Butterfly Phonics

Key Conclusions

1. This evaluation provided evidence of promise; there was a positive, statistically significant effect on the primary outcome measure of reading comprehension. However, this effect size was lower than the minimum detectable effect size of the trial, so we cannot confidently conclude that the effect was due to the intervention and did not occur by chance.

2. The secondary outcome measures indicated positive impacts on children’s literacy skills, but these were not statistically significant.

3. This intervention is recommended to take place during the school day, when it is easier to secure sustained co-operation and support from school staff. Where that support was present, the intervention was able to progress more satisfactorily than in schools where it was lacking.

4. Schools should ensure that people delivering the intervention receive training in the Butterfly method so that it is implemented as intended.

5. Further research could investigate the intervention’s impact on early readers. Its emphasis on larger word units and comprehension skills might enable a more rapid progression in early reading than a pure phonics course.

What is the impact?

The evaluation found that, on average, the reading comprehension skills of pupils who received the intervention improved at a faster rate than those in the control group. This improvement is equivalent to the pupils who received the intervention making an additional five months’ progress over the course of the school year. This estimate is statistically significant, but the observed effect size is lower than the minimum detectable effect size that was specified at the beginning of the study. This means that, although this evaluation provides evidence of promise, we are unable to confidently conclude that the observed effect is real and did not occur by chance.

The evaluation also considered the impact of the intervention on two secondary outcome measures of literacy skills, but did not find statistically significant impacts.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>NO. OF PUPILS</th>
<th>EFFECT SIZE (95% CONFIDENCE INTERVAL)</th>
<th>ESTIMATED MONTHS' PROGRESS</th>
<th>EVIDENCE STRENGTH</th>
<th>COST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention vs control (all pupils)</td>
<td>310</td>
<td>0.43 (0.03 to 0.84)</td>
<td>+5 months</td>
<td>4</td>
<td>£5</td>
</tr>
<tr>
<td>Intervention vs control (FSM)</td>
<td>140</td>
<td>0.16 (-0.18 to 0.49)</td>
<td>+2 months</td>
<td>N/A</td>
<td>£5</td>
</tr>
</tbody>
</table>

How secure is the finding?

The primary analysis in this evaluation is judged to be of weak security and was awarded a security rating of 0 padlocks. The main cause of the low security rating awarded to this evaluation is that the differences between schools’ post-test results were greater than envisaged at the beginning of the trial. This meant that the trial was not large enough to confidently detect an effect as small as the one that was ultimately observed. We cannot therefore confidently conclude that the observed effect is real and did not occur by chance.

A further limitation of the study is that the test administrators reported that the post-tests in two schools were
disrupted by the poor behaviour of the pupils involved. It was judged reasonable for the pupils to re-sit the test in one of these schools and the primary analysis, which was performed on an ‘intention to treat’ basis, used this re-sit data. It was not judged reasonable for the pupils in the other school to re-sit the post-test and therefore the data from the disrupted post-test session was included in the primary analysis. A report about the compromised administration conditions of the post-test in this school is included in Appendix 2. These disrupted post-tests and their inclusion in the final analysis should be considered when interpreting the security of the findings.

It is possible that the positive impact was caused by a ‘confounding’ factor. Most of the participating schools scheduled the intervention to take place outside English lessons and some of the effect was possibly due to the pupils in the intervention group receiving additional English teaching, not the nature of the Butterfly Phonics intervention itself. Also, pupils received the intervention in small groups, while pupils in the control group continued with normal classroom teaching.

The schools involved were located within a small geographical area of London, and care should be taken when applying these findings to schools in different contexts.

To view the project's evaluation protocol click here.

How much does it cost?

The cost of Butterfly Phonics as it was delivered in this evaluation is estimated at £108.50 per pupil. This estimate is based on the assumption that there are eight pupils in each Butterfly Phonics class, and includes the salaries of the Butterfly Phonics teaching staff, training costs and the cost of course books (about £10 each). The cost of the Butterfly teaching staff for a one hour class was £35-40 and the specialist training from Irina Tyk cost £600 for half a day.