Let's Think! Secondary Science

IEE, University of York
Pam Hanley

Education Endowment Foundation

Evaluation Summary	
Age range	Secondary (Year 7 and Year 8)
Number of pupils	10,070
Number of schools	53
Design	Randomised control trial with randomisation at the school level
Primary Outcome	Science

LET'S THINK EVALUATION PROTOCOL

BACKGROUND

Significance

Let's Think! Secondary Science (LTSS) is based on the programme first developed and evaluated in the 1980s, Cognitive Acceleration through Science Education (CASE). The two-year CASE programme has been evaluated over the last 20 years with positive indications across a range of outcomes, including cognitive development and better-than-expected gains in maths and English as well as science. Most of the more robust evaluations have taken the form of quasi-experiments with matched controls (Adey, 2004). Adey cites one exception: a Finnish RCT conducted by Hautamäki, Kuusela and Wikström in 2002. This was a relatively small study (276 students) based in just one town. Although it showed an immediate post-test effect on cognitive development for CASE students, this advantage was eroded over the following three years. Adey argued this was due to control students improving their performance to an unexpected degree because they were in classes where two-thirds of the students had undergone either CASE or its mathematical equivalent (CAME).

Intervention

As an educational intervention, CASE was designed to promote better thinking through science education and was based on teachers delivering an hour-long session every fortnight over two years for a total of thirty lessons. In this intervention, it is proposed to reduce the number of lessons in the programme from 30 to 19 across the two years. The structured programme challenges students' thinking through cognitive conflict, reflects the social construction of knowledge by promoting collaborative working, and encourages students to reflect on their own thinking and learning (metacognition). The teachers are provided with training and resources. The proposed training will be delivered face-to-face and through video-conferencing to clusters of schools over the two years of the project. Each teacher will teach each LTSS lesson twice (to two different classes) so that lesson delivery can benefit from a practice effect on the second occasion.

RESEARCH PLAN

Research Questions

Questions the evaluation is designed to answer are:

- What is the impact of Let's Think Secondary Science on student achievement in:

    - science
    - maths
    - English?
- What are the effects on the development of cognitive reasoning?
- Are any gains in cognitive reasoning correlated to gains in student achievement?

Design

The impact evaluation is a clustered randomised control trial (RCT) of 53 secondary schools to assess the effects of LTSS. The process evaluation will assess the quality of implementation of the LTSS components and teacher and student response/engagement.

Randomisation will be at school level, using matched pairs of schools stratified by average percentage GCSE (A*-C) results over three years, percentage of students eligible for free school meals, percentage EAL, and number of students registered at the school. Random number generation will be used to identify which school in each pair will be the intervention school and which the control. An intent-to-treat design will be used, whereby schools will be asked to provide evaluation data if they withdraw from the programme after randomisation.

Participants

Eligible schools are secondary schools (having Year 7 and Year 8 classes) with teachers having no previous experience of CASE in the last 10 years. To facilitate training and running of the intervention, schools will be recruited and randomised in geographical clusters. It is anticipated that all teachers of science to Year 7 will be trained and that the training will include teaching assistants if applicable and possible. The provider will be responsible for recruiting appropriate schools to the project, and a member of the evaluation team will be invited to any recruitment conferences to explain the concept and practice of RCTs. A sizeable proportion, though not all, schools need to draw from a high poverty catchment area.

Outcome Measures

The primary outcome at the end of the study will be a science assessment. Secondary outcomes will be online numeracy, English and cognitive ability tests.

In order to minimise the costs and disruptions of data collection, routinely collected English and Maths KS2 scores (obtained from the National Pupil Database) will be used as a pre-test. These standardised measures are high in contextual validity, as they constitute the main indicators of primary school pupils' academic performance.

