Education Endowment Foundation

Response to Intervention

Evaluation Report and Executive Summary

February 2014

Independent evaluators:

Durham University

Professor Stephen Gorard

Dr Nadia Siddiqui

Dr Beng Huat See

The Education Endowment Foundation (EEF)

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.

We aim 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;
- Encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was founded in 2011 by lead charity The Sutton Trust, in partnership with Impetus Trust (now part of Impetus – The Private Equity Foundation), with a £125m grant from the Department for Education. With investment and fundraising income, the EEF intends to award as much as £200m by 2026. Together, the EEF and Sutton Trust are the Government-designated What Works Centre for Improving Education Outcomes for School-Aged Children.

Literacy Catch-up

In May 2012 the Department for Education awarded the EEF a further £10 million for a grants round dedicated to literacy catch-up projects for children at the primary-secondary transition. The projects funded within this round aimed to identify effective ways to support pupils who do not achieve Level 4 in English by the end of Key Stage 2.

For more information please contact:

Robbie Coleman
Research and Communications Manager
Education Endowment Foundation
9th Floor, Millbank Tower
21-24 Millbank
SW1P 4QP

p: 020 7802 1679
e: robbie.coleman@eefoundation.org.uk
w: www.educationendowmentfoundation.org.uk

About the evaluator

The project was independently evaluated by a team from Durham University led by Professor Stephen Gorard.

Stephen Gorard is Professor of Education and Well-being, and Fellow of the Wolfson Research Institute at Durham University. He is a Methods Expert for the US government Institute of Education Science, member of the ESRC Grants Awarding Panel, and Academician of the Academy of Social Sciences.

His work concerns the robust evaluation of education as a lifelong process, focused on issues of equity and effectiveness. He regularly advises governments and other policy-makers, including oral and written advice to the House of Commons Education Committee every year since 2003. He is also a widely read and cited methodologist, involved in international and regional capacity-building activities, and used regularly as an adviser on the design of evaluations by central and local governments, NGOs and charities. He is currently an evaluator for the European Commission Directorate-General for Education and Culture, the Department of Work and Pensions, the Food Standards Agency, the Learning and Skills Information Service, and the Education Endowment Foundation. He is author of nearly 1,000 books and papers.

Contact details:

Professor Stephen Gorard
School of Education
Durham University
Leazes Road
DH1 1TA

p: 0191 334 8419
e: s.a.c.gorard@durham.ac.uk

Contents

Executive Summary

The project

Response to Intervention (RTI) is a targeted programme that uses a tiered approach to identify the needs of low achieving pupils. The approach begins with whole class teaching (Tier 1), followed by small group tuition (Tier 2) for those who need more attention, and one to one tutoring (Tier 3) for those who do not respond to the small group instruction.

In this evaluation, the programme was delivered to Year 6 pupils who were at risk of not achieving Level 4 English at Key Stage 2 (KS2). RTI was delivered in the summer term in preparation for their transfer to secondary school. The development of the intervention, training and materials was led by the Centre for Use of Research Evidence in Education (CUREE). Achievement for All 3As (AfA3As) provided support as schools recruiter and the ongoing contact with schools. AfA3As Achievement Coaches worked closely with schools to support their use of RTI within the AfA3As framework. The two organisations worked together to deliver the intervention.

What impact did it have?

The overall impact is based on the New Group Reading Test post-test, which sought to compare the outcomes of pupils who were supported using the RTI approach to other similar pupils who were not. The first had an estimated effect size of +0.19 . This headline result suggests that this could programme have an impact equivalent to 3 months of additional progress in one year.

However, this result is based on a trial which should be considered 'spoilt' due to the level of school dropout post-allocation, the number of schools which did not carry out post-testing especially in the control group, and confusion over which pupils were eligible for the intervention. Therefore this effect size should not be taken as indicative of impact.

Even further caution must be taken when reading the results for specific sub-groups of pupils such as boys only and pupils eligible for free school meals. These show larger effect sizes but these results are based on very small numbers, and are therefore not deemed secure.

The process evaluation suggested that the intervention would have been more effective had it been started at the beginning of the year and run over a longer period of time, rather than as a catch-up intervention in the busy period at the end of Year 6. Some schools were only able to implement a small number of sessions, which made it impossible to implement, monitor and adjust the intervention as intended.

The approach was popular with teachers and pupils, and appears likely to have had positive benefits on wider outcomes such as confidence and self-esteem.

Group	Number of pupils	Effect size	Estimated months' progress	Evidence strength*
All pupils	**385**	**+0.19**	**+3**	★☆☆☆☆
FSM-eligible	96	+0.48	+6	
Boys only	210	+0.26	+3	

*Evidence ratings are a new measure under development based on a number of factors including study type, size and drop-out.
Ratings are provisional and are not given for sub-group analyses, which will always be less secure than overall findings.
For more information about ratings visit: **www.educationendowmentfoundation.org.uk/evaluation**.

How secure is this finding?

The evaluation was set up as an efficacy trial to test the impact of RTI as delivered with the developers, CUREE and AFA3As, leading the training and overseeing the provision of the intervention. Efficacy trials seek to test evaluations in the best possible conditions to see if they hold promise, but they do not seek to demonstrate that the findings hold at scale in all types of schools.

The findings are based on a randomised controlled trial using a simple waiting list design. All volunteer schools would receive the RTI intervention; half would be randomly allocated to the immediate intervention group and half would receive the intervention the following school year.

A total of 61 schools agreed to take part and were randomised to treatment (30) or waiting-list control (31). All Year 6 pupils completed a pre-test before the allocation of schools to treatment or control, and before identification of the pupils eligible or suitable for intervention in each school. Schools were required to identify their eligible Year 6 pupils by using the results of the pre-test, teacher judgement, and the RTI diagnostic tool. The identification of these pupils in both treatment and control schools, and the sharing of this information with the evaluator, was important in establishing a valid comparison group prior to the start of the intervention.

Ideally, schools would have identified eligible pupils prior to randomisation in order to ensure that the pupils in the control group were identified under the same conditions and timing as those in the intervention group. However, though this sequencing was originally proposed by the evaluator, it was agreed that pupil identification could take place after randomisation due to the short timeframe in which the intervention was to be delivered. This would not have compromised the trial if the eligible control pupils had been identified early, and this data shared with the evaluator. In practice, however, the evaluator did not receive data on the eligible control pupils until after the delivery of the intervention. The timing and unreliability of this data meant that the findings of the evaluation were substantially weakened.

The security of the findings was further undermined during the trial by the withdrawal of 12 schools between randomisation and post-test.

Due to the combination of problems noted above the trial was irrevocably damaged and consequently the findings are deemed too insecure to assess the impact of RTI as a catch-up intervention.

Results were also analysed using a second measure, the Progress in English test. In contrast to the New Group Reading Test, this test suggested that the approach may have a negative impact on outcomes (an effect size of -0.09). These mixed results make interpretation of the results more challenging and highlights the sensitivity of outcomes to what is tested.

Though prior research, mostly from the US, suggests that RTI is an intervention with some promise, this evaluation has not answered the question as to whether the approach can be effective in English schools.

How much does it cost?

The cost of the approach is estimated at £117 per pupil, based on 15 pupils being eligible per school. This estimate includes resources (costed at £20 per pupil), initial training and supply cover (£68), plus on-going monitoring (£29). This does not include any additional cost of teacher or teaching assistant time to deliver Tier 2 or 3 interventions.

Key Conclusions
1. Due to the weak control group and problems with recruitment and retention, this trial has not added much to the existing evidence on the impact of Response to Intervention.
2. The work of the developers and the results of the process evaluation suggest that the approach is feasible and welcome in schools.
3. Increasing the time between training and completion would enhance the chances of the intervention having an impact.
4. Running the intervention at the very end of Year 6 in preparation for Year 7 made it necessarily brief, and put it in competition with too many other end-of-school activities. This meant that some schools reported delivering very few sessions in practice, compared to previous studies which involved 20 or more hours of tuition.

Introduction

Intervention

This is a report of an evaluation of Response to Intervention (RTI). The programme was offered to Year 6 pupils who were at risk of not achieving Level 4 English at Key Stage 2 (KS2). The intervention was administered in the last few weeks of the 2013 summer term in 61 primary schools in England. The aim of RTI is to provide an individualised intervention targeted at the specific needs of each child in the form of a whole class approach as preventive teaching (Tier 1), followed by small group remediation (Tier 2) for those who need more attention, and one to one tutoring for those who do not respond to the small group instruction (Tier 3). In line with the RTI approach, the duration, intensity and frequency of treatment was allowed to vary from school to school. The programme being evaluated was led by CUREE (the Centre for Use of Research Evidence in Education), supported by AfA (Achievement for All), itself a UK-based school improvement intervention to improve the attainment of children with special educational needs. The AfA3As framework has some elements which are similar to RTI.

Background

RTI is a school-wide multi-tier programme that measures pupils' response to research-based instruction. It is a personalised and targeted intervention developed in the United States. Early evidence from the US suggests that this approach is effective with pupils in the transition period, defined in England as the stage when pupils move from Year 6 (final year of primary school) to Year 7 (first year of secondary school). It was identified as one of the more promising approaches to literacy for pupils in this age group in a prior review for the EEF (See et al. 2012).

Existing evidence for the intervention

Early evidence from some quasi-experimental studies and RCTs in the US suggests that RTI has a positive impact on literacy outcomes for pupils in the transition age, however the quality of evidence is mixed, and the scale of most studies is far from satisfactory.

In one of the larger studies, Vaughn and Fletcher (2012) examined the efficacy of RTI as a remedial intervention for 784 middle school pupils in the 6th, 7th and 8th Grades (English Years 7, 8, 9). Treatment pupils were randomly assigned and those that received both Tier 1 and Tier 2 interventions made improvements in decoding, fluency and reading comprehension ($d = 0.16$) compared to those who only received Tier 1 intervention. However, many of the Tier 1 only students also received other interventions from their schools (Fuchs et al. 2010).

Another randomised study, but with only 30 pupils in Grades 6 to 8, found differences between intensive small group treatment and a control (Leroux et al. 2011). The intervention was administered to treatment students every day in a 45 to 50 minute period. The results suggest that the intervention had a positive effect for pupils with severe reading difficulties, but the gains were not large enough to close the gap with typically performing pupils.

Graves et al. (2011) conducted a quasi-experimental study that compared small group intensive reading instruction (Tier 2) with a control group ('business as usual') for 6th Graders with and without

learning disabilities. The study was conducted in a large urban school where all of the pupils were on free or reduced price lunch and where 90% of the children were not first language English speakers. All were 'below' or 'far below' basic level in literacy. The duration was 30 hours over 10 weeks. The study reported that the treatment was more efficacious for pupils with learning disabilities, and for oral reading fluency, but less so for reading comprehension.

Faggella-Luby and Wardwell (2011) examined the effects of the small group (Tier 2) component of the intervention for 86 at-risk students in the 5th and 6th Grades in an urban middle school. Students were randomly assigned to one of the three instruction practices: (a) experimental (Story Structure to improve reading comprehension), (b) comparison (Typical Practice) and (c) control (Sustained Silent Reading). Each session was 30 minutes and was administered 2-3 days per week for 18 weeks. The results suggest that the programme benefitted older children (Grade 6) more than younger children (Grade 5).

