Graphogame Rime

NFER
Jack Worth

Education Endowment Foundation

Evaluation Summary	
Age range	Year 2 (6 – 7)
Number of pupils	400
Number of schools	10
Design	Pupil level randomised controlled trial
Primary Outcome	Reading

Background

Intervention

Graphogame Rime is a *computer game developed to teach pupils to read by developing their phonological awareness. Originally developed by a Finnish University, the* Graphogame group of computer programs employ algorithms that analyse a child's performance and constantly adjust the difficulty of the content so that the challenge matches the learner's ability. The English version of Graphogame Rime was developed by educational neuroscientist Usha Goswami, building on research into "rhyme analogy".

Pupils sit at a computer, laptop or tablet with headphones on and play the game for around 10 minutes a day. *Instruction is focused on helping children to match auditory signals with groups of letters (rimes) displayed on the screen. The game first focuses on rimes that are most common in English. Each child has a personal log-in and the game offers increasingly challenging levels as they improve their skills.*

Significance

Graphogame Rime has been the subject of two, small non-randomised studies, which suggest promising effect sizes on specific measures such as spelling (0.9) and phonemic awareness (over 1). However, the studies have limitations and further research is needed to test these results.
Graphogame was developed to maximise the development of the "visual word form area" (VWFA) in struggling readers. Neuroimaging suggests that this area is most active as children learn to decode words. Graphogame is designed to develop a targeted response in this area. It focuses initially on rimes that are common and consistent (i.e. groups of letters that normally represent the same sound), mimicking the natural exposure that competent readers are likely to have received.

Research plan

Research question

The primary research question is whether playing Graphogame Rime as part of literacy lessons improves pupils' development of reading ability, compared to 'business as usual' participation in literacy lessons.

Design

The evaluation will be a pupil-randomised controlled trial with two arms: a group of pupils that spend 10-15 minutes each day playing on the Graphogame Rime computer games during literacy lessons and a 'business as usual' control group of pupils from the same classes who will have business-as-usual tuition in literacy lessons.

The randomisation will be stratified by class so that the *number* of pupils that will be in the intervention and control groups respectively will be known by teachers long in advance, while the *identity* of the pupils that are allocated to each group will not be revealed until after the pre-testing has been completed to prevent any bias.

Participants

Ten primary or infant schools in Cambridgeshire that have Year 2 pupils will be recruited by the University of Cambridge team. The trial will run over two years with the same schools, with two cohorts of pupils participating, each when they are in Year 2.

Year 2 pupils that scored below a certain score on the phonics screening check will be eligible for participation in the evaluation. Using the government's required standard of 32 marks out of 40 as the eligibility criteria is an option, though it might lead to a sample that is too small. In 2014, 26% of Year 1 pupils in Cambridgeshire did not meet the required standard and the proportion has been decreasing each year.

NFER will select an eligibility threshold in the region of 32-35 marks that ensures a large enough sample size for the evaluation while also including the pupils with the lowest level of phonological skill in the trial. The eligibility criteria will be the same for each school and, once decided, will remain the same over the two years[1].

Outcome measures

The primary outcome measure will be the raw score on the New Group Reading Test (NGRT) Level 2. The paper version of the test will be used, provided by GL Assessment. The tests will be administered by independent NFER test administrators, who will not know the group allocation of the pupils being tested. The secondary outcome measures will be the GL Assessment single-word spelling test and, subject to availability through the National Pupil Database, the spelling component of the end of Key Stage 1 grammar, punctuation and spelling test.

The NGRT Level 1 test will be administered by the University of Cambridge team as a pre-test measure. The outcome of the random allocation of pupils to the intervention and control groups will only be communicated to the research team and to teachers once the testing has been completed to eliminate the possibility that pre-test scores are biased.

