Primary Science Quality Mark (PSQM)

Evaluation Report

July 2022

Miriam Broeks, Lucy Gilder, Giulia Lanfredi, Elena Rosa Brown, Frederico Cardoso

Funded by

The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents.

The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus - Private Equity Foundation) and received a founding £125m grant from the Department for Education. Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.

This project was funded as part of the Improving Science Education scheme, which was jointly funded by Wellcome and the Education Endowment Foundation and launched in December 2017. This scheme aimed to generate new evidence about science teaching, with the particular aim of closing the science attainment and progression gap that exists between disadvantaged pupils and their more affluent peers.

For more information about the EEF or this report please contact:

Jonathan Kay

5th Floor, Millbank Tower

Education Endowment Foundation
21–24 Millbank
SW1P 4QP

0207 802 1653

jonathan.kay@eefoundation.org.uk

www.educationendowmentfoundation.org.uk

Contents

About the evaluator 3

Executive summary 4

Introduction 6

Methods 18

Findings from the implementation and process evaluation 35

Conclusion 65

References 68

Appendix A: EEF cost rating 71

Appendix B: Security classification of trial findings 72

Appendix C: Memorandum of Understanding 73

Appendix D: Participant information sheet 86

Appendix E: Privacy notice 88

Appendix F: Withdrawal form 90

Appendix G: Documentary review of PSQM submissions 91

Appendix H: PSQM evaluative criteria 94

Appendix I: Pilot for measures of pupil attitudes to science 96

Appendix J: Thematic coding for interview analysis 107

About the evaluator

The project was independently evaluated by a team from RAND Europe:

Miriam Broeks, Lucy Gilder, Giulia Lanfredi, Elena Rosa Brown, and Frederico Cardoso.

The lead evaluator is Elena Rosa Brown (2019 to present). Previously this was Dr Emma Disley (June 2019 to November 2019) and Dr Alex Sutherland (inception to June 2019).

Contact details:

Name: Elena Rosa Brown

Address: RAND Europe, Westbrook Centre, Milton Road, Cambridge, CB4 1YG, UK

Tel.: +44 1223 353 329

Email: **erbrown@randeurope.org**

We would also like to thank Alex Sutherland, Amelia Harshfield, Andreas Culora, Eleftheria Iakuvidou, Sashka Dimova, Sonia Ilie, and Yulia Shenderovich (formerly RAND Europe) for their support at various stages of the project.

Executive summary

The project

Primary Science Quality Mark (PSQM) is a one-year professional learning programme for primary science subject leaders (SSLs). It focuses on improving school self-evaluation of science. It is available to primary schools in the U.K. and internationally.[1] PSQM seeks to improve science teaching and learning throughout the school by improving the skill, quality, and confidence of science subject leadership.

PSQM comprises training and mentoring led by 'hub leaders' for groups of science subject leaders (SSLs)[2] from schools in a locality (the PSQM hub). The hub leaders are supported and trained by the PSQM 'headquarters' at the University of Hertfordshire. The programme also involves a structured set of activities carried out by the subject leaders in school. The quality mark is achieved following the submission and assessment of a reflective portfolio of evidence of the impact of the improved subject leadership. Existing qualitative research suggests that PSQM can benefit schools in multiple ways.[3] However, to date there is no experimental evidence on whether PSQM accreditation leads to improvements in pupil outcomes in science or related subjects.

This evaluation was planned as a two-arm, cluster-randomised controlled trial. In total, 155 schools were recruited and allocated either to participate in the PSQM programme or to carry on with 'business as usual'. Delivery was scheduled to take place between September 2019 and June 2020.

This project and its evaluation were affected by the 2020 and 2021 partial school closures and social distancing regulations caused by the coronavirus (COVID-19) pandemic: the trial and PSQM submission date was extended and intervention schools were invited to receive additional delivery intended to partly compensate for lost time. Implementation and process evaluation data collection occurred throughout the evaluation until July 2021. Ultimately, the collection of pupil outcome data was not possible. The impact evaluation element was foregone and, therefore, we cannot estimate the impact of the project on science subject attainment.

Table 1: Key conclusions

Key conclusions
1. Due to COVID-19, the primary outcome for this evaluation was not collected and so no measure of impact on science attainment is reported and there is no security rating for this trial. Key conclusions are based on qualitative data from the implementation and process evaluation.
2. Overall, headteachers and SSLs attributed several changes that were in line with PSQM's theory of change. For example, they reported that taking part in PSQM had raised the profile of science in their schools, with more time being dedicated to science. They also reported that those in science leadership in their school had become more confident and credible.
3. The training was well received by SSLs. There was over 90% attendance at all five sessions and SSLs reported finding all sessions highly relevant. Hub leaders reported that SSL engagement was high throughout the programme.
4. Headteachers and SSLs in PSQM schools were more likely to report a shared understanding of the importance and value of science across the school compared to other schools. However, headteachers and SSLs in PSQM schools reported similar levels of engagement with continuing professional development (CPD) and with science enrichment activities compared to other schools.
5. Interviews and surveys suggest that willing and enthusiastic staff are a key facilitator for the implementation of the PSQM process. Time pressures on staff were identified as a key barrier. Collating evidence for submissions and the administrative work associated with the programme were considered time-consuming by headteachers and SSLs.

[1] International Schools | PSQM
[2] A science subject leader (SSL) is the person in primary schools with responsibility for science teaching and learning. They are generally a classroom teacher who has additional responsibility for the subject—in much the same way that all curriculum subjects in a primary school have a subject leader. Other names for this role include 'science coordinator' and 'science leader'.
[3] See, for example, White et al., 2016; White et al., 2015.

Additional findings

Survey responses showed that hub leaders felt prepared in their role and found it easy to communicate with the PSQM team at the University of Hertfordshire and with other hub leaders. All hub leader respondents (13/13) agreed or strongly agreed that (i) the training they received from PSQM prepared them to fulfil their role as hub leader, (ii) they felt prepared to provide relevant support to subject leaders in their hub to complete the PSQM process, and (iii) they felt prepared to use PSQM's assessment criteria to determine their school's PSQM award at the end of the programme. All hub leaders (100%) strongly agreed that it was easy to maintain regular communication with the PSQM HQ team; 83% (10/12) found regional meetings and the hub leader conference extremely useful.

There was less agreement on the extent to which hub leaders exchanged ideas with other hub leaders, with only five out of 13 strongly agreeing with the statement 'I exchanged ideas with other hub leaders about how to implement PSQM more effectively'. The impact of COVID-19, which meant that there were no face to face meetings between hub leaders, may have affected this.

The programme was well received by SSLs: they reported finding all training sessions highly relevant, with two thirds of respondents describing the first three sessions as 'extremely useful' and the final two sessions as 'extremely useful' by over half. Despite the significant disruption of COVID-19, over 90% (70/76) of schools completed the PSQM submission.

Over 90% of surveyed teachers in PSQM schools felt that science teaching had improved because of participating in PSQM. In addition, staff in schools participating in PSQM were more likely to report that they had a shared understanding of the values and principles of science and that they had developed a range of effective strategies for teaching and learning science compared to control schools. While this suggests that PSQM implementation was successful, evidence on impact is limited due to the rescoped design. Without a counterfactual impact estimate it is difficult to conclude that staff-reported changes in practice led to the hypothesised impact on student attitudes and attainment.

School staff and hub leaders were asked about the factors that enabled them to implement PSQM effectively. The most frequent enablers mentioned were support from the senior leadership team in the school, support from the hub leader, and the positive attitude of staff involved in the programme. Other enablers mentioned by interviewees include flexibility within the school timetable and the provision of teaching cover to be able to attend the training and implement the programme.

Staff time and the workload of the programme were considered barriers to implementation. SSLs and hub leaders reported that one key barrier was the limited time that SSLs had to engage with training and activities. All interviewed headteachers and one teacher mentioned that the administrative work associated with the programme was burdensome and time consuming, with one SSL suggesting that it detracted from what they were trying to achieve through the programme. The need to facilitate a more manageable PSQM workload was a suggested improvement made by a few hub leaders in interviews.

Regarding COVID-19, respondents felt that the pandemic negatively impacted on a range of aspects of science teaching, and on practical science learning opportunities in particular. But there was no clear consensus on the impact of COVID-19 on the implementation of PSQM. Headteachers who responded to the survey were approximately equally divided between those that felt that COVID-19 had decreased their school's ability to implement PSQM and those that felt their ability to delivery PSQM had increased (39% versus 37%). Teachers, though, were more likely to respond that that their ability to engage with PSQM had decreased after COVID-19 rather than increased (39% versus 11%). Hub leaders reported little differences in SSL engagement as a result of the pandemic, though there was a perceived loss of 'community feeling' due to moving activities online.

Introduction

Background

A recent review of evidence suggests that high-quality science education depends on effective subject and school leadership, including allocating sufficient time to teach the science (Ofsted, 2021). However, research shows this does not always happen, particularly in primary schools. A review of primary science in 2019 found that disproportionate amounts of curriculum time in primary schools were being spent on English and mathematics, which led to significantly reduced curriculum time devoted to science (Ofsted, 2019). This is supported by a 'state of the nation' report for primary science education that found that 31% of respondents said their leaders saw science as 'very important' compared to 88% for English and 86% for mathematics (CFE, 2017).

Key challenges reported in relation to science education in primary schools include lack of teaching time, lack of quality monitoring, and limited access to science expertise (Wellcome Trust, 2014; Ofsted, 2019). It has been suggested that effective science teaching and learning depends on leaders ensuring that science teachers have access to regular, high quality, subject-specific continuous professional development (CPD) (Ofsted ,2021).

Primary Science Quality Mark (PSQM) is aimed at improving school-wide science teaching and raising the profile of science in U.K. primary schools through (i) effective science leadership and (ii) supported school self-evaluation. PSQM is already widely used: more than 3,700 schools have previously completed the programme (11.8% of all U.K. primary schools) and more than 583 are currently engaged.[4] PSQM has also been endorsed by Ofsted (Ofsted, 2013) and is one of the only national award schemes for science in English primary schools.[5]

Existing qualitative research suggests that PSQM can benefit schools in multiple ways, such as contributing to raising the profile of science in primary schools and providing schools with a framework and professional support for developing science leadership, teaching, and learning (White, et al., 2016). Previous evaluations of PSQM drew on interview, focus group, and survey data from participating science leaders and hub leaders from eight schools: perceptions of participants were that PSQM improved the profile of science as a subject and the quality of science teaching within schools and facilitated dissemination of relevant good practices between schools (White et al., 2016; White et al., 2015).

However, to date, there is no experimental evidence on whether PSQM accreditation leads to improvements in pupil outcomes in science or related subjects. Previous evidence is also rather limited regarding the impact of accreditation programmes in primary and secondary education. Evidence to date comes mostly from higher education, which suggests that accreditation programmes can improve the quality of teaching (Hanbury et al., 2008; Volkwein et al., 2006; Blouin et al., 2018). There is also evidence from survey data that accreditation translates into better outcomes for university students (Volkwein et al., 2006). Research from collaborative professional development programmes are mixed, with some suggesting they have positive effects on instructional practice and student outcomes (Opfer, 2016; Darling-Hammond et al., 2017; Gore et al., 2017) while others suggest these do not always lead to improvements in pupil outcomes (for example, Garet, 2011; 2016; Sims and Fletcher-Wood, 2018).

Evaluation overview

This evaluation was planned as a two-arm, cluster-randomised controlled trial (cRCT) across 140 schools. In total, 155 schools were recruited with 77 randomly allocated to carry on with business as usual (the control group) and 78 randomly allocated to receive PSQM training (the treatment group). The aim of the efficacy trial was to help determine whether PSQM leads to observable improved outcomes in children's science outcomes. However, as a result of the COVID-19-related lockdowns and social distancing measures introduced in England from March to June 2020 and from January to March 2021, activities for the evaluation of PSQM were revised.

The first lockdown from March to June 2020 closed schools to almost all students (only children of key workers were given the opportunity to attend) and attendance was optional from June to July 2020. This made pupil testing unfeasible at the end of the 2019/2020 school year. The second national lockdown from January 2021 to March 2021 closed schools again to almost all students (apart from children of key workers). During this period, schools continued to take

[4] http://www.psqm.org.uk/about-us
[5] http://www.psqm.org.uk/__data/assets/pdf_file/0006/78405/PSQM-flyer-July-2017.pdf

part in PSQM, attending training and taking part in the activities described below (see Intervention section), however, the extent to which schools could implement and evaluate action plans was limited.

A number of elements of the PSQM programme were adapted to support schools (see Changes to the Evaluation section). While these changes allowed the programme to continue despite lockdowns, these adaptations made the programme significantly different to what would be delivered under 'normal circumstances'. There was also some uncertainty over the extent to which pupils' outcomes would have been affected given the relatively limited length of time PSQM would have been embedded in schools (due to extensions of submission deadlines) and the limited time pupils would have been in classrooms (due to the repeated lockdowns). A joint decision between the EEF, PSQM and RAND Europe was made to cancel the testing at the end of the 2020/2021 school year. Given COVID-19 disruptions to education, it was considered that it would be more difficult to capture changes at the pupil level using standardised tests. In addition, minimising burden on schools was prioritised and therefore the evaluation activities were overall reduced.

Rescoped implementation and process evaluation (IPE) activities continued. Endline surveys, interviews, compliance analysis, and document reviews were carried out at the end of the 2020/2021 school year to better understand implementation and delivery of PSQM (in treatment schools) and business as usual (in control schools). However, surveys and interviews planned for midline (end of 2019/2020 school year) were foregone. To maximise insights obtained through the rescoped data collection, in addition to what was originally planned the evaluation was rescoped to include additional questions in the survey (i) to understand the impact of COVID-19 on science teaching (both treatment and control), (ii) how it may have moderated the delivery and impact of PSQM (treatment settings only), and (iii) indicative differences in practice across both arms drawn from additional survey items and a document review (both treatment and control). Further details can be found below (see Changes to the Evaluation section).

The evaluation was funded by the Education Endowment Foundation (EEF) and Wellcome. The PSQM programme was led by The University of Hertfordshire and was independently evaluated by RAND Europe.

Intervention

This section outlines the underlying rationale and structure of the PSQM programme. A visual overview of this can be seen in the theory of change (Figure 1). To refine the theory of change, the evaluation team worked with PSQM to elicit further details on what distinguishes PSQM practice from typical school practice so that the evaluation team could better understand what behaviours to explore in more detail. These are described in Table 2.

Table 2: PSQM practice

	Intended changes to practice
School-wide practices in science teaching and management	**Required to meet PSQM award criteria** The PSQM CPD programme aims to facilitate/promote: • a clear vision for the teaching and learning of science; • a shared understanding of the importance and value of science; • an agreed set of principles for good science teaching and learning; • appropriate and active goals for developing science; • a clear development cycle linked to an established monitoring programme—learning walks, lesson observations, book scrutiny, pupil voice, and data analysis. **Optional activities**: • science visible on the website and in communal areas; • a science display in every classroom; • science featuring in school newsletter; • a designated budget for science resources and CPD; • science subject leader (SSL) actively engaged in external networks to develop science; • school membership of ASE; • regular science staff meetings; • reports to governors include science; and • links with others schools to share and develop science.
Types of pedagogies and practices that teachers	**Required to meet PSQM award criteria** The PSQM CPD programme aims to facilitate/promote: • engagement with professional development to improve science teaching and learning: regular CPD for all teachers led by SSL who is proactive in seeking out new pedagogies/resources;

	Intended changes to practice
use with their classes	• a range of effective strategies for teaching and learning science that challenge and support the learning needs of all children, strategies are identified in accordance with development needs, not ad hoc, and effectiveness evaluated, and the use of up to date resources, for example, Explorify, BAES; • range of up to date, quality resources for teaching and learning science, which are used regularly and safely—including IT, for example, data loggers, IPADs etc.; • a shared understanding of the purpose and process of science enquiry: children use full range of enquiry types, asking and answer own question, investigate independently, enquiry part of almost every lesson; • a shared understanding of the purposes of science assessment and current best practice: formative assessment used regularly to improve lesson, and summarised periodically for reporting, progression is tracked; • a commitment to developing all children's science capital: every class has science visits and visitors, teacher know the importance of developing positive attitudes towards science, parents are involved, science weeks take place, children engaged in global scientific issues, children care for their environment, and children are engaged in national scientific initiatives; • appropriate links between science and other learning; and • appropriate links with families, other schools, communities, and outside organisations to enrich science learning. **Optional activities**: • regular outdoor science, children grow fruit and vegetables, get to know their local environment well.

Both the theory of change and the 'template for intervention description and replication' (TIDieR) checklist (Hoffman et al., 2014) provide structure for the rest of this section. Both the TIDieR and theory of change were developed in discussions between RAND, PSQM, and the EEF at early set-up meetings and reflect the intervention as delivered in the trial and changes made as a result of the pandemic.[6] The theory of change was updated by RAND Europe in consultation with PSQM as the evaluation progressed.[7] The theory of change was updated during the report phase by the evaluation team in the following ways: the inputs were updated to better reflect the programme as opposed to the evaluation, science leaders were removed from the outputs and impact as these are not considered outcomes or impact, and 'creation of a hub community of practice' was changed to 'quality of support offered by PSQM and hub leaders' to better reflect the quality of this contextual factor. It is important to note that some elements represented on the theory of change, including the training of Year 5 teachers and changes as a result of COVID-19 restrictions, have not traditionally been elements included in typical delivery of the training or award process.

Owing to the closure of schools to most children as a result of national lockdown measures, adaptations were made to the programme by PSQM. To ensure clarity, the following sections outline what was intended (see PSQM usual practice) as well as record all changes that were made (see Unplanned Adaptations Made to Delivery sub-section). Further details on changes to the evaluation can be found in the **Changes to the Evaluation** section.

Who—recipients of the programme

PSQM was initiated in 2008 at the University of Hertfordshire to raise the profile of science in primary schools in England and promote professional development in science teaching and leadership.[8,9] PSQM is a developmental accreditation programme available to primary schools in the U.K. and internationally.[10] PSQM training is led by hub leaders to groups of schools, 'hubs', with a mean of approximately ten schools per hub. PSQM is a whole-school improvement programme—the aim being that all teachers and pupils benefit through more skilled and confident science subject leadership. It is compulsory for the science subject leaders (SSL) to attend training. Given the fact that pupils in Year 5

[6] Specifically, the addition of session 3b and the delivery team dropping the requirement for Year 5 teachers to be trained after March 2021.
[7] Both versions of the theory of change can be found on the EEF website: Primary Science Quality Mark | EEF (educationendowmentfoundation.org.uk). The original can be seen in the June 2019 protocol while the update version can be found in the August 2020 protocol.
[8] http://www.psqm.org.uk/what-is-psqm ; http://www.psqm.org.uk/about-us
[9] http://www.psqm.org.uk/__data/assets/pdf_file/0010/123130/Primary-Science-May-2016-PSQM-update.pdf
[10] International Schools | PSQM

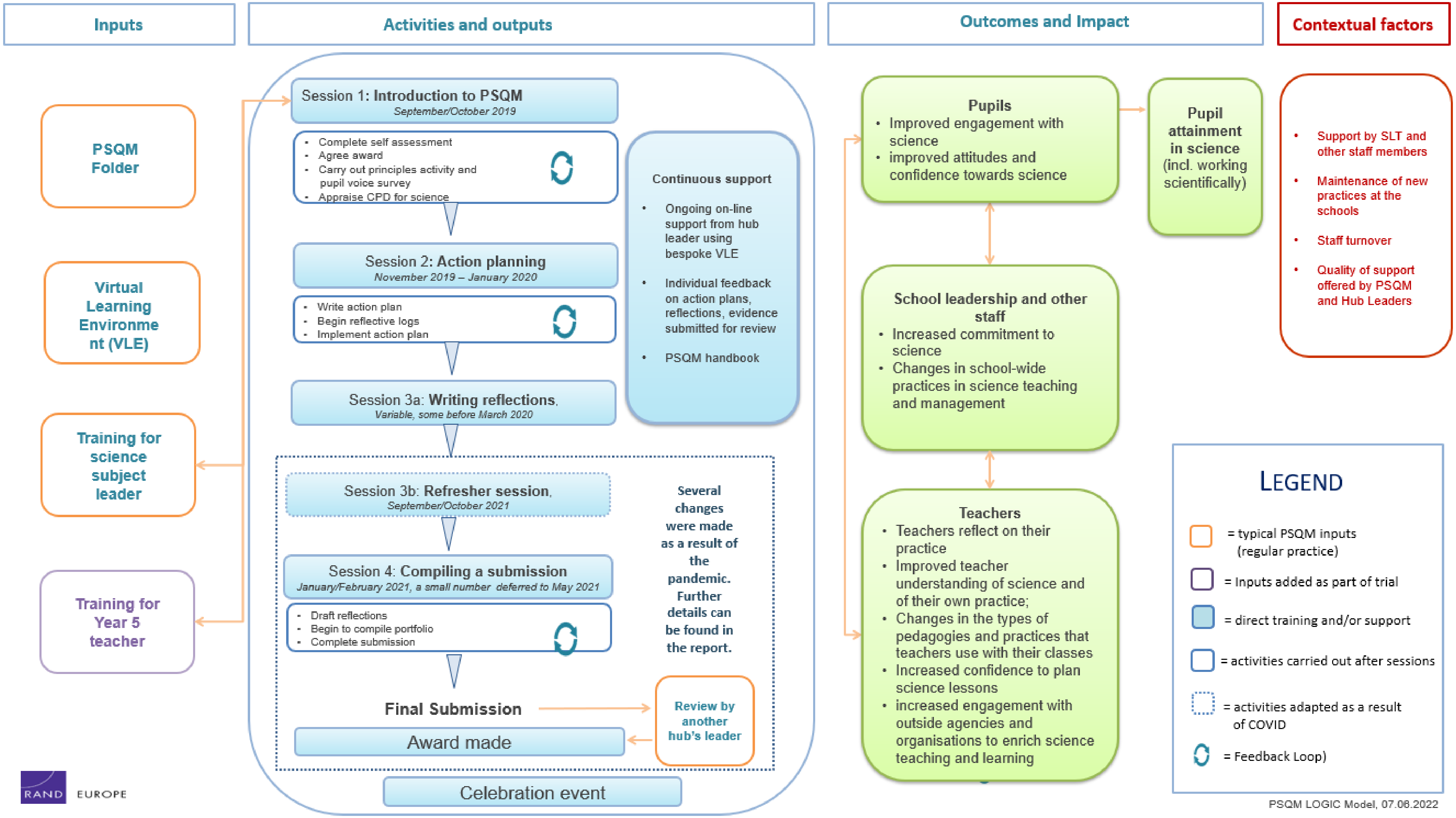
were the key group of interest for the attainment impact analysis in this evaluation, Year 5 teachers were invited to attend training. In March 2021 the delivery team sent a letter to schools asking for SSLs to return to training (see Appendix C: Follow-up communication with schools on further changes, March 2021) but the request was not extended to teachers due to the impact of COVID-19 on schools and the fact that there was no longer an impact evaluation.