At the end of Year 8, assessments will be administered for science, maths and English. To reduce the financial implications for the evaluation, and the testing burden for the schools, it is recommended that a matrix sampling model is used. All students would complete the science test, as this forms the primary outcome measure. Half the classes would be randomly assigned within each school to receive either the maths or English assessment (ie approximately half the student sample completing each). For preference, if the technology works sufficiently smoothly to enable this by summer 2015, students could be randomly assigned to one test or other at an individual level. Tests would be administered concurrently, so that only one lesson slot would be needed for testing each class of students. The Progress in Maths and Progress in English tests (GL Assessments) would be used, and the science assessment will consist of age-appropriate questions extracted from optional KS3 tests.

It is argued that the effectiveness of this intervention is related to improvements in cognitive reasoning. To judge this claim, it will be useful to have an assessment of cognitive development at the end of the two year period. A measure that is independent of the intervention needs to be identified and agreed. The tasks used in the original research (Adey & Shayer, 1993), based on Piagetian scientific reasoning, are inappropriate evaluative measures as they are too close to the treatment. The recommended test is CAT4 (GL Assessments), to be administered at the end of Year 8. Again, to reduce costs, this would be administered as part of a matrix to 25% of the sample, alongside the maths and English tests. A possible design is illustrated in the table below:

	science	maths	literacy	*CAT4*
.25 of N	X	X	X	
.25 of N	X		X	
.25 of N	X	X		
.25 of N	X			*X*

Impact Analysis plan

The impact evaluation will use multilevel modelling, whereby students are nested within schools and school means are compared. The test scores would be analysed in summer 2015, comparing schools randomly assigned to the Let's Think programme to those in the randomly assigned control group, controlling for KS2 pre-test scores. Planned missing data analysis would be used to account for the matrix sampling.

After the main analyses including all students, subgroup analyses will be carried out for boys and girls, and for high, average and low achievers. Additionally, EAL and FSM would be used as variables in analysis. If used, cognitive ability scores would be correlated with science, English and maths achievement scores.

Process Evaluation

The process evaluation will aim to develop an understanding of how LTSS is being received and implemented in schools. By means of triangulation, the data collected will also help to inform and explain some of the impact evaluation findings.

Teacher and student feedback will be gathered from teacher and student on-line survey questionnaires in all schools, along with some visits to intervention schools to observe LTSS lessons. We suggest one visit to each of 6 schools from the 25 intervention schools over the course of the programme (during the second year). Two schools will be selected at random from each of the three regions and asked to allow an evaluation visit. Subject to the agreement of the EEF and the project provider, we would also request that a member of the research team attends at least one training session (ideally one in each region), so as to have first-hand experience of the Let's Think teacher professional development.

The **teacher questionnaire** will be delivered to science teachers in both intervention and control schools (for comparison purposes). The survey will be administered towards the end of the two years (spring/summer term 2015) to compare responses as the programme becomes more familiar to teachers in the intervention schools. Topics to be covered by the questionnaire will be finalised after discussion with the developers about their expectations of the programme, but will probably include:

- details of the school context and context for science teaching
- age, experience, roles and qualifications of teachers
- teaching strategies used in science lessons (type and frequency)

- views on teaching and learning approaches in science
- effects of any training /CPD/programmes on their teaching practices

LTSS schools only:

- specific views on effectiveness of the professional development and resources
- any challenges in terms of transferring the professional development and resources to the classroom
- views on the impact of the LTSS programme: (a) on the teachers themselves; (b) on the students
- whether the approach appears to be more or less suitable for any student groups (eg disadvantaged students)
- any information about knock-on effects across the curriculum

The **student questionnaire** will be short, visually appealing and student-friendly. It will be administered during the spring/summer term, 2015. Students will be asked what kinds of things they learn about in science, what activities they undertake, what teaching strategies are used and how frequently, and how capable and confident they feel in science. Similar questions will be asked about English and maths as a point of comparison. Students in the intervention group will be asked to compare Let's Think lessons with their ordinary science lessons.