Explanation of the stage of development of the intervention

RTI is a widely used intervention in the US and the concept was first developed for tackling problems associated with pupils with learning disabilities (Fuchs et al. 2010, Fletcher et al. 2004) and as an early reading intervention (National Reading Panel 2000, Snow et al. 1998). Evidence on RTI so far has been mixed, incomplete and largely from the US. Most studies reported have small samples and sometimes involve only those with learning disabilities. There is no clear evidence that RTI works, neither is there valid evidence that it does not work. It is not yet clear, therefore, whether RTI is the way to go when dealing with 10 and 11 year olds struggling with literacy in the UK. In the US, despite the numerous resources, books, materials and training programmes being developed and made available to the public, there is little guidance on how RTI could be implemented within the framework of the classroom. In the UK, the situation is even more under-developed as such resources are not generally available, and the programme has not been tested in any large RCTs in classroom conditions. The programme is thus appropriate for an efficacy trial in the UK.

Relevant policy and practice context

In the US the RTI programme gained popularity from 2004 when the government initiated an inclusion policy to bring together regular classroom and special education programmes. Programmes like RTI were used to identify children with learning disabilities, and so provide a differentiated strategy of intervention within regular classroom settings. The aim of the RTI programme being evaluated here is different to that in the US in that it does not provide a differentiated programme within a regular school system for children who would otherwise be in special schools. In other words, it is not an inclusion policy. It is being developed as a literacy 'catch-up' programme offered to pupils in their last year of primary school who are at risk of not achieving the expected level, to enable them to reach the functional literacy necessary for success at secondary school.

In the UK, concern for the dip in pupil progress during transition from primary to secondary education has led to a plethora of induction activities to facilitate a smooth transition. These have tended to focus on the social and emotional issues related to transition. Hargreaves and Galton (1999) found that one of the reasons for the hiatus in pupil progress at transition was the change in emphasis from literacy skills, such as reading, writing and comprehension in primary school, to an emphasis on response to literature in secondary school. They suggested that more should be done to ensure curriculum continuity as well as continuity in teaching and learning practices (Galton et al. 2000). In September 2000, a pilot study of the Key Stage 3 Strategy was conducted in 150 schools across 17 LEAs, the aim of which was to promote continuity between KS2 and KS3 (Goodwin 2002), however it was only in

2001 that a policy initiative was introduced to ensure continuity and progress in pupils' literacy between the two phases with the extension of the National Literacy Strategy to Key Stage 3. Since then there have been many studies looking into ways to improve literacy and numeracy for pupils in the transition phase.

Catch-up literacy projects are a set of educational interventions intended for pupils struggling to reach what are officially deemed the age-appropriate levels in reading. They are founded on evidence that struggling pupils entering secondary school are more likely to remain behind their classmates, or fall further behind their classmates, which can also lead to other issues such as disruptive classroom behaviour. In May 2012 the government in England made £10 million available to the Education Endowment Foundation (EEF) via the Department for Education for a grants round dedicated to literacy catch-up projects for children at the primary-secondary transition. It was intended to benefit pupil premium children who enter secondary school with below Level 4 in literacy (https://www.gov.uk/government/news/10-million-to-boost-literacy-for-year-sevens). In 2011 the EEF commissioned a review to identify the most promising literacy catch-up programmes for pupils from disadvantaged backgrounds (See et al. 2012): one of the most promising programmes was RTI, having many features of successful literacy catch-up interventions, including individual case analysis, one to one and small group support, and specially prepared teaching and learning materials.

Rationale for conducting the evaluation

RTI is a widely used intervention in the US and the concept was first developed for tackling problems associated with pupils with learning disabilities (Fuchs et al. 2010, Fletcher et al. 2004) and as an early reading intervention (National Reading Panel 2000, Snow et al. 1998). Evidence on RTI so far has been mixed, incomplete and largely from the US. Most studies reported have small samples and sometimes involve only those with learning disabilities. There is no clear evidence that RTI works, neither is there valid evidence that it does not work. It is not yet clear, therefore, whether RTI is the way to go when dealing with 10 and 11 year olds struggling with literacy in the UK. In the US, despite the numerous resources, books, materials and training programmes being developed and made available to the public, there is little guidance on how RTI could be implemented within the framework of the classroom. In the UK, the situation is even more under-developed as such resources are not generally available, and the programme has not been tested in any large RCTs in classroom conditions. The programme is thus appropriate for an efficacy trial in the UK.

The impact evaluation should therefore provide evidence of the average effect of RTI on literacy performance of struggling readers in transition. It can help determine the effect of such targeted intervention for children identified as not achieving the expected Level 4 at KS2. This is the first randomised controlled trial of the RTI programme in the UK at this scale. Given the underdeveloped resources and previously patchy application of RTI in the UK, a process evaluation of this programme was also warranted to assess the appropriateness of the training, the tools and protocols used, as well as fidelity of treatment in real classroom conditions. The process evaluation provided formative evidence on all aspects and phases of the implementation. This helps to assess whether teachers, given some basic guidance, can develop their own resources for use in the classroom for the different tiers of application. The evaluation also assessed the usefulness, ease of use and practicality of the protocols and tools developed for identifying areas of need and for monitoring and tracking the progress of pupils. The process evaluation looked at the training of staff in the proper use of such protocols and the fidelity of implementation of the programme in classroom conditions.

Objectives

The impact evaluation aimed to estimate the effects of the RTI programme on the literacy performance of Year 6 pupils who are at risk of not achieving the expected Level 4 in English at Key Stage 2 (KS2). The process evaluation aimed to provide formative evidence on all aspects of the programme: the initial selection and retention of schools, the subsequent training of staff, through to the fidelity of treatment and testing in schools.

Project team

The programme was developed by the Centre for Use of Research Evidence in Education (CUREE) in collaboration with Achievement for All 3As (AfA3As). CUREE was responsible for developing the instructional materials and the design and conduct of the training; AfA3As provided support, primarily acting as recruiter and as ongoing contact with schools. AfA3As Achievement Coaches worked closely with schools to support their use of RTI within the AfA3As framework.

Methods

Trial

The evaluation approach used here was a randomised controlled trial based on a waiting list design. All volunteer schools would receive the RTI intervention, and all had previously agreed to be randomised to immediate intervention or intervention the following school year. The project involved school-level randomisation of an originally estimated 80 primary schools, using Year 6 pupils after their KS2 assessments. 40 schools would receive RTI immediately, and 40 schools a year later. The developers were confident that they could recruit 80 schools, as many were already working with AfA3As. A pseudo-random number generator was used to select treatment or waiting group status for each school, and the results were revealed after the pre-test for both groups.

A total of 85 schools were recruited initially but 24 then dropped out before randomisation. In general, these schools dropped out when they realised the expectations of the project. Reported issues included the timing, timescale, and uncertainties around being a control group school, while introducing an intervention in the busy final weeks of the summer term for pupils who were about to leave was perceived as a burden to some schools. Some also found the initial online testing an obstacle, especially primary schools with limited IT resources. This level of dropout before allocation, while undesirable and leading to a smaller trial, is not a major concern in terms of the internal bias of the eventual results.

After recruitment of schools, there were three key steps prior to the intervention beginning:

1. Pre-test: all Year 6 pupils in all 61 schools (2,352 pupils in total) sat the New Group Reading Test.
2. Randomisation: 30 schools were allocated to treatment, 31 to control.
3. Identification of eligible pupils: both treatment and control schools were required to identify pupils that were “eligible” (i.e., most suitable) for the intervention.

Schools were required to identify their eligible Year 6 pupils by using the results of the pre-test, teacher judgement, and the RTI diagnostic tool. The identification of these pupils in both treatment and control schools, and the sharing of this information with the evaluator, was important in establishing a valid comparison group prior to the start of the intervention. Ideally, schools would have identified eligible pupils prior to randomisation in order to ensure that the pupils in the control group were identified under the same conditions and timing as those in the intervention group. However, although this sequencing was originally proposed by the evaluator, it was agreed that pupil identification could take place after randomisation due to the short timeframe in which the intervention was to be delivered.

This would not have compromised the trial if the eligible control pupils had been identified early, and this data shared with the evaluator. In practice, however, the evaluator did not receive data on the eligible control pupils until after the delivery of the intervention, and the timing and unreliability of this data meant that the findings of the evaluation were substantially weakened.

Between randomisation and post-test a further 11 schools dropped out (and a further school provided invalid post-test results, involving pupils from a different year group). One school experienced organisational problems and four schools dropped out of the project for similar reasons as above, or because RTI was no longer a priority (perhaps because of changed circumstances or leadership). The timing of drop-out ranged from almost immediately after allocation to refusal to complete the post-test at the end. Of these, three were treatment schools and eight were control. A further control school conducted the post-test for NGRTB but on a different cohort of children. Drop-outs were not promptly

reported to evaluators. All were pushed as far as possible by the evaluators to complete the post-test correctly and so allow a fuller intention-to-treat analysis. This even went so far as to allow schools to complete the test in the following term of the next school year, and particular attention was given to the school that tested the wrong cohort and to one large all-age school where the original Year 6 would still be present as the new Year 7. No results from secondary results were forthcoming but testing of the target pupils in a middle school (control group) was carried out. The chain of communication between schools and evaluators was complex, which contributed to the problems in completing post-tests.

There were a number of changes and late confirmations regarding the number of pupils identified as being in the target group leading up to post-test. The main explanation for this was that, in line with the RTI approach, target pupils were identified only after the pre-test since the test was used as a diagnostic tool. Before administering the tests all schools were required to forward the number of Year 6 pupils and the approximate number of target pupils, however it was not possible for schools to give exact target pupil numbers at that stage since their selection drew on the test results. Schools were then tested and randomised. Schools in both the treatment and control groups were asked to select their target pupils using the same criteria (drawing on the pre-test data). In order to confirm numbers for the end of summer testing, both treatment and control groups were asked to send in their target pupil numbers. In some cases this data had not been received from schools in time for the ordering of tests so their original estimates of target pupil numbers were used. As a result the number of pupils initially identified did not tally with the number of pupils registered for the post-test.

In combination, given the scale of post-allocation non-response (20%) and its prevalence in the control group (29%), plus a lack of clarity about which pupils were initially deemed eligible for the treatment in the control schools, the results of this evaluation must not be read as those of a true randomised controlled trial. They may be interesting or indicative of impact, but cannot be used to provide a secure estimate of impact.

Eligibility

Targeted schools were those with a high level of pupils eligible for free school meals. A standardised test of literacy (GL New Group Reading Test A, NGRTA) was administered to all Year 6 pupils in schools which agreed to be part of the project. Pupils targeted to receive the intervention were meant to be identified in all schools using a combination of teacher's judgement about which child or group of children would benefit from the treatment, coupled with the NGRT data as well as the RTI diagnostic tool. These were meant to be pupils who were at risk of not achieving Level 4. In general, six to eight vulnerable target pupils were to be identified for each Year 6 class.

Letters to schools were sent out by AfA3As to get schools' agreement to take part in the programme and to agree to be randomised to either control or treatment. Letters to parents were also sent out by schools to inform them of the literacy transition programme that the school was involved in.

Intervention

In this project, CUREE piloted the RTI intervention within the pre-existing AfA3As approach. CUREE developed the specialist tools, resources and training plans to support research-based choices about literacy approaches, while AfA3As worked directly with the project schools providing support by participating in the initial training, and on an ongoing basis through in-school coaching.

Randomisation was carried out by the evaluators immediately after schools submitted their pre-test results. Schools were informed of their allocation to either Phase 1 (immediate treatment) or Phase 2 (receiving treatment a year later).