Who—providers of the programme

PSQM is led by hub leaders who provide training and support for schools; hub leaders, in turn, are supported and trained by the PSQM 'headquarters' at the University of Hertfordshire. Hub leaders have backgrounds such as local authority advisers, consultants, university lecturers, Advanced Skills Teachers (AST), Specialist Leader of Education (SLE), and teachers who have achieved Primary Science Quality Marks as subject leaders in the past.

Figure 1: PSQM theory of change

Note: An additional session was added to the programme in response to COVID-19 disruptions. This took place after Session 3 (called 3b) and brought the total number of sessions to five. Also, the requirement for Year 5 teachers to attend training was foregone by the delivery team in March 2021. The timeline reflects the timing in the evaluation with delays caused by restrictions imposed by COVID-19.



PSQM usual practice

Schools can work towards one of three Primary Science Quality Marks: PSQM, PSQM Gilt, and PSQM Outreach. PSQM is for 'schools which demonstrate how effective science leadership is *beginning* to have an impact on science teaching and learning across the school', whereas PSQM Gilt requires the demonstration of a '*sustained* impact', and PSQM Outreach is for schools that meet Gilt criteria and also impact science leadership and teaching in other schools. Hub leaders were supported by a dedicated team at the University of Hertfordshire.

Regardless of the mark, in usual practice PSQM involves the following activities over the course of one year:

- SSL training, provided by the hub leader, completed over two full days or four half-days—topics include: introduction to the PSQM process, assessing needs, creating and executing an action plan, evaluating impact, writing a reflective submission, and collating appropriate supporting evidence—plus linking to other sources of professional support and quality resources.
- The subject leader works with colleagues across the school to audit existing provision in science and agree the appropriate quality mark to work towards.
- The subject leader creates an action plan to develop aspects of science teaching, as specified in the PSQM framework and works with colleagues to implement it.
- Subject leaders are supported by the hub leader through ongoing online mentoring provided via the PSQM virtual learning environment (VLE) and access to resources such as the PSQM handbook and information on relevant CPD offers.
- The subject leader collates and submits the evidence for the relevant PSQM criteria, which requires them to show the impact on teaching and learning of actions taken to meet the criteria which is reviewed by a hub leader from another hub as part of the quality assurance process. This process is supported by training for reviewers and moderation.
- Hub leader reviewers use PSQM evaluative criteria to consider whether a school has achieved the requirements to gain the chosen Primary Science Quality Mark (see Appendix H).

Awards are made to schools following an analysis of a series of documents that detail how the activities implemented during the intervention year have impacted on the science teaching and learning across the school and how the school meets the PSQM criteria. There are 13 PSQM criteria covering (1) primary school science leadership, (2) teaching, (3) learning, and (4) wider opportunities—plus two additional criteria for schools working toward the PSQM Outreach. Rather than the award itself being central, *the focus of the programme is on the development cycle, the process of self-evaluation, reflection, and development.*

What—materials

At the time of the trial, the standard implementation of PSQM involved the following inputs.[11]

PSQM folder

At the beginning of the process, science subject leaders were given a PSQM folder; this contained:

- information about the responsibilities of the SSL, SLT, and hub leader, a guide to the VLE, a calendar of PSQM activities—including training sessions (see below)—and an application form for the Association of Science Education (ASE);

[11] These have since been updated.

- the PSQM framework, the initial self-assessment audit, and the criteria with information about each one, including recommended reading, resources, and activities plus information about PSQM Outreach;
- self-led CPD activities for SSL's to use in school; and
- a writing frame to support reflective writing, submission dos and don'ts, FAQs, and reaccreditation guidance.

There is a wide range of additional supporting resource for SSLs on the VLE.

Training sessions

The PSQM training available to SSLs and Year 5 teachers consisted of five sessions lasting half a day.[12] Normally these would be delivered in one academic year, however, as a result of disruptions caused by COVID-19 these were adapted after March 2020.

- Session 1: Getting started on the PSQM journey (September/October 2019)

  In Session 1 of the PSQM training, 'Introduction to PSQM' (half term 1), SSLs were introduced to the initial **self-assessment questions.** After the training, SSLs completed the self-assessment with their colleagues in schools. The self-assessment helped them evaluate their starting position against the criteria and was aimed at helping them judge how many of the indicators for PSQM required work at the start of their participation in the programme and how many were already established and evidencing impact.

- Session 2: Action and Reflection (November 2019 to January 2020)[13]

  During Session 2, 'Action Planning' of the PSQM training, SSLs were given guidance on how to write and implement an **Action Plan** to develop aspects of science teaching such as subject leadership, learning, teaching and wider opportunities in order to meet all the PSQM criteria, and plan how they would measure and evidence the impact of these actions. During this session, SSLs were also asked to begin **reflective logs** about their subject leadership activity.

- Session 3a: Writing a reflective submission (variable, some before March 2020)[14]

  During Session 3, 'Writing Reflections' (half term 4), SSLs had a chance to discuss and reflect on the activities introduced at their schools as a result of the PSQM training.

- Session 3b: Refresher session on 'writing a reflective statement' in light of COVID-19 disruption (September/October 2021)[15]

  In light of COVID-19 disruptions, the PSQM team created an additional training Session (3b) as a refresher on content of Session 3.

- Session 4: Completing your submission—identifying the PSQM legacy.

---

[12] Once training was moved online as a result of the restrictions imposed by COVID-19 the PSQM team suggested that HLs split these into shorter sessions.
[13] HLs organised these according to their own calendars - but within a defined time frame set by the PSQM team.
[14] Regarding the date, some hubs completed this before the pandemic and others did not. In some cases, HLs did it again after lockdown and others did a refresher.
[15] Regarding the date, this was organised by HLs to reflect their calendars and the appropriate timing for the schools involved in their hubs.

During Session 4, ‘Compiling a Submission’ (January/February 2021, with a small number of schools deferring to May 2021) subject leaders were supported to finalise all compulsory documents for the submission—the PSQM task. These included:

- an ‘about your school’ questionnaire;
- self-assessment;
- action plans;
- core documents—science leader and CPD log, statement of principles of good science teaching and learning, calendar of events, and school development plan;
- reflections; and
- final reflective questions.

Virtual learning environment

Subject leaders are supported by the hub leader, with ongoing online mentoring provided via the PSQM virtual learning environment and access to resources such as the PSQM handbook and information on relevant CPD offers. The VLE provides examples of the core documents schools need to submit for PSQM and a range of additional resources. An area of the VLE under ‘schools’ allows subject leaders to upload the key tasks and documentation for PSQM submission:

- the ‘about your school’ questionnaire;
- self-assessment;
- needs analysis and action plans;
- core documents;
- reflections; and
- final reflective questions.

Planned adaptations—tailoring

All SSLs must complete the same initial self-assessment of their schools and themselves, write and implement an action plan, and compile a reflective submission comprising compulsory documents to show how the criteria are met and the impact of the actions taken to meet them. Each school’s action plan, implementation, and final submission will be different, tailored in response to the needs of the school.

Unplanned adaptations—modifications due to COVID-19 restrictions

To support successful participation in the PSQM programme, the delivery team made the following adaptations:

- extended submission deadline from June 2020 to March 202, plus offer for schools to defer to June 2021;
- provided an additional training session (3b);
- facilitated school participation with all training sessions from March 2020 onwards, delivered online;
- provided PSQM headquarter-led webinars (led by PSQM team at University of Hertfordshire) during the summer and autumn terms in 2020;
- created a COVID-19 guidance document to help science subject leaders adapt planned actions and identify impact;
- reviewed assessment expectations and provided updated moderation training for reviewers;

- made additional hub leader support available post submission to replicate the level of support schools would have normally received; and
- no longer compulsory for Year 5 teachers to attend training after Session 2.[16]

Changes to the evaluation

COVID-19-driven changes

As a result of the COVID-19-related lockdowns and social distancing measures introduced in England from March to June 2020 and from January to March 2021, activities for the evaluation of PSQM were revised twice.

The first lockdown closed schools to almost all students (except children of key workers) and attendance was optional from June to July 2020. This made pupil testing unfeasible at the end of the 2019/2020 school year: a joint decision between the EEF, PSQM, and RAND Europe teams was made to cancel the testing at the end of the 2019/2020 school year. In addition, it was decided to rescope the IPE to avoid capturing data from schools in 2019/2020 (that is, IPE activities for 2019/2020 were cancelled). An amended protocol was made available on the EEF's website.[17]

The second national lockdown closed schools again to almost all students (apart from children of key workers). During this period, some schools continued to take part in PSQM, attending training and taking part in the activities described below (see Intervention section), however, the extent to which schools could implement and evaluate action plans was limited. Other schools made the decision to defer their submission to June 2021. A number of elements of the PSQM programme were adapted to support schools (see Unplanned Adaptations Made to Delivery above). While beneficial for delivery, these adaptations made the programme significantly different to what would be delivered under 'normal circumstances'. There was also some uncertainty over the extent to which pupils' outcomes would have been effected given the relatively limited length of time PSQM would have been embedded in schools (due to extensions of submission deadlines) and the limited length of time pupils would have been in classrooms (due to the repeated lockdowns).

A joint decision between the EEF, PSQM, and RAND Europe was made to cancel the testing at the end of the 2020/2021 school year. Given COVID-19 disruptions to education, it was considered that it would be more difficult to capture changes attributable to the PSQM programme at the pupil level using standardised tests. In addition, minimising the burden on schools and pupils was prioritised and therefore the evaluation activities were overall reduced. A second, amended protocol was made available on the EEF's website.[18] Table 3 summarises the changes made to the original protocol.

[16] Noting that training for Year 5 teachers was the only required part of the trial and is not part of typical PSQM practice.
[17] https://d2tic4wvo1iusb.cloudfront.net/documents/projects/PSQM_protocol ol_Amended_June_2020.pdf
[18] https://d2tic4wvo1iusb.cloudfront.net/documents/projects/PSQM_protocol_June_2021_v3.pdf

Table 3: Changes to PSQM evaluation as a result of disruption to school activities

Original plans	June 2020 revised plans	April 2021 current plans
- Pupil outcome testing at end of 2019/2020: science attainment and attitudes towards science (secondary outcomes). - Pupil testing at end of 2020/2021: science attainment (primary outcome) and attitudes towards science (secondary outcome). - Wave 1 IPE activities: staff surveys and case study interviews in summer of 2020. - Wave 2 IPE activities (as above) in summer of 2021.	- Pupil outcome testing at end of 2020/2021 only: science attainment (primary outcome) and attitudes towards science (secondary outcome). - IPE activities previously planned for June 2020 shifted to December 2020 to February 2021. - IPE activities in summer of 2021: ○ surveys in all schools; and ○ case studies—interviews and document reviews of PSQM submissions in five PSQM schools.	- All pupil outcome testing cancelled. - IPE activities planned for the period of December 2020 and February 2021 cancelled. - IPE activities in summer of 2021: ○ planned surveys remain the same; ○ number of case studies expanded from five PSQM schools to ten PSQM schools and ten control schools; and ○ number of schools to review science documentation/PSQM submission review expanded to 30.

Evaluation objectives

The core research questions this project originally sought to answer were distributed across the impact and IPE evaluations.

The **impact evaluation** was designed to investigate the following research hypotheses:

- **Hypothesis 1**. Year 5 pupils in randomly allocated primary schools participating in PSQM (intervention schools) will have higher levels of science attainment than the pupils in the comparison schools one year following the end of PSQM implementation, 2020/2021 (summer 2021; primary outcome).
- **Hypothesis 2**. Year 5 pupils in primary schools participating in PSQM (intervention schools) will report higher levels of enjoying science than the pupils in the comparison schools in 2020/2021 (summer 2021; secondary outcome).
- **Hypothesis 3**. Year 5 pupils in primary schools participating in PSQM (intervention schools) will have higher levels of science attainment than the pupils in the comparison schools at the end of the school year when the intervention takes place, 2019/2020 (summer 2020; secondary outcome).

Update. None of these hypotheses could be tested due to pupil outcome data collection being cancelled. However, details on the design of the trial are provided in the Methods section as randomisation was completed. Further details of the rescoped research can be found in the Implementation and Process Evaluation (IPE) section below.

For the **IPE**, the original research questions were:

- Was the intervention implemented with fidelity for PSQM schools?
- What was practice as usual in the control schools?
- What appear to be the necessary conditions for success of the intervention?
- What were the barriers to delivery?

Update. In April 2021 the IPE was amended to include an additional question on the impact of COVID-19 . The goal was to better understand the impact of COVID-19 on science teaching across all schools in the trial. Changes can be seen in the updated protocol.[19] The research question guiding this is:

- To what extent has COVID-19 impacted on science teaching?

Ethics and trial registration

The trial has been registered on the ISRCTN registry, which stands for 'International Standard Randomised Controlled Trial Number' and is used to describe RCTs and efficacy trials at inception. The trial has been assigned an ID registration number: ISRCTN50771738.

The ethics and registration processes are in accordance with the ethics policies adopted by RAND Europe. The evaluation has been reviewed by RAND U.S. Human Subjects Protection Committee (HSPC).

Parents were sent information and withdrawal forms that had been developed by the evaluation and delivery teams. The parental information sheets and withdrawal forms were sent out to parents by the schools after the school representative signed the Memorandum of Understanding (MoU) describing what was involved in the trial. Parents could withdraw their children at any time from the research, but the initial withdrawal forms could be returned by parents within two weeks (see Appendix F).

If participants chose to withdraw their children from the study later on, their data was collected or was deleted as appropriate. For interviews, RAND Europe provided participant information sheets explaining the goals of the research and collected consent forms from school staff who volunteered to participate in an interview. Furthermore, the cover page for each survey informed respondents that participation in the survey was entirely voluntary and a second page contained a privacy notice for respondents outlining how their data would be handled.

None of the evaluation team had any conflicts of interest and all members of the study team approved the study protocol and statistical analysis plan (and their updates) prior to publication.

Data protection

RAND Europe obtained personal data from schools and pupils as data controller. Memorandums of Understanding and privacy notices provided to schools prior to enrolment in the RCT outlined in detail what data was to be collected, how it would be used in the evaluation, and how teams would share data. Interview participants were also provided with information on the research project in participant information sheets. These documents can be found in Appendices C to E.

Basic pupil information, including Unique Pupil Number (UNP), date of birth, gender, name, free school meal (FSM) status, and children with English as an additional language (EAL) was obtained from schools at baseline by the PSQM team and shared with RAND Europe using a secure file sharing service (Syncplicity). This data was collected on the basis of legitimate interests from schools under GDPR. All processes were handled in accordance with RAND's data protection policy. RAND Europe is registered with the Information Commissioner's Office (ICO), registration number Z6947026 and is certified for adhering to ISO 9001:2015 quality management practices.

[19] **https://d2tic4wvo1iusb.cloudfront.net/documents/projects/PSQM_protocol_June_2021_v3.pdf**

Once the final report has been published, anonymised data will be shared with the EEF for archiving. No individual child, teacher, staff member, or school will be identified in any reports arising from the research.

Project team

Delivery team: PSQM (University of Hertfordshire)

Project leader and PSQM Director: Associate Professor Jane Turner.

PSQM Deputy Director: Helen Sizer.

PSQM team: Claire Harman.

Evaluation team: RAND Europe

Overall project and evaluation lead: Elena Rosa Brown. Previously: Dr Emma Disley, June 2019 to November 2019; Dr Alex Sutherland, inception to June 2019.

Project manager: Miriam Broeks. Previously: Amelia Harshfield, April 2019 to March 2020; Dr Yulia Shenderovich, inception to April 2019.

Core fieldwork and analysis team: Lucy Gilder, Giulia Lanfredi, Frederico Cardoso, Eleftheria Iakuvidou, Andreas Culora, Sashka Dimova, and Sonia Ilie.

Methods

Trial design

As outlined in Table 4, the PSQM evaluation was designed as a two-group, parallel, stratified, cluster-randomised trial with school as the unit of randomisation. To ensure comparability between schools in the PSQM arm and those in the control arm, and ensure 'exchangeability' (Oakes, 2013), schools were randomised within hubs, which served to balance the study arms on geographical location and, therefore, any regional differences.[20]

During the recruitment period (2018/2019), schools were asked to nominate their SSLs and the Year 5 teacher, or one Year 5 teacher in cases where there were multiple Year 5 classes, to participate in PSQM. The Year 5 teacher's class was to be considered the focal class for the evaluation and would be assessed in summer 2020 following participation in the PSQM process. In summer 2021, the Year 5 class taught by the same teacher would be assessed for a second time. If the teacher left or moved to another year group, the evaluators planned to assess the Year 5 class or randomly select another Year 5 class (if there were more than one Year 5 classes).

To minimise the burden on pupils and schools, the evaluation team planned to use administrative data for baseline measurement, with schools providing pupil identifiers (UPNs), which could be linked to the National Pupil Database (NPD). After schools were recruited and the pupil and teacher information collected, the evaluation team randomised schools to one of two arms: PSQM or control.

PSQM schools received the programme for free and each received a payment of £1,500 towards teaching cover and £120 towards travel costs. Control schools were not allowed to participate in PSQM during the study period and were eligible to receive a payment of £1,500, which they could put towards buying PSQM after the trial if they wished on completion of the trial if they participated in all data collection tasks

Table 4: Trial design

Trial type and number of arms		Two-arm, stratified, cluster-randomised controlled trial, randomised at the school level
Unit of randomisation		School
Stratification variables (if applicable)		Region (hub) School size (single- versus multiple-form entry)
Primary outcome	variable	Science attainment
	measure (instrument, scale)	An adapted version of the Year 5 Science Assessment (developed by the University of York)[21] or the GL Assessment Progress Test in Science (PTS).[22] The psychometric properties of Thinking, Doing, Talking Science (TDTS) were being explored by a team at York University before the impact evaluation was foregone.
Secondary outcome(s)	variable(s)	Pupil attitudes to science and science teaching

[20] That is, if one were to swap the intervention and control groups the results from the trial should be the same.
[21] https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluation/eef-outcome-measures-and-databases/year-5-science-assessment?utm_source=/projects-and-evaluation/evaluation/eef-outcome-measures-and-databases/year-5-science-assessment&utm_medium=search&utm_campaign=site_search&search_term=year%205
[22] https://www.gl-assessment.co.uk/assessments/progress-test-in-science/

Baseline for primary and secondary outcomes	measure(s) (instrument, scale)	A newly-derived, piloted instrument with items from TDTS and Trends in International Mathematics and Science Study (TIMMS) was being piloted by RAND Europe before the impact evaluation was foregone
	variable(s)	Standardised and pooled scores using teacher assessed maths, reading and writing scores at KS1 taken from the NPD
	measure(s) (instrument, scale)	Maths, reading, and writing (MATH_OUTCOME, READ_OUTCOME, and WRIT_OUTCOME, respectively); categorical measure with five categories

Update. Due to the issues outlined above, testing was foregone, thus, the data for primary and secondary outcomes listed in the table above were not collected. Control schools were still offered £1,500 on completion of IPE data collection activities.

Participant selection

Schools

Schools were recruited by the PSQM team and PSQM hub leaders, based on the following eligibility criteria:

Inclusion criteria:

- The school cannot have received a PSQM award in the last 3 years (that is, a school has not participated in PSQM in 2017, 2018 or 2019). This was checked by the PSQM team.
- The school must be a state primary, junior or all-through school.
- Schools with mixed Year 5/6 or another combination were eligible if they have Year 5 pupils taught separately by one teacher for science.
- Schools had to have a Science Subject Leader.

Exclusion criteria:

- Infant or first schools, private schools, special schools, Pupil Referral Units (PRUs) or middle schools were not eligible.

Schools were recruited from the following 16 broad geographical areas:

- Buckinghamshire
- Cambridgeshire
- Cheshire
- Chorley
- Cumbria
- Devon
- Hereford, Gloucester, and Worchester
- Kent
- Leicester
- Merton
- Oxfordshire
- Portsmouth
- Staffordshire
- Waltham Forest
- Warrington
- Yorkshire

Teachers

One science subject leader and one Year 5 teacher were asked to attend the training. PSQM is a whole-school improvement programme—the aim being that all teachers and pupils benefit through more skilled and confident science subject leadership. It is compulsory for the SSLs to attend training, and another member of staff may attend. Given the fact that pupils in Year 5 were the key group of interest for the attainment impact analysis in this evaluation, Year 5 teachers were invited to attend. The protocol stated that a Key Stage 1 (KS1) teacher should be included in cases where the subject leader was also the Year 5 teacher.

Update. In practice, it did not make sense to replace the Year 5 teacher with a KS1 teacher as the latter did not have knowledge of the programme. Instead, if the SSL was also a Year 5 teacher we asked them to complete an extended survey specifically designed to incorporate teacher and SSL questions while avoiding duplication. No Key Stage 1 teachers were explicitly invited to attend training (unless they attended in their SSL role). This change was not reflected in the protocol. In March 2021 the delivery team sent a letter to schools asking for SSLs to return to training (see Appendix C: Follow-up Communication with Schools on Further Changes, March 2021) but the request was not extended to Year 5 teachers due to impact of COVID on schools and the fact that there was no longer an impact evaluation.

Pupils

There were no inclusion or exclusion criteria at the pupil level as PSQM is delivered at the school level. To minimise burden on schools, pupils enrolled at the time of school recruitment in 2019 were included in the study but pupils who joined the schools at a later time were not included in the evaluation as this would require additional information to be collected from schools.

Outcome measures

Baseline measures

The evaluation team planned to collect pupils' KS1 maths, reading, and writing scores from the NPD—data from teacher assessments in Year 2. The statistical analysis plan specified that these KS1 baseline scores would be standardised and pooled.[23]

Update. In line with the theory of change, the team had planned to pursue the collection of pupil baseline data from the NPD to understand if randomisation was successful. However, in light of having no outcome data, exploring baseline equivalence was seen by the EEF and the evaluators as being less meaningful. As a result, EEF and RAND Europe agreed that they would include all available data on treatment and control arms at baseline using publicly available school-level data on Ofsted rating, FSM, and school size and data collected at baseline on pupil level FSM and EAL. However, attainment data (KS1) from the NPD would not be included in these descriptive analyses.

Primary outcome

The evaluation team originally planned to use a test on science attainment developed by the University of York and co-funded by the EEF and Wellcome. This test was developed during 2020 and drew on a previous test of knowledge, thinking, and reasoning in science used in the EEF evaluation of TDTS (Hanley et al., 2015). It includes process- and inquiry-based questions, concept-based questions, and open-ended, conceptually-based questions.

