I CAN

IFS
Claire Crawford

Evaluation Summary	
Age range	Primary (Nursery and Reception)
Number of pupils	c. 360
Number of schools	30
Design	Randomised controlled trial, randomised at the pupil level
Primary Outcome	Literacy

BACKGROUND

Significance

It is well known that children with oral language difficulties often struggle to read well. In particular, poor language skills are often associated with poor reading comprehension, which is fundamental to the development of functional literacy. Yet there is relatively little evidence of whether and how early interventions to improve oral language skills may impact upon subsequent literacy development. Given the importance of basic skills such as literacy for an individual's subsequent educational attainment and labour market success, there is substantial policy interest in understanding how to improve children's literacy skills. This study will investigate the effectiveness of one particular intervention – the Nuffield Early Language Intervention programme, implemented by the charity ICAN – designed to increase children's oral language skills around the time that they start school.

Intervention

The Nuffield Early Language intervention focuses on improving spoken language skills, such as vocabulary knowledge, story-telling and listening skills, amongst children with relatively poor language skills around the time that they start school. Two different programmes are being tested: one starts in the final term of nursery and lasts for 30 weeks (i.e. it continues through the first two terms of reception); the second starts in the first term of reception and lasts for 20 weeks. Both interventions are delivered by teaching assistants (TAs) working with selected children individually or in small groups. At nursery, children will participate in three 20 minute group sessions per week. At school, they will participate in three 30 minute group sessions and 2 15 minute individual sessions per week.

RESEARCH PLAN

Research questions

- What is the impact of receiving 20 or 30 weeks intensive language support at age 4 or 5 on children's vocabulary, reading, spelling and comprehension skills one year later?
- How much more effective is it to receive 30 rather than 20 weeks support?
- To what extent do differences in the effects of the 30 and 20 week interventions continue or fade out once the programme has finished?

Participants

A selection of primary schools with attached nursery schools or classes from London and Sheffield were recruited to participate in the trial. All children in the selected nurseries who are due to start school in September 2012 will be tested using the CELF (Clinical Evaluation of Language Fundamentals) Pre-School Expressive Vocabulary and Sentence Structure tests. The parents of the

12 children with the lowest composite language scores will be invited to consent to participate in the programme. Children who do not opt in will be replaced with the next lowest scoring children.

Design

Randomisation will occur at the pupil level. The 12 children recruited from each nursery/school will be randomly allocated to one of the two treatment groups (i.e. will receive either 20 or 30 weeks language support) or to a control group, who will receive a different intervention designed to improve literacy skills at the end of the ICAN trial. Minimisation will be carried out largely on the basis of a baseline composite language score (comprising an average of the CELF tests described above, plus the British Picture Vocabulary Scale), but we will also seek to ensure a balance of children by gender and month of birth as well. Children allocated to the 30 week intervention programme will also all have to attend nursery during either the morning or the afternoon session for administrative purposes.

Within-school randomisation buys a certain amount of statistical power and hence enables the sample sizes to be lower than would be necessary if the randomisation occurred between schools, but means that we cannot account for the possibility of spillovers, e.g. as a result of teaching assistants applying the ICAN techniques to other pupils in the class with poor language skills. If such spillovers exist, then we might expect the programme to have some effect on pupils in the control group as well as the treatment groups, such that the impact estimates produced would *underestimate* the impact of ICAN. We will seek to address these issues via our process evaluation.

Outcome Measures

- The primary outcome is a composite language score, based on the CELF Expressive Vocabulary test, the Action Picture Test information and grammar scores, and a measure of listening comprehension (adapted from the YARC reading comprehension measure).
- The secondary outcome is a literacy composite based on tests of letter knowledge, early word reading and spelling.

The tests will be "blinded" (i.e. carried out by assessors who do not know which pupils are in the treatment and control groups).