The intervention went through the following stages:

- After allocation, teachers identified the target pupils. Control group schools were asked to identify pupils using the same criteria as intervention group schools. As noted above, the evaluators did not receive any information on which pupils were eligible for treatment until after the post-test. This problem endangered the whole trial.
- Teachers in the intervention group then determined the specific needs and pervasiveness of the problems across key literacy areas (e.g. comprehension, spelling, vocabulary or other areas) for each of the focus students.
- Using this information, teachers established the degree of intensity of delivery. There are three tiers or levels of intensity of delivery. Tier 1 is the whole class approach (designed to involve all pupils in a class), Tier 2 is small group delivery and Tier 3 is a more intensive one to one or paired work. Tiers 2 and 3 form the bulk of the innovation. For this project, determining each of the steps described was supported by a 'Close-case Analysis Tool'.
- Teachers then identified the most appropriate research-based practices to tackle these problems using the project menu of interventions tool.
- Teachers monitored the progress of students by measuring the students' response to instruction. In this project, this is shaped using Progress Tracking Tools.

Each school received regular visits from their Achievement Coach as part of the AfA programme, and three additional RTI-specific visits from the Achievement Coaches (AC).

Prior to implementation of the programme, three days of training were organised for participants. One day was devoted to the training of Achievement Coaches and two for School Champions (SCs), the latter are the school heads, deputy heads SENCOs or literacy heads responsible for overseeing RTI in their schools. Much of the training was conducted using a carefully planned combination of presentation, interactive discussion and an exploration of how to use the diagnostic tools in the specific context of the needs of identified pupils. These activities were debriefed to develop an understanding of the purpose of tools, protocols and coaching. The training included a general introduction to the intervention and the research background of the programme. Teachers were taught how to use the range of tools and protocols for screening pupils for eligibility and for assessing their needs, and how to select appropriate research-based approaches – planning for their use and then using them to monitor progress.

The key elements of RTI include the following:

a) Needs analysis

This involves first identifying the needs of the individual pupils and then identifying the level of intensity. The individual needs of the pupils are assessed using the Close-case Analysis Tool. The tool helps staff to determine the literacy areas to focus on for each pupil, such as phonics, fluency or comprehension.

A series of questions helps teachers to determine the degree of intensity for the intervention. Staff members are invited to discuss each pupil's literacy needs and barriers to progress, using a questioning framework. For example, they consider whether issues have been targeted in the past and whether there are additional factors that are likely to slow the child's progress in a group or other learning environment. Colleagues assess patterns of needs for the group as a whole. They note areas that present a serious issue for particular children – these are potential Tier 1 (whole class) or Tier 2 (small group) intervention areas. Staff members are asked to target an issue at whole class level (Tier

1) if it is identified as a barrier to (nearly) all target children, as this suggests that the whole cohort is likely to need support in this area. Teachers are advised to avoid selecting Tier 2 for children who have already had an intervention targeting the same/related barrier and where their progress was slow or insufficient. If a very small number of target children have a serious issue with a particular area of literacy, and/or the pupils involved have already had interventions targeting this area, or are unlikely to make progress in a group setting, teachers are invited to consider Tier 3 (one to one or paired work). Tier 2 and Tier 3 are not substitutes for the whole class environment, and all pupils attend class in addition to their intervention groups.

b) Targeted intervention

The next level involves making choices about appropriate interventions to use to meet the needs of the target pupils. All are described as research-based instructions with prior evidence of impact. There is a menu of such interventions targeting different literacy areas from which teachers can choose, and they are supported in evaluating other interventions using an 'Evaluating Existing Interventions Tool'. For each intervention on the menu there are suggested levels of intensity or tiers for which the intervention is most appropriate. In summary there are:

- four strategies for promoting reading comprehension, e.g. Peer-Assisted Learning Strategies (PALS)
- learning strategies including word identification, visual imagery, self-questioning and paraphrasing, inference training and repeated reading
- two spelling interventions using paired cued spelling and teaching morphemes
- one grammar intervention using peer-assisted sentence combining
- two paragraph/text structure, e.g. self-talk using writing frameworks and collaborative reasoning
- one fluency intervention through repeated reading
- two vocabulary interventions employing intensive vocabulary instruction and teaching morphemes
- four interventions for phonics using embedded phonics, auditory discrimination in depth, rapid phonics and scholastic.

A set of 'Making Choices' information sheets is provided to enable schools to work with Achievement Coaches or School Champions to decide which interventions are the most appropriate for their pupils. For each aspect of literacy there is a tool to guide teachers in tracking pupils' progress to assess how pupils are responding to the intervention: these are called Tracking Pupil Progress Tools.

c) Monitoring progress

Pupils' progress is monitored at two points throughout the programme, once in mid-intervention and once at the end of the intervention, based on teacher assessments and judgements. The mid-intervention monitoring encourages teachers to consider making adjustments to the intensity of the intervention to meet the needs of the pupils (e.g. to increase the intensity from small group to one to one).

Issues arising in implementation

As the interventions were implemented in the summer term after the KS2 assessments (i.e. from the last week of May) schools did not feel that they had enough time to carry out the intervention as thoroughly as they would have liked to. There were some reports that the intervention was not given sufficient time to develop and to complete the full cycle from needs analysis to monitoring and

adjustments. There were also concerns about the validity of the GL tests: teachers reported that pupils often scored lower on GL assessments than on APP and teacher assessments and SATs. This time constraint was largely due to the somewhat limited timescale of the funding conditions rather than the preferences of the developer or the schools themselves.

Outcomes

One outcome measure was the pupils' performance in the NGRT (New Group Reading Test) which is a standardised test of reading (www.gl-assessment.co.uk). Version A was given as the pre-test and version B as the post-test. This was converted into a Hedges' Effect Size using the pooled standard deviation of both groups. Effect sizes were calculated for post-test only and then for gain scores (to assess the impact of any imbalance in the initial groups). A second outcome was the equivalent comparison of the performance of both groups on the Progress in English Test, again using NGRTA as a kind of pre-test score. All results are presented for the raw-scores in each test. An equivalent analysis was conducted using age-standardised scores for each test, leading to the same substantive findings. The age-standardised results are not presented here.

Further planned analyses included consideration of the impact for sub-groups such as those eligible for free school meals (FSM) or boys only. To achieve this, additional data was collected, largely uploaded from SIMS to the GL test system, creating a record of individual background data plus the pre-test scores linked via unique pupil identifier numbers (UPNs).

The post-test was delivered in schools by staff of each school. Evaluators made sample visits to treatment and control schools to observe the administration of the post-test, and report on any irregularities.

Sample size

The project was planned as a school-level randomisation of a target of 80 primary schools, using Year 6 pupils after their KS2 assessments: 40 schools would be randomly assigned to receive RTI immediately, and 40 schools a year later. Assuming an average of 10 eligible pupils per school, this would provide around 400 pupils per arm of the trial.

Vaughn and Fletcher (2012) reported an effect size (d) for RTI of 0.16 after two tiers working with struggling pupils in Grades 6 to 8. This is quite modest. Using Lehr's approximation for an 80% chance of detecting a presumed effect size of 0.2 with 5% alpha, the minimum sample size needed per arm is 400 individual cases (Gorard 2013). However, this trial involves randomising schools rather than individual pupils. Assuming an inter-cluster correlation (ICC) of 0.2, then the estimated design effect of school level allocation for these 400 individuals will be 2.8. This calculation suggests that approximately 1,120 (2.8 times 400) would be needed. Of course, in reality the ICC could be lower and the effect size higher. Also the correlation between the pre- and post-test scores for individuals will tend to be high which can make the effect somewhat easier to detect. Nevertheless, there is a danger that this trial would not have had full power even if the intended sample had been achieved. In the event, 61 schools were successfully recruited by the developers, and 12 of these dropped out before or at the post-test. The drop from 80 to 61 considerably reduced the a priori power of the trial. However, as noted above, the bias caused by the subsequent loss of three treatment and nine control schools makes considerations of power irrelevant.

Randomisation

The 61 schools with pre-test scores were randomised by the evaluators to Phase 1 (immediate intervention) or Phase 2 (intervention a year later). To ensure blinding and to minimise bias due to knowledge of group allocation, evaluators carried out the randomisation using a mechanically shuffled pack of cards to select the treatment or waiting school, and revealed the result after the pre-test for all schools. There was one card for each school plus one, and an equal number of red and black signifying treatment or control. The last card was black, so the control had one school more than the treatment group. The treatment group had 30 schools.

Analysis

The impact of the trial is represented by the effect size (Hedges' g) for the post–test only results of NGRTB and PiE. It is also assessed by the effect size for gain scores from NGRTA to NGRTB and PiE. Gain scores refer to the average difference in scores between the pre-test and the post-test. This is relatively straightforward for the NGRTA to NGRTB test score difference, but less satisfactory for the NGSTA to PiE test score difference. For the purposes of creating gain scores, for one analysis, the PiE post-test scores were re-scaled to be in the same range as NGRT scores.

The analysis used here is 'intention to treat', meaning that all pupils originally identified as eligible/suitable in both groups of schools (using the pre-test score and whatever other criteria schools applied) should have been post-tested and their outcomes analysed, regardless of the time actually spent on the intervention. In practice, this was hindered by whole schools dropping out of the post-test, and by inconsistent identification of suitable/eligible pupils by control schools. Further analyses include using the same 'effect' sizes for sub-groups such as FSM pupils or boys only.

Two regression models were created – one for each outcome score. The predictors were pupil background characteristics, including sex, ethnicity, first language, SEN and FSM, plus prior attainment. These variables were all entered in one step. The binary variables representing allocation to treatment or control was entered in a second step.

Process evaluation methods

The aim of the process evaluation was to provide formative evidence on all aspects of the intervention from the recruitment and retention of schools, through the training of teachers to evaluating the outcomes. The process evaluation took the form of observation visits to training sessions, and field visits to schools to observe delivery of intervention and post-testing in both treatment and control schools to ensure that no special attention was given to either group that would bias the results. These visits were informal and non-intrusive involving participant observations, face-to-face interviews with pupils, trainers (Achievement Coaches, CUREE/AfA3As trainers) and teaching staff. Pupil interviews were either in pairs or small focus groups. They were loosely structured to get a sense of pupils' perceptions of the programme. Resources and tools used in the programme were also collected. The purpose of these visits was:

- to evaluate the quality of the training delivery and to get a sense of whether teacher participants understood the RTI approach and whether there was any resistance or resentment to the project
- to have an understanding of the intervention and how it should be implemented; this was to help assess the fidelity of treatment in the classroom

- to collect feedback from pupils and teaching staff regarding the effectiveness or otherwise of the programme, the benefits and challenges faced; this would enable the evaluator to make recommendations for improvement in future scale-up, and help identify the key features of success as well as to pre-empt potential problem areas if the intervention is rolled out
- to observe the conduct of post-tests (as teachers are then aware of treatment allocation. This was to ensure that the tests are conducted in reasonable exam conditions and with no special attention given to either treatment or control pupils that would bias the results.

The visits were conducted by trained evaluators and doctoral researchers who were also trained teachers and CRB checked.

Schools were first approached by CUREE to see if they would agree to the evaluation team visiting them. The names of Achievement Coaches were given to the evaluators who then negotiated access to their schools. Five schools were visited. It was the original intention to visit the schools twice, once at the beginning of the programme and once at the end, to get a sense of the progress that children had made, however as many schools only managed to conduct five sessions, it was not always possible to schedule two visits. With the cooperation of Achievement Coaches visits were arranged, and with the agreement of the schools, evaluators also visited the schools to observe the conduct of the post-test. Two additional visits were made to control schools for the sake of comparison.