The evaluation team planned to evaluate the effect of the intervention on pupil science attainment and attitudes following the 2019/2020 implementation year with the second Year 5 cohort, in 2020/2021 (labelled Cohort B) in all schools. The test would have been paper based, administered and marked by a third-party, the National Foundation for Educational Research (NFER). This approach would have allowed for blinding to allocation as a list of schools could be supplied to the assessors without revealing allocation.

The Progress Test in Science (PTS) was to be considered as an alternative test if the TDTS proved to not meet psychometric specifications.[24] Changes to the trial were made in March 2021 given the disruptions brought by the January 2021 national lockdown. It was decided to forego all outcome testing for the trial and to only collect IPE data.

Secondary outcomes

To assess changes in pupils' attitudes towards science, the team planned to carry out a post-test survey at the same time as the primary outcome assessment. The attitudinal measure at post-test was also going to be administered by NFER using paper forms. The attitudinal measures were going to be compiled in machine-readable forms to allow scanning, data entry, and scoring by RAND Europe.

Enjoyment of science and confidence in science were going to be measured using the 'enjoyment of science' subscale adapted from the Trends in International Mathematics and Science Study (TIMMS) Grade 4 surveys from TIMMS 2015.[25] If possible, it was planned to capture science enquiry skills from relevant items in the science attainment test by Hanley and colleagues.

To ensure that this adapted attitudes to science measure was valid, a small validation pilot was conducted prior to the planned outcome measure collection in trial schools. This pilot used a sample of Year 5 pupils (in non-trial schools) to test the measurement structure of the variable (in exploratory factor analysis). The pilot would have yielded a final measure to be used in the trial.

Initial plans were to administer the science attainment test and the 'attitudes to science and science skills' test among Year 5 pupils in the nominated classes.

Update: Owing to reasons outlined above, secondary outcome testing was foregone despite this test being piloted. Results of the pilot can be found in Annex I.

Sample size

Minimum Detectable Effect Size calculations

Power and minimum detectable effect size (MDES) calculations were performed using the PowerUp tool for main effects (Dong and Maynard, 2013) and moderators (Spybrook, Kelcey and Dong, 2016; Dong, et al., 2017). Based on EEF guidelines (EEF, 2018) and on a previous evaluation using science outcomes in this age group (Kitmitto, 2018),[26] the amount of variation explained by covariates for 140 schools with an average of 25 pupils each was assumed to be 0.40 (equivalent to correlation of 0.63) for level 1 (pupils) and 0.00 for level 2 (schools). The efficacy evaluation of Thinking, Doing, Talking Science (TDTS), which used a similar primary outcome

[24] These were to be predetermined but given that testing was foregone, the specifications were not determined.
25 https://timssandpirls.bc.edu/publications/timss/2015-methods.html
[26] The effectiveness evaluation of TDTS (Kitmitto, 2018) found variance explained at level 1 to be 0.40 for the same primary outcome as in the current trial and KS1 reading, writing, and mathematics as baseline, so we expect the current trial to have at least the same variance explain as a minimum.

(Hanley et al., 2015) reported an intra cluster correlation (ICC) of 0.15 in the analyses. With one class per school included in the evaluation, we assumed an average cluster size of 25 pupils. We also assumed an alpha of 5% and an intended 80% power to detect effects. We used two-level clustered designs, assuming a continuous, normally distributed (Gaussian) outcome.

The team believes it would be important to power 0.2 even though this is an efficacy trial because the nature of the intervention is likely to result in comparatively smaller effect sizes. Using the parameters above and with equal allocation to PSQM and control, the MDES in the protocol calculation of 140 schools is 0.197 (see Table 5). When calculating the MDES for the actual randomised data, the MDES comes to 0.188.

In the end, 155 schools were recruited—15 more than had been outlined in the protocol. At randomisation, using baseline data, the actual number of pupils was 4,098. At protocol stage it was assumed that there were four FSM pupils per class, based on the average number of FSM pupils in U.K. primary schools—14% in 2018.[27] However, PSQM recruitment for the trial focused on high-FSM areas, so the actual number was slightly higher at five pupils per class, which led to a total of 765 FSM pupils. Using the same assumptions as the main analysis, MDES difference regarding Cohen's d was intended to be 0.251 at protocol stage; at randomisation the MDES was 0.161.[28] The number of FSM pupils was 765.

Table 5 presents the full breakdown of the power calculations both at protocol stage and at randomisation stage.

Table 5: Power calculations at protocol stage and randomisation

		Protocol		Randomisation	
		OVERALL	FSM	OVERALL	FSM[29]
Minimum Detectable Effect Size (MDES)		0.197	0.251	0.186	0.228
Pre-test/post-test correlations	level 1 (pupil)	0.63	0.63	0.63	0.63
	level 2 (class)	N/A	N/A	N/A	N/A
	level 3 (school)	0	0	0	0
Intracluster correlations (ICCs)	level 2 (class)	N/A	N/A	N/A	N/A
	level 3 (school)	0.15	0.15	0.15	0.15
Alpha		0.05	0.05	0.05	0.05
Power		0.8	0.8	0.8	0.8
One-sided or two-sided?		Two	Two	Two	Two
Average cluster size		25	4	27	27[30] (5FSM)
Number of schools	PSQM	70	70	78	78
	Control	70	70	77	77

[27]**https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/719226/Schools_Pupils_and_their_Characteristics_2018_Main_Text.pdf**
[28] Using the same parameters but applying the sub-sample calculation as per protocol, the FSM sub-group MDES at randomisation would be 0.230.
[29] Analysis for subgroup run as moderator (interaction term). If subgroup analysis done on sub-sample, MDES for same parameters = 0.230.
[30] Analysis for subgroup run as interaction, therefore group size remains the same as in overall analysis.

		Protocol		Randomisation	
		OVERALL	FSM	OVERALL	FSM[29]
	Total	140	140	155	155
Number of pupils	PSQM	1,750	1,750	2,071	2,071 (405 FSM)
	Control	1,750	1,750	2,027	2,027 (360 FSM)
	Total	3,500	3,500	4,098	4,098 (765 FSM)

Randomisation

Randomisation occurred as planned in July 2019 and was conducted in Stata by the evaluation team's primary investigator. Hub was the main stratifying variable, with 16 hubs recruited. In addition, the evaluation team stratified on school size (single-entry versus multiple-entry school) as reported by the school. The trial allocation was recorded and communicated to the implementation team and the EEF in a password protected Excel file to prevent editing. In total, 155 schools were randomised with 78 assigned to treatment and 77 to control.

In the protocol, it was specified that we would assess imbalance at baseline by means of cross-tabulations that assess the distribution of each characteristic within the control and PSQM groups. The evaluation team had originally planned to use pupils' KS1 maths, reading, and writing data as baseline data to assess baseline equivalence of the intervention and control groups after the randomisation process of the schools. These data were also going to be used as covariate(s) in outcome analyses.

Update. The evaluation team carried out a descriptive analysis of the two groups at randomisation using pupil-level variables collected directly from schools prior to randomisation (that is, FSM and EAL status). We also have included school-level variables (Ofsted ratings and school-level FSM) to understand similarities at baseline across the two groups. However, we have not included an analysis of KS1 scores. This was agreed with the EEF and is in line with similar trials that had impact evaluations foregone because of COVID-19 restrictions.

Statistical analysis

Update. This section outlines the analyses that were initially planned for the trial. However, due to the reasons outlined, it was no longer possible to conduct most of these planned analyses.

Primary analysis

The primary outcome was intended to be science attainment as measured by a test of science (see Outcome Measures above). PSQM and control arms were going to be compared in terms of the difference in means between groups at follow-up, conditional on baseline measures (KS1 maths, reading, and writing) and stratification variables (area and school size).

The unit of analysis here would have been pupils. There is an ongoing discussion about how 'best' to analyse results from RCTs that involve clustered data. One approach, 'analyse how you randomise' (Senn, 2004), suggests that one should explicitly account for clustering via multilevel models (AKA 'random effects'). This approach assumes that the schools in the study are a random sample of all schools—which is often a source of contention—but one benefit of this approach is being able to explicitly partition variance and more flexibly handle complex variation within schools (Snijders and Bosker, 2012). Our approach would have been to conduct sensitivity analyses to assess results against different model specifications.

The primary outcome was intended to be change in science attainment, as measured by pupil-level science test (standardised with mean of zero and standard deviation one, and then pooled). As outlined in the SAP,[31] we proposed to use a two-level multilevel model to account for clustering of pupils in schools. Multilevel approaches assume that the schools in the study are a random sample of all schools and the multilevel modelling framework can flexibly handle complex variation within or between schools (Snijders and Bosker, 2012).

The main analysis would have modelled outcomes of pupils nested in schools, which is:

$$Y_{ij} = \beta_0 + \text{PSQM}_j\tau + Z_j\beta_1 + X_{ij}\beta_2 + u_j + e_{ij} \quad (1)$$

where Y_{ij} is the science achievement of student *i* in school j (the primary outcome measure); PSQM_j is a binary indicator of the school assignment to PSQM [1] or control [0]; Z_j are school-level characteristics, here the stratifying variable of geographical location (hub) used for randomisation; X_{ij} represents characteristics at pupil level (pupil *i* in school j), specifically standardised baseline pupil scores (KS1 maths, reading, and writing), standardised and pooled as for the outcome measure); u_j are referred to as school-level residuals ($u_j \sim i.i.d\ N(0, \sigma_u^2)$) and e_{ij} are individual-level residuals ($e_{ij} \sim i.i.d\ N(0, \sigma_e^2)$). In relation to X_{ij}, this assumes that the KS1 measures are not too highly correlated with each other. If the pair-wise correlation between measures exceeds r = 0.7 (that is, the shared explained variance is over 50%), then only one measure will be included (KS1 maths score, treated in the same manner as above).

Equation (1) is known as a ‘random intercepts’ model because $\beta_{0\text{j}} = \beta_0 + u_j$ is interpreted as the school-specific intercept for school j and $\beta_{0j} \sim i.i.d\ N(\beta_0, \sigma_u^2)$ is random (as in it can take any value). The total residual variance can be partitioned into two components: the between-school variance σ_u^2 and the within-school variance σ_e^2.

Our target parameter (the focal result of the trial), τ, would have been the average effect of the intervention on pupil outcomes compared to control schools. The τ coefficient refers to the relationship between PSQM allocation and the outcome for Cohort B represents the main result of the trial. All analyses will be performed in Stata, versions 15.1 onwards.

The outcome analysis would have been intention-to-treat (ITT). This method compares outcome means for the treatment and comparison groups, and subjects are analysed according to their randomised group allocation. The ITT approach is inherently conservative as it captures the averaged effect of offering the intervention, regardless of whether or not the participants comply with the assignment.

Secondary analysis

The planned secondary outcome was ‘attitudes to science and science skills’ among Year 5 pupils. The attitudinal measure at post-test was to be administered by NFER using paper forms. The attitudinal measures were to be compiled in machine-readable forms to allow scanning, data entry, and scoring by RAND Europe. Enjoyment of science, confidence in science and engaging teaching in science were to be measured using the ‘enjoyment of science’ subscale adapted from the Trends in International Mathematics and Science Study (TIMMS) Grade 4 surveys from TIMMS used in 2015.

To ensure that this adapted attitudes to science measure was valid, a small validation pilot was conducted in the summer of 2019. This pilot used a sample of Year 5 pupils in non-trial schools to test the measurement structure of the variable in an exploratory factor analysis. The pilot was to yield a final measure to be used in the trial. The plan was for this measure to generate a continuous score that would have been entered into Equation (2) for the purposes of analysis.

[31] **https://d2tic4wvo1iusb.cloudfront.net/documents/projects/PSQM_SAP.pdf**

$$Y_{ij} = \beta_0 + \text{PSQM}_j\tau + Z_j\beta_1 + X_{ij}\beta_2 + u_j + e_{ij} \quad (2)$$

where Y_{ij} is the science attitude of student *i* in school j (as captured with the science attitude measure above); PSQM_j is a binary indicator of the school assignment to intervention [1] or control [0]; Z_j are school-level characteristics, here the stratifying variable of geographical location (hub) used for randomisation; X_{ij} represents characteristics at pupil level (pupil *i* in school j), specifically standardised baseline pupils scores (KS1 maths, reading, and writing), standardised and pooled as for the outcome measure); u_j are referred to as school-level residuals ($u_j \sim i.i.d\ N(0, \sigma_u^2)$) and e_{ij} are individual-level residuals ($e_{ij} \sim i.i.d\ N(0, \sigma_e^2)$).

The unit of analysis for the secondary outcome analysis was to be the pupil.

Subgroup analyses

With the number of schools at randomisation and an interaction term approach to undertaking the subgroup analysis, the study was powered for meaningful subgroup analysis, such as stratifying by FSM pupils. The plan had been to report mean outcomes by sub-categories of FSM as a basic descriptive step. As an exploratory analysis we had planned to do subgroup analyses for FSM, acknowledging that this analysis may have been underpowered. As an exploratory modelling approach, EverFSM was to be incorporated into the multilevel model as a binary variable: 1 if EverFSM, 0 otherwise. The EverFSM indicator was then to be interacted with treatment allocation to assess the conditional impact of PSQM on FSM pupils.

We had planned to conduct the FSM subgroup analysis for the primary outcome only (that is, using Cohort B data), using the model in Equation (3) below:

$$Y_{ij} = \beta_0 + (\text{PSQM}_j \times FSM_i)\tau_s + Z_j\beta_1 + X_{ij}\beta_2 + u_j + e_{ij} \quad (3)$$

where Y_{ij} is the achievement of student *i* in school j; $(\text{PSQM}_j \times FSM_i)$ is a binary indicator of the school assignment to intervention [1] or control [0] interacted with the pupil-level binary eligibility for FSM variable; Z_j are school-level characteristics, here the stratifying variable of geographical location (hub) used for randomisation; X_{ij} represents characteristics at pupil level (pupil *i* in school j), specifically pupils' standardised baseline scores (KS1 maths, reading, and writing), standardised and pooled as for the primary outcome measure analysis); u_j are referred to as school-level residuals ($u_j \sim i.i.d\ N(0, \sigma_u^2)$) and e_{ij} are individual-level residuals ($e_{ij} \sim i.i.d\ N(0, \sigma_e^2)$).The sum of the coefficient for the treatment and the coefficient for the interaction of the treatment variable with FSM would have represented the result for the subgroup analysis.

A second set of subgroup (moderator) analyses were also planned, again for Cohort B, using gender (operationalised here as a binary variable, coded 1 if female, 0 if male) as the moderator. Equation 3 was to be applied in the same manner as above, replacing the FSM indicator with a binary gender indicator.

As these analyses were to be exploratory and potentially underpowered, we would have reported point estimates and confidence intervals transformed into effect sizes but would not report significance tests or p-values.

Additional analyses

The evaluation team had also planned to carry out an exploratory analysis in relation to teacher retention for the duration of the trial. This was to be undertaken in a two-stage least squares approach, for the main outcome of the trial only.

Missing data

Missing data can arise from item non-response or attrition of participants at school, teacher, or pupil levels. The evaluation team had planned to determine the proportion of missing data in the trial and had proposed to explore attrition across trial arms as a basic step to assess bias (Higgins, et al., 2011). The evaluation team was planning to provide cross-tabulations of the proportions of missing values on all baseline characteristics at both pupil and school level, as well as on the primary outcome measures.

To assess whether there are systematic differences between those who drop out and those who do not—and thus whether these factors should be included in analysis—the team had intended to model missingness at follow-up as a function of baseline covariates, including treatment. The analysis model for this approach would have mirrored the multilevel level model given above (pupils clustered in classes), but the outcome would have been a binary variable identifying missingness (yes/no).

For less than 5% missingness overall, a complete-case analysis may have sufficed (that is, assuming data is Missing Completely at Random, MCAR), but our default would have been to check results using approaches that account for missingness but that rely on the weaker Missing at Random (MAR) assumption. Our preference was to use Full-Information Maximum Likelihood (FIML) over multiple-imputation (MI) because FIML can be estimated in a single model and simulation studies show that it can reduce bias as well as MI (for a discussion of FIML versus MI, see Allison, 2012). For missingness on outcome variables only, standard statistical packages such as Stata use ML for estimating parameters so FIML would not be necessary.

Compliance

The main framework of analysis for this trial was intended to be ITT. This means that we would use the original school allocations in the final analyses, regardless of the level of compliance or implementation fidelity. However, the evaluation team had also planned to explore the effect of the intervention on schools that implemented the intervention with a high degree of compliance.

In collaboration with the delivery team and drawing on the logic model for the intervention, a binary compliance measure was developed. We defined 'compliance' as the fulfilment of a set of minimum criteria that determine whether a school had effectively participated in PSQM. This is a binary measure, indicating whether a school was compliant or not. It involved scoring separate elements of the implementation using binary variables to capture whether an identified dimension was met by the school. Missing data on any measure was to be scored as zero. Schools did not have sight of the compliance measure checklist.

Table 6 specifies the three dimensions that were considered for the compliance metric. These are training attendance—distinguishing between science subject leader and Year 5 teacher attendance—and PSQM task completion. These dimensions are equally important and receive equal weighting in the final compliance score per school.

Table 6: Dimensions contributing to the compliance metric

Compliance dimension	Variable (options)	Data source
Science subject leader attendance at training sessions	Did the science subject leader attend all four training sessions (yes/no)?	Attendance logs from PSQM
Year 5 teacher attendance at training sessions	Did the Year 5 teacher attend at least two training sessions (yes/no)?	Attendance logs from PSQM
School task completion	Did the school complete all ten tasks (yes/no)?	School task completion logs (looking into submission of common core documents in VLE)

The evaluation team had planned to receive training attendance logs collected by hub leaders to assess on a school-by-school basis whether the science subject leader and Year 5 teacher attended training sessions. There were four sessions in total; it was mandatory for SSLs to attend all four sessions and Year 5 teachers needed to attend at least two sessions.

The completion of PSQM tasks was also considered a key aspect of compliance. For this, the evaluation team accessed PSQM's virtual learning environment platform to check the extent to which each required task was completed by each school. Each school was required to complete and upload a total of ten tasks (16 documents); a school was considered compliant if it submitted all documentation for the ten tasks in the VLE.

Update. It was not possible to obtain a complete record of attendance from hub leaders. Attendance data was provided by 66 of 76 PSQM schools (87% of all PSQM schools). The evaluation team acknowledges that missing data may not equal lack of attendance. For example, changes made to facilitate participation during the pandemic means that not all sessions may have been recorded or that alternative approaches to training were taken (that is, more ad-hoc and bespoke). Given the nature of the submission process, the PSQM delivery team believes that it is impossible to complete a submission without the SSL taking part in all of the training. Furthermore, attendance for Year 5 teachers became non-compulsory during the changes made by the PSQM team as a result of the pandemic, which is likely to have affected attendance rates. However, the analysis was completed in line with the protocol using available data and the extent of the missing data was specified in the 'Fidelity ' IPE findings section. We note that this is unlikely to be a true reflection of attendance.

Effect size

The evaluation team planned to use the effect sizes for cluster-randomised trials given in the EEF evaluator guidance; an example, adapted from Hedges (2007), is given below:

$$ES = \frac{(\bar{Y}_T - \bar{Y}_C)_{adjusted}}{\sqrt{\sigma_S^2 + \sigma_{error}^2}}$$

where $(\bar{Y}_T - \bar{Y}_C)_{adjusted}$ is the mean difference between intervention groups adjusted for baseline characteristics and $\sqrt{\sigma_S^2 + \sigma_{error}^2}$ is an estimate of the population standard deviation (variance).

In the multilevel models, this variance would have been the total variance—across both pupil and school levels, without any covariates, as emerging from a 'null' or 'empty' multilevel model with no predictors. The effect size (ES), therefore, would have represented the proportion of the population standard deviation attributable to the intervention. A 95% confidence interval for the ES, that takes into account the clustering of pupils in schools, would have also been reported. Effect sizes would have been calculated for each of the regressions estimated.

Implementation and process evaluation—research methods and analysis

Update. In line with the decision to cancel testing for the trial and with the intention to minimise burden on schools while maintaining the quality of data collected for the evaluation, the planned IPE activities were revised. IPE data collection was reduced to only one wave in June 2021 (removing planned data collection for 2020). Therefore, IPE data is the only source of information for the project.

To strengthen the evaluation, the scope of the planned case studies was updated to include control schools (initially planned only for PSQM schools). While the extension of case studies to include control schools may seem like an increase in school burden, overall, burden on schools from the trial was reduced as there were no longer requirements for testing. In addition, with the revised IPE plans, input was requested from case study schools at only one point in time (as opposed to several timepoints as originally envisioned). Exact details on the updated activities are provided in

Table 7, with more detail provided in the text that follows.

Table 7: Overview of IPE data collection (updated plan)

Data type		Participant	When	Who collected the data	Topics
Online Surveys	**Headteacher survey 1**	Headteacher/ SLT	Year 1, Sep 2019	RAND Europe designed, and shared with schools	Experience with other trials/research school status (baseline); usual practices around teaching science (baseline).
	Teacher survey	Y5 Teachers (teacher selected for PSQM)	Year 2, May–Jun 2021	RAND Europe to design, PSQM to share with schools	Experience with intervention activities (PSQM schools only); school's commitment to science (all schools). Teacher's background (highest education in science, years teaching, years in the school) (all schools); whether Y5 teacher who took part in PSQM continued teaching the Y5 or not (PSQM schools only); sustainability of changes related to PSQM (PSQM schools only).
	Science subject leader survey	Science subject leaders	Year 2, May–Jun 2021	RAND Europe to design, PSQM to share with schools	Experience with intervention activities (PSQM schools only); school's commitment to science (all schools); leader's background (years as subject leader, all schools); whether Y5 teacher who took part in PSQM continued teaching the Y5 or not (PSQM schools only); sustainability of changes related to PSQM (PSQM schools only).
	hub leader survey	hub leader	Year 2, May–Jun 2021	RAND Europe to design, PSQM to share with schools	Interactions with the schools, perceived level of school engagement, perceived barriers and enablers (PSQM schools only)
	Headteacher survey 2	Headteacher/ SLT	Year 2, May–Jun 2021	RAND Europe to design, PSQM to share with schools	Usual practices around teaching science (all schools); sustainability of changes related to PSQM (PSQM schools only).
Case studies	**Interviews**	Teacher, headteacher, science subject leader (3 people per school)	Year 2, May–Jul 2021	RAND Europe to design/conduct,	Experience with intervention activities (PSQM schools only); usual practices around teaching science (all schools); school's commitment to science (all schools); sustainability of changes related to PSQM (PSQM schools only).
		Hub leader	Year 2, MayJul 2021	RAND Europe to design/conduct, PSQM to help contact participants	Experience delivering programme and perception of school engagement (barriers and enablers).
	Documentary review	Schools	Year 2, May–Jul 2021	PSQM provides VLE documentation to RAND Europe. RAND Europe requests other relevant documentation from schools	Assess evidence of science presence/relevance in school plans and communications.
Monitoring data from PSQM		Schools	From trial start up until PSQM submission due February 2021	PSQM Europe to share with RAND	Non enrolment numbers/reasons (all schools); post-randomisation drop-out/reasons for drop-out (all schools); training attendance logs (PSQM schools only); school task completion logs (PSQM schools only).

