# The LIT Programme

IFS
Haroon Chowdry and Claire Crawford

| Evaluation Summery | |
|---|---|
| **Age range** | Year 7 |
| **Number of pupils** | c. 1,100 |
| **Number of schools** | 40 |
| **Design** | Randomised controlled trial, with randomisation at the school level. |
| **Primary Outcome** | Reading |

**Evaluation Protocol: The LIT Programme**

## Background

### *Significance*

The LIT programme aims to help children in Year 7 who start secondary school with below average literacy. LIT aims to fill a major gap in early secondary school provision: it gives weak students who have been drilled for Key Stage 2 tests a chance to improve their fundamental comprehension skills, and it gives secondary school staff the training and support they need to deliver basic instruction in literacy. It aims to equip children with a core skill that will help them across all subjects in secondary school. As it is a simple, inexpensive intervention that makes good use of existing resources (i.e., teaching assistants), it could be scaled up quickly if proven effective.

LIT uses a mix of evidence-based approaches, including meta-cognitive methods and effective feedback, but at its core is a method called reciprocal teaching. Some previous non-experimental evidence on reciprocal teaching suggests that it can be a promising intervention and is worth exploring more rigorously. A 1994 meta-analysis of 16 controlled studies of reciprocal teaching found an average effect size of 0.32 standard deviations on standardised test performance, and 0.88 standard deviations on experimenter-developed comprehension tests.[1] These are large effects, but it should be noted that the underlying studies all assessed slightly different treatments, and were not of uniformly high quality. There is still some uncertainty over the precise impact of reciprocal teaching.

### *Intervention*

The LIT programme consists of staff training and detailed lessons that can be delivered by teachers or teaching assistants, to whole classes or small groups. It is generally used as a

[1] Meister, C. & Rosenshine, Barak (1994). Reciprocal Teaching : A Review of the Research. Educational Research, 64(4), 479-530.

supplement to ordinary English lessons; schools are advised to do three to four hours of LIT per week.

In the core approach of reciprocal teaching, the teacher and pupils take turns to lead discussion of a text adapted from children's literature, with the teacher showing children how to engage with the writing in increasingly sophisticated ways. Specifically, the teacher shows the children how to apply four comprehension strategies: summarizing, clarifying, questioning, and predicting. These strategies force children to check that they understand the content, to identify exactly what the problem is when they don't understand, and to make inferences based on what they have read.

**Research Plan**

***Research questions***

The impact evaluation will seek to address the question of whether the LIT intervention had a significant impact on the literacy skills of the children who take part, as measured by a reading comprehension test (Access Reading Test) taken at the end of Year 7. Where sample sizes allow, the impact evaluation will also explore potential variation in the impact of LIT, in order to understand for whom the programme is most effective. This will be done by looking at the impact of the programme across different types of pupil. This could be done according to pupil characteristics (such as ethnicity, SEN, EAL status, etc.) or according to pupils' scores prior literacy ability.

The process evaluation will provide a detailed understanding of how the LIT programme works, and to identify the elements that are critical to its success. It will investigate:

The content of the LIT programme and the design and delivery of the training for schools;

How the delivery of LIT varies and the extent to which teachers adhere to a programme (i.e. fidelity);

How the LIT programme works, who it seems to work best for and what are the critical features;

The value, role and efficacy of LIT programme;

What other literacy initiatives may be operating alongside LIT (in treatment schools) and instead of LIT (in control schools), e.g. reading weeks, library resources, curriculum time, visiting authors, etc.

***Design***

The evaluation is a randomised controlled trial (RCT) with randomisation at the school level. Schools were randomly allocated to either a control group or a programme group. In the programme group schools, Year 7 pupils would participate in the LIT programme if they

were deemed eligible for it. Their eligibility was assessed according to a baseline reading comprehension test (Access Reading Test) issued to the whole year group.

The randomisation was conducted according to an iterative routine in order to find an optimal randomisation that ensured the programme and control group were as similar as possible in terms of the make-up of their pupils. Stratifying first by LEA, schools within an LEA were allocated to either the programme group or the control group, ensuring that each LEA in the study had a school that implemented LIT and a school that did not implement it.

The characteristics used to judge similarity of the two groups were the ethnic composition of the pupil body, the proportion of pupils eligible for Free School Meals and the proportion of pupils with English as an additional language. We carried out the randomisation 1,000 times and kept the repetition that minimised the difference in these characteristics between the programme and control groups. There were no statistically significant differences in these characteristics.