Sample size calculations

The aim is to recruit a total of 360 pupils across 30 schools, to be split equally across two treatment and one control group. This means that there will be 8 treated pupils (4 receiving each intervention) and 4 control pupils per school. Table 1 below shows the total sample size required for 80% test power, at a 5% level of significance, depending on the estimated effect size (measured in standard deviations, SDs) and the within-cluster correlation in test scores. It assumes that we can explain 60% of the post-test variation in test scores. The left hand panel illustrates the total sample size required (across both treatment and control groups), while the right hand panel focuses on the comparison between a single treatment and control group (which is the relevant sample size for calculating the effect of the intervention).

Table 1 Power calculations

	Within cluster correlation					
	Total sample size (two treatment and one control group)			**Sample size required to detect impact (one treatment and one control group)**		
Effect Size	0	0.05	0.1	0	0.05	0.1
0.05 SDs	7,535	N/A	N/A	5,023	N/A	N/A
0.1 SDs	1,884	N/A	N/A	1,256	N/A	N/A
0.15 SDs	837	11,401	N/A	558	7,601	N/A
0.2 SDs	471	938	N/A	314	626	N/A
0.25 SDs	301	430	821	201	287	548
0.3 SDs	209	259	352	140	173	235

Notes: "N/A" in a particular cell means that the number of clusters (schools/nurseries) is insufficient for the power required, given the effect size and degree of within-cluster correlation.

Our power calculations suggest that an effect size of 0.22 SDs could be detected – as long as there is no evidence of within-cluster correlation (which the project team found to be minimal in their previous studies). Even if the intra-cluster correlation was actually 5% or even 10%, we would still be able to detect an effect size of 0.3 SDs, which was the minimum found in the project team's previous study.

Analysis plan

The Institute for Fiscal Studies (IFS) is responsible for the quantitative impact evaluation.

As with all randomised control trials, assuming that the treatment and control groups are well-balanced (i.e. statistically indistinguishable and, given the small sample sizes involved, quantitatively similar) at baseline, it should be possible to obtain unbiased estimates of the impact of the programme by simply comparing the average outcomes of pupils in the treatment and control groups after the intervention. However, given the relatively small sample sizes involved, together with the need to maximise the variation in test scores that we can explain (for the purposes of the power calculations), we are instead likely to use regression or matching techniques in order to be able to account for baseline characteristics of both individuals and schools (potentially via fixed effects). To do so we hope to use individual and school information from the National Pupil Database (NPD).

Given the small sample sizes involved, it may not be possible to detect significantly different impacts amongst different subgroups. However, we will investigate whether and to what extent the impact appears to vary according to both baseline language skills and a range of characteristics available in the NPD, such as gender, free school meal eligibility, ethnicity and language status.

Beyond the scope of this evaluation, EEF are also hoping to use information on Key Stage test scores available in the NPD to monitor the longer-term progress of pupils who receive the ICAN intervention.

Process evaluation methods

NatCen Social Research is responsible for the process evaluation, whose aim is to provide a detailed understanding of the ICAN programme – including how it fits into the school's overall literacy strategy, whether TAs adhere to the programme model (i.e. fidelity), the ways in which support varies across schools and which elements are viewed as being critical to its success – in order to help understand and explain the impact estimates produced by IFS.

We will gather broad details about each of these elements via a short online survey administered to the ICAN lead in all participating nurseries/schools. We will then use this information to develop a typology of schools from which we will purposively select eight for more detailed follow-ups. In the first four schools we will carry out face-to-face interviews with the TA implementing the programme, as well as the relevant class teacher or key stage coordinator. In the remaining four schools we will conduct telephone interviews with the TAs delivering the intervention in nursery and reception.