The updated process evaluation activities address the following research questions:

- Was the intervention implemented with fidelity for the intervention schools?
- What was practice as usual in the control schools?
- What appear to be the necessary conditions for success of the intervention?
- What were the barriers to delivery?
- To what extent has COVID-19 impacted on science teaching?

A multi-stage, mixed-methods IPE data collection plan was developed. The evaluation team collected data through monitoring data surveys for all schools in the trial (PSQM and control) in addition to a documentary review, interviews, and school visits (to observe science displays) for selected case study schools (see

Table 7). Each data collection activity is discussed in turn below. Details of responses obtained for surveys and interviews, and the number of documents reviewed, is included in the Data Collection and Response Counts section.

Surveys

Surveys were conducted at two different timepoints. **Baseline headteacher online surveys** were conducted in autumn 2019 to understand headteacher motivations for joining the trial and their school's current practice related to science teaching.

Endline online surveys were sent to headteachers, nominated teachers, science subject leaders, and hub leaders in summer term 2021. The focus of the surveys was to collect information on perceptions of, and attitudes to, science teaching and learning, changes in science practice (in both PSQM and control schools) and the extent to which they were attributed to PSQM (in PSQM schools), and usual practice (in control schools). The questions were based on relevant expected outcomes and practices as defined in the theory of change (Figure 1) and typical PSQM practice (Table 2). Where possible survey items were replicated across control schools and PSQM schools so that comparison could be made. The survey also asked about how COVID-19 affected running science-related activities. In addition, in the PSQM arm, intervention-specific questions were included based on the expected intervention outcomes outlined in the logic model. Descriptive quantitative analyses will be used to analyse survey data using Stata. Hub leader surveys focused on their experiences of working with schools and any barriers and facilitators to implementation.

Surveys were prepared and distributed by RAND Europe, after review by PSQM and the EEF. Further details on survey administration and response rates can be found in Data Collection and Response Counts.

Case studies

The evaluation team originally planned to approach five PSQM schools to carry out in-depth case studies. However, this was expanded in the end to cover twenty schools (ten PSQM schools and ten control schools). The aim was to use the control case studies to get a better understanding of business as usual and to compare perceptions, activities, and practice between PSQM and control schools. Selection criteria were used to ensure a range of schools were included in the study. For PSQM schools these were:

- type of award the school was working towards;
- whether they are single or multi-form entry; and
- location (hub).

In the case of control schools, the same criteria applied with the exception of award. Once schools had been selected, we reviewed Ofsted ratings to ensure they were broadly proportionate to national averages (that is, not all 'outstanding' or 'requires Improvement').

Data collection activities were then arranged within the case study school as outlined below.

Interviews with staff

Original plans were to conduct interviews in the five case study schools with school teaching staff and headteachers as well as school governors. There were going to be two rounds of interviews, one in December 2020, another in June 2021. For the latter, initial plans were for a member of the evaluation team to conduct school visits to do the interviews in person and to observe whether there are any science boards displayed around the school and document this. This was also intended to increase the chances to interview other relevant school stakeholders such as governors. However, as a result of COVID-19 disruptions, this was reduced to one round of remote interviews in June 2021. No interviews with school governors or site visits were conducted.

Headteachers, science subject leaders, and Year 5 teachers were interviewed in the case study schools. Interviews were used to gain a more in-depth understanding of what PSQM involved in practice and explore the mechanisms of change as a result of the intervention. Interviews also sought to gain a general understanding around science practices in case study schools (PSQM and control). The questions were based on relevant expected outcomes and practices as defined in the theory of change (Figure 1) and typical PSQM practice (Table 2) with a particular focus on:

- experience with intervention activities (PSQM schools only);
- usual practices around teaching science (all schools);
- school's commitment to science (all schools); and
- sustainability of changes related to PSQM (PSQM schools only).

Where possible, questions were kept similar across different respondents to compare findings and understand perceptions of different staff. Thematic coding was used to analyse responses with relevant answers being compiled in a coding framework (See Annex J): this allowed themes to be explored across respondents and between trial arms (intervention and control).

Interviews with hub leaders

Interviews with hub leaders were conducted to understand their experience of delivering the programme and their perception of school engagement. Questions were designed to be similar between school staff interviews and hub leader interviews to compare findings and understand perceptions of different people working in PSQM. The same analysis approach discussed above was used to code and analyse hub leader responses.

Documentary review

PSQM submissions from the selected case study schools and an additional 20 randomly selected PSQM schools were reviewed (resulting in a total of 30 submissions). Each submission consisted of 16 documents. After an initial screening to identify the most informative documents (those containing most relevant information in light of the theory of change, most information contained within the document, and ease of comparability across schools), four documents were selected for an in-depth review:

- calendar—illustrating schools' programme of science events;
- CPD Log—schools' record of CPD activities for SSLs and teachers;
- portfolio—summarising the approach to science teaching and how staff were implementing changes as a result of PSQM; and
- reflections form—one out of the four 'PSQM Action to Reflection Plan' forms was randomly selected; these reflections were written by SSLs and submitted as part of the PSQM submission.

These documents contained information on key PSQM aspects such as enriching the science learning experience, improving the shared understanding of the value of science across the school, a focus on CPD, the centrality of science enquiry, guiding principles for science teaching, a better understanding of science assessment and monitoring of teaching practices, and a commitment to expand children's science capital, among other aspects. Each submission was cross-checked against each dimension and frequencies logged to understand the extent to which the documents reflected the different dimensions.

This resulted in 120 documents being reviewed in-depth using an extraction template to assess whether they provided evidence on the 26 dimensions outlined in Table 2. The goal of the review was to identify whether there was evidence on changes to practices as expected from participating in PSQM (Table 2) and as set out in the theory of change (Figure 1).

Control school documents

To understand business as usual, the following documents were requested from control schools to capture equivalent evidence of science-related activities and practice:

- Ofsted reports;
- lesson observation notes;
- feedback on school improvement plans (SIPs);
- reports for, and communications with, governors; and
- newsletters to parents.

While acknowledging that these documents from control schools do not translate to a direct comparison to PSQM submissions, they were chosen as they contained information that could be used to assess differences and similarities between control and PSQM case study schools around science practice. That said, the evaluation team is aware of the limitations around the comparisons that can be drawn. On one hand, the volume of documents reviewed is notably different: the review of PSQM submissions covered 120 documents from each of the 30 selected schools while the review of documentation from control schools only covered 16 documents from four schools. On the other hand, as reported in the Findings from the Implementation and Process Evaluation section, the review of control schools' documentation yielded little evidence on the 26 dimensions assessed. The evaluation team is aware that differences may be explained by the fact that PSQM schools were required to provide evidence on these dimensions and were given guidance on how to do this and hence were more likely to focus on these in their submissions. There was also a proportionately smaller volume of documents submitted by control schools, which reduces the number of documents that could potentially provide evidence.

Monitoring data

Monitoring data—including training attendance data and PSQM task completion logs—was collected to better understand fidelity and other aspects related to implementation. In addition, it was planned to collect additional data on the number of schools that were successful in gaining the quality mark they aimed for and the numbers of cases when a submission was sent for a second review and a school was asked to submit additional evidence to get the quality mark or when a school had a deferral or extension.

Two of the 70 engaged PSQM schools (that is, those that did not withdraw and remained responsive) were awarded a different quality mark from the one they had originally aimed to achieve. In addition, due to COVID-19-related disruptions, all schools were granted an extension to deliver their final PSQM submissions by the end of March 2021, with four receiving an extension until the end of June 2021. All engaged schools obtained a PSQM by the end of the programme extensions. Only training attendance and rates of completion of the PSQM programme were taken into consideration to assess fidelity. Analysing data on whether a submission was sent for second review, whether a school was asked to provide additional evidence, or if their submission was deferred or extended was foregone as these would provide limited insights in light of the different context COVID-19 disruptions brought about during the evaluation period.

Examining continuity of Year 5 teachers

The protocol outlined that the team would descriptively compare programme outcomes for those schools where the same teachers are working in Year 5 in 2020/2021. The information on continuity would be based on self-report by teachers in Year 2 of the trial as part of their surveys.

Given that no pupil outcome data was collected, a meaningful analysis of data on Year 5 teacher continuity was not possible (for example, to assess whether pupils having the same teacher across the years obtained better outcomes). Hence, an analysis of this data was not done given the lack of clarity on how such findings would be interpreted in light of IPE data only.

Costs

Given that there is no impact estimate and that costs would not be representative of ‘usual’ delivery practices, data on costs was not collected. This also helped reduce the data collection burden on schools.

Cost data was intended to be gathered through online surveys as well as through the interviews in the IPE (see above). Questions were going to be targeted at assessing any prerequisite costs (such as training costs and materials) and any direct and marginal costs directly attributable to schools’ participation in the intervention (printing, staff time, cover, and so on). This information would have been used to estimate cost per-pupil, following EEF guidelines (EEF, 2015).

The main costs of the intervention relate to training, materials, and the time of teachers and subject leaders to complete the programme activities. To calculate the cost of training and materials, the evaluation team would use data provided by the delivery team. RAND Europe also intended to also take into account the cost of the time of hub leaders, headteachers, teachers, and other staff in delivering the programme.

Timeline

Details of the timeline are outlined in Table 8.

Table 8: Final evaluation timelines

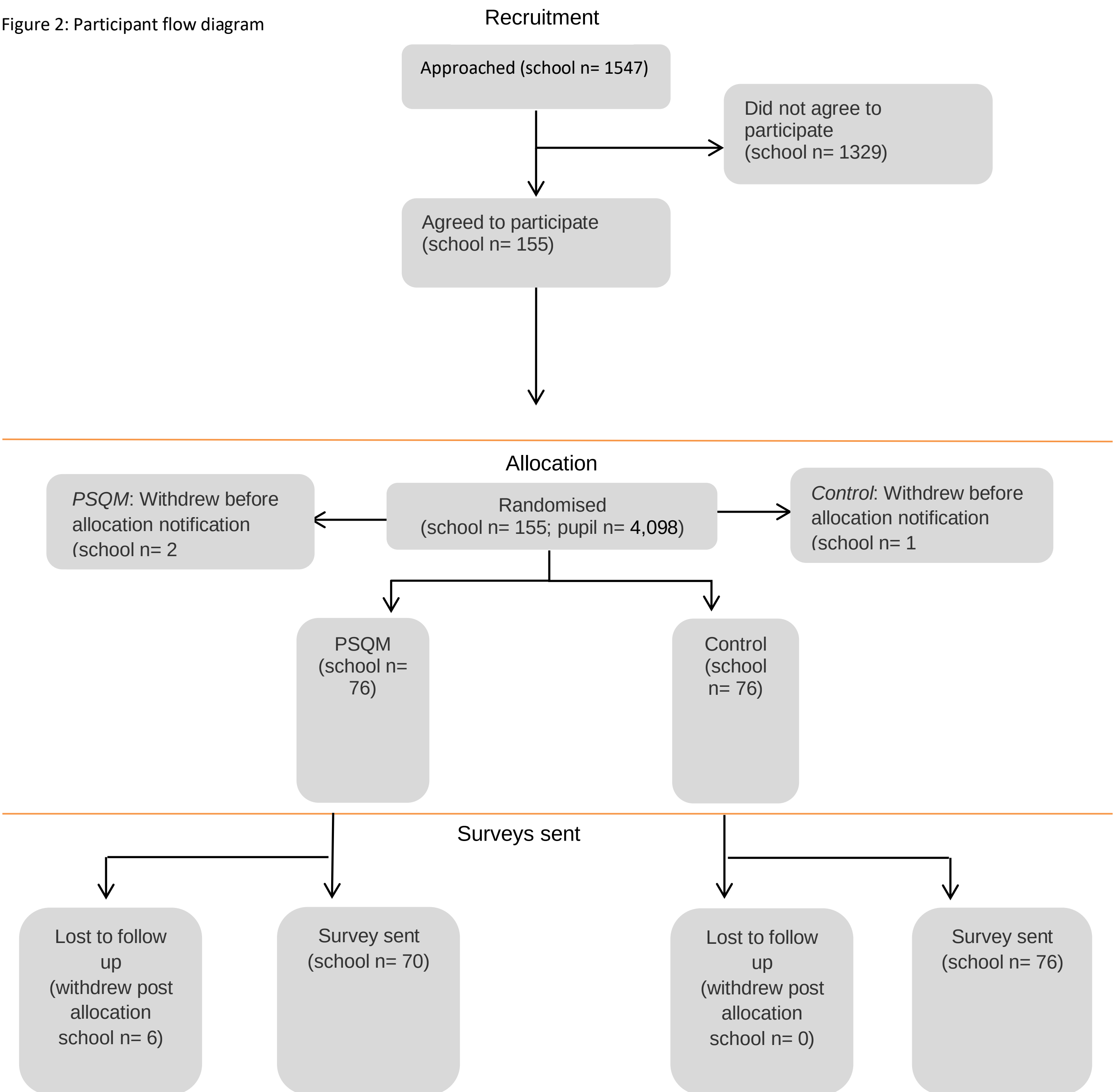
Timeline Dates	Activity	Staff responsible/ leading
Oct 2018	IDEA workshop	RAND Europe
Jan–Apr 2019	Recruiting schools and teachers	University of Hertfordshire
Jan–Apr 2019	Opt out forms to be sent to parents	Schools
Mar–May 2019	Collection of pupil information	Schools
May–Jun 2019	School and pupil information to be collected sent to RAND	University of Hertfordshire
May–Jun 2019	Randomisation	RAND Europe
Sep–Oct 2019	Baseline survey of headteachers all schools	RAND Europe
Sep 2020	Completion of Statistical Analysis Plan	RAND Europe
Sep 2019–Feb 2021*	Programme implementation	University of Hertfordshire
Feb–Apr 2021*	Compilation of CPD attendance records, task completion and other intervention data for compliance measure/IPE	University of Hertfordshire
May–Jul 2021	Interviews in case study schools (teachers, subject leaders, headteachers) and documentary review	RAND Europe
May–Jun 2021	Surveys of headteachers, subject leaders, Year 5 teacher (all schools) and hub leaders	RAND Europe
30 Dec 2021	Draft EEF report	RAND Europe
April 2022	Final EEF report	RAND Europe

Findings from the implementation and process evaluation

Participant flow diagram

Figure 2 presents details of the participants' flow through various stages of the evaluation up to responses on the survey. In total, 1,547 schools were invited to take part. Of those, 155 agreed to participate and were randomised. Before allocation was revealed three schools withdrew—two that had been allocated to PSQM and one that had been allocated to control. As a result, the trial retained 152 schools (76 intervention and 76 control). By the summer of 2021, five PSQM schools had formally requested to be withdrawn from the trial and one PSQM school had stopped responding to emails from the PSQM and evaluation teams and was considered withdrawn. Of the remaining schools, staff in 70 PSQM schools and 76 control schools were sent surveys in June 2021.

Figure 2: Participant flow diagram



Data collection and response counts

Findings from the IPE are informed by the analysis of surveys of, and interviews with, hub leaders, headteachers, science subject leaders, and Year 5 teachers from across PSQM and control schools. The tables below summarise the survey response rates. It should be noted that there were issues with the administration of the survey early in the process. The link to the survey was accidentally sent out to contacts from the 2019/2020 academic year as opposed to the 2020/2021 academic year. This meant that some staff did not initially receive the email with the request to complete the survey. In total, 109 of 146 schools were not affected (that is, staff emails were consistent across the two academic years); in the remaining schools the RAND team contacted the relevant individual (the person whose details were updated in 2020/2021). Given the error, the deadline for the survey was extended to 30 July and teachers that got in contact up to September 2021 were allowed to complete the survey. Though the deadline was extended, the evaluation team acknowledges that this may have impacted on response rates.

Table 9: Survey responses in PSQM schools

	Headteacher baseline	Headteacher endline	SSL*	Year 5 teacher	Number of schools with complete responses at endline**
Completed in/by	63	48	68	46	40
Total number of schools			70		

Counts under 'completed' related to the number of respondents that accessed the survey. Not all survey questions have complete responses as some were skipped.
* 12 SSLs indicated having a dual role (SSL and Y5 teacher) when completing the survey.
** Schools with responses to the headteacher, SSL, and teacher surveys.

Table 10: Survey responses in control schools

	Headteacher baseline	Headteacher endline	SSL*	Year 5 teacher	Number of schools with complete responses at endline**
Completed in/by	63	61	63	51	44
Total number of schools			76		

Counts under 'completed' related to the number of respondents that accessed the survey. Not all survey questions have complete responses as some were skipped.
* 23 SSLs indicated having a dual role (SSL and Y5 teacher) when completing the survey.
** Schools with responses to the headteacher, SSL and teacher surveys.

IPE findings are also based on the analysis of interviews and school documents related to science teaching and learning in both PSQM and control schools. This analysis compares data of 120 documents from 30 PSQM schools and 16 documents from four control schools.

Table 11: Interviews in PSQM schools

	Headteachers	SSL	Year 5 teacher	Total
Number of interviews	4	7	5	16
Number of schools			7	

Ten PSQM schools were randomly selected to be approached to participate in the evaluation case studies; ten schools were selected as back-up. Following contact to all schools, including all in the back-up list, only seven PSQM schools agreed to participate.

Table 12: Interviews in control schools

	Headteachers	SSL	Year 5 teacher	Total
Number of interviews	7	6	4	17
Number of schools			7	

Ten control schools were randomly selected to be approached to participate in the evaluation case studies; four were selected as back-up. Following contact to all schools, including all in the back-up list, only seven control schools agreed to participate.

Table 13: Hub leader survey and interview counts

	Hub leader survey	Hub leader interviews
Number respondents	14	7
Total number	16	10*

* Only hub leaders of schools participating in case studies (ten) were approached.

Not all schools that were randomised were approached to take part in the survey. It was decided by the evaluation team not to contact five of the PSQM schools that withdrew from the trial after allocation/ during the intervention and the one school that was no longer responsive to RAND or PSQM's attempts at contact. Firstly, we could not obtain updated contact details for any individual from these schools (they remained unresponsive and in some cases contacts from the previous school year had moved on). Secondly, as schools were not responding to contact from PSQM or RAND it was felt best to respect the difficult context in which schools were operating in summer 2021 (that is, managing catch-up as a result of school closures and continuing COVID-19 outbreaks). It is acknowledged that by not contacting schools that may have found it hardest to engage with PSQM it has potentially limited the ability of the evaluation to understand what barriers affected these schools. However, we also wanted to respect the fact that after numerous attempts at contact these schools did not want to engage with the evaluation.

Before the trial

Key findings

- PSQM and control schools were similar on a range of characteristics including proportions of FSM and EAL pupils, school size, and Ofsted rating.
- A higher proportion of headteachers in PSQM schools had a university-level degree in science compared to their control group counterparts (37% compared to 30%), although more headteachers in control schools held an A level in science (19% compared to 6%).
- Prior to PSQM, headteachers across both arms expressed similar levels of agreement with regard to science teaching provision in their schools, such as appropriateness and adequacy of science teaching resources, budget, and leadership attitudes to science.
- There was little variation in the time spent by the SSLs on science-related CPD as reported by headteachers in the PSQM and control schools in the baseline survey: 42% of headteachers in PSQM schools reporting that SSLs had one or two days of CPD, compared to 44% in the control group.

This first section of the IPE findings provides an overview of the context and usual practice of the PSQM and control schools prior to the start of the trial. First, baseline characteristics of schools on a range of school and pupil characteristics are presented. Then, findings from the headteacher baseline surveys to understand the usual practice of science teaching and learning in schools before the trial began are discussed. The aim is to understand if there were any key differences that may have impacted participation and engagement such as headteacher ‘buy-in’ to science teaching and learning, engagement in science CPD, or active engagement with whole-school teaching and learning.

PSQM and control schools had similar proportions of FSM and EAL pupils

As can be seen in Table 14, the average proportion of FSM pupils across the schools was 22.5% in PSQM schools and 20.1% in control schools.[32] When FSM at the pupil level was analysed using data collected prior to randomisation, 9.5% of pupils in PSQM schools were classified as FSM compared to 8.6% in control schools. In the case of pupil-level EAL proportions, this was slightly higher in PSQM schools (7.9%) than in control schools (5.7%).

Table 14: Baseline characteristics of PSQM and control schools’ FSM and EAL levels

Characteristic	PSQM schools (%)	Control schools (%)
Average FSM school level across all trial schools	22.5	20.1
Average FSM at pupil level across Y5 classes in trial	9.5	8.6
Average EAL at pupil level across Y5 classes in trial	7.9	5.7

The FSM average proportion is based on information from 75 of the 78 PSQM schools and 76 of the 77 control schools. The FSM pupil-level percentage is based on information from 388 PSQM pupils and 355 control pupils. The EAL percentage is based on information from 324 PSQM pupils and 234 control pupils.

As can be seen in Table 15, there were no marked differences between PSQM and control schools on school-level characteristics. The numbers on roll were very similar across both groups. Meanwhile, a similar proportion of schools with an ‘outstanding’ Ofsted rating were observed (13% PSQM, 12% control). More variation on the proportion of schools with a ‘good’ Ofsted rating is seen between trial arms: 56% of PSQM schools had a ‘good’ rating, control schools 65%. However, a higher proportion of PSQM schools (22%) had missing Ofsted rating information than control schools (14%) owing to being academy converters.