All pupils in the schools taking part will sit a baseline Access Reading Test (ART) in order to establish initial literacy skills and potential eligibility for LIT. At the end of Year 7, the pupils will all sit a follow-up ART in order to determine how much progress they have made since the baseline ART.

***Participants***

In LIT schools, pupils are deemed to be eligible to participate in LIT if they score below the 25$^{th}$ percentile on the baseline ART. While not all of these pupils will end up receiving the LIT intervention (because of constraints in delivering the programme), nearly all the pupils who do participate will have scored below the 25$^{th}$ percentile. In the control group schools, pupils who score below this threshold – who would have been eligible for LIT if their school provided it – will be used as a potential comparison group in order to determine the impact of LIT.

***Outcome Measures***

The primary outcome is the score achieved on the follow-up ART. This is conducted either on a computer or a paper version is marked. There is no face-to-face interaction with an assessor.

In the programme group schools, the LIT participants will also undergo a single word reading test (Hodder Oral Reading Test). However this is not being collected in control group schools and will not therefore inform the impact of the programme; rather, it will be used to better understand any progress in literacy that has been made.

***Sample size calculations***

The tables below show the total sample required (number of children across programme and control schools) for 80% statistical power. Within each table, the required sample varies

according to the estimated effect size (measured in standard deviations) and the within-school correlation in test scores.[2] Each table assumes a different variance for the test scores: for a simple t-test using standardised scores, the variance will be 1. The *residual* variance will be lower if a regression is fitted using other explanatory variables (or baseline measures from the pre-test). Tables 2, 3, and 4 assume that the regression model explains 25%, 50%, and 75% respectively of the variance of the test score.

Table 1. Variance of test score = 1

| | | Within-school correlation | | | |
|---|---|---|---|---|---|
| | | 0 | 0.1 | 0.2 | 0.3 |
| Effect size | 0.1 | 3,136 | N/A[3] | N/A | N/A |
| | 0.2 | 784 | N/A | N/A | N/A |
| | 0.3 | 348 | 4,213 | N/A | N/A |
| | 0.4 | 196 | 368 | N/A | N/A |
| | 0.5 | 125 | 169 | 301 | 219,520 |

Table 2. Residual variance of test score = 0.75

| | | Within-school correlation | | | |
|---|---|---|---|---|---|
| | | 0 | 0.1 | 0.2 | 0.3 |
| Effect size | 0.1 | 2,352 | N/A | N/A | N/A |
| | 0.2 | 588 | N/A | N/A | N/A |
| | 0.3 | 261 | 769 | N/A | N/A |
| | 0.4 | 147 | 217 | 537 | N/A |
| | 0.5 | 94 | 113 | 150 | 263 |

Table 3. Residual variance of test score = 0.5

| | | Within-school correlation | | | |
|---|---|---|---|---|---|
| | | 0 | 0.1 | 0.2 | 0.3 |
| Effect size | 0.1 | 1,568 | N/A | N/A | N/A |
| | 0.2 | 392 | N/A | N/A | N/A |
| | 0.3 | 174 | 292 | 1,872 | N/A |
| | 0.4 | 98 | 119 | 164 | 313 |
| | 0.5 | 63 | 68 | 75 | 88 |

Table 3. Residual variance of test score = 0.25

| | | Within-school correlation | | | |
|---|---|---|---|---|---|
| | | 0 | 0.1 | 0.2 | 0.3 |
| Effect size | 0.1 | 784 | N/A | N/A | N/A |
| | 0.2 | 196 | 368 | N/A | N/A |
| | 0.3 | 87 | 102 | 130 | 199 |
| | 0.4 | 49 | 51 | 53 | 56 |
| | 0.5 | 31 | 31 | 30 | 29 |

[2] The power calculations also assume a significance level of 5%, 40 schools (clusters), split equally between treatment and control groups, and a coefficient of variation for the cluster size of 0.25.
[3] "N/A" in a particular cell means that the number of clusters (schools) is insufficient for the power required, given the effect size and within-school correlation.

Where detection at 80% power is feasible, most of the numbers are under 1,000, meaning that there is a very good chance that an overall impact can be detected given that the evaluation is following 40 schools. Under some circumstances, it will also be possible to detect impacts for sub-groups as well. Under a basic t-test, 368 pupils would be required in the sample to detect an impact of 0.4 standard deviations (with 80% power) if the within-school correlation is 0.1. This points to sub-group analysis being feasible for approximately three groups if around 1,000 pupils are selected in total. If a regression model is used which explains 25% the variance of the test score (Table 2), the required samples are smaller and approximately five sub-groups more could be analysed under the same assumptions.