Post-test visits were deemed necessary as schools then had knowledge of group assignment: the purpose of these visits was thus to check that the tests were conducted properly, and that there was no evidence of any behaviour from the school that might indicate bias. It also helped with understanding the conditions under which the pupils took the test. Experience from previous projects and the pre-test suggested that a number of primary schools had technical difficulties using online tests.

Impact evaluation results

Participants

The participating schools were recruited by AfA3As, many of which were already working with AfA3As. Targeted schools were those with a high level of pupils eligible for free school meals. Letters to schools were sent out by CUREE and AfA3As to get their agreement to take part in the programme and to agree to be randomised to either control or treatment. Letters to parents were also sent out by schools to inform them of the literacy transition programme that the school is involved in. There is no record of a parent not agreeing to participation. In addition, structured conversations with parents of target group pupils were also held with the support of Achievement Coaches prior to the implementation of the programme. The evaluators had no role in this (see above). In total, 91 schools were approached, 85 agreed to take part initially, and 61 continued to take the pre-test. The 31 control schools were more variable in size and test scores than the 30 treatment schools (see Appendix).

After pre-testing and allocation to groups, a further 12 schools were lost, mostly in the control group (Figure 1).

Figure 1. Participant flow diagram – Sample allocation and attrition

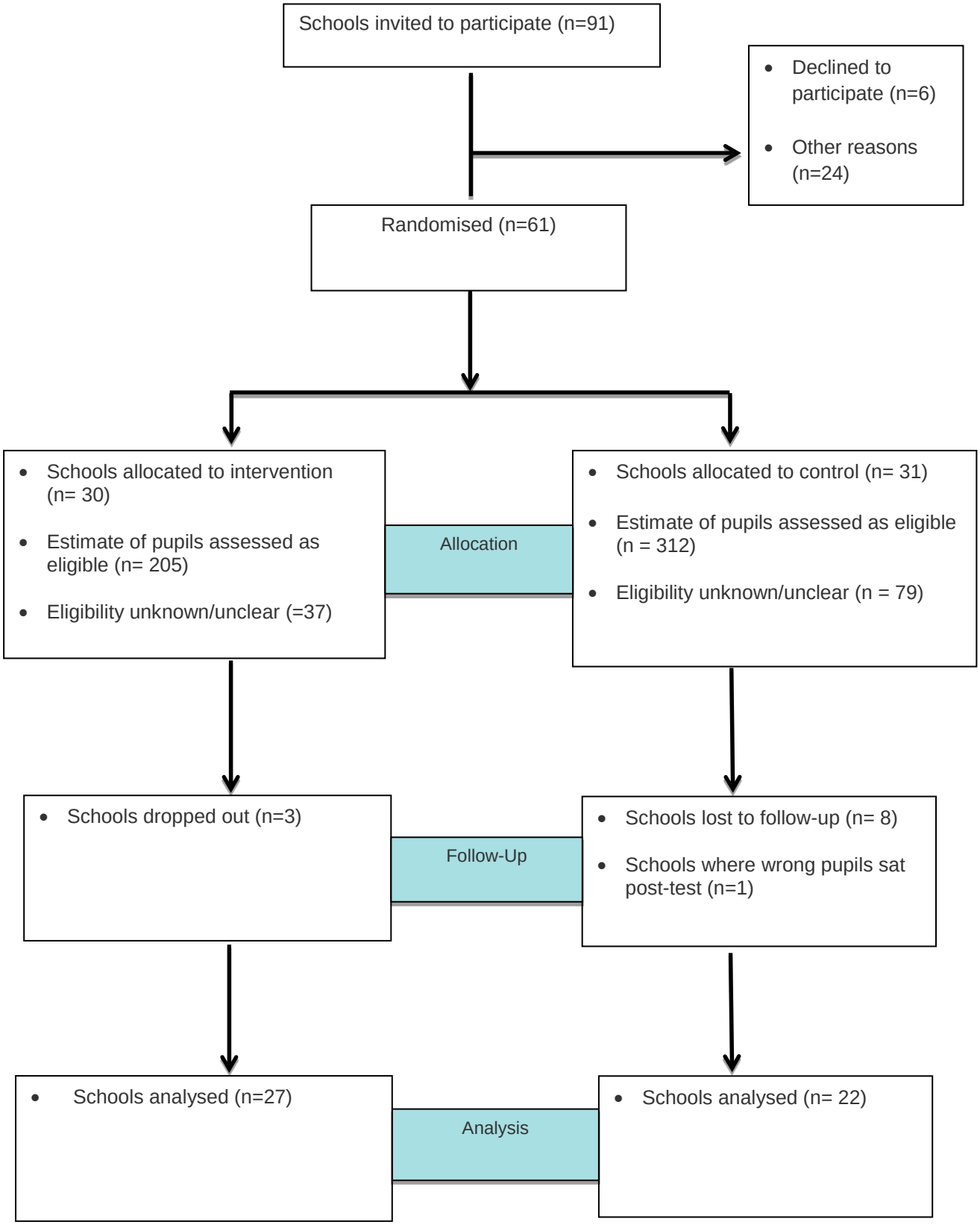

Table 1 shows the pre-test results for all schools dropping out. Again, the control schools are more variable, and the dropouts include the largest school of all, and the schools with the original highest and lowest test scores.

Table 1. Pre-test results for all schools dropping out after allocation

Group allocated	N	Pre-test result
Treatment	37	96.69
Treatment	27	102.19
Treatment	24	106.96
Control	18	80.17
Control	43	87.63
Control	79	89.28
Control	47	97.38
Control	29	97.66
Control	85	99.16
Control	117	103.04
Control	42	105.00
Control	13	109.31
All schools in the study	2,352	97.32

Outcomes and analysis

As noted above, the evaluators were not provided with a complete list of pupils deemed eligible/suitable for the intervention until after the post-test had been conducted. The approximation of eligibility/suitability with the largest number of cases, therefore, comes from a consideration of all of the achieved post-test scores for either post-test measure (on the assumption that the control schools tested only those deemed eligible/suitable). Taken together and at face value, these post-test scores provide no good evidence that the programme was successful. There is a small positive 'effect' size in terms of the New Group Reading Test, and a small negative 'effect' size in terms of the Progress in English Test (Tables 2 and 3). This is the headline finding. Given the problems of control identification and differential dropout it would be unsafe to treat these results as being distinguishable from zero. At best, the headline findings for this programme suggest an impact equating to two months of additional progress in one year. At worst, this programme may have zero impact or even a slight negative impact.

Table 2. Effect size based on all achieved NGRTB post-tests

	N	Mean	Standard Deviation	Effect Size
Intervention	181	287.85	53.02	+0.19
Control	204	276.49	66.98	-
Overall	385	281.83	61.00	-

Table 3. Effect size based on all achieved PiE post-tests

	N	Mean	Standard Deviation	Effect Size
Intervention	131	12.96	5.58	-0.09
Control	123	13.55	7.65	-
Overall	254	13.24	6.66	-

This post-test only analysis is reasonably justified by the initial balance in the pre-test scores between the two groups defined in this way (Tables 4 and 5). Both groups had similar NGRTA pre-test scores, whether their achieved post-test score was NGRTB or PiE. However, the treatment group with post-test scores for NGRTB initially had slightly higher scores for NGRTA (and then a post-test only positive outcome). The control group with post-test scores for PiE initially had slightly higher scores for NGRTA (and then a post-test only negative outcome). This means that the effect sizes taking pre-test scores into account will be slightly nearer zero than the headline figures.

Table 4. Pre-test scores based on all achieved NGRTB post-tests

	N	Mean	Standard Deviation
Intervention	178	264.07	54.00
Control	194	261.36	60.09
Overall	372	262.66	57.19

Note: the N differs slightly from Table 1 as some pupils did not have a pre-test score for a number of reasons (including being newly arrived in the school).

Table 5. Pre-test scores based on all achieved PiE post-tests

	N	Mean	Standard Deviation
Intervention	131	271.12	54.94
Control	121	268.79	60.56
Overall	252	270.06	57.60

Note: PiE was conducted by fewer schools than NGRTB. Several reported not being aware that PiE was part of the post-test and most of these refused to test again.

An effect size based on gain scores from pre- to post-test also gives a similar small positive result for NGRTB (Table 6). There was no specific pre-test for PiE. Using NGRTA as a baseline, the ‘gain’

scores for the rescaled PiE scores are again negative (Table 7). However, this result must be treated with caution as NGRTA is not so clearly a suitable pre-test for PiE as for NGRTB

Table 6. Effect size based on all achieved gain scores – NGRTA pre-test to NGRTB post-test

	N	Mean	Standard Deviation	Effect Size
Intervention	178	25.01	50.21	+0.15
Control	195	18.01	42.01	-
Overall	373	21.35	46.17	-

Note: the N differs slightly from Table 1 as some pupils did not have a pre-test score for a number of reasons (including being newly arrived in the school).

Table 7. Effect size based on all achieved 'gain' scores – NGRTA pre-test to PiE post-test

	N	Mean	Standard Deviation	Effect Size
Intervention	131	-101.78	68.54	-0.16
Control	123	-90.32	71.94	-
Overall	254	-96.23	70.31	-

The original intention had been to look at the difference due to treatment for all cases identified as eligible/suitable for the intervention at the outset. Some schools never provided this information, so this list is incomplete, which means that further cases are lost for both post-tests when data is considered in terms of identified eligibility/suitability. It is not clear why some pupils subsequently reported as ineligible by the same schools were told to take the post-test, and it was never resolved whether any of these cases did in fact receive the intervention despite being reported as ineligible. It is noticeable that some of those pupils reported by schools as eligible/suitable had the highest pre-test scores in their schools, and even across all schools, making them unlikely candidates for RTI. The lists are therefore presumed not accurate. Using these very limited figures, the 'effect' size rises to +0.25 for NGRTB and to +0.02 for PiE (Tables 8 and 9). The issue of power is no longer relevant as these are incomplete samples, but again there must be considerable doubt that the result is safely distinguishable from zero.

Table 8. Effect size based on achieved NGRTB post-tests, designated eligible for treatment

	N	Mean	Standard Deviation	Effect Size
Intervention	177	287.84	53.42	+0.25
Control	183	273.18	65.28	-
Overall	360	280.39	60.11	-

Table 9. Effect size based on achieved PiE, designated eligible for treatment

	N	Mean	Standard Deviation	Effect Size

	N	Mean	Standard Deviation	Effect Size
Intervention	129	12.88	5.58	+0.02
Control	114	12.78	7.18	-
Overall	243	12.84	6.37	-

One further way of identifying any eligible cases is to consider only those graded at Level 4C or below in the pre-test. For this analysis, the evaluators ignored the incomplete list of reported eligibility provided by the developer. This means it is not at all clear whether cases in the 'treatment' group labelled below did in fact receive the treatment. Again, the number of cases drops. Using these limited figures, the 'effect' size rises to +0.29 for NGRTB and to +0.17 for PiE (Tables 10 and 11). Again, it would be unsafe to assume that these results are distinguishable from zero. What the results may show is merely that pupils with initial low scores tended to show greater improvement over time (since it is not known that those labelled 'intervention' here actually were allocated to it by their schools).

