# Helping to Raise Primary School Attainment in Disadvantaged Schools Through the Use of Evidence

**Stephen Gorard, Nadia Siddiqui, Beng Huat See**

**Durham University Evidence Centre for Education**

**DOI: https://dx.doi.org/10.47772/IJRISS.2025.910000339**

**Received: 12 October 2025; Accepted: 20 October 2025; Published: 12 November 2025**

## ABSTRACT

This paper describes a use-of-evidence intervention implemented by eight primary schools in and around Durham County, in the North East of England. The researchers provided schools with a menu of evidence-based approaches to improve attainment, especially for disadvantaged pupils. Collectively, the schools chose Learning by Questions (LbQ) software to improve maths for Year 4, and Student Tutoring to improve English in Year 5. Half of the schools were randomised to each approach for each year group. Data collection, especially of the prior attainment scores, was affected by Covid lockdown. LbQ was the most feasible approach, with schools given help by the developers, and appreciated by teachers and pupils. But the approach showed little or no impact on test scores in the short term. The Tutoring intervention was more complex with local university students having to travel to schools in remote areas, fit in with timetables, and develop activities according to school priorities. The students also had other commitments. Nevertheless, there is evidence that the treatment group improved more than the control. Overall, the use-of-evidence intervention is promising based on this first pilot trial. The next step is to scale it up with a more varied set of options, a larger scale, and over a longer time period.

## INTRODUCTION

It sounds sensible to say that education should be informed by research evidence. This is not to say that evidence should dominate. But that once the policy/school context has determined the issues to be faced, and before the professional judgement of policy-makers/practitioners is used to help carry out any plan, good evidence should presumably inform the decision. However, plausible as this sounds, there is as yet little convincing evidence that using research evidence does improve real-life educational outcomes (Gorard and Chen 2025).

In England, as in many other areas, there is considerable pressure for educators to be research-informed, and for their practice to be evidence-led. However, even 20 or more years since the inception of a “what works” revolution, the actual evidence is sparse on how to implement evidence use successfully (Flynn 2019, Nuttley et al. 2019). A recent very large-scale review of how to get evidence into use found that there was no clear answer on how evidence transfer from research to use could be best achieved (Gorard et al. 2020a). It has not even been convincingly demonstrated that using research evidence in education does actually improve educational outcomes.

Many primary schools in the North East region of England have pupil intakes with high levels of disadvantage, including many pupils who will go on to have been persistently disadvantaged throughout their school careers. This is then linked to their lower average attainment at each Key Stage, and therefore to a serious poverty attainment gap.

Schools North East, a charity with universal membership of schools across the NE region, reports that many teachers in schools would appreciate advice from education specialists to help identify and implement high quality, evidenced-based interventions that show promise of raising attainment overall, and especially for disadvantaged pupils. They report that teachers are confused over the strength of evidence for, and the impact

of, any activity, for example. Therefore, schools might unwittingly, and in good faith, choose inappropriate interventions that will not be suit their pupil's needs.

Staff from the Evidence Centre for Education (ECE) have therefore been using their research and evaluation expertise to help teaching staff in local primary schools to identity, implement and assess promising evidenced-based interventions for raising primary attainment and reducing the poverty attainment gap. The interventions were selected and agreed with schools according to their needs, from a toolkit of most promising approaches (or "best bets"), devised by our staff and based on our own evaluations, extensive structured reviews of evidence, and work done by/with the Campbell Collaboration, Education Endowment Foundation, and the US Institute of Education Science.

**The intervention(s)**

We offered participating local primary schools a simple "menu" of evidence-led approaches or programmes. These were selected because they are tested, feasible and show some promise of enhancing pupil outcomes, especially for disadvantaged pupils. They are also cheap or free to implement. We asked schools to pick a few options each.

Table 1 presents a summary of eight possible approaches to improving overall attainment in literacy or numeracy in primary school.

**Table 1** – the menu of possible evidence-informed approaches.

| Theme or programme | Outcome | Year group(s) | Estimated impact | Evidence strength | Cost per pupil |
|---|---|---|---|---|---|
| Accelerated Reader | Literacy | 1-5 | 0.20 | 3□ | £10 |
| LbQ | Literacy, numeracy | 2-6 | 0.10 | 2□ | 0 |
| Enhanced oral feedback | Literacy, numeracy | 1-4 | 0.20 | 3□ | 0 |
| Student tutoring | Literacy, numeracy | 1-6 | 0.20 | 3□ | 0 |
| Peer tutoring | Literacy, numeracy | 1-6 | 0.20 | 3□ | 0 |
| Dialogic teaching | Literacy, numeracy | 1-5 | 0.20 | 3□ | £10 |
| Texting parents | Literacy, numeracy | 1-6 | 0.10 | 3□ | 0 |
| Self-affirmation | Literacy, numeracy | 2, 5, 6 | 0.10 | 2□ | 0 |