PERSONNEL

Claire Crawford	Institute for Fiscal Studies	Overall evaluation manager and head of impact evaluation
Elaine Kelly	Institute for Fiscal Studies	Working on the impact evaluation
Mehul Kotecha	NatCen Social Research	Day-to-day lead on the process evaluation
Amy Skipp	NatCen Social Research	Overall lead on process evaluation

TIMELINE

Date	Activity	Organisation
February-March 2013	Coordinate with project team and EEF to obtain ethical approval and permission from headteachers to link to pupils' NPD records	IFS
March 2013	Prepare and consult on content of online questionnaire to be sent to all schools to gather information about implementation of programme and other literacy strategies in operation	NatCen

March-April 2013	Receive information from project team on screening/pre-test and carry out randomisation of pupils into treatment and control groups; feedback these details to the project team	IFS
April-May 2013	Submit NPD data access request	IFS
May-July 2013	Invite schools to complete online questionnaire	NatCen
September 2013	Analyse responses to online questionnaires; develop a typology of schools from which to select some for more detailed follow-up	NatCen
September 2013 to March 2014	Contact eight schools to invite them to participate in further qualitative research; carry out face-to-face/telephone interviews with TAs, teachers/coordinators.	NatCen
May-July 2014	Receive data from first post-test and analyse; present verbal feedback on results to EEF and project team	IFS
December 2014 to February 2015	Receive data from second post-test and analyse; present verbal feedback on results to EEF and project team	IFS
End March 2015	Deliver final evaluation report to EEF	IFS/NatCen

Risks and 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.

Risk	Likelihood	Impact	Mitigation and contingency plan
Imperfect randomisation	Low	Low	We have successfully randomised treatments for previous evaluations, using statistical techniques to ensure that the characteristics of pupils in the treatment and control groups are well-balanced. To further mitigate these risks, we will use non-experimental methods to ensure that our treatment and control groups are as similar as possible at baseline. We have considerable expertise in using such methods.
Sample sizes too small to detect significant effects	Low	Medium	The agreed sample sizes should be sufficient to detect effects that are slightly smaller than those anticipated by the project team, even if there is a higher than expected degree of intra-cluster correlation. This means that there is also some flexibility to cope with smaller than anticipated recruitment of pupils or higher than expected attrition.
Control pupils receive some form of treatment as a result of spillovers	Medium	Medium	As outlined above, there is some risk that spillovers across pupils with low language skills within schools may mean that we would underestimate the overall impact of the ICAN programme on children's language skills. The project team has found little evidence of such spillovers in previous analysis; however, we could consider constructing a comparison group of pupils in non-intervention schools.
Difficulty in recruiting schools, teachers and stakeholders for the process evaluation and keeping them engaged	Medium	Medium	We have considerable experience of working with schools to conduct qualitative research and would draw on this expertise to ensure that the evaluation agenda is useful for participants, thus keeping them interested and engaged. We will minimise the burden on participants by ensuring they are recruited early and given clear information in advance about what we expect from them. We will also give them a choice about the best way to interview them and will work flexibly around their time commitments.
Delay in project team providing data to the evaluation team	Low	Medium	The proposed timetable should allow sufficient time for the project team to collect the data, clean it and pass it on to the evaluation team in a timely fashion. The main danger is likely to be following the final post-test in December 2014. We will prepare our code in advance to minimise the risk that the final report will not be delivered on time.

Cannot access pupil NPD data	Medium	High	The risks arise both from the fact that headteachers may not grant us permission to link to children's data, and that the DfE may not grant us permission to access NPD data, conditional on receiving consent. We will seek to mitigate these risks by ensuring that we agree appropriate wording with the NPD team before seeking consent.
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 who would be capable of taking over work on this project. Systems and procedures will be adequately documented to ensure handover can occur smoothly if necessary.
Loss of or damage to data	Low	High	Both NatCen and IFS have high levels of IT security. All evaluation team members have extensive experience of working with data, and are well aware of the importance of keeping data safe and of using the necessary security procedures. Back-ups are located off-site and can be retrieved within one working day. EEF and the project team will be immediately notified if data is accidentally damaged or stolen; contingency plans can then be put into place.