Table 15: School size baseline characteristics of PSQM and control schools

	PSQM schools	Control schools
Average enrolment (number of pupils) across schools	314	315
Information based on n number of schools	74	76
Total number of schools in trial	78	77

Table 16: Ofsted rating characteristics of PSQM and control schools school size

	PSQM schools		Control Schools	
Ofsted ratings	Count (n of schools)	%	Count (n of schools)	%
1. Outstanding	10	13	9	12
2. Good	44	56	50	65
3. Requires improvement	7	9	6	8
4. Special measures	0	0	1	1
5. Inadequate	0	0	0	0
6. Missing —academy converter	17	22	11	14

[32] Taken from publicly available government data: **https://www.get-information-schools.service.gov.uk/**

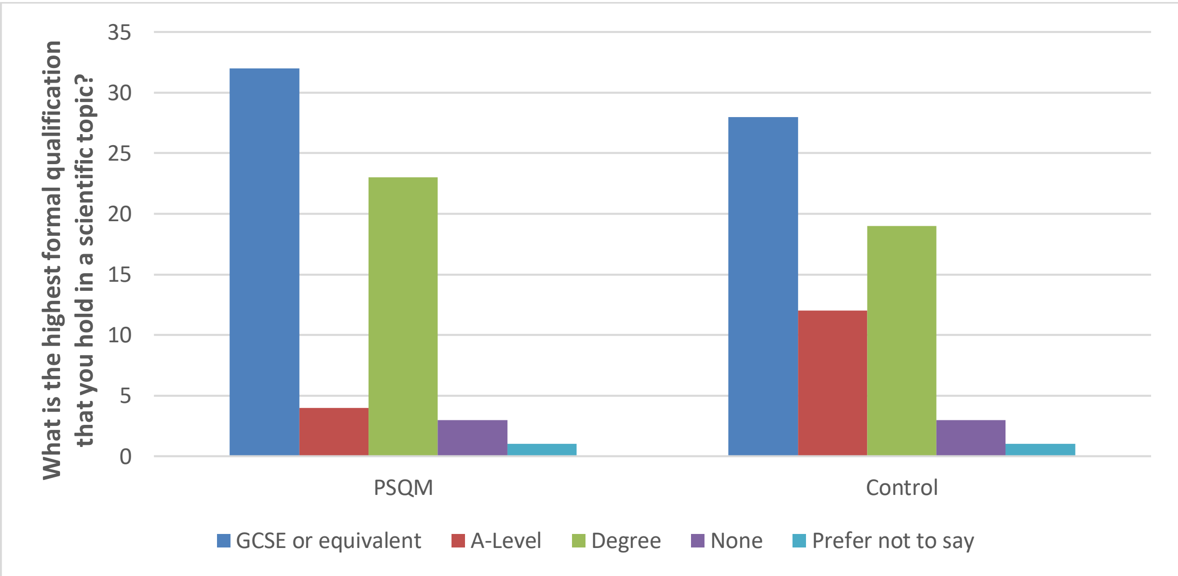
Slightly more PSQM schools reported being a research school before the trial compared with schools in the control group

Being a research school or working closely with a research school implies that schools are familiar with evidence-based practice and could be more keen to implement such practices in their settings. Before the start of the trial, 17.4% (11/63) of PSQM schools were research schools compared to 10% of control schools (6/63). However, very few headteachers across both groups reported that their school had previously worked with a research school on science related projects (1/63, PSQM; 2/63, control).

A higher proportion of headteachers in PSQM schools were educated to degree level while a lower proportion of headteachers in PSQM schools held an A level in science

Headteachers were asked about their highest science qualification to understand if there were any differences in potential predisposition towards science in schools. The hypothesis was that schools with headteachers who had a science degree may be more likely to support school-wide science teaching and learning initiatives compared to those that do not. As Figure 3 shows, headteachers in PSQM schools were more likely to have science degrees (37%; control, 30%) but headteachers in control schools were more likely to have A level science qualifications (44%; PSQM, 6%). Overall, this indicates that, before the start of the trial, headteachers in PSQM schools had higher level qualifications in a science subject than headteachers in control schools as measured by degree-level qualifications, but this was unlikely to be significant as substantially more headteachers in control schools had science A-level qualifications.

Figure 3: Headteacher survey respondents' response on their highest level of formal qualification in a scientific topic



Based on responses to the question, 'What is the highest formal qualification that you hold in a scientific topic? (please, select only one option)' in the baseline headteacher survey asked both to PSQM and control headteachers. PSQM schools: total response n = 63; control schools: total response n = 63.

Headteachers across PSQM and control schools expressed similar levels of agreement with regard to science teaching provision at baseline, although a higher proportion of headteachers in PSQM schools reported that enough time is spent on science teaching

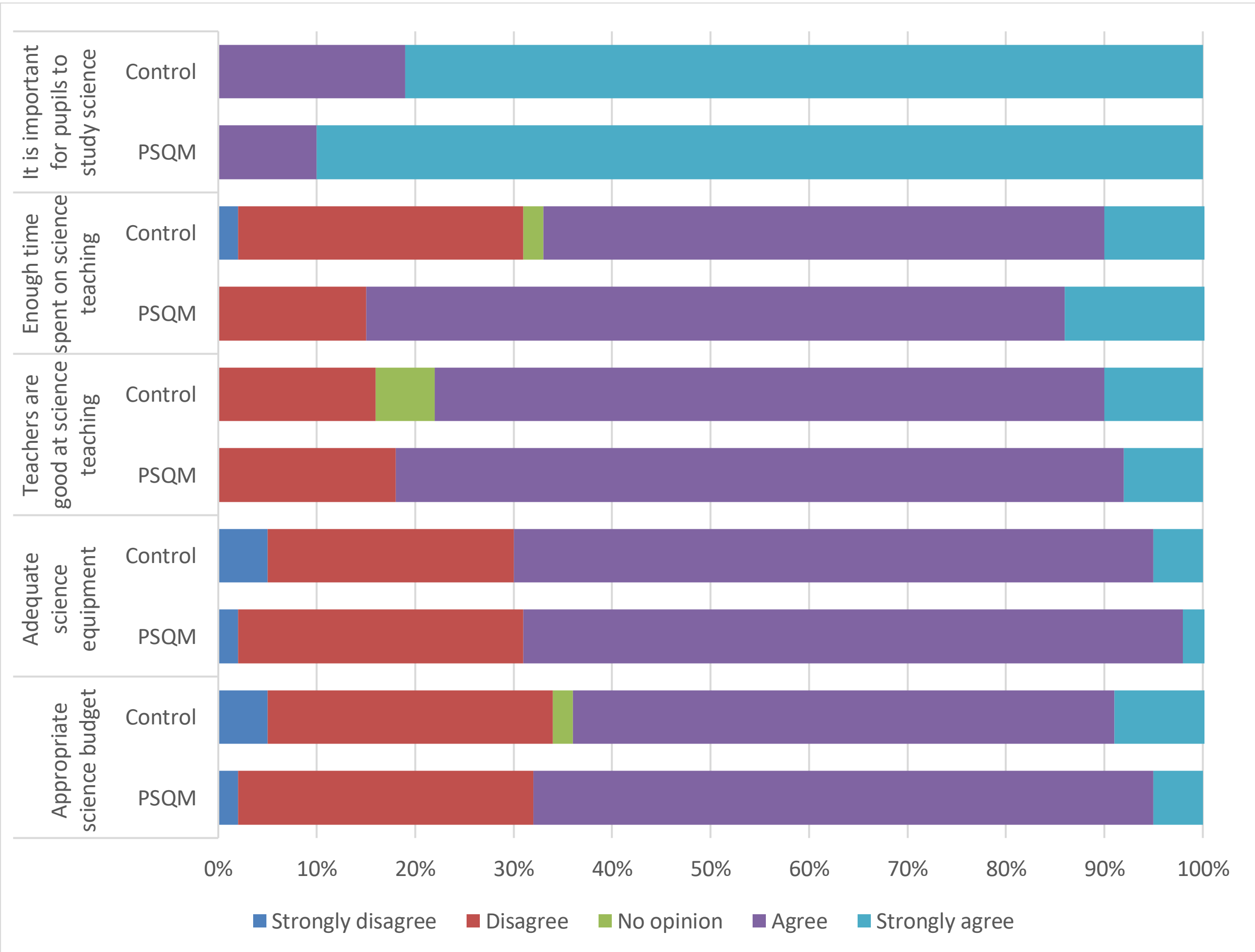
To understand the extent to which schools differed prior to the start of PSQM delivery, in the baseline survey headteachers were asked about a number of factors relating to science provision, including the appropriateness and adequacy of science teaching resources and budget, science teaching, and leadership attitudes towards science. As can be seen from Table 17: and Figure 4, the majority of headteachers in PSQM and control schools reported similar levels of agreement to relevant statements. One notable difference is that more PSQM headteachers agreed or strongly agreed that enough time was spent teaching science within their school compared to headteachers in control schools (86%, PSQM; 68% control).

Table 17: Comparison of headteacher baseline responses relating to science provision across PSQM and control schools

Survey statements	Q9.1. My school has an appropriate budget for science teaching		Q9.2. My school has an adequate range of science equipment to carry out hands-on science investigations		Q9.3. Teachers in my school are good at science teaching		Q9.4. Enough time is spent teaching science within my school		Q9.5. It is important for pupils to study science	
	PSQM	Control	PSQM	Control	PSQM	Control	PSQM	Control	PSQM	Control
Strongly agree	5%	10%	3%	5%	8%	10%	15%	11%	90%	81%
Agree	63%	55%	67%	65%	74%	68%	71%	57%	10%	19%
No opinion	0%	2%	0%	0%	0%	6%	0%	2%	0%	0%
Disagree	30%	29%	29%	25%	18%	16%	15%	29%	0%	0%
Strongly disagree	2%	5%	2%	5%	0%	0%	0%	2%	0%	0%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Responses to the question, 'With regard to science teaching in your school, to what extent do you agree or disagree with the following statements?' Baseline headteacher survey, PSQM headteachers (n = 63) and control headteachers (n = 63).

Figure 4: Headteacher survey respondents' views on science teaching in their school



Responses to the question, ‘With regard to science teaching in your school, to what extent do you agree or disagree with the following statements?’ Baseline headteacher survey, PSQM headteachers (n = 63) and control headteachers (n = 63).

There was little variation in time spent on science-related CPD as reported by headteachers in PSQM and control schools in the baseline survey

There were similar reports at baseline on time spent by SSLs on science-related CPD in the previous 12 months:[33] 42% (20/48) of PSQM headteachers reported that SSLs had one or two days of CPD; 33% (16/48) estimated that they had more than two days of CPD. This compares to 44% (24/54) and 31% (17/54) in the control group, respectively. Additionally, 19% (9/48) of PSQM headteachers and 17% (9/54) in the control group selected ‘somewhere between half a day and one day’; the corresponding figures for ‘half a day’ were 4% (2/48) in the PSQM group and 7% (4/54) in the control group. Finally, 2% (1/48) of headteachers in the PSQM group and 0% in the control group stated their SSL had not spent any time at all in science-related CPD in the previous 12 months.

Fidelity

Key findings

- SSLs attended all training sessions and found the training useful.
- Both Hub Leaders and SSLs in PSQM schools appreciated the flexibility of the PSQM training, however some felt that the networking aspect was lost during the online transition.
- Hub Leaders felt well prepared, trained and supported in their role, and the PSQM Conference and Regional Hub Meetings were instrumental to this. Many of them also drew on the support of PSQM HQ and Senior Regional Hub Leaders.

In this section, findings relating to the research question ‘Was the intervention implemented with fidelity for the PSQM schools?’ are reported. We examine what facilitated or hindered implementing PSQM as intended. We draw on the interviews with staff, surveys with school staff and hub leaders, and the PSQM submission review. To complete the analysis, the evaluation team looked at whether the findings presented in this section are in line with the set-up, activities, and outputs described in the theory of change, focusing on (i) SSL and teacher training, (ii) ongoing support by hub leaders to schools, (iii) training, preparedness, and support for hub leaders, and (iv) compliance.

PSQM training

SSLs attended all training sessions and found the training useful

Compulsory training sessions for SSLs are a core component of the PSQM programme activities and outputs, as illustrated by the theory of change. In the endline survey almost all SSLs (61/62) reported attending PSQM training sessions. SSLs’ attendance rate was also constant across all five sessions, ranging between 94% and 98%. SSLs reported finding all sessions highly relevant: two thirds of respondents described the first three sessions as ‘extremely useful’; half had the same opinion about the final two sessions. Interestingly, both SSLs

[33] In surveys, ‘CPD’ was defined as science-focused staff meetings or workshops as well as externally-provided professional development in and out of school, attendance at courses, workshops, or conferences, participation in science networks or clusters, one to one support (external or internal), team teaching, lesson study, and so forth.

and HLs rated Session 3a, 'Writing a Reflective Submission', as the most useful (SSLs: 44/61, 72%; HLs: 10/13, 77%) suggesting this is an area of the programme where SSLs welcome support.

When asked to expand on why they found some sessions more useful than others, responses ranged from the superiority of face to face sessions, the added value of having an additional session funded by the Royal Society of Chemistry, and the first meeting being particularly useful as it outlined expectations for the programme. These findings were supported by interviews with SSLs where they were particularly positive about the initial face to face training they received as it clarified expectations, broke down the process, and enabled them to interact with SSLs from other participating schools.

SSL engagement was high but somewhat varied across sessions

According to hub leaders' endline survey responses, SSL's engagement with PSQM training varied across sessions. Session 2 was found to be most engaging (11/13 respondents found SSLs engaged to a 'great extent'), followed closely by sessions 1, 3a, and 4 (10/13 respondents found SSLs engaged to a 'great extent'). Session 3b, in contrast, stands out for having a much lower level of engagement: only five out of 13 hub leaders found SSLs engaged to a 'great extent'. It should be noted that this session was optional given it was a repeat of training previously provided to support schools once they reopened after COVID-19 restrictions. Interestingly, four out of 12 hub leaders explained in the survey open-ended responses that the engagement of schools was hindered by increased workload levels caused by the COVID-19 related disruption, which led to some schools being worse off than others. Furthermore, interviews with hub leaders suggest that most SSLs were very engaged in many aspects of the PSQM programme, evidencing the use of the chat function during online training and the trainings attendance rate. This suggests that SSL engagement was high overall, with some variation likely linked to pressures caused by COVID-19.

More teachers attended the first two PSQM training sessions and perceived them to be more useful than later sessions

Ninety-five percent of all teachers attended the first training session but the figure decreased to 59% for Session 3b and Session 4. This is perhaps unsurprising given that teachers were no longer required to attend after Session 3a as a result of the adaptations made to the training as a result of the pandemic. Given this, teachers were still asked to reflect on their experience of training. When asked to reflect in the endline survey on the usefulness of the training received, teachers reported finding the first two sessions substantially useful. In fact, 49% (18/37) considered the first session to be 'extremely useful' and 46% (17/37) felt the same about the second session. However, teachers rated Sessions 3a, 3b, and 4 as somewhat less useful: only 30% (39) felt that the sessions were extremely useful.

When asked whether they had found some of the training sessions more useful than others, teachers were divided: 49% (17/35) found them 'equally useful' whereas 51% (18/35) found some sessions more useful. Teachers felt that the initial in-person sessions were more useful, allowing them to network with local science teachers, making people more likely to share their ideas (in comparison with the later online sessions), and were, overall, more informative and productive. Sessions focused on sharing practical lessons were also deemed more useful than others: those focused on sharing specific ideas that would enhance science throughout the school—such as workshops, resources, and practical strategies to take back to the classroom—were particularly appreciated. Finally, some respondents also mentioned finding sessions around completing the assignments especially useful as they could receive feedback from the course leader before the submission.

In contrast, SSLs were overwhelmingly positive about both the in-person and online training—the latter, note, was adopted in response to social distancing measures as a result of COVID-19 and was not a planned aspect of delivery (see Unplanned Adaptations Made to Delivery section).

Ongoing support by hub leaders to schools

The majority of SSLs were engaged with the virtual learning environment, which was used by hub leaders alongside other approaches to work with schools

As can be seen in the theory of change (Figure 1), a number of activities were in place to support implementation including ongoing support using a bespoke virtual learning environment (VLE). The majority of hub leaders (8/13) indicated in the endline survey that SSLs were 'greatly engaged' with the VLE, while the remaining (4/13) reported that SSLs were 'a little' or 'somewhat' engaged. In an interview, one hub leader (INT-HL-04) noted that the engagement with the VLE is very varied, with a burst of activity close to deadlines. They also emphasised the importance of encouraging SSLs to engage with the VLE and post questions on the platform. The relatively high level of engagement with the VLE is especially relevant as this is indicated in the theory of change as the only way for SSLs to complete the tasks and final submission.

In practice, hub leaders found many effective ways of working with schools. They highlighted in interviews several examples of working with SSLs that were particularly effective such as sharing good practice between SSLs, developing a sense of community among SSLs (through, for example, group email exchanges), and one to one conversations between hub leaders and science subject leaders.

Hub leader training, preparedness, and support

Hub leaders felt prepared in their role and the PSQM conference and regional hub meetings were instrumental in this

Among the hub leaders, 83% (10/12) found regional meetings and the hub leader conference extremely useful. These meetings provided an opportunity for hub leaders to exchange ideas and gain further knowledge around best practice for delivering the PSQM programme. Ensuring that hub leaders felt prepared and supported in programme delivery, to which these meetings contributed, is an important element for successful implementation. Hub leaders played an instrumental role in providing continuous support to participating schools.

Hub leaders also were confident about the level of preparation provided by the programme to fulfil their role: all respondents (13/13) agreed or strongly agreed that (i) the training they received from PSQM prepared them to fulfil their role as hub leader, (ii) they felt prepared to provide relevant support to subject leaders in their hub to complete the PSQM process, and (iii) they felt prepared to use PSQM's assessment criteria to determine their school's PSQM award at the end of the programme. Some hub leaders (4/7) mentioned that these training initiatives provided a good opportunity to talk about experiences in smaller groups and were helpful to prepare for reviewing submissions. A number of hub leaders (3/7) also felt that the PSQM conference was especially useful in explaining recent changes in science learning and teaching.

Many hub leaders drew on the support of PSQM HQ[34] and senior regional hub leaders

In interviews with hub leaders, the PSQM team and senior regional hub leaders were particularly commended for their accessibility, even responding to issues after business hours.

In the survey, 100% of hub leaders strongly agreed that it was easy to maintain regular communication with the PSQM HQ team. In addition, 12 of the 13 hub leaders strongly agreed that they had received adequate support from the PSQM HQ team to deliver PSQM and that it was easy to maintain regular communication with the PSQM senior regional hub leaders. There was less agreement on the extent to which hub leaders exchanged ideas with other hub leaders: only five out of 13 hub leaders strongly agreed with the statement, 'I exchanged ideas with other hub leaders about how to implement PSQM more effectively.' It should be noted that as a result of COVID-19 there were no face to face meetings between hub leaders. Regardless, these findings suggests

[34] This expression refers to the team at the PSQM headquarters.

that hub leaders did not use one another as a form of support much (perhaps as a result of lack of face to face contact due to COVID-19 restrictions) but rather turned to the PSQM senior leadership.

Compliance

As outlined in the protocol, the evaluation team intended to measure compliance according to three metrics: (1) SSL attendance at all PSQM training sessions, (2) Year 5 teacher attendance at a minimum of two training sessions, and (3) school completion of all of the PSQM tasks needed for submission. Unfortunately, hub leaders were not able to provide a complete record of attendance. For intervention schools, attendance data was only available for 39 of 76 schools (51%) in relation to SSLs; for teachers, the corresponding figure was 27 of 76 (36%).[35] We completed the analysis using available data and present results below including noting the extent of the missing data.

As can be seen in Table 17, 33% of all PSQM schools (22/76) were considered compliant. However, this number is unlikely to be an accurate figure given the extent of the missing training attendance data and the fact that the PSQM team made Year 5 teacher attendance at training non-compulsory. The evaluation team also understands that the eight schools that were considered 'missing' on 'school completed all ten tasks' at the time of analysis have since completed their submission. Given these considerations, it seems more likely that compliance was 100%, however, the evaluation team reports on what was planned as part of the evaluation.

Table 18: Compliance rates for PSQM according to attendance at the training and submission of tasks

Compliance measure	Number compliant	Number not compliant	Missing	Proportion compliant of total sample (%)
SSL attended all four training sessions	35	4	37	46
Year 5 teacher attended at least two training sessions	25	2	49	33
School completed all ten tasks	62	6	8	2
Schools considered compliant	22	17	37	29

The RAND team only obtained training information for 66 PSQM schools, which included data on schools that did not respond to the PSQM team. The RAND team obtained PSQM award information for 62 PSQM schools. This information was used to infer the number of schools that completed all tasks (only schools completing all tasks could be granted a PSQM award). The 'proportion compliant' is calculated based on 76 schools that were considered to be part of the trial.

[35] As noted earlier, this is likely a result of the fact that post March 2020 training became more ad-hoc and bespoke to support schools given the extra pressures of delivering in a COVID-19 environment.

Changes to practice

Key findings

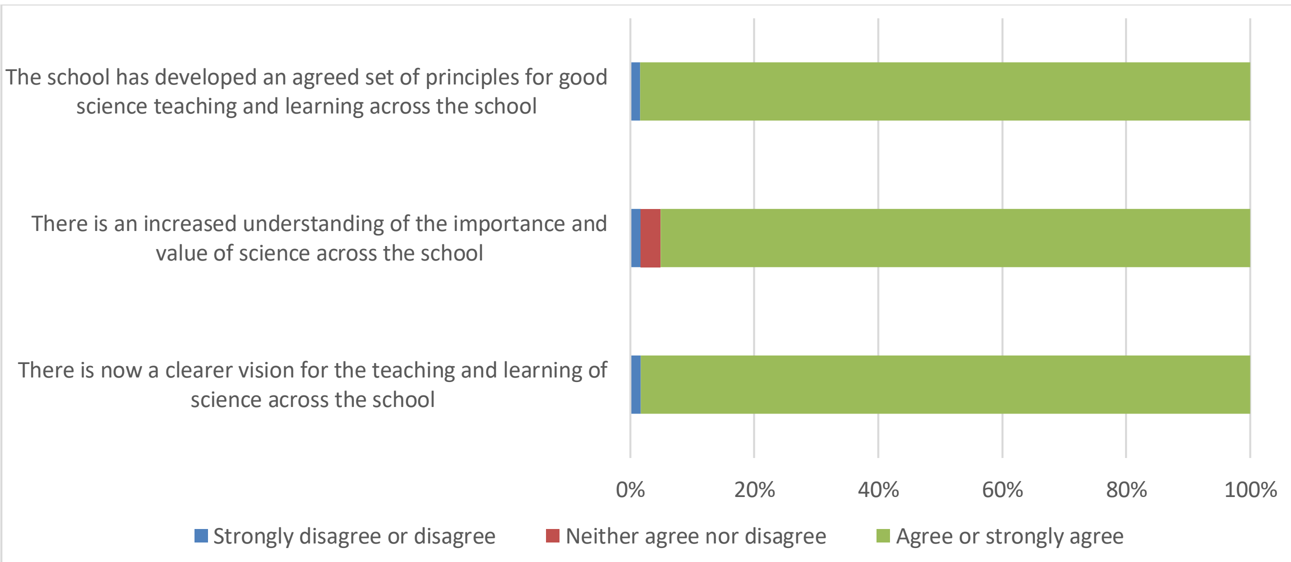
- Survey and interview data indicates that almost all school staff felt that science teaching at their school had improved as a result of participating in PSQM.
- Many school staff also reported in surveys and interviews that the science leadership in their school had become more confident and credible.
- Participants felt that the monitoring of science teaching and learning improved across the majority of schools as a result of participating in PSQM
- Survey data suggests that PSQM led to more communication with other schools, families, and the wider community.
- Almost all school staff agreed that pupil engagement and confidence improved after participating in PSQM.