Each approach was linked to the following attributes:

Name or description. Some approaches are generic, while some are more specific protocols or pieces of software. The assumption is that schools will pick options that they are not already doing, or not doing systematically. The outcome of interest. All approaches concern improving literacy, and all but one also concern improving numeracy. These could be assessed in terms of KS results, bespoke tests, and attitude or enjoyment surveys. The year groups suggested. In general, it is more efficient to use standardised test scores as one outcome, or for a baseline figure. So, Years 2, 3 and 5 (if there is a formal KS2-based mock) or 6 are preferred. But we can also provide additional tests in literacy/numeracy for some programmes. One programme (self-affirmation) is designed to be used only with pupils who are approaching a high stakes test. Estimate of promise. The likely impact or benefit is presented as an estimated "effect" size based on the strongest prior evaluations. In education, an "effect" size of around 0.2 is common for those programmes that do seem to make a difference. Here, we only propose ideas that are promising. Many other programmes, often promoted to schools, make no difference, and a few have been found to be harmful. There are, of course, programmes with larger recorded impacts but many of these may be expensive or otherwise not feasible here. Strength of evidence. The padlock rating from 0 to 4 gives an indication of how strong the prior evidence is for each approach. 4 would mean the evidence is as strong as could be imagined in real-life, and 0 means there is no trustworthy evidence at all. This rating is independent of the likely impact of any intervention. Cost per

pupil. Some interventions are naturally free of charge. Some can be offered free as part of this menu, due to the co-operation of the developer with the University Evidence Centre for Education. Others would require use of funds, perhaps via school Pupil Premium funds.

There are, of course, many other evidence-informed approaches. Some would be more expensive than those listed here, some require extensive training or are complex to implement in the time available. For example, while promising, some interventions for reading comprehension, metacognition, formative assessment require teacher training and are difficult to get right.

The schools were most keen on using Student Tutoring, followed by Learning by Questions (LbQ). What follows is an outline or pen portrait of each of these.

**Learning by Questions (LbQ)**

LbQ is an online tool with curriculum-aligned Question Sets for maths, English and science. Here we propose use only for maths and/or English. Each of these questions comes with immediate feedback. LbQ thus provides continuous formative assessment resources to give teachers insights into learning. It also provides immediate feedback to students, which is personalised to the individual student. LbQ states that it reduces teacher workload and improves student learning, with automatic marking, and instant insight for effective interventions. See their website: https://www.lbq.org

**How does it work?**

Teachers access a cloud-based repository of 60,000 + questions arranged into 1,800 structured Question Sets and organised by subject, topic and year group. Up to three Question Sets can be selected simultaneously and set as tasks. Teachers select and launch Question Sets which students work through during lessons, or by themselves.

Pupils work at their own pace and can retry questions after receiving system-generated instant feedback, hints and reminders, where answers are wrong. The higher the ability, the faster LbQ moves them forward to more challenging questions, so that everyone is working at an appropriate level of pace and challenge. Answers are analysed in real-time and relayed to the teacher's device where struggling pupils and challenging questions are easily identified. Teachers can intervene, teach and plan ahead, without marking. LbQ has built-in tools to support adaptation of Question Sets re-teaching. Lesson data is stored automatically to aid planning and interventions.

**What do you need?**

Internet connection - Pupils equipped with (almost) any internet connected device can access and work through tasks. LbQ will run on devices such as tablets, Chromebooks, laptops, and desktops using iOS, Windows and Android operating systems. Tasks can be run in a web browser or the free LbQ App available for Apple, Android and Windows.

The video clip below shows how it works.

https://www.youtube.com/watch?v=yyo8xYIFqOU

What does it usually cost?

The package is free for schools in this project, but the usual subscription fees are:

- £250 per teacher per year for access to all subjects and all years.
- £625 per teacher for 3-years for access to all subjects and all years.

What training is available?

- Online and telephone support is included with the subscription.
- After signing up, the school receives an administration dashboard enabling management of teacher accounts in the school.
- Teachers are able to run the product without specialist training

Suggested advantages of LbQ

- Reduces teacher workload in terms of lesson planning and marking
- Improves or supports feedback in real-time
- Learning is individualised
- Provides banks of relevant example questions
- Allows comparisons of data about students/classes
- Supports non-specialist teachers
- Tracks pupil achievement and progress
- Reporting and analysis of pupil/class performance

**Prior evidence**

There is as yet little strong evidence on LbQ. There have been a number of school-led evaluations of LbQ on a range of topics, many of which are small-scale and conducted by individual schools. There are two more robust studies using a randomised controlled design relevant to LbQ, one for maths and one for grammar (Sheard and Chambers 2011, Sheard, Chambers and Elliott 2012). Both studies were for Year 5 pupils. The predecessor of LbQ is Questions for Learning (QfL). Note that there is a link between these evaluators and the LbQ developer, which may represent a conflict of interest.