Table 10. Effect size based on NGRTB post-test, cases initially below secure Level 4

	N	Mean	Standard Deviation	Effect Size
Intervention	171	286.60	51.32	+0.29
Control	180	270.31	59.47	-
Overall	351	278.25	56.17	-

Table 11. Effect size based on PiE post-test, cases initially below secure Level 4

	N	Mean	Standard Deviation	Effect Size
Intervention	124	12.67	5.29	+0.17
Control	109	11.74	5.75	-
Overall	233	12.24	5.52	-

For interest and completeness, Tables 12 and 13 present the post-test results for the FSM-eligible pupils only. The number of cases, especially for Table 13, is now very small. But if there is any impact there is an indication it may be more important for FSM-eligible pupils who, as a group, also started with somewhat lower pre-test scores on average.

Table 12. Effect size based on achieved NGRTB post-tests, FSM only

	N	Mean	Standard Deviation	Effect Size
Intervention	44	284.64	70.48	+0.48
Control	52	251.31	68.29	-
Overall	96	266.58	70.92	-

Table 13. Effect size based on achieved PiE post-tests, FSM only

	N	Mean	Standard Deviation	Effect Size
Intervention	16	14.25	5.70	+0.43
Control	30	11.50	6.52	-
Overall	46	12.46	6.32	-

Tables 14 and 15 present the post-test results for boys only. Again, both outcomes are positive for the intervention group.

Table 14. Effect size based on achieved NGRTB post-tests, boys only

	N	Mean	Standard Deviation	Effect Size
Intervention	97	281.69	52.27	+0.26
Control	113	265.56	69.31	-
Overall	210	273.01	62.40	-

Table 15. Effect size based on achieved PiE post-tests, boys only

	N	Mean	Standard Deviation	Effect Size
Intervention	73	12.81	5.75	+0.11
Control	68	12.07	7.49	-
Overall	141	12.45	6.63	-

Table 16 presents the R values for two regression models, each based on two steps. The first model is used to try and explain variation in the NGRTB outcomes, and the second in the PiE outcomes. In Step 1, the pupil background data and NGRTA pre-test scores are included. Step 2 adds the binary variable for being in the treatment group or control. For both models the bulk of the variation that is 'explained' by the variables in the model is explained at Step 1. Once pupil background and prior attainment is accounted for, very little difference is made by knowing whether a pupil was in the treatment group or not. The model for NGRTB is substantially better, probably because NGRTA is a

better predictor of subsequent NGRTB results than of PiE results. These models are not, in themselves, any test of causation but do provide a caution about the strength and importance of the intervention in relation to prior pupil characteristics.

Table 16. Variation explained (adjusted R-squared) in two-stage regression model, using two possible outcomes

	NGRTB outcome scores	Progress in English outcome scores
Step 1 – background and prior attainment	0.48	0.34
Step 2 – intervention or not	0.48	0.34

Table 17 presents the standardised coefficients for all variables retained in either model. The largest of these by some way is the pre-test score for both models. This is the best single predictor of the post-test score. As would be expected from the analysis so far, there is slight positive standardised coefficient for being in the treatment group, as assessed by NGRTB (+0.05) and slight negative one assessed by PiE (-0.06).

Table 17. Standardised coefficients for the regression model in Table 15

	NGRTB outcome scores	Progress in English outcome scores
Step 1		
FSM	+0.03	+0.00
Sex (female)	+0.06	+0.04
SEN	-0.02	-0.02
EAL	-0.05	+0.06
Ethnicity (White UK)	+0.05	+0.00
NGRTA score (pre-test)	+0.65	+0.60
Step 2		
Treatment (or not)	+0.05	-0.06

No measure of actual pupil attendance (i.e. dosage) was provided consistently by all schools via the developers. The developers collated some data on the planned number of sessions in each school (see Appendix).

Cost

As with any intervention, the cost of setting up RTI in a new school differs slightly from the cost of delivering it annually to each new batch of students. For set-up it is assumed that two members of staff will attend initial training for two days each, and as here there will be one and a half further days of meetings with ACs or similar. If each day costs £200 for supply cover, the total per school is £700. The actual cost of delivering training was £322 per school. Ongoing monitoring and support by ACs cost £428. Printing and stationery costs for resources are estimated at £300. This gives a total per

school of £1,750. If there were 15 eligible pupils in each school, this would amount to a cost of just over £117 per pupil, rising to £175 if only 10 pupils were eligible. In addition, the running costs would include any extra staff needed to meet the requirements of Tiers 2 and 3.

Process evaluation results

Implementation

The programme began with one day's training for Achievement Coaches. These were staff of AfA3As specially trained to work with teachers and to support schools in their programme to improve the attainment of low-achieving pupils (i.e. beyond RTI) and had already been working with AfA3As schools. A two-day training session was then conducted for representatives of the schools randomised to treatment. School representatives were known as School Champions. These consisted of heads, deputy heads, heads of SEN or literacy who would either have to train their staff to implement the intervention or who would implement the intervention themselves.

Training delivery

The training was well delivered. Participants were guided through the stages of RTI, including the use of the protocols, and given opportunities to explore and discuss the use of these. Participants received a transition starter pack with suggested strategies for each component of reading and writing.

Several participants were apprehensive about RTI at the start of the training, but became more receptive to the idea towards the end of the second day. The teachers interviewed were all positive and expressed excitement about the intervention. Some were initially sceptical about the suggestion of using the Close-case Needs Analysis to identify individual pupil needs. Initially some teachers felt that this stage was not necessary as they would already know the needs of their pupils. But feedback from schools at the end of the process showed that they found the Close-case Analysis very useful, demonstrating the value of more in-depth diagnostic work. It was also suggested that schools may use APP (Assessing Pupil Progress), but some of the participants felt that there was strong resistance from teachers against this.

Programme delivery

Observation visits were made to five schools (A to E). As schools were very busy in the last few weeks of term, there was sometimes only one observation visit per school.

In School B, ten students were on the programme. Three types of interventions were selected:

- Repeated Reading strategy was used for two students most in need using a one to one approach.
- Peer-Assisted Learning Strategies (PALS) for the ten students in a whole class approach, with mixed ability pairing.
- Pair cued spelling for 10 students, also in a whole class approach, pairing the least able with the most able.

The students were split into three groups. Two of the students were identified for the most intensive one to one intervention. They worked with the teaching assistant on the repeated reading intervention. The rest of the class was divided into two groups. One half of the class did PALS working on inference, the other half did the pair cued spelling. The intervention lasted half an hour at a time.

In School E, there were 14 pupils on the programme taken by two teachers. The school decided on the small group approach, with seven in each group. Although the school planned to conduct the

intervention every day, they only managed to have two or three sessions every week. This was partly due to staffing constraints as the small group intervention was taken by the two regular literacy teachers during literacy lessons in separate rooms. The rest of the class was taken by an experienced and trained teaching assistant

The intervention started with identification of target pupils and their needs. This was carried out by the deputy head, who was also the School Champion, using a combination of the Close-case Analysis Tool and the GL assessment as well as teacher assessment of the pupils. Each session lasted 20 minutes to half an hour. Children were taken out of their classes during literacy sessions. These were pupils who were at Level 3 but below Level 4: those below Level 3 were not eligible for the programme but were supported by a higher level teaching assistant. The literacy component targeted was *inference* as the teachers felt that this would enhance the pupils' writing and comprehension. Both teachers used the six steps suggested for teaching sequence in inference. These were:

- Consider prior knowledge
- Define and elaborate on words
- Ask questions
- Fill in missing sentence
- Create images of their reading
- Summarise and predict.

The lesson was carried out over two sessions as they had only allocated 20 minutes to the intervention per day instead of the recommended 40 minutes.

In School A – adopting a holistic approach – RTI was delivered across all subject areas, not just in the literacy hour: it was introduced in Art, Maths, Literacy and Creative Writing. Three intervention strategies were selected to address the needs of the pupils. These were PALS, comprehension and phonics. Different tiers were employed for different strategies. The lesson observed was the comprehension activity based on the book *Kensuke's Kingdom* by Michael Morpurgo, a story particularly relevant to Year 6 pupils' experience of transition from primary to secondary school.

The observed lesson was well structured and there appeared to be many activities that attempted to engage children in finding their own answers. At the same time there was lots of scaffolding from the teacher. For example, pupils had opportunities to explore the text, examining the sentences, paragraphs and appropriate sentence fillers needed to make the text cohesive and comprehensible. The teacher then asked pupils to underline words that they did not understand. She clarified the meaning of the words, exploring these words further through questioning. The sentence-filling activity consisted of children making sentences in the gaps in a passage, and the teacher checking pupils' spelling, grammar and punctuation using the Literacy Board. Each pupil then read out the paragraph they had made and corrected their own mistakes with guidance from the teacher. It was clear that some pupils were struggling with commas as they could not find the connectives in compound sentences. The exercise enables the teacher to identify this issue. The teacher then got them to break down the story down into short sentences and summarise it.

In School C, the school had only started the intervention 10 days prior to the first visit. The number of sessions conducted was five at the time of visit, which took place two months after the training. The reason given was the number of transition activities in that term. The intervention was delivered up to four times a week and each session lasted 20 to 25 minutes. The teacher, together with the Achievement Coach, decided on two interventions. One was the small group intervention focusing on reading comprehension because it was felt that reading comprehension was a particular problem. The suggested activity was the use of drawings to teach comprehension. Children read a passage and then drew what they thought the passage was about. The teacher chose descriptive texts with lots of visual images and also texts that were relevant and interesting to the pupils. Using questioning the teacher checked pupils' understanding of the passage. Children also made predictions from the text. There were discussions about what images certain words conjure. For example, words like 'jostling', 'perplexed' and 'rabbit warren'. Pupils were taught strategies for reading using the mnemonics: SCENE ASK IT. For example:

Search for picture words
Create or change scene
Enter lots of details
Name parts
Evaluate picture

Ask a question
Say some questions
Keep prediction in mind

Identify answers
Talk about the answer

The teacher also implemented a whole-class intervention for vocabulary and grammar.

In all schools there were already a number of other literacy interventions going on. In some schools these were suspended when RTI was introduced, in others it was not clear. For example, in one school the pupils were exposed to other literacy interventions, such as Rapid Reading, although this was suspended when they started on RTI. The children told us that the school also had other reading programmes like ERIC (Everyday Reading in Class) where the class read silently for 15 minutes. They also had Guided Reading where they read with the teacher in class. Children were placed in different reading groups colour-coded by ability where they read books appropriate for their level. This was a weekly session. On other days they did different topics which might, for example, be reading or comprehension. It was not clear, however, if these other programmes were still running or whether the children were talking about the programme that they had been on early in the year. If the former, there may be a risk of diffusion.

Conditions for successful implementation

Delivering RTI depends on a number of considerations:

- The quality of training enabling teachers to use the tools and protocols as they are intended.
- Teachers' competence and ability to use the protocol/tools for (a) identifying the area of focus/needs, (b) determining the degree of intensity and (c) in tracking and monitoring the progress of pupils to adjust the intensity.
- A sufficient duration for the intervention to allow for the full cycle of RTI to be observed, from identifying target pupils to determining the level or intensity of intervention and implementation. Time is needed to allow the intervention to develop before teachers can track, monitor the progress of the pupils and adjust the intensity. Elapsed time is therefore essential for a proper round of tracking and monitoring.
- Whether there is sufficient time for teachers to plan lessons, and for the preparation, collection and development of instructional materials.
- The availability, accessibility and suitability of materials for teaching the different components of reading and writing. Many teachers shared the same view that it was time-consuming to look for appropriate and interesting materials. However, once developed these resources could be used in subsequent years. It would be a good idea if all the schools could come together and share these resources and talk about successful lessons. In future a library or bank of resources could be developed although supporting access to the resources would need to be carefully considered.
- The commitment of teachers and school as a whole to the intervention. It is essential that teachers have a positive attitude and are willing to adapt to changes in routines. The successful implementation of intervention relies on the general support given by the head and the ability of schools to use the data (NGRT pre-test data, teacher assessment data and SATs data) to inform decisions. It was very apparent, for example, in School B, how the

achievement data was feeding into the general work and the interventions that were already in place.