We now report on whether staff practice in PSQM schools changed in line with the outputs and outcomes outlined in the theory of change (Figure 1 and Table 2) and whether staff attribute changes in science teaching to PSQM. To address these aspects we draw on findings from interviews and surveys with school staff and hub leaders and also data from the PSQM submission review.

SSLs reported perceiving that the value and vision for science teaching and learning improved as a result of their participation in the PSQM programme

As can be seen in Figure 5, the vast majority of participating SSLs agreed or strongly agreed that a range of school-wide teaching practices inherent to PSQM's expected outputs—as outlined in the theory of change—were being employed by teachers. These included a clearer vision for the teaching and learning of science (98%, 60/61), an increased understanding of the importance and value of science (95%, 58/61), and that there was an agreed set of principles for good science teaching and learning across the school (98%, 61/62).

Figure 5: SSL survey respondents' views on how participating in the PSQM programme has changed some elements of science teaching in their school



Based on answers to the question, 'Reflecting on science teaching at your school, to what extent do you agree that the following elements changed as a result of your participation in the PSQM programme?' SSL endline survey, asked to PSQM schools. SSLs total response on the different answers: 'There is now a clearer vision for the teaching and learning of science across the school', n = 61; 'There is an increased understanding of the importance and value of science across the school', n = 61; 'The school has developed an agreed set of principles for good science teaching and learning across the school', n = 62.

Responses from all interviewed headteachers and SSLs support these survey findings. Staff reported that taking part in PSQM had raised the profile of science in their schools to that of a core subject, with more time being dedicated to science. Others suggested that their schools had developed a clearer vision for science, including having science as part of the school development plan.

This was supported by the views of several hub leader interviewees who reported that since schools participated in PSQM, teachers' perceptions about the relevance of science had improved, which in turn raised the profile of science at their schools. Hub leaders also mentioned that some schools had developed a stronger vision and principles in science teaching and that staff were provided with better ownership over science teaching. In this respect, SSLs felt more equipped to talk about science and to explore different science teaching strategies. It is worth highlighting reports that PSQM participation helped develop the leadership skills of SSLs, improving both their resilience and understanding of their role (INT-HL-02, INT-HL-07).

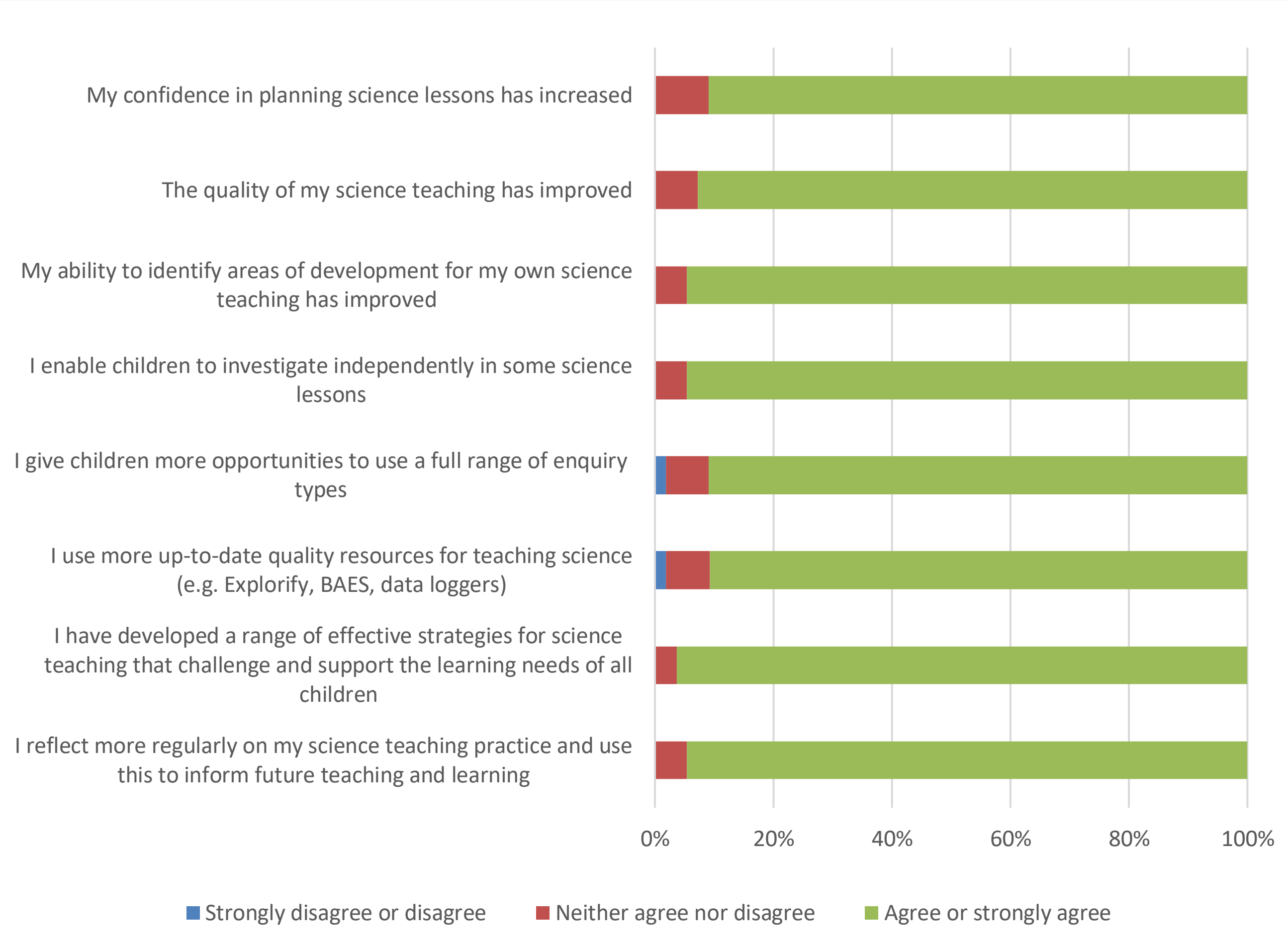
Data from the PSQM submission review corroborates the interview and survey findings (see Appendix G for details). Of the 26 dimensions related to typical PSQM practice (Table 2), seven were clearly evidenced by all 30 PSQM schools.[36] In particular, (i) a shared understanding of the importance and value of science, (ii) an agreed set of principles for good science teaching and learning, (iii) appropriate and active goals for developing science, (iv) a shared understanding of the purpose and process of science enquiry, (v) a shared understanding of the purposes of science assessment and current best practice, (vi) formative assessment used regularly to improve lessons and summarised periodically for reporting such that progression is tracked, and (vii) there are appropriate links with families, other schools, communities, and outside organisations to enrich science learning.

In addition, over 90% of surveyed Year 5 teachers felt that science teaching had improved as a result of participating in PSQM

One key outcome of the PSQM programme, according to the theory of change, is for science teaching practices to improve in participating schools. As shown in Figure 6, over 90% of Year 5 teachers agreed or strongly agreed that their teaching ability had improved in line with the dimensions of typical PSQM practice, including developing a range of effective strategies for science teaching that challenge and support the learning needs of all children (giving children more opportunities to practice a range of science enquiry skills such as asking questions, making predictions, recording data, observing, and measuring. Furthermore, teachers felt that their own skills had improved, including their ability to identify areas of development for their own science teaching and the quality of their science teaching.

[36] As illustrated in the section Implementation and process evaluation—research methods and analysis, the 26 dimensions in the data extraction sheet reflect key PSQM aspects such as enriching the science learning experience, improving the shared understanding of the value of science across the school, a focus on CPD, the centrality of science enquiry, guiding principles for science teaching, a better understanding of science assessment and monitoring of teaching practices, and a commitment to expand children's science capital, among other aspects. Each school submission was cross-checked against each dimension and frequencies logged to understand the extent to which the documents reflected the different dimensions.

Figure 6: Teacher survey respondents' views on how PSQM has changed some elements of their science teaching



Based on answers to the question, 'Reflecting on your science teaching, to what extent do you agree that the following elements changed as a result of your participation in the PSQM programme?' Endline teacher survey, PSQM schools. Teachers' total responses across all the answers, n = 55.

Interviews with staff from PSQM schools, including headteachers, SSLs, and Year 5 teachers, offer more insight into the many ways in which PSQM had improved science teaching. For example, staff reported that science teaching had become more consistent across schools due to the adoption and application of new teaching methods and tools—methods such as scientific enquiry and the use of tools such as floor books and science wheels—and confidence in planning processes and monitoring, all of which was leading to better and more transparent assessment. Interviewed teachers corroborated the changes in teaching methods being implemented, explaining that PSQM had improved and expanded the way teachers use different strategies to allow all children to access content at their own level—drawing pictures, using word maps, and writing floor books (INT-TR-04. INT-TR-06 INT-TR-07). Three teachers reported that their confidence in science teaching had improved since taking part in PSQM (INT-SSL-03, INT-TR-04, INT-TR-07) while two teachers were positive about the freedom and flexibility afforded by the programme (INT-TR-06 and INT-TR-07); one also raised concerns that the overwhelming focus on experiments could sometimes be detrimental and they would like more of a balance between knowledge-based learning and exploratory learning (INT-TR-05). The quotes below provide additional detail as to SSL and teacher views on this matter.

> *'Teachers now use the "plan assessment" documents to aid planning. They also use the portfolio documents to assess their own pupils are meeting standards. They plan for science enquiry in most lessons (where appropriate) as it is now at the heart of our science curriculum. Reach Out Reporter is used within lessons as well as Explorify' (INT-SSL-03).*

'My teaching has changed dramatically. I let the children take much more of a lead. They think about how they want to run it rather than me running it all the time. Allowing them to do that has changed me and how I teach. Children are more adventurous than I would have been' (INT-TR-04)).

'Since we have been doing "everyday materials", we have been doing too much hands-on… more of a split, a balance rather than testing and experimenting [would be better]. I worry about gaps—experiments for the sake of it … [There are] times when the approach is beneficial versus when it isn't' (INT-TR-06).

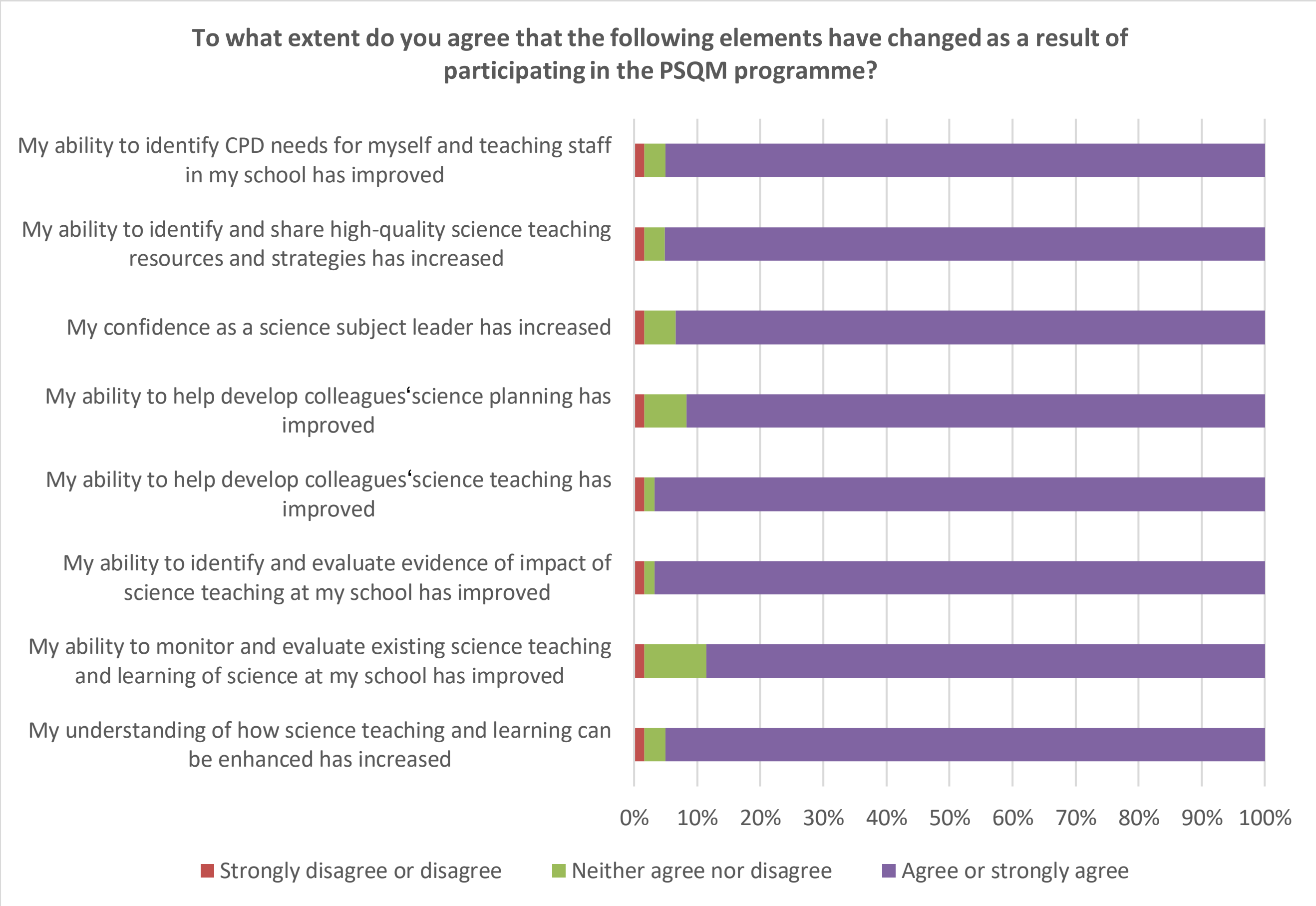
Many school staff also reported that those in science leadership had become more confident and credible

In interviews, headteachers noted changes in their SSL's ability to lead science since implementing PSQM. For example, one headteacher felt that their SSL had become more confident as a science leader and was taking a more systematic approach to science leadership (INT-HT-07). Another noted that their SSL was providing more support to the other teaching staff through CPD sessions as well as focusing more on the monitoring of science teaching (INT-HT-05). These observations were echoed by the SSLs, who all spoke about feeling more confident and credible as leaders and explained how their practice had developed, for example, through being more reflective (INT-SSL-05, INT-SSL-07), providing teacher support (INT-SSL-06, INT-SSL-05), and taking a 'bigger picture' approach to science teaching (INT-SSL-07).

Exemplifying these reported improvements in SSL leadership are several findings from the SSL survey, illustrated in Figure 7. For example, 95% or more of SSLs either agreed or strongly agreed with the statement relating to improvements in science teaching and learning.

Similarly, other statements in the surveys had high agreement levels, with 89% (54/61) of SSLs either agreeing or strongly agreeing that their ability to monitor and evaluate existing science teaching and learning of science at their school had improved specifically as a result of taking part in PSQM, and 93% (57/61) agreeing or strongly agreeing that their confidence as a science subject leader had increased. Taken as a whole, these findings demonstrate notable changes in school-wide practices in science teaching and management being perceived a result of participation, including improved SSL confidence—another key outcome outlined in the PSQM theory of change.

Figure 7: SSLs survey respondents' views on how participating in PSQM has changed some elements of their role as science subject leader



Based on responses to the question, 'Reflecting on your role as a science subject leader, to what extent do you agree that the following elements have changed as a result of participating in the PSQM programme?' Endline SSL survey, PSQM schools. SSLs total response on the different answers: 'My understanding of how science teaching and learning can be enhanced has increased', n = 60; 'My ability to monitor and evaluate existing science teaching and learning of science at my school has improved', n = 61; 'My ability to identify and evaluate evidence of impact of science teaching at my school has improved', n = 61; 'My ability to help develop colleagues' science teaching has improved', n = 61; 'My ability to help develop colleagues' science planning has improved', n = 60; 'My confidence as a science subject leader has increased', n = 61; 'My ability to identify and share high-quality science teaching resources and strategies has increased', n = 61; 'My ability to identify CPD needs for myself and teaching staff in my school has improved', n = 60.

Participants felt that the monitoring of science teaching and learning had improved

The theory of change outlines that, after participating in PSQM, teachers would reflect more on their science teaching practice and monitor their progress. In the SSL survey, 87% (54/62) of SSLs agreed or strongly agreed that the way in which science teaching and learning was monitored had improved as a result of participation in PSQM; similarly, in the teacher survey, 94% (63/67) of teachers claimed that they 'reflected more regularly on their science teaching practice and used this to inform future teaching and learning'. Likewise, SSL and Head Teacher interviewees reported improvements to the monitoring of science teaching and learning. For example, a teacher explained that their SSL created a checklist for all the topics that should be covered in science teaching throughout the year. The compiled checklist was shared among teachers and used to identify gaps in science teaching. This tool provided will be used to monitor science teaching over time (INT-TR-06).

In addition, the submission review found that the dimension 'a clear development cycle linked to an established monitoring programme' was evidenced by 29 of the 30 schools under analysis and over 60% of the 120 documents provided displayed evidence of monitoring in science (see Appendix G).

An increase in the number of planned science enrichment activities

Based on the review of PSQM submissions, 95 of the 120 documents reviewed across 30 schools (79%) provided evidence that 'school/classes/pupils participated in science enrichment activities' (see Appendix G). This was further supported by the average number of science-related activities reported in schools' science calendars. During the period of the PSQM programme delivery,[37] schools took part in an average of 17.4 science activities per school. This compares to one science activity per school being reported prior to PSQM, with the vast majority of submitted documents not recording any science events being hosted prior to participation in the programme. However, it should be noted that the number of science activities reported by schools varies from school to school and ranges from a minimum of zero (lowest value; only one out of 30 schools did not host any events during PSQM) to a maximum of 70.

The interview analysis supports the above finding that schools increased their provision of science activity. All the interviewed headteachers and SSLs confirmed their school was participating in events run by external providers (for example, the Big Science Event) as well as developing their own science enrichment activities aimed at developing science capital (for example, making links with the local community, including companies and individuals, and virtual visits). This is in line with the PSQM outcome identified in the theory of change that aims at teachers gaining an increased engagement with outside agencies and organisations to enrich science teaching and learning. However, it should also be noted that in the reviewed calendars there was explicit mention of at least 86 events being cancelled due to COVID-19 across schools—but not all schools provided an accurate number. Some simply mentioned 'many' or 'several' events being cancelled. The majority of these cancelled events are in the PSQM implementation phase (spring 2020), which coincided with the first COVID-19 lockdown.

A large proportion of staff in PSQM schools agreed that their school had developed links with other schools to share ideas and develop science teaching and learning

At the endline survey, 62% (32/48) of headteachers in PSQM schools agreed or strongly agreed that their school had developed links with other schools to share ideas and develop science teaching and learning, compared to control headteachers agreeing with this statement in the endline survey (44%, 26/59 of control headteachers). SSLs concurred with this PSQM-related change in the endline survey with 71% (43/61) of respondents agreeing that their school had developed better links with other schools. This was supported by the submission review, with 30 schools showing evidence for links to outside organisations and 26 providing evidence of linking to other schools (see Appendix G).

PSQM schools reported engaging in more CPD activities than before participation

The review of CPD logs as part of the documentary review showed that there were, on average, 38.2 CPD activities per school during the PSQM phase, a considerable increase from the 1.3 CPD activities reported prior to PSQM, on average.[38] While this does suggest a movement in line with the theory of change, this comparison should be interpreted with caution as not all schools provided equally detailed records of their activities prior to PSQM and one school did not record any CPD activities during the PSQM phase.

In interviews, school staff linked positive changes in the provision of CPD to PSQM. One headteacher (INT-HT-06) noted that their SSL was providing more support to the other teaching staff through CPD sessions as well as focusing more on monitoring. One headteacher and a science subject leader (INT-HT-06 and INT-SSL-06) mentioned that some additional CPD was accessed through local providers and involvement in groups, for example, science education centres, local authority networks, primary education conferences, and local school clusters. One SSL (INT-SSL-07) also reported accessing online provision such as Massive Open Online Course

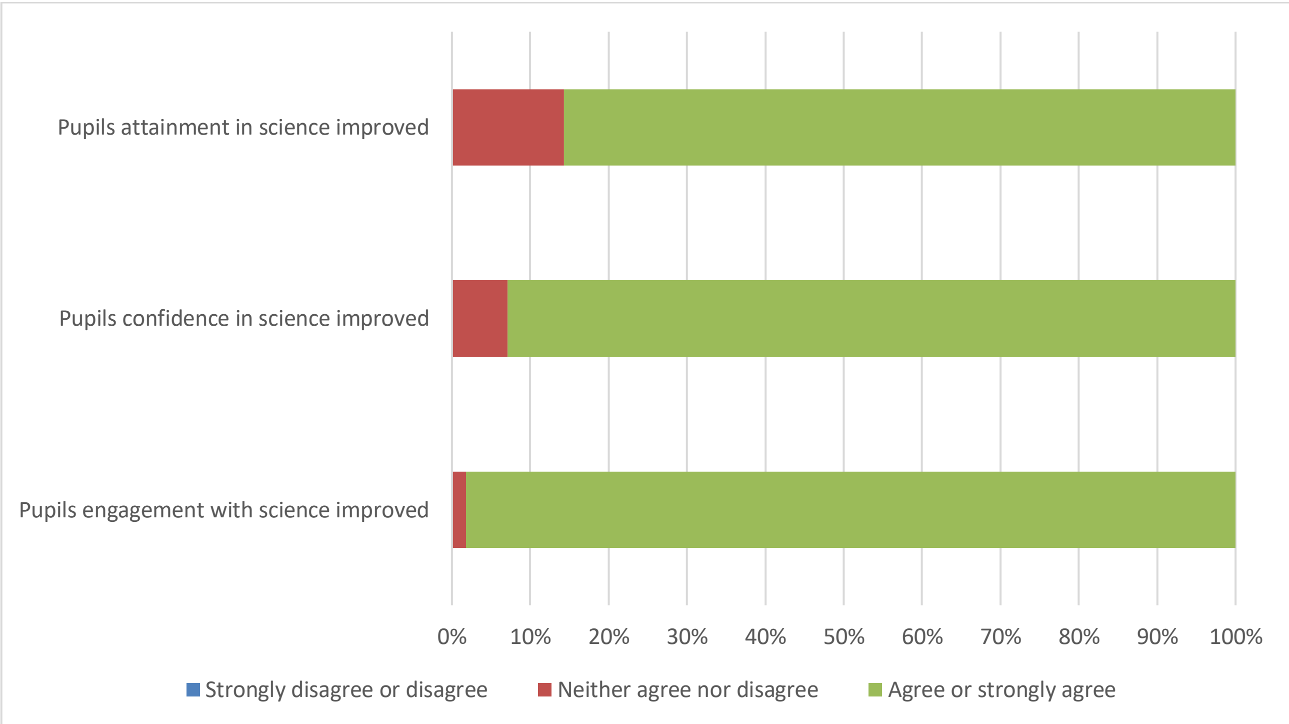
[37] The programme delivery took place between September 2019 and June 2020 (school year 2019/2020). Events that took place prior to September 2019 are indicated as 'pre-PSQM' and those that happened after June 2020 as 'post-PSQM'.
[38] Examples of CPD were taken from the SSL and CPD logs submitted as part of the PSQM submission. Examples provided by SSLs include activities such as LBQ Training for Science, SeeSaw training, staff meetings to cascade new resources or teaching to staff, and training in resources such as Explorify and Mentimeter.