| Report | Year group | Duration | Outcome | Effect size |
|---|---|---|---|---|
| Sheard et al. (2012) | Year 5 | 12 weeks | Grammar<br>Writing | +0.16<br>But schools using the device as recommended by the developer produced better results. +0.27<br>No effect on writing |
| Sheard & Chambers (2011) | Year 5 | 12 weeks | Maths | +0.39 (equivalent to perhaps an additional 3-month progress over a year) |

**Student tutoring**

Student tutoring involves trained university students providing additional academic support to pupils who are struggling in mathematics. Low-achieving students receive one-to-one or small group tuition by paid volunteers. Student tutoring assists students in academic achievement by providing tailored attention and support to address individual academic requirements and difficulties.

**How does it work?**

Student tutoring involves recruiting and training undergraduate students from Durham University, as part of their widening access programme. The programme evaluates applicants' communication, interpersonal, and teaching skills and provides two full-day training sessions to tutors.

Before tutoring begins, class teachers are required to identify pupils who are working insecurely at or below age-related expectations in English to receive tutoring. Tutors use materials provided by the class teacher or design their own session plans under the guidance of the class teacher to cater to individual needs.

Targeted pupils receive one hour tuition per week on a 1:1, 1:2, or 1:3 basis during one semester, with the timing of the sessions based on the schools' requirements. Tutoring sessions are held in the participating schools' libraries, resource rooms, and other common areas. Tutors and teachers work closely together throughout the tutoring period, frequently communicating to ensure that the expected goals are met and any necessary alterations to the tutoring sessions are implemented.

Example project: Tutor Trust https://www.thetutortrust.org/

What do you need?

- Contextual information about the pupils (gender, Pupil Premium/FSM/LAC/EAL status, attendance).
- Baseline attainment data: KS1 attainment, mock SATs score, aim of tuition (e.g. Age-Related Expectations, ARE).

What does it usually cost?

This is free for schools in the project (based on Durham volunteers, and paying for their travel), but the usual tuition fee could be:

- £108 per pupil at a 1:3 basis at a block of 15 hours of tutoring in mainstream schools.
- up to a maximum of £10.80 per pupil per hour in mainstream schools.

What onboarding/training is available?

No additional training for class teachers.

Suggested advantages of student tutoring

- Affordable tutoring.
- High-quality tutors.
- Provides individualised assistance for mathematics attainment.
- Reduces workload of class teachers by improving low-achieving students.
- Bridges the achievement gap inside class.

**Prior evidence**

Many studies have evaluated the effectiveness of student tutoring, most of which are small scale at the school level. Overall, there are signs of promise (Slavin 2021).Two studies conducted randomised control trials where at-risk pupils randomly received tutoring or not. The first study randomised 105 primary schools in England, involving more than a thousand students from Year 6. The second study randomised 550 students from grade 4 to 8 in 12 schools. Both studies have promising results in improving mathematics achievement of low-achieving students.

| Report | Year group | Duration | Outcome | Security | Effect size |
|---|---|---|---|---|---|
| Torgerson et al. (2018) | Year 6 | One hour per week for 12 weeks | Mathematics | 4□ | +0.20 Overall<br>+0.25 FSM |
| Parker et al. (2019) | Grade 4-8 (U.S.) | One hour per week for 12 weeks | STAR Mathematics | 3□ | +0.20 Overall<br>+0.40 Grade 4<br>+0.00 Grade 5<br>+0.20 Grade 6 |

**Design and methods used in this new evaluation**

This pilot trial of a use-of-evidence intervention involved eight primary schools in and around County Durham, who joined the School Membership Scheme run by Durham University Access and Engagement Group. The intention had been to start the project earlier in the school year, but changes to staff in the Access Group meant that the interventions were run for less than a complete term before the final test was delivered.

The schools were given the menu and information summarised above. Most schools selected student tutoring as being of interest. There was also some interest in LbQ given that it was provided free. Other programmes were less popular or fit less well with school priorities at this stage. So, Tutoring and LbQ were selected as cross-controlled interventions.

Four schools were randomised to use Student Tutoring in their Year 5 class(es) to improve English. Their Year 4 classes acted as a control group for the other intervention. The other four schools were randomised to use Learning by Questions (LbQ) in their Year 4 classes for maths. Their Year 5 classes acted as the control for Tutoring.