Challenges to implementation

- The biggest challenge to successful implementation was the very short duration given to the intervention. This was primarily due to the funding requirement that it should be implemented in the last few weeks of the summer term. This concern came up regularly in interviews with teachers and achievement coaches. School C planned to have a session every day but only managed two to three sessions per week. In School B, they planned to have five sessions per week, but only managed three. This was not deemed long enough for the impact of the intervention to be realised.
- The short duration of implementation also meant that teachers could not go through the complete process from needs analysis, to implementation of intervention and tracking and monitoring. The last two stages have been given the least attention. With only five sessions altogether, teachers could not realistically monitor and make adjustments in the middle of the intervention as this would have meant teachers having to make such assessments by the second session. For example, one teacher said there was *'not enough time to do it properly*.' If the intervention was carried out on a daily basis over a 4-week period, it might be possible to monitor and track pupil progress. It is not clear that all teachers managed to use the tracking and monitoring protocol to adjust the intensity of the intervention.
- The timing of the intervention was less than ideal. This was an issue raised by all the five schools visited. As the intervention was implemented in the last few weeks of term, teachers had to compete with end-of-term activities, such as sports days, prize-giving days and rehearsals for end-of-year productions. Year 6 pupils also had a number of transition programmes lined up, for example, visits to secondary schools, school camps and secondary school integration programmes. In one primary school the pupils spent a week in their secondary school, thus losing a whole week of instruction. Almost all the teachers and Achievement Coaches spoken to felt that the programme would have had greater impact if it had started in Year 5 or at the beginning of the year to allow time for adjustments and for the intervention to develop.
- Schools reported a high rate of absence in the last term after SATs. There had been lots of absences due to transition visits and sports day. Consequently, monitoring the impact of the intervention had been difficult.
- Finding appropriate short texts for teaching was also a challenge, and a time-consuming activity. There was also a need for flexibility in considering what the teachers and what the children want to do.
- Another challenge was the time needed for the literacy teacher to plan the lessons, especially where there was more than one tier of intervention. Even if there was only one tier, teachers still had more than one preparation for each lesson. In one school the teacher had three preparations: one for the whole class, one for the RTI pupils and one for the weaker pupils, in other words, for each session the teacher effectively had to prepare for three lessons. So each literacy period can involve two to three sets of materials and instructional strategies. Teachers appeared overwhelmed to begin with, but many seemed to manage it well. The head teachers and deputy heads in the schools seemed very supportive. Teachers were given the necessary support in terms of use of rooms and teaching assistants. In School E, the deputy head herself took on the role of identifying the targeted pupils and in organising the testing. In School A, the teacher remarked that she was overwhelmed with paper work. In schools where the deputy head took the responsibility of organising the RTI tests, testing went smoothly. In schools where this was done by adjunct staff, there were several issues with testing, most common of which was the misplaced password.
- The toolkit of interventions was relatively easy to use, but some teachers noted that the Close-case Analysis and Monitoring Tool might be a challenge if the classes were large and if the number of pupils needing help was large. For this reason, project guidance highlighted the importance of a small target group of pupils.

- Finding space for different groups of children receiving different intensities of treatment within one period can be a barrier. Some schools may have decided to have only one tier because of the logistics of planning and finding rooms to accommodate the intervention.
- Disruption during the lesson. Given that the intervention period had been shortened quite substantially, it was important that every lesson was conducted as intensively as possible. In one school, the teacher was called out twice during the lesson to attend to another pupil with behavioural problems. This was disruptive to children receiving the intervention.

Attractiveness of intervention to stakeholders

Despite the issues with time, space and the logistics of planning, the teachers, pupils and Achievement Coaches were positive about the intervention. This was evidenced in the way schools were willing to work out the complex logistics of arranging classes, with some pupils receiving small group intervention and others one to one attention. Additional staff were enlisted for this. One school asked a higher-level TA to deliver one to one tuition, and in another school they had a TA take the whole class (these were the more able pupils), while the regular literacy teachers conducted the small group intervention. Schools made special efforts to find the space for these different groups which is a testimony to their enthusiasm.

What do teachers like about the programme?

There was some apprehension among teachers to begin with, but many soon found that they actually liked the programme. One teacher in School B, for example, said they initially thought the intervention was no different to others the school had tried before, but he quickly realised how successful these tasks were. He admitted that his initial scepticism of the intervention was misplaced**.**

- Teachers and Achievement Coaches all found the Close-case Analysis useful. One teacher remarked that the way the Close-case Analysis results were recorded was very clever. He liked its simplicity and said he would definitely be using it again, remarking that the Close-case Analysis also prompted teachers to think about how to link the different roles of TAs and SENCOs. He liked the student sheets and intended to use something very similar, however he did note that they were hard to print, especially the initial data in Excel.
- One school also particularly liked the Close-case Analysis which enabled them to pin-point the pervasive problems with their Year 6. For example, knowing that comprehension was a major problem, they intend to focus on this aspect in earlier years before the children reach Year 6.
- Some teachers liked the suggested activities which also gave them the impetus to generate more interesting lessons. In School B the teacher liked the idea of pairing the middle ability children with the higher and lower ability children. He found this worked well with the PALS activity, but pairing the highest with the lowest ability child worked better for the pair-cued spelling intervention because in the '*middle there wasn't much of a difference between the students, therefore it was harder to make progress there*'. In the past he used to pair students randomly. He said with RTI he now understood the advantage of strategic pairing. However, there were disadvantages with pairing the more able with the less able in that the more able pupils did not feel challenged. Nevertheless, the teacher felt that all the pupils got something out of it.
- Some teachers liked the online test. One school suggested that they would use such a test again if they had the budget for it. The test enabled them to diagnose the problem areas, so they now knew, for example, that comprehension was a major problem for Year 6 children, and became aware of the need to target this issue in earlier years.

What do pupils like about the programme?

- Pupils particularly liked the small group interaction. They liked the fact that they could work among pupils of similar levels to themselves. This meant that they did not have to worry about making mistakes in front of their peers. They also liked the fact that they had more attention from their teachers. When asked what they liked about the programme, some pupils said it was the '*small group', 'better interaction', 'not afraid of making mistakes', 'being with people of the same level'* and *'can concentrate more*'.
- Pupils liked the fact that they were learning more about spelling and writing. They also said that they enjoyed learning to write better sentences and use more powerful words. Before they were not using the best words. They also said they now enjoyed work more.
- Pupils also reported that they liked the activities, for example pupils in School A said they liked the use of mirrors to see how their mouths shape words to help them learn how to pronounce and spell words.
- Pupils liked the fact the there was fun in learning:

'Helped me because you do fun stuff mixed with learning.'

When asked if there was anything they did not like about RTI, all the pupils interviewed said there was nothing that they did not like, although they would like to be able to take the books home to read.

Another child said they would like to keep reading in bigger chunks rather than in small bits.

Outcomes

Development of new teaching resources and teaching strategies

One important outcome was the development of new teaching resources. Teachers also showed innovation in their teaching strategies, using interesting and age-appropriate texts and props.

Reported improvement in pupils' perceived learning

Teachers, Achievement Coaches and head teachers/school leaders reported improvements in pupils' learning. In School B, for example, the teacher claimed that pupils had made an equivalent of five months' progress in comprehension and spelling in the four weeks. In some cases, pupils made a year's progress. These claims were based on the data collected by the school. Two students in particular went from reading 200 words in three minutes to nearly half of that time. This has significantly boosted pupils' confidence. Both the teacher and head teacher felt that the four weeks had made a massive impact on their pupils' reading. At their recent assessment, all their targeted pupils achieved Level 4. According to them:

'This has never happened in the history of the school before.'

However, teachers also added that such progress could be due to the increased attention given.

Pupils also remarked on how the activities had helped them. For example:

'I remember things that I learn now.'

'I can do a lot more – better with spelling, reading and writing.'

The programme also helped develop independent learning. Pupils said they were taught to mark their own work and improve. When asked if it was the teacher or the book that helped them learn, these were their answers:

'Both – books good, teacher helped.'

'Both – teacher helped me 99% and the book 1%.'

'Both – books to read more and the teacher made it fun.'

Pupils were now more confident in their ability. They said that if they were to do the GL assessment again they would be able to answer some of the questions which they could not before. Pupils expressed no reservation or resentment about being taken out of the classroom for lessons. This was a common practice in the school. Teachers also did not consider this an issue as the top set of pupils was also taken out of classes. Different pupils were taken out for different subjects.

Reported improvement in reading

A number of schools reported improvements in reading. One pupil said she used to hate reading but that this had changed recently. Some pupils said they now preferred books to films as they could use their imagination and they were eager to find out what happened in the books. Some claimed that they enjoyed reading more this year and were reading more books. Below is what some pupils thought of the reading intervention:

'I would give 10 out of 10 for PALS. 5 out of 5 for the Kenusuke's Kingdom.'

'Others won't laugh at me because I can read better.'

The teacher in one school attributed the marked progress made by the two pupils on the programme to the repeated reading intervention. The two pupils were now able to read faster achieving better results with each new text. On the day of the observation, one read the text in 1.4 minutes, aiming for 1.2 the following day. The other showed improvement in spelling, which boosted his self-esteem.

Reported improvement in spelling

One teacher reported how pupils had made such a big improvement in their spelling that he had had to switch to 'harder' spelling lists. In School A, the teacher gave an account of how one of the pupils, who could not speak properly, had no concept of decoding at all and was unable to spell properly, but had shown remarkable improvement in spelling using the phonic intervention activity despite the short duration of the intervention.

'Now if I don't know a word I slow down and think.'

The pupils explained how the use of mirrors helped them to see how their mouths shaped the words and how seeing the shapes of the word had added 'texture' and that when they read they could feel the word. This, they said, helped them with their spelling as they could see all the letters when mouthing the words.

'With a tricky word, now I say it and I write sounds down.'

In another school, a pupil reported how the RTI strategy had helped him learn to spell, for example he realised that to spell the word 'sweets' was only a matter of changing one letter of the word 'tweets'.

Reported improvement in comprehension

In School B, the teacher and Achievement Coach reported a dramatic improvement in pupils' comprehension even within a short time-frame.

Reported improvement in writing

In School E, pupils said they enjoyed the inference intervention, and that their writing had improved. They were now able to write in a more interesting and descriptive way. When they looked back at their Year 5 writings they could see how they could have done better. It is unclear to what extent this suggested change/awareness was due to this intervention or to others. Pupils said that feedback from their teachers had suggested that their writing had improved. For example:

'The teacher said my writing is better this year.'

'I have more comments from my teacher like "well done".'

Pupils also said that the inference intervention had enabled them to understand what they read in story books. Their writing had improved through the use of inferences. Pupils also reported that the intervention gave them more ideas and helped them to be more imaginative in their writing.

School B pupils reported improvements in the use of better expressions and a wider vocabulary, which had helped them with their writing. Similar sentiments were expressed by pupils in other schools.