***Analysis plan***

The primary method of statistical analysis will be least squares regression, comparing the test score outcomes among treated pupils in LIT schools with the test score outcomes of low-literacy pupils in the non-LIT schools. Since the design has been randomised and the randomisation minimised the differences in school characteristics between LIT and non-LIT schools, a simple comparison of test scores at the end of Year 7 can be expected, on average, to provide a reasonably valid estimate of the impact of the LIT programme on literacy skills. However, we will control for school and pupil characteristics in the statistical analysis in order to improve the robustness and precision of the impact estimates.

We will also control for the pupils' scores on the baseline ART to take into account their previous literacy skills and measure the impact of LIT on the progress made over the year. Statistically, we would compare the change in literacy scores among the treated pupils with the change in literacy scores among the comparison group pupils. This is known as a 'difference in difference' estimate of the impact.

The comparison above recovers an estimate of the 'average treatment on the treated' (ATT). However, to the extent that scoring below the 25$^{th}$ percentile may not guarantee treatment in LIT schools, and that other pupils who score above this threshold could still be treated for other reasons, additional analysis will be undertaken. First, propensity score matching (PSM) can be carried out to identify eligible pupils in control schools who are the most 'similar' in terms of their individual characteristics to the treated pupils in LIT schools. The statistical comparison will then focus on these pupils in order to refine the estimate of the ATT. Second, we can compare the literacy outcomes (or progress made in literacy) among pupils in control and LIT schools who scored below the 25$^{th}$ percentile on the LIT programme. This measures the average impact among those who are eligible for LIT, rather than among those who actually receive it. It is known as the 'intention to treat' (ITT) impact.

***Process evaluation methods***

The aim of the process evaluation is to explore views and experiences of delivery and fidelity to the Programme. The findings will help understand and explain the impact analysis. The process evaluation involves:

A short web-based survey administered to all participating schools to gather information on how LIT is being delivered;

15 follow-up depth interviews with treatment schools to gather in-depth data on school views and experiences of LIT, including how closely schools adhere to the programme and the nature of any differences in the way it is delivered;

10 follow-up depth interviews with control schools to explore the nature of other literacy support provided to pupils that may have a bearing on the impact analysis.

The interviews will be based around a topic guide to ensure systematic coverage of key issues, but will also be flexible and interactive, allowing issues of relevance for individual respondents to be covered through detailed follow up questioning. The interviews will be digitally recorded with the respondents' permission, which is essential for the generation of data of sufficient quality for detailed and rigorous analysis and to prevent selective reporting. It also enables the interviewer to concentrate completely on the respondent, picking up essential non-verbal cues (in the face-to-face contacts) and engaging fully in exploratory and responsive questioning. This data will be analysed using Framework, a systematic approach to qualitative data management developed by NatCen Social Research and now widely used in social policy research.

**Personnel**

The evaluation will be led by Haroon Chowdry, a Senior Research Economist in the Skills sector at IFS. Haroon has previously directed projects investigating young people's engagement in risky behaviours on behalf of DfE; he was also responsible for the IFS's contribution to the evaluation of the Every Child a Reader evaluation. Haroon is the main point of contact for the project and will also lead on the randomisation, data analysis and quantitative impact analysis. Claire Crawford, Programme Director of the Skills sector at IFS, will provide expert guidance and overall quality assurance for the impact evaluation.

Meg Callanan, a Senior Researcher in the Children and Young People Team at NatCen, will take overall responsibility for the process evaluation. She will be supported by Sarah Haywood and by Amy Skipp, a Research Director in the Children and Young People Team, who will provide expert consultation and quality assurance for the process evaluation. Meg and Sarah will also liaise with schools to monitor the progress of the baseline and follow-up ARTs.

## Timeline

| Task | Month/Year | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 06/12 | 07/12 | 08/12 | 09/12 | 10/12 | 11/12 | 12/12 | 01/13 | 02/13 | 03/13 | 04/13 | 05/13 | 06/13 | 07/13 | 08/13 | 09/13 | 10/13 | 11/13 | 12/13 |
| Assist LT with selection of schools and assess randomisation | | | | | | | | | | | | | | | | | | | |
| Assess selection and balancing of pupils based on pre-test | | | | | | | | | | | | | | | | | | | |
| Process evaluation:<br><br>Profiling work and Typology of schools | | | | | | | | | | | | | | | | | | | |
| Process evaluation:<br><br>Interviews | | | | | | | | | | | | | | | | | | | |
| Impact analysis based on post-test | | | | | | | | | | | | | | | | | | | |
| Sub-group analysis based on post-test | | | | | | | | | | | | | | | | | | | |