We asked schools for some background details about all pupils in Years 4 and 5, and for their prior Key Stage 1 results. These were to be used as baseline data to assess the comparability between schools in the two groups for each intervention evaluation. The Year 5 pupils were one of the age cohorts affected by Covid lockdown during their KS1 Assessment period, and so two schools had no KS1 scores. This severely limits our ability to compare pre- and post-test scores for Tutoring especially. Other schools sent the results using their own classifications and codes. We converted all scores or codes to a common score, using the website https://smartgrade.zendesk.com/hc/en-gb/articles/15905913158802-Standardised-grades-in-Smartgrade, and then converted all to standardised z-scores. The pre-test scores would be improved by using the National Pupil Database in any future trials.

The outcome measures were based on GL Assessment's Progress Towards English (PTE) for Year 5, and Progress Toward Maths (PTM) for Year 4. All pupils in the relevant year, and in both treatment and control groups, sat the test. In the Tutoring intervention only a small number of pupils, selected as needing improvement, were put forward by the schools. However, only some schools were prepared to link the results to these tutees, meaning that the headline analysis compares all pupils in treatment and control schools. This would dampen the scale of any effect size if the intervention were found to be effective

**Analysis**

The headline results for each programme are presented in terms of prior (KS1) attainment where this is available, post-intervention PMT or PME scores, and progress scores based on z-score gains from pre- to post-test. The differences between groups are converted to Cohen's d effect sizes. The overall results for the use-of-evidence intervention, regardless of year or subject, are presented only in terms of post-test scores.

**The intervention**

This use-of-evidence (menu) intervention involved pupils in Years 4 and 5.

Participating schools completed an IT questionnaire for LbQ, so that the software could be installed and tested. Participating teachers were trained either on-line or face-to-face on how to use the features of LbQ.

The eight student tutors were MA and BA students from Durham University who were supposed to visit one school each week to help a small group of low-attaining pupils. Not all visits were conducted, but some tutors visited more often. They worked with pupils chosen by the schools, on topics also chosen by the schools. Sometimes the schools offered ideas and resources, but more often the students created their own.

## Results

Eight schools participated in the programme, with a total of 418 pupils in Years 4 and 5. There were 165 in the schools selected for LbQ, and 253 in the schools selected for Tutoring (Table 2).

**Table 2 -** Number of pupils by intervention school and year group

| School group | Year 4 | Year 5 | Total |
|---|---|---|---|
| LBQ intervention/Year 5 control | 83 | 82 | 165 |
| Tutoring intervention/Year 4 control | 124 | 129 | 253 |
| Total | 207 | 211 | 418 |

## Baseline LbQ

In the LBQ trial, the two groups are reasonably well balanced in terms of background characteristics (Tables 3 to 7). Where there is imbalance, this is largely due to one or more schools having missing or unknown data on a specific pupil characteristic, and this being predominantly in one group. For LbQ, the groups are well-balanced in the proportion of pupils with SEND, but the treatment group clearly have more FSM-eligible pupils. This could affect the outcomes.

**Table 3 –** Percentage of each sex in LbQ treatment groups

| Sex of pupils | LbQ | Control | Total |
|---|---|---|---|
| Female | 39.8 | 34.7 | 76 |
| Male | 43.4 | 44.4 | 91 |
| Other | 16.9 | 21.0 | 40 |
| Total | 83 | 124 | 207 |

**Table 4 –** Percentage of each ethnicity in LbQ treatment groups

| Ethnicity | LbQ | Control | Total |
|---|---|---|---|
| White British | 80.7 | 74.2 | 159 |
| Unknown | 16.9 | 21.0 | 40 |
| Total | 83 | 124 | 207 |

**Table 5 –** Percentage of FSM status in LbQ treatment groups

| | LbQ | Control | Total |
|---|---|---|---|
| FSM eligible | 56.6 | 41.1 | 98 |
| Not FSM | 43.4 | 29.0 | 72 |
| Unknown | 0 | 29.8 | 37 |
| Total | 83 | 124 | 207 |

**Table 6 –** Percentage of EAL status in LbQ treatment groups

| | LbQ | Control | Total |
|---|---|---|---|
| EAL | 1.2 | 2.4 | 4 |
| Not EAL | 31.3 | 17.7 | 48 |

| | | | |
|---|---|---|---|
| Unknown | 66.3 | 79.8 | 154 |
| Total | 83 | 124 | 207 |

**Table 7 –** Percentage of SEND status in LbQ treatment groups

| | LbQ | Control | Total |
|---|---|---|---|
| SEND | 30.1 | 29.8 | 30.0 |
| Not SEND | 69.9 | 49.2 | 119 |
| Unknown | 0 | 21.0 | 62 |
| Total | 83 | 124 | 207 |

The difference in FSM-eligibility between the two groups is reflected in the lower prior attainment of the LbQ treatment group, with a pre-test "effect" size of -0.22 (Table 8). A relatively large number of prior KS1 scores are missing, and this reduces the strength of the comparison.