Improvement in wider outcomes

Improvements in self-esteem and confidence were reported in all schools observed. Beyond the actual outcomes, it seems that the intervention empowered some of the students which has had an impact on pupils' self-esteem.

However, it became clear that when teachers talked about the impact of RTI some were actually referring to AfA3As. For example, one teacher told us that a parent was so confident in what AfA was doing they said the school could do anything with their child.

General increase in teachers' enthusiasm

Teachers and head teachers/deputy heads said they could see the changes in the pupils and were excited about the programme. One school perceived the intervention as such a success that the teacher and head teacher were considering a home version of these programmes to monitor further results at Year 7. They had discussed this with the pupils' secondary school. Most teachers said they would definitely continue with the intervention the following year: one teacher, for example, said she was going to use the interventions with her new class the following September, planning to use the comprehension, sentence-combining and phonics interventions.

The head teacher of School B was so confident in the intervention, having seen the results it had produced among their Year 6 pupils, that she said she would roll out part of RTI to other years. In School A, the teacher felt that the impact she saw was very encouraging and had already prepared to introduce the intervention to her new Year 6 class in September 2013 focusing on comprehension, sentence-combining and phonics. They liked the intervention so much that they had asked if a similar programme was available for numeracy. They also said they would like to buy the NGRT test for other years and use it as a diagnostic tool.

Fidelity

In the schools visited, there was evidence that teachers were using the needs analysis tools in identifying target pupils and their needs. They also adhered quite closely to the instructional strategies suggested in the protocols. Some schools managed to implement all three tiers but on different days, while some managed only one. The major issue was with the number of sessions and their duration. Some schools were able to implement only five sessions of 20 to 30 minutes each in total. This was too short for proper monitoring and tracking and adjustments to intensity. The intensity and duration of the trial is not deemed to be sufficient for impact to be realised. This was largely due to the timing of the programme being introduced in the last week of term after SATs in line with the funding constraints. Many teachers in the schools observed were not able to conduct as many sessions as they had intended. It is also not clear whether every teacher used the monitoring and progress tracking tools. None of the schools observed had completed enough sessions for this. However, 21 schools posted completed tools to CUREE, and these showed variations in the intensity of use which seemed to correspond to teacher observed effectiveness of outcomes for pupils. The tools that were used in more detail and depth highlighted high levels of progress in a variety of areas, including pupil confidence, and some of these schools reported increases in spelling and reading ages of 9 to 10 months.

There is also a potential issue with diffusion in that control schools in the same cluster, under the same Achievement Coaches, may have had access to the RTI resources despite Achievement Coaches being given clear guidance not to share the resources with other schools. In schools visited it was clear that Achievement Coaches did not share these resources and RTI tools with control schools.

At the initial training, Achievement Coaches highlighted issues arising from exposure to numerous interventions. They explained that because they were doing a number of interventions, it was not always clear which intervention or combination of interventions had the impact. There was also the possibility of a Hawthorne Effect. Interviews with pupils suggest that some schools were already involved in a number of interventions prior to their SATs. In some but not all schools these activities were suspended when the RTI was introduced. Also control schools may have had similar programmes going on at the same time. These were not controlled for.

As noted above, it was difficult to get schools to provide the names of eligible pupils via the developers. Initial data used to determine pupil eligibility for the treatment only started to be assembled long after the post-test. This made it difficult to analyse the results as intended.

Formative findings

It is clear that the programme, if used, should be introduced at the beginning of a school year, running through the whole year to allow the full cycle of the intervention to be implemented.

The intervention lessons should ideally be carried out on a more frequent and regular basis.

Control group activity

Control schools continued as per normal including any interventions that they may have already been undertaking.

Feedback on testing

At the initial training, Achievement Coaches and School Champions expressed concern over the discrepancy between GL tests (often lower scores) and teacher assessment (using APP) and teacher assessment SATs-type tests. Although this might not have affected which pupils received the intervention, it had implications on the needs identified for the pupils.

Almost all of the schools claimed that the GL assessment was more difficult than the SATs. In one primary school, almost all pupils were of the view that the GL test was hard – harder than their SATs – however there were instances where pupils performed better on GL assessment than the SATs, for example one pupil achieved a reading level equivalent to a 17 year-old, something that was not evident in his SATs. He was expected to get a Level 6 in SATs, but in the event he only obtained a Level 5a.

Teachers were also concerned that the children were not familiar with online testing: in some cases teachers had to go round telling pupils what they were supposed to do, and some claimed that some pupils were simply clicking at random and hitting the return button. One pupil, for example, just ticked the boxes without bothering to read the words, and as a result performed worse than she would have done normally. One of the trainers suggested that NGRT might not be the right test, but it had been chosen because similar assessments are used across all EEF projects.

Participants also suggested giving pupils more time to practise before the actual test because the children had no idea what was happening. Teachers felt that the practice session that was built in was not enough.

A concern was also raised about children with dyslexia – with black words on white screen it was difficult for children to read. A request was made to use the pen and paper version in any future studies.

Because the NGRT is adaptive, pupils felt that it started at too high a level for them although it soon adjusted to their level. They would have preferred if the questions were easier to begin with. Similar observations were made in other schools. A teacher and the Achievement Coach commented on how pupils tended to score much lower on the NGRT test than the teacher assessment. They gave an example of pupils scoring Level 3A on NGRT, but close to Level 5 on teacher assessment.

One Achievement Coach observed that some pupils with learning needs performed better on online tests, achieving a much higher grade that the pen and paper test. He gave the example of a pupil with SEN scoring Level 6 on the NGRT test but graded as Level 3a on teacher assessment.

In School B, the teacher found that the test was ‘rushed’ and that they ‘were not entirely prepared for the technicalities’. Some students said they hated the test.

The developers supported schools during the online testing, and during this process gathered feedback about how it could be developed to ensure that schools and pupils were not lost through the testing process. Areas for development are: ease of administration, support, access to specific tests, access to paper tests and running of the tests. For future evaluations the following points could be noted:

- Whilst the tests were relatively easy to administer, the schools experienced a lot of difficulties. Primary schools do not have a great deal of ICT support and class teachers may find the set-up complicated.
- One school mistakenly took the wrong test paper. The online system should be set up so that schools can only access the correct test.
- When ICT systems fail, it is important to have paper tests sent through quickly.
- Some schools highlighted concerns about tests freezing and computers logging pupils out. One school expressed concerns that the results did not accurately reflect pupils' performance as a consequence of this.

Conclusions and implications

Limitations of study

The limitations of this study have been made clear. The level of school dropout post-allocation (20%) and the incompleteness of lists of eligible pupils, especially for the control group, mean that the trial findings, as a trial, are not safe. Issues such as measurement error and generalisability are therefore largely irrelevant.

There were five parties involved in the study: the EEF, the evaluators, the developers (CUREE), AFA3As and the schools. Ideally future studies will not have five parties as this added an additional layer of communication between the evaluators and the schools which was not helpful in conducting the trial. AFA3As Achievement Coaches supported both control and treatment schools, and this was also a possible route for diffusion of the intervention.

Interpretation

Unfortunately this study adds little to knowledge of the impact of RTI. The headline results show no clear benefit, even though the problems encountered in the study make the findings, as a trial, unsafe. The work conducted by CUREE/AfA3As has helped considerably to develop ideas and resources for the implementation of RTI as a catch-up intervention in the UK, and this, combined with the process evaluation, has shown that more time is needed for successful implementation than allowed by the short and busy period from KS2 assessments to the end of the summer term.

Future research and publication

Given the poor quality of the sample achieved post-test here, and the ambiguity over who was initially eligible for the treatment, this study does not yield any clear questions for future research, other than whether the intervention works as intended. In the near future, the evaluators may prepare a summary paper describing the evaluation for a peer-reviewed journal, or use the data in a paper discussing the results of several trials.

References

DfEE (2000). Transforming Key Stage 3: National pilot, Sheffield.

Faggella-Luby, M. and Wardwell, M. (2011). RTI in a Middle School: Findings and Practical Implications of a Tier 2 Reading Comprehension Study, Learning Disability Quarterly, Vol. 34, No. 1, pp. 35-49.

Fletcher, J., Coulter, W. A., Reschly, D. J., and Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. Annals of Dyslexia, Vol. 54, No. 2, pp. 304–331.

Fuchs, L.S., Fuchs, D., Compton, D.L. (2010). Rethinking Response to Intervention at Middle and High School, School Psychology Review, Vol. 39, No. 1, pp. 22-28.

Galton, M., Gray, J. and Ruddock, J. (1999). The impact of school transitions and transfers on pupil progress and attainment. DfEE Research Report No. 131. Norwich: HM's Stationery Office.

Goodwin, A. (ed.) (2002). *Improving literacy at KS2 and KS3*. London: Sage.

Graves, A.W., Brandon, R., Duesbery, L. McIntosh, A. and Pyle, N.B. (2011). The Effects of Tier 2 Literacy Instruction in Sixth Grade: Toward the Development of a Response-to-Intervention Model in Middle School, Learning Disability Quarterly, Vol. 34, No. 1), pp. 73-86 (full paper not available, analysis based on abstracts).

Hargreaves, A. and Galton, M. (eds) (1999). *Moving from the Primary Classroom: 20 Years on*, London, Routledge.

Leroux, A., Vaughn, S., Roberts, G. and Fletcher, J. (2011). Findings from a Three Year Treatment within a Response to Intervention Framework for Students in Grades 6 with Reading Difficulties, Paper presented at the Society for Research on Educational Effectiveness Conference, (no details about place and date of conference) http://www.eric.ed.gov/PDFS/ED518866.pdf.

National Reading Panel (2000). Report of the national reading panel: Teaching children to read. Washington, DC: National Institute for Literacy.

See, B.H., Gorard, S., Cooke, S. and Siddiqui, N. (2012). Improving literacy in the transition period: What do we need to know about what works?, London: Educational Endowment Foundation, 57 pages.

Snow, C., Burns, S., and Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academies Press.

Vaughn, S. and Fletcher, J.M. (2012). Response to intervention with secondary school students with reading difficulties, Journal of Learning Disabilities, Vol. 45, No. 3, pp. 244-256.

Vaughn, S., Cirino, P. T., Wanzek, J., Wexler, J., Fletcher, J. M. and Denton, C. D. (2010). Response to intervention for middle school students with reading difficulties: Effects of a primary and secondary intervention. School Psychology Review, Vol. 39, pp. 3-21.