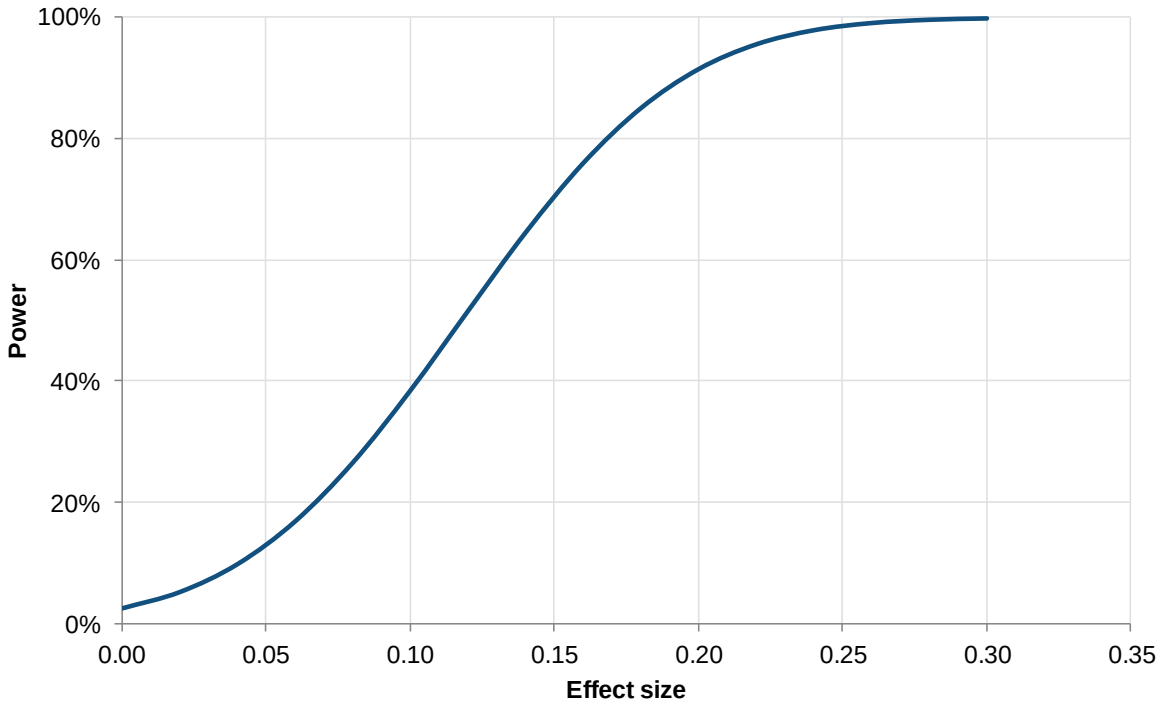
Sample size

The evaluation aims to recruit ten primary or infant schools and have an average of 20 pupils in each school in each of the two years of the study. In total, 400 pupils will participate in the trial and be allocated half to the intervention group and half to the control group in each year.

Assuming the correlation between the pre-test reading measure and the outcome measure is around 0.8, the design would be able to detect an effect size of 0.17 with 80% power. Attrition of ten per cent would still yield a minimum detectable effect size at 80% power of 0.18. The intra-cluster correlation is assumed to be zero for the power calculations because randomisation occurs within schools. The sample of schools is a small convenience sample and it is therefore not appropriate to generalise to a

[1] Using the same threshold mark for both years assumes the test has remained the same difficulty. Depending on whether DfE alter the mark needed to obtain the required standard, the eligibility threshold may be changed to ensure children of equivalent standard are eligible.

wider population of schools. Analysis will use a single-level regression model with dummy variables for school, to take into account of the restricted randomisation.



Analysis plan

All statistical analysis of the programme's impact will be conducted at pupil level, comparing average pupil reading (or spelling) scores in the intervention group with average scores in the control group. The average difference will be measured in a linear regression model that takes account of pre-test score, gender and age in months by including them as covariates. A set of school indicator variables will be included as covariates to take account of average score differences between schools. We will perform a two-tailed t-test of the significance at the 95% level of the group indicator coefficient to determine whether the intervention has had a significant impact on reading ability.