**Table 8 –** Maths KS1 pre-score by LbQ treatment groups

| | LbQ | Control | Overall standard deviation | Effect size |
|---|---|---|---|---|
| Maths pre-score KS1 | -0.28 | -0.01 | 1.22 | -0.22 |
| N | 46 | 80 | | |

**Baseline Tutoring intervention**

The groups are slightly less balanced in terms of pupil characteristics for Tutoring than for LbQ (Tables 9 to 13). Most importantly, there are around twice as many FSM-eligible pupils in the control group as in the treatment. This is likely to have implications for the attainment results.

**Table 9 –** Percentage of each sex in Tutoring treatment groups

| Sex of pupils | Tutoring | Control | Total |
|---|---|---|---|
| Female | 44.2 | 30.5 | 82 |
| Male | 31.8 | 42.7 | 76 |
| Other | 24.0 | 26.8 | 53 |
| Total | 129 | 82 | 211 |

**Table 10 –** Percentage of each ethnicity in Tutoring treatment groups

| Ethnicity | Tutoring | Control | Total |
|---|---|---|---|
| White British | 73.6 | 67.1 | 150 |
| Unknown | 24.0 | 26.8 | 53 |
| Total | 129 | 82 | 211 |

**Table 11** – Percentage of FSM status in Tutoring treatment groups

| | Tutoring | Control | Total |
|---|---|---|---|
| FSM eligible | 31.0 | 59.8 | 89 |
| Not FSM | 45.0 | 40.2 | 91 |

| | | | |
|---|---|---|---|
| Unknown | 24.0 | 0 | 31 |
| Total | 129 | 82 | 211 |

**Table 12 –** Percentage of EAL status in Tutoring treatment groups

| | Tutoring | Control | Total |
|---|---|---|---|
| EAL | 0 | 4.9 | 4 |
| Not EAL | 24.8 | 79.3 | 97 |
| Unknown | 75.2 | 15.9 | 110 |
| Total | 129 | 82 | 211 |

**Table 13 –** Percentage of SEND status in Tutoring treatment groups

| | Tutoring | Control | Total |
|---|---|---|---|
| SEND | 19.4 | 22.0 | 43 |
| Not SEND | 56.6 | 78.0 | 137 |
| Unknown | 24.0 | 0 | 31 |
| Total | 129 | 82 | 211 |

The difference in FSM-eligibility between the two groups is not reflected in the slightly lower prior attainment of the Tutoring treatment group, with an "effect" size of -0.02 (Table 14). Even more KS1 scores are missing here, with only two schools providing scores, seriously affecting the strength of the pre-test comparison.

**Table 14 –** English KS1 pre-score by Tutoring treatment groups

| | Tutoring | Control | Overall standard deviation | Effect size |
|---|---|---|---|---|
| English pre-score KS1 | 0.12 | 0.14 | 0.71 | -0.02 |
| N | 32 | 46 | | |

Outcomes for LBQ

Using all available post-test scores, the LbQ group has lower attainment than the control, with an "effect" size of -0.23 (Table 15). This is roughly the same as the pre-test difference, suggesting that LbQ has had no impact on maths scores for these Year 4 pupils. However, the number of cases with valid scores pre- and post is different, making comparison difficult.

**Table 15 –** Progress toward Maths scores by LbQ treatment groups

| | LbQ | Control | Overall standard deviation | Effect size |
|---|---|---|---|---|
| PTM score | 94.38 | 97.88 | 15.43 | -0.23 |
| N | 58 | 60 | | |

**Outcomes for Tutoring**

At post-test the Tutoring treatment group was ahead, with an "effect" size of +0.28, having been behind at the outset. This suggests that Tutoring had a beneficial impact for these Year 5 pupils (Table 16). However, as with LbQ, the number of cases with valid scores pre- and post is different, making comparison difficult. Only some of the pupils in the treatment group actually took part in tutoring, meaning that any effect size for the whole group is likely to under-estimate any impact. However, having some lower attaining pupils tutored by students would also give the class teacher slightly more time to work with the remainder.

**Table 16 –** Progress toward English scores by Tutoring treatment groups

| | Tutoring | Control | Overall standard deviation | Effect size |
|---|---|---|---|---|
| PTE score | 96.61 | 92.92 | 13.029 | +0.28 |
| N | 82 | 62 | | |

**Outcomes for the overall use-of-evidence intervention**

Combining the scores for both maths and English scores across Years 4 and 5, we can address the issue of whether the treatment groups (i.e. those whose teachers took part in any use-of-evidence) performed better than the overall control groups (Table 17). There is a small positive effect size for the post-test, which is promising given that pupils in both LbQ and Tutoring treatment groups were behind their respective controls at KS1.