Appendix

Table A1. General characteristics of intervention schools

Type of school	Age range	No. pupils	% boys	%SEN*	%EAL	%FSM	%Ethnic minority	KS2 Av point scores 2012
Community	3 to 11	214	49.5	1.9	1.9	5.3	1.2	30.4
Community	3 to 11	245	49	2	4.2	11	2.2	29.6
Community	4 to 11	262	49.6	9.2	21.7	11.5	3.2	27.1
Community	4 to 11	408	49.3	8.3	1.1	18.4	5.1	28.2
Community	4 to 11	246	49.2	13	7.6	8.9	13.8	31.6
Community	4 to 11	199	47.7	6.5	5.9	20.1	15.3	28.2
Academy converter	4 to 11	214	54.7	4.7	4	30.9	7.7	28.7
Community	4 to 11	126	57.1	8.7	5.7	15.1	7.1	27.4
Community	3 to 11	456	54.2	11.8	15.6	21.3	23.7	26.3
Academy sponsor	3 to 11	264	50.8	14	1.6	54.3	7.3	26
Community	4 to 11	282	53.9	2.5	7.7	18.8	12.1	28.8
Community	3 to 11	216	54.6	4.2	11.6	34.1	12.1	26.3
Community	4 to 11	204	53.4	14.7	5.7	19.1	13.5	26.5
Community	3 to 11	266	45.9	4.1	24	41.9	32.6	25.8
Community	3 to 11	471	50.7	10.4	3.7	20.2	6.5	27.3
Academy sponsor	3 to 11	201	49.3	7.5	9	51.8	12.8	24.2
Community	3 to 11	281	47.7	10.7	19.8	34.6	28.8	25.6

Type of school	Age range	No. pupils	% boys	%SEN*	%EAL	%FSM	%Ethnic minority	KS2 Av point scores 2012
Community	4 to 11	410	50.7	7.8	6.3	9.3	13.5	27.1
Community	5 to 11	197	46.2	8.1	0	5.1	4.1	27.7
Community	7 to 11	351	47	1.7	96.6	21.9	0	27.4
Academy sponsor	3 to 11	185	46.5	6.5	22.8	33.3	30.2	24.5
Academy converter	4 to 11	148	53.4	19.6	49.6	45.3	79.4	29.5
Community	3 to 11	224	51.8	12.5	0	17.6	6.1	26.2
Academy converter	4 to 11	196	48.5	7.7	8.3	5.6	24.7	28
VA	4 to 11	184	59.7	16.3	11.5	4.9	15.3	27.4
Academy converter	5 to 11	146	43.8	4.8	58.6	35.6	81.2	28.4
Foundation	3 to 11	437	49.4	13	5	29.6	16.1	28.8
VC	3 to 11	140	43.6	2.1	4.3	23.7	6.9	27
Academy converter	3 to 11	410	52.2	8.3	0	45.2	1.3	28
Community	3 to 11	458	50.9	3.5	39.2	27.3	47	27.7

Table A2. General characteristics of control schools

Type of school	Age range	No. pupils	% boys	%SEN*	%EAL	%FSM	%Ethnic minority	KS2 Av point scores 2012
Community	7 to 11	208	54.8	5.8	7.2	11.5	21.1	28.1
Community	4 to 11	228	52.2	3.1	2.3	9.4	7.6	29.6

Type of school	Age range	No. pupils	% boys	%SEN*	%EAL	%FSM	%Ethnic minority	KS2 Av point scores 2012
Academy converter	3 to 11	479	54.3	11.9	61.8	23.2	73.3	27.8
Foundation	4 to 11	165	58.8	6.7	0	14.5	3	28.6
Community	4 to 11	607	52.7	9.9	52.9	49.6	82	25.5
VA	7 to 11	194	55.7	18	0	13.4	6.9	27
Community	4 to 11	311	55.6	10	8	21.5	12.8	26.9
Community	4 to 11	44	56.8	36.4	0	14	>	28.1
Community	7 to 11	194	46.9	20.6	3.1	37.6	4.6	27.7
Academy converter	3 to 11	206	5.4	11.2	2.1	28.7	8.7	26.3
Community	3 to 11	454	54.8	3.3	27.7	29.8	38.4	28.8
Foundation	4 to 11	468	51.1	4.9	3.2	8.3	7.4	28.5
Community	3 to 11	326	56.1	6.1	5.7	16.4	7.4	29.3
Community	5 to 11	418	48.1	20.3	77.7	32.3	94.7	26.7
Community	3 to 11	377	52.8	10.6	58	35.8	78.6	25.8
Academy converter	3 to 11	173	49.7	6.4	4	10.1	2.6	28
Community	5 to 11	303	50.5	4.6	7.9	18.8	24.3	27.1
Academy sponsor	3 to 11	439	45.8	6.4	2.4	45.6	5	29.5
Community	7 to 11	470	50.4	3	5.3	7	15.5	29.4
Community	4 to 11	173	52	2.3	4.2	41.6	3.5	26.3
VA	4 to 11	117	53.8	11.1	0	6	7.5	29

Type of school	Age range	No. pupils	% boys	%SEN*	%EAL	%FSM	%Ethnic minority	KS2 Av point scores 2012
VA	3 to 11	326	50.6	11.7	4.3	5.7	20.5	30
Community	3 to 11	360	55	6.4	29.5	34.6	37.6	27.1
VC	5 to 11	153	49	8.5	0	15.7	3.1	27
Academy converter	9 to 13	487	54.6	9.7	0	8	4.1	28.4
Community	3 to 11	242	54.1	4.5	72.2	40.2	80.5	23.5
Community	4 to 11	269	48.7	12.3	9.2	13.4	14.3	28.2
Community	3 to 11	428	51.4	17.8	7.9	45.1	16.8	26.1
Community	3 to 11	555	48.1	10.8	79.2	32.3	94.9	31.3
Community	5 to 11	156	54.5	19.2	0	36.5	6.1	28.3
VC	7 to 11	321	51.4	10.6	2.2	19	5.4	29.5

Table A3. General characteristics of schools that dropped out before allocation

Type of school	Age range	No. pupils	% boys	%SEN*	%EAL	%FSM	%Ethnic minority	KS2 Av point scores 2012
Community	4 to 11	444	49.8	7	2.5	4.8	11.3	29.8
Community	7 to 11	166	53.6	14.5	3	27.7	Not available	26.2
Academy converter	4 to 11	451	49.7	2.4	5.2	23.3	11.2	29.6
Community	3 to 11	419	53.5	9.1	0	34.2	3.4	26.9
Community	3 to 11	251	47.8	4.8	2.3	47.1	11.2	25.4
VC	5 to 11	116	62.1	10.3	4.1	7	12.6	26
Community	3 to 11	462	52.8	8.7	1.5	37.7	14.2	28.2
Community	7 to 11	230	46.5	11.7	32.3	35.2	40	26.5
Community	3 to 11	275	52	9.8	4	58.8	2	27.1
VA	3 to 11	95	51.6	5.5	13.3	0	18.2	SUPP
Community	7 to 11	280	48.6	17.9	16.4	30.7	33.3	29
Community	5 to 11	241	48.1	11.2	4.5	31.5	8.4	26.4
Community	3 to 11	327	51.1	5.2	36.3	35.1	52.1	27.1
Community	4 to 11	159	48.4	10.1	19	32.1	26.8	27.2
VA	7 to 11	235	50.2	8.1	2	3.8	9.2	29.8
Community	5 to 11	343	46.6	12.5	0	6.7	11.8	28.7
VA	5 to 11	273	48	10.6	5.2	12.6	13.6	28.5
Community	3 to 11	351	51.3	9.7	8.8	35	17.1	25.8
Community	5 to 11	172	50.6	8.1	0	14.5	7.1	19.9
Community	3 to 11	229	44.5	5.2	5.2	14.4	9.3	29.5
Community	3 to 11	362	52.8	16.9	41.4	38.7	62.1	26
VC	4 to 11	201	48.3	5	22.2	17.4	52.3	28.4
Community	4 to 11	279	51.6	2.2	1.7	10.8	5.9	27.6

Type of school	Age range	No. pupils	% boys	%SEN*	%EAL	%FSM	%Ethnic minority	KS2 Av point scores 2012
VA	4 to 11	405	45.2	4.9	9.6	23.5	34.7	28
VA	5 to 11	104	48.1	12.5	3.4	11.5	7.7	29.1
Community	3 to 11	189	50.8	27.5	0	44.8	3.1	25.8
Community	5 to 11	210	52.9	4.3	0	4.8	3.9	28.8
Community	3 to 11	166	53	17.5	0	25	21.1	26.1
Community	3 to 11	440	53.2	15	6.1	17.3	8.4	27.3

Table A4 shows the pre-test NGRTA scores for all schools randomised to treatments. The 31 control schools were more variable in size and test scores than the 30 treatment schools.

Table A4. Pre-test results for all Year 6 pupils in all participating schools

Group	N all	NGRTA all
Treatment	20	260
Treatment	18	266
Treatment	38	266
Treatment	28	270
Treatment	31	278
Treatment	25	287
Treatment	26	288
Treatment	19	293
Treatment	58	295
Treatment	16	295
Treatment	30	296
Treatment	32	296
Treatment	51	300
Treatment	55	304
Treatment	83	308
Treatment	33	313

Group	N all	NGRTA all
Treatment	27	316
Treatment	16	316
Treatment	54	318
Treatment	26	321
Treatment	27	321
Treatment	58	322
Treatment	50	322
Treatment	37	323
Treatment	30	327
Treatment	27	329
Treatment	26	331
Treatment	22	337
Treatment	34	344
Treatment	34	344
Control	22	14.7 (hard copy scale)
Control	52	241
Control	32	250
Control	43	263
Control	79	277
Control	55	282
Control	39	290
Control	57	293
Control	31	294
Control	17	295
Control	20	304
Control	47	305
Control	29	305
Control	48	307
Control	37	308

Group	N all	NGRTA all
Control	38	310
Control	41	311
Control	85	316
Control	27	318
Control	29	318
Control	58	319
Control	39	326
Control	117	329
Control	27	330
Control	58	332
Control	12	333
Control	117	335
Control	17	335
Control	42	337
Control	7	342
Control	13	353

Table A5. Planned number of session per tier in each school

Tier 1			Tier 2			Tier 3		
Number	Length	Pupils	Number	Length	Pupils	Number	Length	Pupils
12	20 mins	58	6	20 mins	3	6	15 mins	2
6	30mins	16	10	15 mins	6	10	10 mins	6
20	60 mins	22	8	30 mins	6	20	60 mins	4
24	60 mins	16	6	60 mins	16	6	30 mins	4
6	30-40 mins	14	6	20 - 30 mins	15	6	20 mins	4
30	60 mins	18	18	25 mins	9	12	45 mins	5
28	20 mins	3	18	20 mins	5	30	30 mins	2

Tier 1			Tier 2			Tier 3		
24	30-60 mins	28	24	10-30 mins	24	24-36	10-20 mins	5
6	60 mins	30	18	20 - 30 mins	2 to 5	12	20 mins	3
12 to 18	20 mins	29	12 to 18	30 mins	18	30	15 mins	30
30	30 mins	whole class	6	30 mins	whole class	6	30 mins	4
12	60 mins	26	12	30 mins	5	0	0	0
8	60 mins	whole class	22	30-60 mins	14	0	0	0
18	30 mins	4	10	30 mins	2	0	0	0
25	60 mins	18	12	60 mins	5	0	0	0
30	20 mins	59	18	60 mins	10	0	0	0
57	20 - 30 mins	60	0	0	0	30	5 mins	2
0	0	0	6	30 mins	4	5	15 mins	1
0	0	0	12	30 mins	5	25 - 30	10 mins	1
0	0	0	2 x 3 weekly	20-30 mins	3	0	0	0
0	0	0	4	20 mins	4	0	0	0
0	0	0	10	40-60 mins	5	0	0	0
0	0	0	5	40-60 mins	7	0	0	0
0	0	0	1	45 mins	4	0	0	0
0	0	0	5	20 mins	3	0	0	0
0	0	0	10	20 mins	6	0	0	0
0	0	0	22	20 - 40 mins	n/k	33	20 - 40 mins	n/k
0	0	0	0	0	0	0	0	0

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 v2.0.

To view this licence, visit **www.nationalarchives.gov.uk/doc/open-government-licence/version/2** 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 **www.educationendowmentfoundation.org.uk**.

Education
Endowment
Foundation

The Education Endowment Foundation
9th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP
www.educationendowmentfoundation.org.uk