The process evaluation will provide information to Let's Think and to the EEF about teachers' (and students') views of the project, and what they might see as strengths and weaknesses in the approach. This information may inform future development of the intervention.

The survey findings will be analysed using both qualitative and quantitative techniques, comparing control and intervention. Data from the surveys will be compared with the outcome measures in each school so that one might be used to explain the other. It is assumed that the provider will also be collecting implementation fidelity data, and we would work closely with them to co-ordinate approaches to schools and ensure that data collection is not needlessly duplicated. We might be able to include their data in our analyses to determine if implementation fidelity modifies the impact on student outcomes.

PERSONNEL

- **Pam Hanley, PhD.** Dr Pam Hanley is a Research Fellow at the IEE and has a solid track record in science education research and the use of both qualitative and quantitative methods. Pam is Principal Investigator with responsibility for data collection for the impact and process evaluations, including liaison with the Let's Think team and participating schools.

- **Robert Slavin, PhD.** Professor Robert Slavin is currently Director of the Center for Research and Reform in Education at Johns Hopkins University in Baltimore and professor at the Institute for Effective Education at the University of York. Bob is Co-investigator on the evaluation, with particular responsibility for evaluation design and statistical analysis.

- **Louise Elliott.** Louise Elliott is the Data Manager at the IEE. Louise is responsible for managing all database organisation, data entry, cleaning and descriptive statistical analyses conducted in the evaluation.

RISKS

Four main areas of risk are identified as follows:

- recruitment and retention of schools: the intervention involves a considerable commitment from schools (affecting the whole of a one-year cohort for their first two years in school) and the delayed treatment necessitates a wait of two years for the intervention. Although it is based on a well-known model (CASE) these two factors, with attendant logistical constraints, may prove a barrier to participation.
  Recommendation: very clear explanation to schools about the nature of their commitment and the justifications for an RCT, and the responsibilities of participants.
- the nature of the intervention: the level of intervention (19 rather than 30 lessons over 2 years) is significantly reduced from the original CASE study, which may not lead to a comparable impact.
- fidelity of implementation: this is known to be a variable that can affect the impact of an intervention.
  Recommendation: providers share with evaluators details they collect on level of fidelity, for potential use in analysis.
- age-appropriate, reliable and valid tests of science and (if possible) cognitive reasoning need to be identified and agreed.
  Recommendation: evaluators have consulted widely to identify appropriate measures, and will continue to do so during the course of the project if new or updated tests are designed.

DATA PROTECTION STATEMENT

All outputs will be anonymised so that no schools will be identifiable in the report or dissemination of results. Data will be handled in accordance with the Data Protection Act (1998). Statistical databases will hold non-identifiable data. Ten per cent of papers will be double-marked and all data will be double entered to assess reliability. Confidentiality will be maintained and no one outside the trial team will have access to the database. The trial database will be securely held and maintained on the University's research data protection server, which is regularly backed up.

REFERENCES

Adey, P. (2004). Evidence for Long-term Effects: Promises and Pitfalls. *Evaluation & Research in Education, 18(1-2)*, 83-102.

Adey, P. and Shayer, M. (1993). An exploration of long-term far-transfer effects following an extended intervention program in the high school science curriculum. *Cognition and Instruction 11*(1), 1-29.

Hautamäki, J., Kuusela, J. and Wikström, J. (2002) *CASE and CAME in Finland: 'The second wave'.* Harrogate: 10th International Conference on Thinking.

TIMELINE

	School Year 2012-2013	School Year 2013-2014			School Year 2014-2015			2015
	Summer Term	Autumn Term	Spring Term	Summer Term	Autumn Term	Spring Term	Summer Term	Autumn Term
Recruit schools								
Randomly assign schools to treatment								
Training								
Observations								
Analyse KS2 data								
Science test (end Y8)								
English & Maths tests								
Cognitive devt test								
Process & analyse scores								
Online teacher survey								
Online student survey								
Analysis phase								
Write final report								