The primary outcome analysis will estimate the average difference in the main outcome measure between the intervention and control groups. The secondary outcome will be analysed in a similar way. An interim analysis will take place at the end of the first year on just the primary outcome. The interim results will be discussed with the EEF Grants Committee, with advice from the Evaluation Advisory Group if required, to determine whether or not the trial should continue. If there is substantial evidence of a positive effect of the intervention then the Grants Committee may advise to move to a larger effectiveness trial, while if there is evidence of negative effects then they may advise to stop the trial. However, the analysis of data from the first year will be underpowered for plausible positive or negative effects and therefore it is most likely that the trial will be completed as planned.

A set of sub-group analyses will also be conducted, looking at:

- the differential impact of the intervention on pupils eligible for free school meals (FSM), using an interaction between the group indicator and FSM eligibility
- the impact of the intervention on the sub-sample of pupils eligible for FSM
- the differential impact of the intervention by a pre-determined measure of programme fidelity, or dosage. Pupil-level game usage data will be recoded into

levels of programme fidelity (shown in the table below), and explore the differential impact by fidelity level. It will also be analysed without recoding.

Level of programme fidelity	Total number of hours spent on Graphogame Rime
Low	0-3
Medium	3-6
High	6-10
Very high	More than 10

Process evaluation methods

The process evaluation will explore the effectiveness of the implementation of the intervention and provide insight into how it is used and received by schools. As the sample is relatively small, we will gather data from all of the schools involved in the trial to better understand the experiences and challenges faced by schools. The process evaluation will collect information on:

- preparation of staff to deliver the intervention
- implementation of Graphogame Rime, including challenges, adaptations and fidelity to the delivery plan
- resources involved in the intervention, including training and staffing costs (e.g. use of teaching assistants, cover time for training)
- perceived outcomes and impacts
- deliverers' views on the suitability, sustainability and potential for roll out of the intervention
- pupil views about the intervention.

We summarise below our proposed methodology for the process evaluation.

1. Negotiate the evaluation requirements with EEF and the University of Cambridge.
2. Attend key events (e.g. project set up meeting, recruitment event, training) and gather documentation about the interventions.
3. Review the data available from the Graphogame Rime system and what it tells us about patterns of use, rates of completion.
4. Visit five schools towards the end of the implementation period (March/April) in each of the two years of the evaluation, to:
    - observe intervention sessions,
    - interview pairs or small groups of pupils using Graphogame Rime
    - interview school staff (likely to include a senior leader, the teacher of the class and/or the person delivering the interventions).

Interview topics for staff will include: what is working well and what isn't; any adaptations made; resources used (including any issues relating to access to technology); perceived outcomes for pupils; perceptions of scalability; the school's 'standard' provision of phonics for the control group.

Focus groups or small group interviews with pupils, combined with the observations, will provide evidence of how pupils are receiving and reacting to the intervention. We will include interactive ways of gathering their views, such as using flash cards.

The fidelity data (see above) will be used to inform the selection of schools for visits such that we will aim to visit schools with varying levels of use of the programme (if this applies).

Visits will take place in both years 1 and 2 of the project to look at intervention implementation and fidelity. Additional questions can be added in year 2 to explore any longer term perceived impact from the intervention's use in year 1.

5. Interview the remaining five school deliverers by telephone (30 minutes) after the end of intervention delivery (i.e. at the beginning of the summer term), and complete shorter, follow-up telephone interviews (15 minutes) with the five case study schools to capture any changes to delivery and perceived impact.
6. Determine the extent of potential contamination between treatment and control groups
7. Review relevant project documentation, such as training materials and outputs from the Graphogame Rime logs (see point 3 above).
8. Analyse the data from the case studies, interviews and training materials to feed into the final report[2].

Consent

Agreement to participate in the trial will be sought from headteachers when the school is recruited. A memorandum of understanding setting out the central commitments and data sharing details that are required will then be shared with, and signed by, the headteacher. We will make parents aware of their right for their child to opt-out of data being linked to the NPD or used in any way in the research by asking schools to send a letter to the parents of eligible pupils.

Registration

The trial is registered at the ISRCTN registry under reference number ISRCTN10467450[3].

