 (60%)	17 (18%)
2 = Not OA/PA and Ofsted-rated 'Requires Improvement' or 'Inadequate'	6 (9%)	16 (17%)
3 = Not OA/PA and Ofsted-rated 'Good' and KS2 reading below national average	20 (31%)	60 (65%)
Total	65	93

Table 5 shows that the two groups of schools are quite different in their make up, based on the matching criteria discussed earlier. If a weighting strategy was adopted to balance these two groups, a significant percentage would have a weighting of above 3.¹¹ If we only used the 56 signed-up schools from similar schools (i.e., those that matched on at least one variable; see Table 4) we would have had proportions of 30, 14 and 55% in the TLIF criteria categories one, two and three respectively. Again, this would be quite different to the percentages for the treatment schools; it would also introduce issues with statistical power and a minimum detectable effect size (MDES)¹² increasing to a level that the evaluation team, and EEF, felt was unsupportable. The lack of comparability of the treatment and comparison samples was the main reason for project closure as EEF, NFER and RMT were in agreement that the robustness of any findings could be compromised.

Additional reasons considered in the decision making process concerned the level of exposure to the *Read Write Inc. Phonics* programme within the treatment schools. The original programme was for a two-year period of phonics delivery and the final design resulted in the Year 1 pupils of interest only receiving one full year of uninterrupted classroom learning. The previous Reception year was interrupted with a period of school closure between January 2021 and March 2021. This would have resulted in a gap in their learning at an important time in the development of reading skills. This reason also contributed to the decision that we would not continue with the IPE, as this disruption would have impacted on a teachers' experience of the programme, as well as some teachers having potentially changed from those who completed the original baseline survey. Our rationale for this decision was also influenced by the fact that a previous programme evaluation had its own IPE . A report of the IPE from the current study, without an accompanying impact evaluation, would potentially confuse teachers and schools, especially as there is a risk that the findings from a free-standing IPE could conflict with those of the RCT.

The gap in phonics data for 2020 and 2021 also introduced problems in the analysis and the underlying assumptions required for the main analysis. For the IPE analysis the original baseline survey was undertaken in the Autumn term of 2018 and so its relevance to what was happening in schools in the 2021/22 academic year was much reduced. The combination of these factors led the evaluation team to the unfortunate decision of cancellation,¹³ as to continue would have placed too much burden on schools and teachers at a time when they had many challenges within their schools. A burden that would not have resulted in an evaluation would have been as robust as originally intended.

¹¹ As a rule of thumb, weights of <0.5 and >2 can be problematic. Whilst larger weights can be considered as acceptable if the percentage of cases with those higher weights is low, Table 4 clearly shows that we would have had a high percentage of pupils with a weight of greater than 3.

¹² A further impact on the MDES came from a further reduction in treatment schools. At the time of the decision to stop the evaluation, only 58 schools had submitted all relevant information. A number of schools were still being contacted to provide the requested information but at the time of a decision needing to be made, only 58 had completed all requests.

¹³ Whilst the evaluation was cancelled, the support to treatment schools offered by RMT continued for the rest of the 2021/22 academic year.

Conclusion and lessons learned

This report summarises the activity that took place to evaluate the impact of *Read Write Inc. Phonics* on Year 1 and Year 2 pupils in 72 primary schools. The funding for *Read Write Inc.* was part of the DFE's TLIF funding opportunity. EEF were charged with managing this evaluation and following project delays provided additional funding to allow this evaluation to continue. NFER were commissioned by EEF to carry out this evaluation in 2018. The project went through a number of delays and design changes due to the Covid-19 pandemic and the resulting school closures.

There are primarily three lessons that can be learned from this project that could assist similar projects avoiding the same pitfalls. These are set out below.

1. Ensure there are more review periods built into the timeline to introduce stopping points, particularly when the evaluation is stopped and restarted.

This evaluation was set up prior to the Covid-19 pandemic. The pandemic resulted in periods of partial school closure with pupils undertaking virtual lessons. It also resulted in the cancellation of statutory assessments in the Summers of 2020 and 2021. The primary outcome for this evaluation was intended to comprise of a combination of the phonics screening test and KS1 reading assessments. The first attempt to restart the evaluation was in September 2020 but the trial was put on hold again following a further wave of partial school closures and the cancellation of national assessments. A second attempt to restart was made in September 2021, which ended in the eventual cancellation in Spring 2022. Given the important changes that took place to the study plan on each occasion, it would have made sense for a third person, or advisory group, to review the new plans. Issues that could have been assessed for a stop/go decision include:

- Has the programme changed?
- Are we still able to evaluate the programme as intended?
- Have the demands on the schools changed?
- Has the likelihood of school recruitment changed?
- Have the risks to the evaluation changed?
- Has the budget changed?
- Has the MDES changed? Is it still within reasonable expectations?
- Are we still able to produce meaningful insights for the sector?

A review of these questions, with a third party asking the evaluator and the EEF project team for justifications, may have resulted in an earlier decision not to continue.