(MOOCs) providers (for example, Future Learn), online science sites, and webinars. SSLs further explained that they passed on their learning to other staff members, for example, through staff meetings or by forwarding training information by email (INT-SSL-05). The registered, positive impact of the programme on CPD and learning opportunities for teachers is in line with the outcomes listed in the theory of change that aim at increasing teachers' understanding of science and their own practice and to promote changes in the types of pedagogies and practices that teachers use.

Teachers agreed that pupil outcomes improved after participation

Improved pupil outcomes in science is another important outcome outlined in the PSQM theory of change, including improved pupil engagement and confidence towards science learning. The combined analysis of surveys (both teacher and SSL), interviews, and the submission review indicates that that the majority of school staff had observed an improvement in pupil outcomes following their participation in PSQM: as can be seen in Figure 8, a large proportion of respondents in the teacher survey noted positive changes in pupils.

Figure 8: Teacher survey respondents' views on how pupil outcomes in their school changed as a result of participation in the PSQM programme



Based on responses to the question, 'Reflecting on pupil outcomes, to what extent do you agree that the following changed as a result of your school's participation in the PSQM programme?' Endline teacher survey, PSQM schools. Teachers' total responses across all the answers, n = 44; SSLs with dual roles, n = 12.

Both SSLs and teachers concurred in interviews that children's outcomes had improved since the implementation of PSQM . As a measure of the increased pupil confidence and engagement, one SSL noted that pupils had started to ask more questions in science lessons. It was noted by one teacher that the hands-on approach to science helps to engage less-academic students (INT-TR-05) and another suggested that the focus on exploration, and pupils realising that there is no 'right' answer to investigations, had increased their confidence (INT-TR-07). Furthermore, as noted by an SSL, pupils were able to envisage having careers in science (INT-SSL-07). While COVID-19 disrupted the ability of SSLs to comprehensively assess pupil progress, one reported that from a sample of pupils, lower ability pupil outcomes had increased by 40%, middle ability by 20%' and higher ability by 18% (INT-SSL-07). One teacher explained that PSQM made it easier for them to assess progress due to the detailed recording, which supported the finding previously discussed about improved monitoring of science (INT-TR-05). The following quote from a teacher provides further insight into how PSQM has heightened pupil engagement and confidence in science lessons.

> *'Reluctant [pupils] that hate science are much more engaged because, often, they are more practical and [now] get to do hands-on [activities] … As a whole, engagement has gone up … confidence improved for a lot of them because a lot of what they do is discussing with each other and the teacher and they are seeing it first-hand for themselves so it is easier for them to explain. Then teachers can extend [their knowledge] by asking questions—[so that they] can understand it more for themselves' (INT-TR-06).*

The submission review also provided evidence of numerous ways in which pupils were engaged in science. This included a number of structured activities focused on science enrichment on top of what a school would normally provide to improve pupils' science learning, including a planet-naming competition and visits to the Bristol planetarium and London Science Museum. One school also cited extra-curricular science clubs, which were added to their science portfolio thanks to PSQM. In addition, 73% of the 120 documents (and 29/30 schools) provided evidence on pupil engagement, confidence, and attainment in science through the use of a number of activities such as 'pupil voice meetings' for pupils to discuss their opinions of science with their teachers.

Schools provided less evidence on having dedicated resources for science, prominence to science in newsletters, or communication with stakeholders

The dimension with the lowest volume of evidence shows that only eight of 30 schools—and 12 of the 120 documents reviewed—provided evidence of the presence of a budget designated for science resources and CPD. The other dimensions with minimal evidence include giving science prominence in the school newsletter and reporting to governors about science activities.

Business as usual

Key findings

- There were some similarities between PSQM and control schools. For example, schools' reported engagement with CPD and opportunities to provide science enrichment activities were similar across control and PSQM schools
- Compared to staff in control schools, a higher proportion of PSQM school staff reported having a shared understanding of the importance and value of science and shared understanding of the principles of good science teaching.
- In addition, compared to respondents in control schools, a greater proportion of respondents from PSQM schools reported developing a range of effective strategies for teaching and learning science.
- Teachers in PSQM and control schools were relatively positive about students' confidence, progress, and engagement in science. Teachers in PSQM schools tended to be more positive than teachers in control schools, though differences were relatively small.

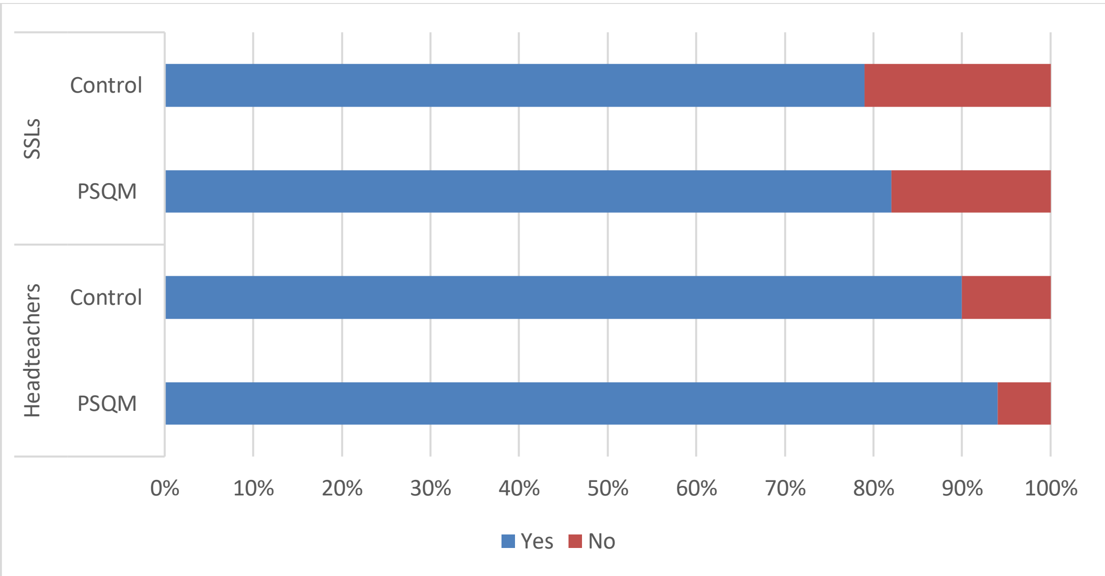
In this section the findings from interviews, surveys, and document reviews from both control and PSQM schools will be examined with the aim of understanding how science teaching and learning in PSQM schools compares with business as usual in control schools. This analysis will describe what business as usual looks like in control schools and how similar or different this is to science teaching and learning in PSQM schools.

Both PSQM and control school reported similar levels of engagement with CPD and opportunities to provide science enrichment activities

As can be seen in Figure 9, headteachers and SSLs in PSQM and control schools reported very similar levels of engagement with CPD and science enrichment activities. The document review supports this: all four control settings provided at least one document demonstrating that regular CPD was being led by SSLs. As for enrichment activities, five of the 16 documents reviewed across the four control schools provided evidence of regular outdoor science and the same number of documents provided evidence on there being a commitment to developing all children's science capital.

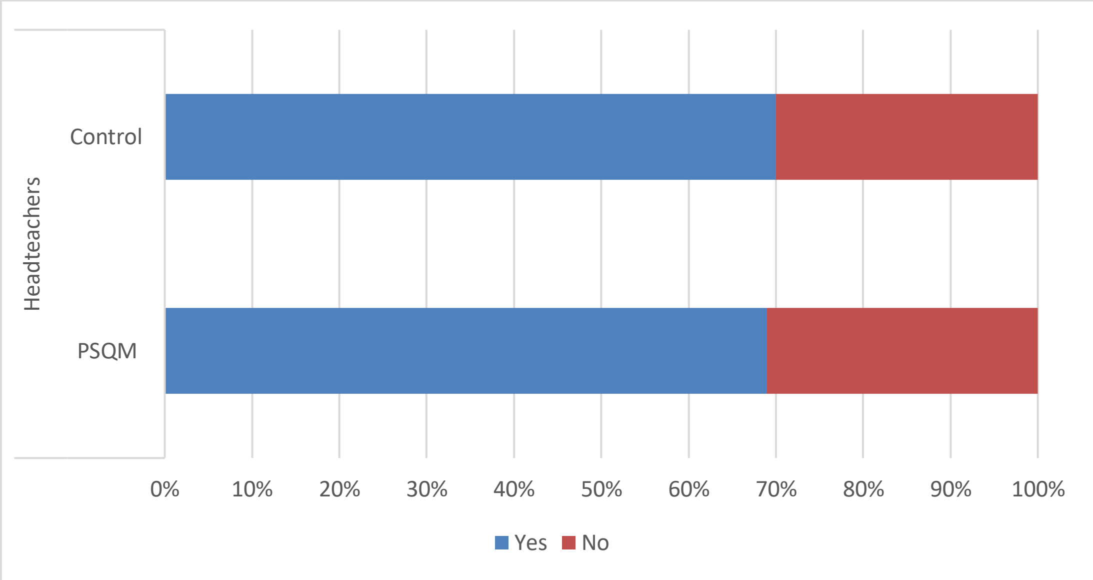
When headteachers were asked in the endline survey whether, in the previous two years, staff from their school had participated in science-related CPD, 94% (45/48) of the participants in PSQM schools reported 'yes' compared to 90% (55/61) in control schools. This reveals a similar attendance to science-related CPD in both groups. Survey responses from SSLs in PSQM schools were similar. When asked whether staff from their school had participated in science-related CPD (apart from PSQM training) in the previous two years, the responses from both groups were mostly affirmative: 82% (51/62) of SSLs in PSQM schools and 79% (45/57) in control schools reporting they had done so. Headteachers were also asked whether, in the previous two years, staff or children at their school had participated in any specific programmes or activities to improve pupils' science enquiry, attitudes, or experience (Figure 10). Similar proportions of PSQM and control headteachers agreed with this statement: 69% (31/45) and 70% (42/60), respectively, said 'yes'.

Figure 9: Headteacher and SSL survey respondents' responses on whether staff has participated in science related CPD in their school (in the past two years)



Based on responses to the question, 'In the past two years, have staff from your school participated in science-related continuous professional development (CPD) in addition to PSQM training?' (endline headteacher and SSL surveys, PSQM schools) and the question, 'In the past two years, have staff from your school participated in science-related continuous professional development (CPD)' in controls schools. Headteacher respondents in PSQM schools, n = 45; headteacher respondents in control schools, n = 61; SSL respondents in PSQM schools, n = 62; SSL respondents.in control schools, n = 57.

Figure 10: Headteacher survey respondents' responses on whether staff or children have participated in programmes to improve science enquiry in their school (in the past two years)



Based on responses to the question, 'In the past two years, have staff and/or children at your school participated in any specific programmes or activities to improve pupils' science enquiry, attitudes, or experience?' Endline headteacher survey, PSQM and control schools. Headteacher respondents in PSQM schools, n = 45; headteacher respondents in control schools, n = 602.

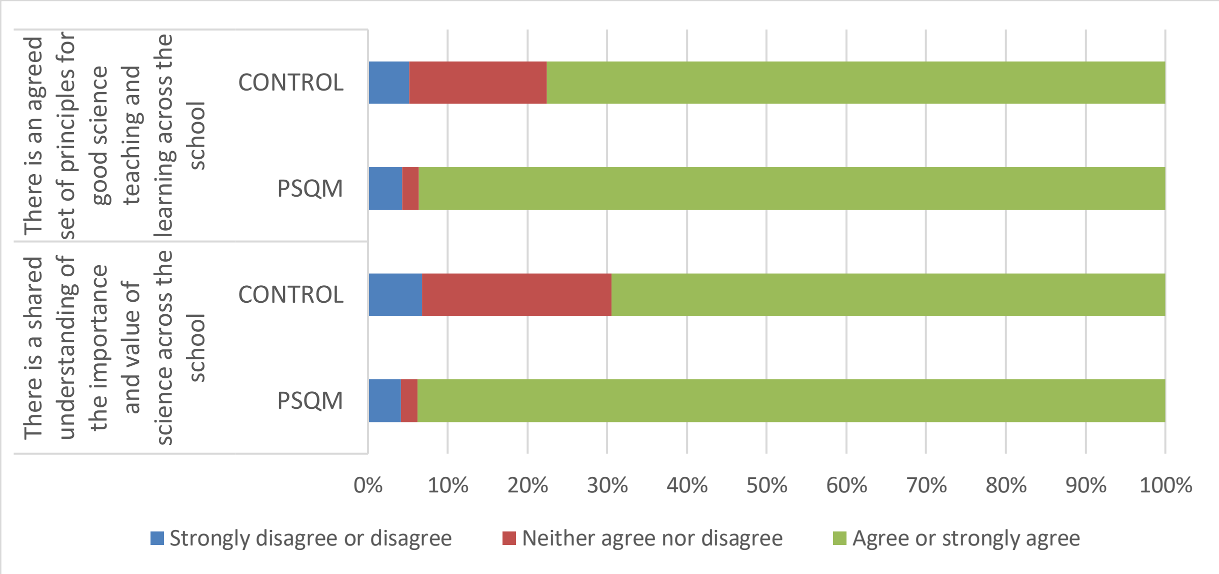
There were some small differences between CPD in PSQM and control schools. Based on headteacher responses on the survey, it appears that PSQM schools took part in more CPD compared to control schools with 43% (13/30) of respondents in PSQM schools reporting their staff took part in three or more programmes or activities compared to 31% (13/42) in control schools. While small, this difference could suggest that while all schools were taking part in CPD, PSQM schools were engaging in more CPD.

A smaller proportion of survey respondents from control schools reported having a shared understanding of the values and principles of science

As can be seen in Figure 11, headteachers in PSQM schools were more likely to agree or strongly agree to having a shared understanding of the importance and value of science and a shared understanding of the principles of good science teaching. A similar pattern is observed in the responses of SSLs where fewer SSLs from control schools agreed or strongly agreed with these statements and more remained neutral compared to SSLs from PSQM schools (Figure 12).

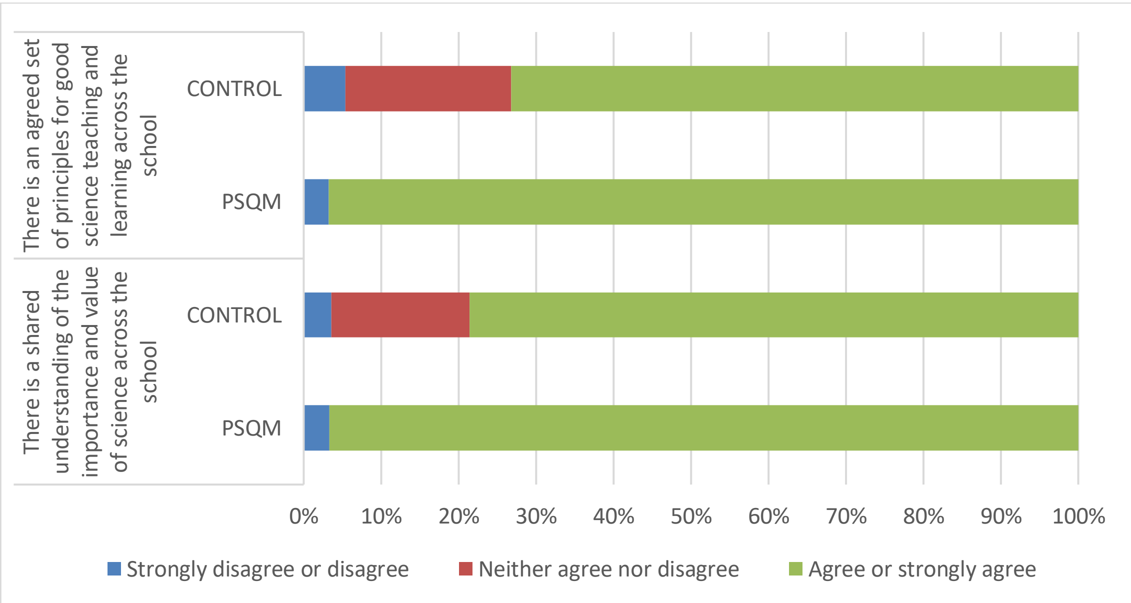
In interviews with staff from control schools, there was a lack of consistency on whether SSLs agreed that their schools have a shared vision for science, ranging from 'very little' through to 'strongly agree'. The control SSLs who felt there was a good or fair shared vision cited initiatives such as having a 'set of non-negotiables', having science boards on walls, having knowledge organisers for each year group, or the SLT being supportive (CNTL-SSL-08, CNTL-SSL-13, CNTL-SSL-11, CNTL-SSL-12).

Figure 11: Headteacher survey respondents' views on the shared values around science teaching in their school



Based on responses to the question, 'Reflecting on science in your school, to what extent do you agree with the following statements?' (endline headteacher survey, PSQM and control schools). On the statement, 'There is a shared understanding of the importance and value of science across the school': PSQM headteachers, n = 48; control headteachers, n = 59. On the statement, 'There is an agreed set of principles for good science teaching and learning across the school': PSQM headteachers, n = 47; control headteachers, n = 58.

Figure 12: SSL survey respondents' views on the shared values around science teaching in their school



Based on responses to the question, 'Reflecting on science in your school, to what extent do you agree with the following statements?' (endline SSL survey, PSQM and control schools). SSL respondents on the statement, 'There is a shared understanding of the importance and value of science across the school': PSQM, n = 61; control, n = 56. On the statement, 'There is an agreed set of principles for good science teaching and learning across the school': PSQM, n = 62; control, n = 56.

A smaller proportion of respondents from control schools reported developing a range of effective strategies for teaching and learning science

When asked to rate their agreement in relation to developing a range of effective strategies for science teaching that challenge and support the learning needs of all children, positive responses were slightly more prevalent in the PSQM schools compared to control schools: 93% (41/44) compared to 78% (38/49). Moreover, 94% (58/62) of SSLs in PSQM schools agreed or strongly agreed that 'teachers use a wide range of effective strategies for science teaching', compared to 88% (49/56). In the document review, half of the 16 documents reviewed across four control schools provided evidence on this dimension. Examples included the use of 'topic overviews' for science topics previously covered in class and the development of different learning journeys such as learning blocks.

When teachers in control schools were asked in interviews to describe the extent to which they felt able to deliver effective science teaching and learning, teachers with positive experiences described the following factors as facilitating this delivery:

- a whole-school approach (such as use of knowledge organisers) with clear plans in place;
- progression between year groups;
- collaboration across MATs;
- time for reflection;
- peer support; and
- access to resources.

The teachers with more negative responses spoke of the overriding focus on English and maths, a lack of time to put learning into practice, a lack of formative assessments, and lack of time to focus enough on science.

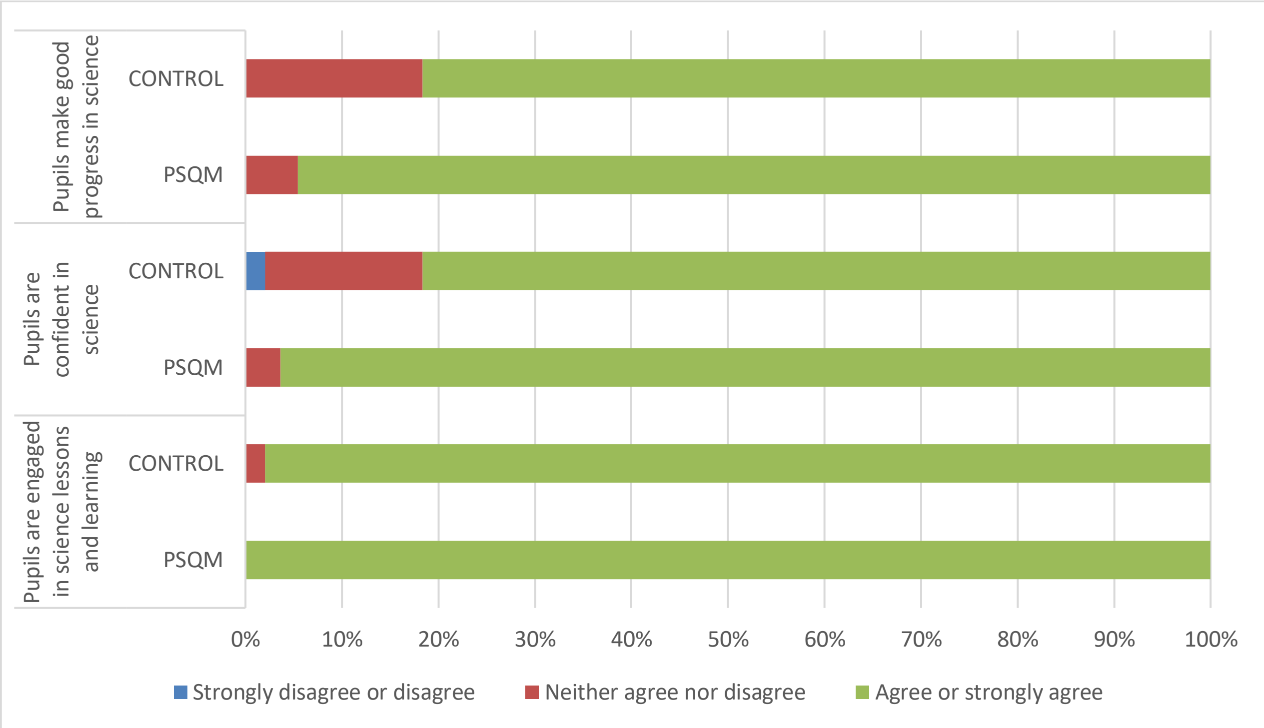
Control schools were less likely to report having regular science staff meetings or making links with other schools to develop science

In the endline headteacher survey, more PSQM headteachers (81%, 39/48) agreed or strongly agreed that 'science was a regular part of staff meetings at the school' compared to control headteachers (53%, 31/58). For the statement 'my school has developed links with other schools to share ideas and develop science teaching and learning', findings were only slightly more positive for the intervention group—PSQM: 50%, 24/48; control: 44%, 26/59. In the document review, none of the schools in the control group provided evidence for either dimension.