Personnel

The project will be led by Usha Goswami (University of Cambridge) and be managed by Mary Anne Wolpert. The evaluation will be overseen by Dr Ben Styles at NFER. The impact evaluation will be led by Jack Worth and the process evaluation will be led by Helen Poet. Calum Davey will have overview of the evaluation at EEF and Eleanor Stringer will oversee the grant.

Risks

Risk	Assessment	How will it be addressed?
School, teacher or pupil attrition	**Likelihood: moderate** **Impact: moderate**	**Clear information for teachers and an induction meeting explaining the principles of the trial and expectations. 'Intention to treat' analysis will be used.** **Attrition will be monitored and reported according to CONSORT guidelines.**

[2] Note that we have not included costs to analyse and report on the process evaluation at the end of Year 1 of the evaluation.

[3] URL: http://www.isrctn.com/ISRCTN10467450 [accessed 29/09/2015].

Risk	Assessment	How will it be addressed?
Intervention is not implemented well	**Likelihood: low** **Impact: moderate**	**Clear information for teachers and an induction meeting explaining the principles of the trial and expectations. Both 'intention to treat' and 'on-treatment' analysis will be used.** **Process evaluation will monitor this.**
Control pupils exposed to elements of the interventions	**Likelihood: moderate** **Impact: moderate**	**Clear information for teachers and an induction meeting explaining the principles of the trial and expectations. 'Intention to treat' analysis will be used.** **Process evaluation will monitor this.**
Delays in training of teachers and commencing interventions	**Likelihood: moderate** **Impact: low**	**Agree a clear timetable with project teams up front.** **Revise timetable for pre and post testing periods if it becomes an issue.**
Insufficient schools recruited to the study	**Likelihood: very low** **Impact: high**	**NFER could help with recruitment for an additional fee if this becomes problematic. Timescale could be revised.**
Evaluation team unable to access fidelity data	**Likelihood: low** **Impact: moderate**	**Project teams to work closely to negotiate appropriate access granted within specified timescales. Data to be handled sensitively and securely.**
Researchers lost to project due to sickness or absence	**Likelihood: moderate** **Impact: low**	**NFER has a large research department with numerous researchers experienced in evaluation who could be redeployed.** **Senior staff can stand in if necessary**
Delivery or evaluation project teams do not follow correct trial protocols	**Likelihood: moderate** **Impact: high**	**Meetings with project teams at start of project.** **Clear information for teachers and an induction meeting explaining the principles of the trial and expectations.**
Developer finds it difficult to work with an independent evaluator	**Likelihood: low** **Impact: high**	**NFER has recent experience of the evaluator role for interventions developed by academics. Some compromises can be made.**
Communication with schools is disjointed	**Likelihood: moderate** **Impact: moderate**	**NFER to share all documents that are being sent to schools with Cambridge and Cambridge given opportunity to suggest changes, and vice versa.**

Timeline

2015	Activity
April – August 2015	**Develop English version of Graphogame Rime for schools**
April – June 2015	**Recruitment of schools**
June – July 2015	**Collect pupil data from schools, including phonics check score**
September 2015	**Induction day for teachers**
October – November 2015	**Pre-testing by Cambridge**
December 2015	**Randomisation allocations determined and communicated to schools and Cambridge**
January – April 2015	**Intervention** **Process evaluation case study visits to 5 schools**
May – June 2016	**Post-testing by NFER test administrators** **Process evaluation follow up telephone interviews**
June – July 2016	**Collect pupil data from schools, including phonics check score**
August 2016	**Interim analysis (impact evaluation only) conducted by NFER**
October – November 2016	**Pre-testing by Cambridge**
December 2016	**Randomisation allocations determined and communicated to schools and Cambridge**
January – April 2017	**Intervention** **Process evaluation case study visits to 5 schools**
May – June 2017	**Post-testing by NFER test administrators** **Process evaluation follow up telephone interviews**
August – October 2017	**Final analysis conducted by NFER. Outcome data sent to Cambridge.**
November 2017	**Delivery of evaluation report to EEF**