**Table 17 –** Post-test scores by overall treatment groups, combined maths and English scores

| | Use-of-evidence | Control | Overall standard deviation | Effect size |
|---|---|---|---|---|
| Test score | 95.69 | 95.31 | 14.11 | +0.03 |
| N | 140 | 122 | 262 | |

**Process data**

**LbQ**

LbQ was implemented in four primary schools. One embraced the software, and it was used at least twice per week. The deputy head who was leading on this changed schools near the end and is now encouraging its use by the new school. Two schools were enthusiastic. One of these was somewhat constrained by being part of another initiative for maths, but they have now rolled out LbQ use to other subjects and all years. As far as they were concerned it was a success. The fourth school clearly had other priorities and did not engage well. The teachers there were more “conscripts” than “volunteers” according to one observer.

In general, the use of LbQ correlated with the range and quality of schools’ IT. Despite being available for a range of platforms, where schools have poor wifi or limited numbers of tablets, LbQ is hard to implement properly. It would have been better to implement over a longer period, as use takes some time to establish faithfully, and there are inevitable delays created by the setup operations. In this pilot, schools really only had access for six weeks of the summer term.

One deputy headteacher who participated enthusiastically felt that pupils really benefitted from instant feedback, and said:

I wish I'd had this throughout my career.

Some children who were difficult to engage were typically on-task and focussed on their learning when completing an activity using LbQ. Teachers appreciated the flexibility afforded by LbQ:there isn't just one way of using it and it doesn't rely on a specific pedagogical approach.

It's not just what they can do in the lesson today, but also what they can do tomorrow. LbQ's Assessment for Learning helps me to see both of these things better, so I can make better decisions as a teacher - I'm not going to rush them on if they're not ready.

One school also cited a benefit from LbQ in terms of inclusivity. They discussed how children can be very aware of the tasks they and other children in the class are being set. As a result, differentiation can have a

negative effect on some children's self-esteem. The ability to set children tasks appropriate to support their next learning steps without having to 'signpost' this within the lesson delivery was felt by some teachers to have a positive impact on the self-esteem of some of the less confident learners in the class.

**Tutoring**

We have tutoring records based on ten pupils tutored by one BA and one MA student, from 22nd April to 16th July 2024. Extracts are summarised here to illustrate the organisation of student tutoring, challenges that student tutors encountered, reaction of tutees, and perceptions of teachers. Of course, the results and the implementation of tutoring are heavily dependent on the skills and energy of the student tutors. These two tutors would appear to be among the most proficient.

Student tutoring supports Year 5 pupils in their English language learning outcomes, mainly for those who were in danger of being left behind. Tutors received a list of pupils who needed tutoring and split them into groups for one-hour tutoring for about eight weeks. Teachers generally showed enthusiasm for the project and supported tutors to better understand students' needs and organise the appropriate activities. In the first session, tutors spent some time with teachers and pupils to know them better.

We began the first session by meeting the teacher and the students and getting to know them and helping them feel comfortable and confident in the new setting.

I spent a significant amount of time with the teacher, asking about the kinds of activities they do and the level of the students.

The tutor mentioned an activity to establish close relationships with pupils:

To break the ice, I made the first meeting fun by asking about their hobbies, the meanings of their names, and their most memorable day and why.

During tutoring, the tutor organised diverse activities to support pupils' English learning with adaptation of the teacher's instructions. They provided targeted support based on pupils' reading abilities and writing. The activities and materials used for tutoring were quite flexible. Tutors could also adjust their teaching focus according to pupils' reaction and organise different activities according to their tutoring themes.

At least three students mentioned they find it difficult to use adjectives and adverbs in sentences, so I plan to work on this area later.

By the end of the session, the students asked for additional support on the use of commas, which I will incorporate into the next lesson. They conducted the lessons progressively, moving from simpler to more complex concepts step by step. The tutor's teaching logs provided detailed examples of how they organised the tutoring. Here is an example extracted

from the teaching log.

Students were divided into teams and tasked with adding commas to a series of sentences. The sentences were:

- The forest is home to deer, squirrels, and birds.
- On our hike, we saw tall trees, colorful flowers, and a clear stream.
- We packed sandwiches, fruit, and water for the picnic.
- The leaves were red, yellow, and orange in the fall.

Then they were given a paragraph and asked to insert commas where needed. The paragraph was:

'While walking through the dense forest, we saw many interesting things. There were tall pine trees, beautiful ferns, and chirping birds. As we walked, we heard the rustling of leaves, the croaking of frogs, and the distant sound of a waterfall. It was a peaceful, quiet, and refreshing experience.'

The tutor started by teaching pupils how to add commas to sentences. When they understood the usage of commas and corrected their mistakes, they moved on to adding commas to a paragraph, which was more complex.