2. Recruiting schools is always a challenge and their willingness to be involved in evaluations is difficult to predict.

The unique point with regards to this project was that the intervention schools were recruited as part of the initial TLIF evaluation and before this evaluation was commissioned. It was therefore not possible to recruit the comparison sample schools until very close to the end of the project, as it was agreed, following a request from RMT, that any schools not already involved in their programme were given maximum opportunity to sign up to their programme during the course of the intervention period. The matched sample approach also meant that we were limited in which schools we were able to contact, along with the fact that these schools were potentially unable to sign up to programmes as they were working on other priorities. In addition, the delays due to Covid-19 and the cancellation of statutory assessments meant that the recruitment and data requirement changed, and therefore recruitment had to be repeated a number of times. The additional pressures put on schools during this period were potentially harder felt by these particular schools.

When signing up to a project, schools will be looking more and more at what are the requirements for them and what are the benefits for them as a school or teacher in participating; for this project, before the thank you payment was introduced, there was little, if any, benefit for a comparison school to sign up and share their PSC and KS1 Reading data. Schools were aware that some of the data we were requesting would be available on NPD.

Providing feedback on the evaluation findings does increase a school's willingness to participate but often, in the case of an RCT/QED, the timeline between participating and publication can be quite long and so there is little immediate benefit.

The Covid-19 pandemic has put additional pressures on schools and in certain project areas it has impacted on the willingness of schools to sign up to research projects, because they need to cover staff absence and implement additional catch-up activities. For these reasons, and for the September 2021 restart, we suggested that an increased incentive should be used for this project. At the time of comparison school recruitment in the Spring term of 2022, schools were keen to sign-up, as identified in the recruitment section of this report. Our assumption is that this was a consequence of the school's knowledge that they would receive £500.

If we were to use this project as evidence of the benefit of an incentive, we could see that after only two weeks in the field for recruitment, we had had a positive response from schools and a substantial number of schools on board. Whilst being a positive response, the make up of the comparison group did not fulfil the requirements for a comparison sample when making robust comparisons and so contributed to the reasons for cancellation. This response achieved through this recruitment activity is in comparison to other national projects that also recruited primary schools, where a rate of less than 10% had been achieved. Although these projects also involved some form of incentive, a response rate of just above 10% can be seen as being above our recent experiences.

3. Use of secondary data .

The NPD is a rich and well-used resource for educational research. The original design for this evaluation specified KS1 Reading results as a secondary outcome. However, the NPD only contained teacher assessments,¹⁴ so the design included collecting scaled scores directly from schools. As the evaluation needed to collect this data, it was decided to also collect the phonics screening test outcomes at the same time, even though they would become available on a future NPD release. In December 2021, just prior to comparison school recruitment, NFER recommended that the evaluation should rely on the data available through the NPD in November 2022. There were significant analytical benefits to using the NPD, although there were also known drawbacks.¹⁵ This recommendation was based on the fact that the decision had already been made to drop KS1 from the planned outcomes and all the required phonics outcome data would have been available through the NPD. The reasons for not following the NPD route were due to an inability to identify the extent of unofficial use of the programme in comparison schools and the inability to access data on current practice in comparison schools, including the extent to which they were providing phonics instruction of a similar nature to Read Write Inc. and information on provision during periods of Covid-19 disruption. Whilst acknowledging the advantages and disadvantages in using the NPD, this does not detract from the known issues in the delivery of the programme during a significant period of disruption to pupil learning. Whilst use of NPD could have been stronger analytically, there would have still been major questions of how to interpret the results to provide meaningful insights to the teacher in the classroom. There would also be no linked IPE evaluation, which is an important part of a programme's overall evaluation.

¹⁴ Teacher assessments are not considered suitable as a primary outcome, due to being a subjective judgement from the teacher, although based on teacher knowledge of the child, and that the number of categories do not differentiate between pupil abilities to allow robust analysis.

¹⁵ The main drawback related to our inability to identify if schools were using *Read Write Inc. Phonics* unbeknown to RMT. RMT provided a list of schools that had purchased the programme in the last five years, and these schools were excluded from the list of possible comparison schools. When the recruitment of comparison schools commenced, all schools were asked not to sign up if they used this programme.

References

- Liu, C., Zhang, L., Luo, R., Rozelle, S. and Loyalka, P., 2010. 'The effect of primary school mergers on academic performance of students in rural China', *International Journal of Educational Development*, [e-journal] 30(6), pp.570–585. Available at: <https://doi.org/10.1016/j.ijedudev.2010.05.003>
- Organisation for Economic Co-operation and Development (OECD), 2002. *Reading for Change Performance and Engagement across Countries – Results From PISA 2000*. [pdf] Available at: [PISA-Reading-eng \(oecd-ilibrary.org\)](https://www.oecd-ilibrary.org/) [Accessed 12 August 2022].
- Read On Get On, 2016. *A Strategy to get England's Children Reading*. [pdf] Available at https://cdn.literacytrust.org.uk/media/documents/Read_On_Get_On_Strategy.pdf [Accessed 12 August 2022].

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0.

To view this licence, visit <https://nationalarchives.gov.uk/doc/open-government-licence/version/3> or email: psi@nationalarchives.gsi.gov.uk


Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned. The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

This document is available for download at <https://educationendowmentfoundation.org.uk>



The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

<https://educationendowmentfoundation.org.uk>

 @EducEndowFoundn

 Facebook.com/EducEndowFoundn