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| and typology | | | | | | | | | | | | | | | | | | |
| Produce draft report | | | | | | | | | | | | | | | | | | |
| Deliver final report | | | | | | | | | | | | | | | | | | |
| | 06/12 | 07/12 | 08/12 | 09/12 | 10/12 | 11/12 | 12/12 | 01/13 | 02/13 | 03/13 | 04/13 | 05/13 | 06/13 | 07/13 | 08/13 | 09/13 | 10/13 | 11/13 | 12/13 |

## Risks

| Risk | Likelihood | Impact | Mitigation and contingency plan |
|---|---|---|---|
| **Insufficient recruitment of schools within time frame** | Medium | High | We would engage in regular dialogue with the Learning Trust and the EEF to monitor the recruitment rate of schools. We would redo any power calculations to decide whether robust analysis remains feasible with a smaller number of schools or whether the programme would need to be delayed for robust impact estimates to be calculated. |
| **Imperfect randomisation** | Low | Low | We would seek to assist the Learning Trust from an early stage with the selection of schools, conducting statistical analyses to ensure that the programme and control groups are well-balanced. If it is not possible to achieve well-balanced groups then we would use a variety of non-experimental techniques during the impact analysis to make it as robust as possible. We have considerable expertise in such methods. |
| **Difficulty in recruiting schools, teachers and stakeholders for the process evaluation** | Medium | High | We would like to discuss with the Learning Trust how we can encourage schools to work with the evaluation from the outset. The key to encouraging participation is orientating the evaluation agenda so that it is relevant and useful for participants. For this reason we would like to promote the evaluation as a co- production with schools – as far as is feasible and also to highlight the benefits and value of the evaluation. We will also consider how best to feed back to schools the findings from the evaluation. We will seek to minimise the burden on all participants by ensuring they are |

| | | | |
|---|---|---|---|
| | | | recruited early and given clear information in advance about what we expect from them. We will also give them choice about the best way to interview them. We will work flexibly around the time commitments of participants |
| **Failure to engage agreed participants in the process evaluation** | Low | High | The qualitative research team have experience of working with teaching staff and other professionals. They are experienced at carrying out in depth interviews and managing group dynamics, creating a compatible environment, preventing people dominating and ensuring all are actively engaged in the process. |
| **Delay in providing data to the project team** | Medium | Medium | It will be necessary for LT to provide the impact evaluation team with the data on participating schools and pupils according to the timetable set out in the tender in order to ensure that the evaluation can be delivered on time. Any delay to the provision of data will delay the preparation and presentation of the impact analysis by an equivalent amount. We will work closely with LT to ensure that the chances of this occurring are minimal. |
| **Unavailability of staff during project** | Low | Low | IFS and NatCen each have a pool of researchers on hand with expertise in programme evaluation, education policy and qualitative research methods respectively. Systems and procedures will be adequately documented to ensure handover can occur smoothly where necessary. |
| **Programme implemented in a variable manner** | Low | Medium | If the treatment is implemented poorly or in a highly variable fashion, it will make it more difficult to identify which elements of the programme are most effective using quantitative methods. The process evaluation will be used to explore any underlying reasons for this and how it could be addressed. |
| **Loss of or damage to data** | Low | High | Both NatCen and IFS have high levels of IT security in place. All members of the evaluation team have extensive experience of working with data, and are well aware of the importance of keeping data safe and with the necessary security procedures. The project team also have wide-ranging experience of data security. Back-ups can be retrieved from off-site and used within one working day. EEF and the project team will be immediately notified if data is accidentally damaged or stolen; contingency plans specified in security policies can then be put into place. |

***Data protection statement***

IFS is registered under the Data Protection Act 1998 (registration number Z5758698) and complies with all its obligations. IFS also ensures that its staff and anyone else involved in its work abide by its Data Security Policy which details the measures that are in place to protect data and to ensure compliance with any legal requirements.

NatCen Social Research operates to extremely high standards in respect of confidentiality and anonymity. We are registered (now termed 'notification') under the Data Protection Act 1998, and comply with all its obligations. In addition, NatCen Social Research is fully accredited to ISO 27001, the international standard which covers information security. To comply with this, our information security procedures are subject to regular external audit to ensure continued compliance.

All data and files held by NatCen Social Research are classified to one of five different levels, with each level having its own specific requirements for how the data are stored, handled and transmitted. In particular, any data containing personal details is deemed to be at level 3, which is 'Respondent Confidential'. For such data, protection against the disclosure of respondent identities is built into all stages of the research process.