For abstract themes such as nouns, adjectives, and adverbs, the tutor tried to use pictures in daily life to help tutees. For instance, the tutor showed a photo of a tree and required pupils to write down nouns, adjectives, and adverbs. In addition, other activities were arranged to better engage tutees, such as brain exercises such as word searches, collective posters, brainstorming ideas about pictures, dice game for discussions (Figure 1), and crossword puzzles (Figure 2). The tutor also invited tutees to reflect on their learning with learning logs (Figure 3).

**Figure 1** Dice game for discussion

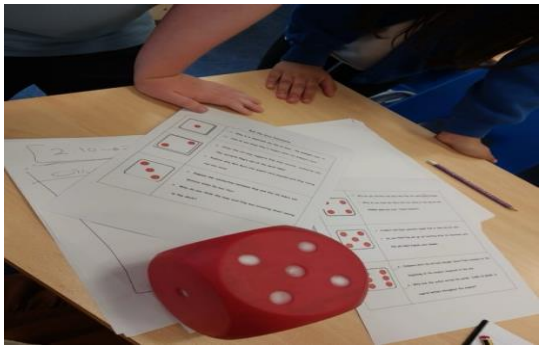
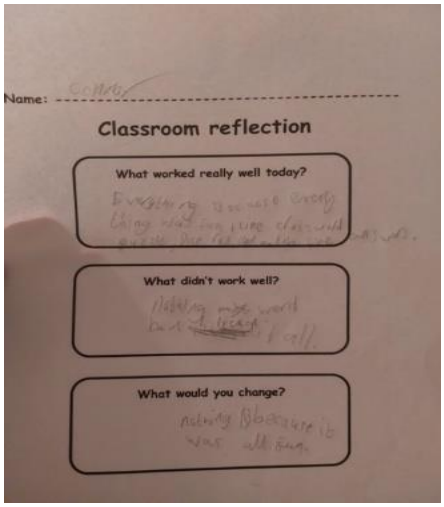


**Figure 2** crossword puzzles


**Figure 3** Example of a learning log




Tutees showed enthusiasm towards student tutoring sessions and joined in the activities promptly. The tutor noted that tutees showed more confidence in 'their understanding of things which they had worked on' and felt

comfortable ‘asking their teacher or another student for help if they did not understand something in the future’.

**Challenges**

The programme encountered multiple challenges in timing and organisation, largely due to scheduling conflicts with holidays, school events, and adjustments to the school’s academic calendar. There was an initial delay preventing many tutoring sessions from beginning until the 22nd of May, leaving only eight weeks available, one of which coincided with a bank holiday.

Some delays and timing issues with holidays meant that the first session could not begin until the 22nd of May, and this left only 8 weeks, one of which was a bank holiday, to work with the student.

There were challenges with timing and scheduling due to holidays and other school events, which affected the number of sessions available. The time for student tutoring was sometimes changed to better fit school events. This adjustment, however, seemed to disturb the students, as the deviation from their usual schedule left them unsettled.

The week before this was the school’s half term, and then they had an inset day on the Monday, so we ran this session on the Tuesday instead, however I think this change in routine did unsettle the pupils slightly.

There was an instance where changes in the classroom environment, such as the presence of a new tutor or different group sizes, affected the focus and progress of the students.

Another tutor came to work with a different group, so we had two groups of three working together in each session. I think this was slightly disruptive to the progress of my groups as it was a different environment to what they were used to and so they were less focused.

In addition to timing and tutor coordination, the tutors themselves faced challenges in providing appropriate support tailored to the pupils’ varying needs and learning levels. Unlike experienced teachers, these tutors needed to coordinate closely with school staff to set clear goals and approaches for each session. In addition to time and tutor management, for student tutors, how to teach pupils and provide adequate support might be a challenge. Pupils had different characteristics, and the tutors should make quick responses to their actions. Tutors were not teachers, so they needed to cooperate with school teachers to settle their teaching tasks. They may also have different understandings and estimations for students’ reading ability.

The tutoring scheme had some issues with setting up in the beginning, mostly because it was new and so expectations of what the students would be working on were undefined and had to be discussed.

Although I expected to start with writing tutoring, the teacher asked me to follow the students' readings and support them. This confused me, but I did it anyway. … When I asked the students about their thoughts on literacy skills, they were enthusiastic, which puzzled me since they needed support in that area.

At the end of the project, both the tutors summarised the challenges and potential improvement of student tutoring. They noted that the initial uncertainty about expectations and the difficulty of establishing routines and a conducive learning environment made the process particularly demanding, especially given the short implementation period. They suggested that closer collaboration and planning with teachers, as well as more experience, could help tutors understand their tutees better and provide more targeted support. Additionally, they emphasised the importance of adequate training in working with this age group to enhance the effectiveness of the sessions.

I would plan more closely with the teacher. I found that I needed to spend more time discussing the students' levels and the materials she found useful. Planning together would be more beneficial. (Student tutor reflection)

I believe it would be more helpful if the children received tutoring for the entire school semester rather than just eight sessions. Time and relationship with students are important to achieve better results. (Student tutor reflection)…working with the students could be a challenge and this could be mitigated through more training and support in how to tutor and work with this age group as it was new to me. (Research assistant reflection)

## DISCUSSION

Use-of-evidence by schools is a difficult area to assess. It must involve schools (and teachers) themselves selecting approaches or programmes that are promising in terms of existing evidence. Schools may make mistakes. Or, as perhaps happened here, they may select some approaches that do not then appear to work better than business as usual in their context. Promising evidence of effectiveness is not a guarantee that a much programme will always work in the future and in all contexts. So, it is important that schools select a range of evidence-informed approaches. The overall treatment here is therefore about schools selecting promising approaches. They were guided by a short menu prepared by the researchers. In wider use, such a menu may not be feasible, or there would be competing sources each offering different choices and different ways of presenting evidence summaries. If evidence-use is seen as desirable, school leaders and other stakeholders therefore need to be better prepared to make appropriate independent judgements about the quality of evidence and the promise of different approaches (Gorard et al. 2020).

The ideal would be to select programmes with the desired outcome (e.g. literacy/numeracy), that have the greatest impact, especially for lower attainers, backed by the strongest evidence, at the lowest cost. It is not clear that this occurred in this pilot trial. In reality, schools will decide on the basis of their priorities and context.

The use-of-evidence intervention was shorter than intended, and was anyway only planned for a part of a school year. In the next phase of research, the study should be longer and larger – both in terms of the number of teachers/schools, and the range of evidenced approaches used.

If schools are happy to provide identifiers for all pupils, in a follow up trial, then we can link these to the National Pupil Database in England to look at the eventual KS2 results for treatment and control groups, and even to later life educational outcomes.

## ACKNOWLEDGEMENTS

The authors would like to thank all of the schools for taking part, the student ambassadors, the executives of Learning by Questions for donating free use of the software for this work, and the Durham University Access Group for their help. The work emerged through a discussion with the Vice-Chancellor at Durham University.

## REFERENCES

1. Flynn, N. (2019) Facilitating evidence-informed practice, Teacher Development, 23(1), 64-82
2. Gorard, S. (2022) What is the evidence on the impact of Pupil Premium funding on school intakes and attainment by age 16 in England?, British Educational Research Journal, 48, 3, 446-468, What is the evidence on the impact of Pupil Premium funding on school intakes and attainment by age 16 in England? - Gorard - - British Educational Research Journal - Wiley Online Library
3. Gorard, S. and Chen, W. (2025) What is the evidence on research-informed education?, Chapter 2, pp.55-76 in Wyse, D., Baumfield, V., Mockler, N and Reardon, M. (Eds.) The BERA/SAGE Handbook of Research-Informed Education Practice and Policy
4. Gorard, S., See, BH and Siddiqui, N. (2020) What is the evidence on the best way to get evidence into use in education?, Review of Education, 8, 2, 570-610
5. Gorard, S., See, BH and Siddiqui, N. (2020) What is the evidence on the best way to get evidence into use in education?, Review of Education, DOI: 10.1002/REV3.3200
6. Nutley, S., Boaz, A., Davies, H. and Fraser, A. (2019) New development: What works now? Continuity and change in the use of evidence to improve public policy and service delivery, Public Money & Management, 39:4, 310-316

7. Parker, D., Nelson, P., Zaslofsky, A., Kanive, R., Foegen, A., Kaiser, P., & Heisted, D. (2019). Evaluation of a math intervention program implemented with community support. Journal of Research on Educational Effectiveness, 12(3), 391-412.
8. Sheard, M. and Chambers, B. (2011) Self-paced learning: Effective technology-supported formative assessment. York: Institute for Effective Education, https://www.lbq.org/Areas/Default/Content/Default/Document/Self-paced%20learning%208%20Aug%202011.pdf
9. Sheard, M., Chambers, B. and Elliott, B. (2012) Effects of technology-enhanced formative assessment on achievement in primary grammar. York: Institute for Effective Education, https://www.lbq.org/Areas/Default/Content/Default/Document/QfL%20Grammar%20Report%20_Final_%20Oct%2002%202012.pdf
10. Slavin, R. (2021) Highlight tutoring among post-Covid solutions, https://robertslavinsblog.wordpress.com/2021/01/28/highlight-tutoring-among-post-covid-solutions/
11. Torgerson, C., Bell, K., Coleman, E., Elliott, L., Fairhurst, C., Gascoine, L., ... & Torgerson, D. (2018). Tutor Trust: affordable primary tuition. Evaluation report and executive summary.