There were small differences between reports of pupil outcomes by teachers in PSQM and control schools, with PSQM teachers reflecting on this more positively

As can be seen in Figure 13, teachers in both PSQM and control schools were relatively positive about students' confidence, progress, and engagement in science. While positive responses were slightly more common in PSQM schools, differences between PSQM and control schools were relatively small—between two and 13 percentage points. It should be noted that disruptions caused by COVID-19 could have impacted on the ability of PSQM schools to deliver activities that would have enhanced these outcomes. However, as delivered in the evaluation these are the findings that were noted by respondents in the survey.

Figure 13: Teacher survey respondents' views on students' confidence, progress, and engagement in science



Based on responses to the question, 'Reflecting on pupil outcomes, to what extent do you agree with the following statements: "pupils are engaged in science lessons and learning", "pupils are confident in science", and "pupils make good progress in science"?' (endline teacher survey, PSQM and control schools). On all three statements: PSQM respondents, n = 55; control respondents, n = 49.

In interviews, all of the control teachers reported that their pupils were largely engaged with science lessons and learning. Indeed, control SSLs reported that science is seen as a popular subject and evidence was cited for this from activities SSLs did to collect pupil voice. However, a few SSLs also noted that, generally, pupil confidence in scientific skills is not as high, despite their strong engagement (CNTL-SSL-08, CNTL-SSL-10). A couple of control SSLs reported that this lack of confidence has been impacted by COVID-19 (CNTL-SSL-08, CNTL-SSL-12). While most teachers felt that the students were confident at doing science, noting that it is a subject that allows less academic pupils to succeed, one teacher noted confident pupils are in the minority (CNTL-TR-11).

Facilitators and barriers

Key findings

- Willing and enthusiastic staff and supportive SLT were considered fundamental to the facilitation of PSQM.
- The PSQM team were perceived to offer a wide range of support and resources, with hub leaders commended by many school staff for their helpfulness.
- Time pressures and collating evidence for the PSQM submission were identified as key barriers by SSLs.

In any process evaluation it is important to ascertain the facilitators and barriers to a particular intervention. In other words, the evaluator must investigate the necessary conditions for the success of the intervention and, conversely, the barriers it might come up against. Moreover, posing these questions enables us to consider the extent to which the contextual factors mentioned in the theory of change effected the implementation of PSQM.

To understand these dimensions of implementation, the evaluation team asked school staff and hub leaders—using interviews and surveys—about what factors enabled them to implement PSQM effectively and what, if any, hindered the implementation.

Facilitators

A higher proportion of PSQM schools identified as research schools compared to control schools and this difference increased at endline compared to baseline

In the endline survey, a quarter of headteachers in PSQM schools (12/48) reported that their school was a research school compared to 8% (5/59) of control schools. Furthermore, when headteachers were asked at endline whether their school had worked with a Research School on a science-related project since September 2019, 15% (7/46) of PSQM headteachers said 'yes' compared to only 5% (3/59) of control headteachers. While it is difficult to draw any firm conclusions, the fact that PSQM schools were more likely to be research schools or work with a research school suggests that they are actively seeking to use evidence to improve teaching and learning. We hypothesise that having research school status or working closely with research schools may act as a facilitator to implementing evidence-based programmes such as PSQM, as these schools may be more attuned to evidence-based practices and the importance of implementing with fidelity. However, without a more robust impact evaluation it is difficult to understand how the prevalence of research schools impacted on outcomes.

Willing and enthusiastic staff were considered fundamental to the facilitation of the PSQM programme

Willing and enthusiastic staff appears to be a key facilitator for the implementation of PSQM. Interviews and surveys both point to the importance of staff support at all levels. SLT 'buy-in', for example, was mentioned by many (INT-SSL-04, INT-TR-04, INT-TR-07, INT-SSL-06). This is a particularly important aspect of PSQM to measure as support by SLT and other staff members was a key contextual factor highlighted in the theory of change. In the PSQM SSL survey, the most frequent enablers mentioned in an open text response were SLT support, hub leader support, and the positive attitude of staff involved in the programme (mentioned in 35/48 or 73% of open-ended responses to this question). Similar responses were provided in an open text question about facilitators in the hub leader survey, where three of 13 respondents mentioned SLT support and the attendance of colleagues at the training. In interviews, half of the hub leaders (4/8) identified SSLs having SLT support as a key enabler.

Other related enablers mentioned by interviewees include flexibility within the school timetable and the provision of teaching cover to be able to attend the CPD and implement the programme (INT-SSL-06, INT-SSL-04). This again emphasises the need for broader support within the school for PSQM, in particular from the decision makers.

The support and willingness of all staff involved was mentioned as well. In interviews with school staff, all of the headteachers underlined the importance of having willing, enthusiastic SSLs and teaching staff who are committed and open to change. A few SSLs also highlighted the role of staff, the necessity of having a supportive and trusting SLT, (INT-SSL-06) and champions within the teaching staff (INT-SSL-07). A teacher noted that to implement PSQM effectively they needed to have strong relationships, good communication, and support within their staff team (INT-TR-07).

SSLs reported that the PSQM team offered a wide range of support and resources and commended hub leaders for their helpfulness

In the SSL survey for PSQM schools, when asked about other elements perceived to be helpful, staff responded that the PSQM festival, networking with schools, hub leader support, and PSQM webinars were also valuable. The Association for Science Education (ASE) conference was also noted by respondents—this was promoted by PSQM but not provided by the team. When asked about barriers, the least reported barrier in the provided list was 'insufficient support from the hub leader' (the response 'not at all' and 'very little' was selected by all respondents). Likewise in the hub leader survey, the flexibility of, and support provided by, the PSQM team

(8/13)—such as the proactiveness of adjusting timelines (1/13), providing quality resources for schools (2/13), and hub leaders providing extra documents as part of COVID-19 support to help schools identify evidence of impact (1/13)—were mentioned as key aspects of the support provided to school staff.

Interviews corroborate these findings. Some teachers and headteachers, for example, praised the large amount of support and resources available throughout the programme, such as the support offered during the online transition and helping schools to identify knowledge gaps among their pupils (INT-HT-06, INT-TR-01, INT-TR-07). One headteacher mentioned receiving support in the form of a bursary (INT-HT-06). The hub leader support model was considered good as it meant SSLs could ask questions and had a sounding board for ideas. The helpfulness of PSQM resources, such as the handbook, was also highlighted by interviewees. Indeed, the theory of change refers to the handbook as a source of continuous support provided to schools for effective implementation. The importance of the hub leader and other PSQM resources in facilitating the implementation also link back to a key activity identified in the theory of change, which is the continuous offer of support available to schools.

Most school staff are optimistic that PSQM-driven changes can be sustained, emphasising the importance of school support

An important contextual factor affecting PSQM implementation, as referenced in the theory of change, is the extent to which new practices in the school can be maintained. While positive about the changes being sustained, interviewees recognised that it would require effort and further development. For example, two headteachers stressed the importance of having an SSL who is willing to champion the programme and ensure it becomes embedded into the school system (INT-HT-06, INT-HT-07). The SSLs explained that further work was needed to embed the programme, for example, ensuring consistency in adoption of new teaching practices across their school, strengthening monitoring processes, and including science in their school action plans. A few teachers recognised they needed ongoing support, training, and learning to sustain the changes (INT-TR-06, INT-TR-07).

Similar to SSL interview reports, SLT support was identified as the most influential factor by hub leaders including continued support and training for the SSL with time allowed outside of usual teaching practices, support for early career science teachers, and opportunities for teachers to self-evaluate and monitor their science teaching. One interviewee mentioned the importance of having systems in place where teachers can identify next steps for their work; this was also identified as a measure to ensure sustainability, which might take the form of a post-PSQM action plan (INT-HL-06). Two hub leaders also pointed out the importance of school staff mentality for sustaining PSQM, specifically mentioning the importance of acquiring an understanding of the systemic, rather than individual, nature of PSQM change (INT-HL-03, INT-HL-04).

> *'It has to be systemic rather than just about one person. You want the quality of science to remain strong even if a teacher leaves. We encourage schools to think about the next steps and planning beyond PSQM.'*

A key element to ensure the sustainability of PSQM-related improvements is, therefore, to ensure that more than one individual within a school prioritises science improvement.

Barriers

Time pressures

In the hub leader survey, when asked about barriers to implementation, the most common response was limited capacity among SSLs to engage with training and keep to deadlines due to lack of release time. Out of those who responded, 69% (9/13) felt that limited capacity was 'somewhat' or 'to a great extent' a barrier to their ability to lead PSQM effectively. Over half of hub leaders (61.5%, 8/13) also responded 'somewhat' to the barrier of 'competing priorities (such as other ongoing programmes) in school'.

Likewise, in the SSL endline survey, 58% (36/62) of respondents reported that a ‘lack of time to drive PSQM activities’ was ‘somewhat’ or to ‘a great extent’ a barrier. SSLs also noted that ‘limited time among staff to engage with PSQM activities’ was a substantial barrier, with 43% (26/61) saying it had somewhat of an impact, and 10% (6/61) saying it was impactful to a great extent. Competing priorities (for example, other ongoing programmes) in school were also perceived as being somewhat of a barrier by 48% (30/62) of SSLs. The need to provide SSLs with a more manageable PSQM workload was a suggested improvement made by a few hub leaders in interviews, further highlighting the problems posed by time-related pressures on school staff (INT-HL-04, INT-HL-06). This is not surprising given the added complications of teaching during the pandemic. Impact of COVID-19 is further discussed in the section below.

Collating evidence for the PSQM submission was perceived as challenging by most SSLs, but not by hub leaders

In the SSL survey, the statement that ‘the process for collating evidence for submissions was difficult or onerous’ was considered somewhat of a barrier for over half (55%, 34/62) of SSLs. Interview findings help pinpoint what specifically the school staff found challenging about the submission. Significantly, all of the interviewed headteachers and one teacher (INT-TR-07) mentioned that the administrative work associated with the programme was burdensome and time consuming, with one SSL suggesting that it detracted from what they were trying to achieve through the programme (INT-SSL-06).

Interestingly, however, in the hub leader survey, 92.3% (12/13) responded that collating evidence for the PSQM submission required little effort on the part of SSLs (that is, ‘not at all’ or ‘very little’). In an open text response asking about other barriers, only one hub leader mentioned difficulties collating evidence during COVID-19 times. In relation to the PSQM submission, one SSL interviewee suggested it might be beneficial to have representatives from PSQM visiting schools to gather evidence themselves, rather than the onus being fully on schools to provide evidence within their portfolios. This, the SSL suggested, might lower the administrative burden (INT-SSL-05).

A minority of school staff and hub leaders pointed to areas for improvement relating to communication and best practice models

In interviews, a small number of school staff identified some ways in which communication could have been improved. A couple of teachers from the same school suggested the need for more clarity around the process and timeline for submission (INT-SSL-07, INT-TR-07). The quote below illustrates this need for greater clarity:

> *‘The main thing we had trouble with was we weren’t sure how it [PSQM] was going to end. If we were told how the last lot of [training] slides were going to impact on what we were doing at the beginning, we would have had a clearer understanding of what was expected. It was further into it [the programme] that we found out’ (INT-TR-04).*

Another minor problem concerned the ‘best practice model’ offered to schools. An SSL suggested that the model provided could be improved (INT-SSL-07) but did not elaborate on how. Similarly, one school would have liked *more* ‘best practice models’ to draw upon to help the implementation (INT-SSL-07, INT-TR-07). In addition, two hub leaders recommended improved signposting to CPD and other PSQM resources rather than directing school staff to the VLE on the basis that some school staff experienced difficulties navigating the vast amount of information on the platform (INT-HL-06, INT-HL-08; see Fidelity section: Ongoing Support by Hub Leaders to Schools).

Hub leader interviewees suggested a range of improvements to the PSQM programme

Shared suggestions for improving PSQM going forwards included a blended approach to training (mixing both face to face and online; INT-HL-01, INT-HL-03), providing schools with a more manageable workload (INT-HL-04, INT-HL-06), offering teachers more time to provide evidence for PSQM Gilt submission (INT-HL-02), and providing further support for teachers to critically evaluate, rather than just describe, what they have done throughout the implementation (INT-HL-08). A couple of hub leaders elaborate on these changes as follows:

'We need to change the mentality in schools of what effective CPD looks like' (INT-HL-08).

'I'd be interested to look at blended approaches—what might be successful to carry on digitally. For Round 19, most of the delivery was face to face. Round 20 has all been digital because of COVID-19. I think there are pros and cons. You can't replace face to face, but digital is good from a time perspective' (INT-HL-01).

Impact of COVID-19

Key Findings

- Most school staff perceived the impact of COVID-19 on science teaching as being negative, mostly impacting their ability to provide practical science learning opportunities.
- Headteachers held mixed views about changes in staff ability to implement the learning from the PSQM programme after COVID-19, but hub leaders were neutral about the pandemic negatively affecting the engagement of SSLs.
- New opportunities provided by online learning and an increased focus on pupils' wellbeing are mentioned among some of the positive implications of COVID-19.
- Some schools found the PSQM workload an additional burden in the context of COVID-19, and reported struggling to collect the evidence required.
- There was a perceived loss of 'community feeling' caused by PSQM activities being moved online due to COVID-19.

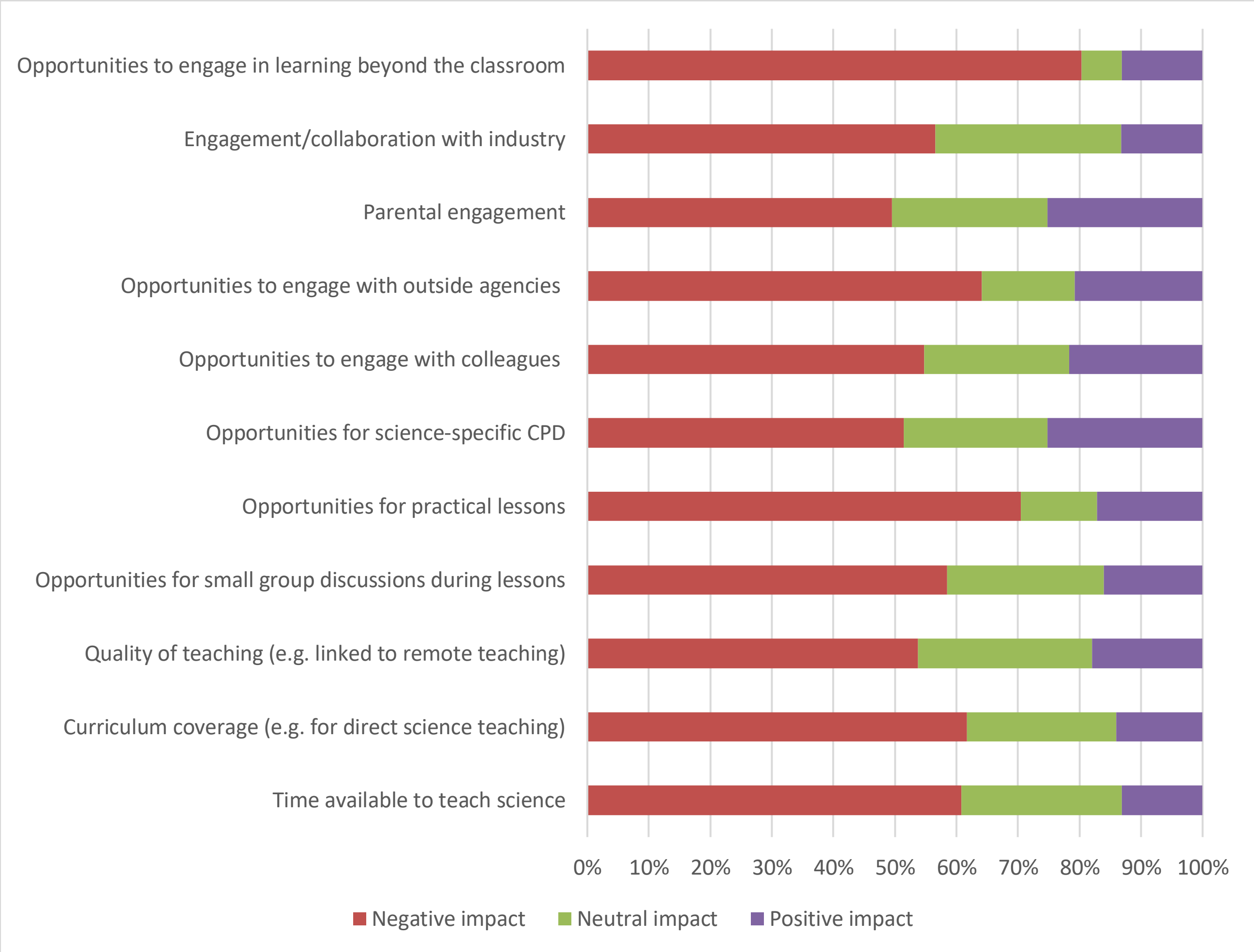
In the following section we aim to answer the key question, 'What was the impact of COVID-19 on science teaching?' Alongside this, the team was interested in looking at the extent to which schools' ability to engage in PSQM activities was affected by the COVID-19 pandemic. In this section we draw on findings from interviews and surveys with staff in PSQM and control schools as well as data from the PSQM submission review. To complete the analysis, we looked at the extent to which the findings presented in this section may have impacted on the theory of change and the contextual factors that would have been functional to the successful implementation of the intervention.

Impact of COVID-19 on science teaching and learning

Respondents felt that COVID-19 negatively impacted on a range of aspects of science teaching—especially practical science learning opportunities in particular.

As can be seen in Figure 14, the majority of headteachers across both PSQM and control arms reported that the overall impact of COVID-19 on science teaching was negative (extremely negative or somewhat negative) across a number of dimensions. In particular, opportunities to engage in learning beyond the classroom or with outside agencies were affected, as was the ability to engage in practical lessons.

Figure 14: Headteacher survey respondents' views on the impact of COVID-19 on elements of science leadership, teaching, learning, and enrichment in their school



Based on responses to the question, 'To what extent do you feel that the following elements of science leadership, teaching, learning, and enrichment at your school have been impacted (either positively or negatively) by school closures as a result of COVID-19?' (endline headteacher survey, PSQM and control schools). The number of responses to each statement were as follows: 'time available to teach science', n = 107; 'curriculum coverage (e.g. for direct science teaching)', n = 107; 'quality of teaching (e.g. linked to remote teaching)', n= 106; 'opportunities for small group discussions during lessons', n = 106; 'opportunities for practical lessons', n = 105; 'opportunities for science-specific CPD', n = 107; 'opportunities to engage with colleagues about science teaching and learning', n = 106; 'opportunities to engage with outside agencies to enrich science teaching and learning', n = 106; 'parental engagement', n = 107; 'engagement/-collaboration with industry', n = 106; 'opportunities to engage in learning beyond the classroom, e.g. extracurricular activities, science clubs, competitions, visits', n = 107.

Interviews provided some more context to the survey findings, in particular, in relation to how the pandemic impacted all schools' science teaching activities. One major theme that emerged is that the pandemic had a negative impact on the practical aspects of science learning, which got lost when online. For example, one hub leader (INT-HL-02) mentioned in an interview that schools experienced difficulties in conducting science practicals due to the need to sanitise and quarantine equipment in compliance with the newly introduced COVID-19 safety measures. Two interviewees went further, suggesting that curriculum coverage was less complete as a result of the pandemic, particularly in relation to developing practical skills and hands-on science (CNTL-TR-09, CNTL-HT-08). This was due not only to having to remotely deliver activities that would have normally been done in-person, but also to the lack of time to focus on these areas given to the need for schools to get pupils to 'catch up' on several topics. Furthermore, two interviewees highlighted that, despite attempting to keep a focus on investigation, there was an increase in knowledge-based learning and a decrease in scientific enquiry and independent learning (INT-HT-06, CNTL-SSL-09).

In addition to a reduced ability to provide practical science learning opportunities, interviewees from PSQM and control schools felt that COVID-19 negatively affected the quality of teaching by reducing opportunities for monitoring (of teaching and children's progress), as well as peer support and learning among the teaching staff.

Impact of home environment and resources

Two teachers (CNTL-TR-08 and CNTL-TR-10) and a science subject leader (CNTL-SSL-09) reported in interviews that during COVID-19 restrictions there was variation in the extent to which pupils were able to access the curriculum, depending on the level of parental support they received and the access to resources and materials they had at home. These disparities were not as evident when children were able to attend school in-person. Similarly, a headteacher said in an interview:

> *'It is difficult to do it all at home. When children are working at home we don't know how much parents are helping, so we can't do progress assessments … To do aspects around working scientifically at home we had to be mindful of what resources that could be accessed at home. We don't want to exclude any children. It is important to provide equal opportunities while developing working scientifically skills' (INT-HT-05).*

Positive aspects of school closures

Almost one fifth of all headteachers that were surveyed reported that school closures had a 'somewhat positive' or 'extremely positive' impact on some aspects of schools' activities (Figure 14). This was particularly notable in the cases of (i) opportunities for science-specific CPD (reported by 25%, 27/107), (ii) opportunities to engage with colleagues about science teaching and learning (22%, 23/106), (iii) opportunities to engage with outside agencies to enrich science teaching and learning (21%, 22/106), and (iv) parental engagement (25%, 27/107).

In interviews, a few headteachers and SSLs noted an increase in the use of online resources to support staff knowledge development and using online video conferencing such as MicrosoftTeams to communicate with each other (INT-HT-06, INT-SSL-01, INT-SSL-07, CNTL-SSL-08). One SSL went further suggesting that COVID-19 led to some positive changes for science teaching and learning, due to the increased focus on collaboration and hands-on learning in the recovery curriculum (CNTL-SSL-08). One interviewed teacher also commented that opportunities to develop knowledge via online learning had increased (CNTL-TR-08).

In addition, one hub leader (INT-HL-01) mentioned in an interview that COVID-19 led schools to prioritise pupil's wellbeing over science learning. This focus on wellbeing (and the recovery curriculum) was highlighted also by one SSL from the control group (CNTL-SSL-08).

Impact of COVID-19 on the implementation of PSQM

Headteacher views were mixed on the extent to which COVID-19 impacted staff's ability to implement the learning from the PSQM programme

While 35% (17/48) of headteacher survey respondents felt that COVID-19 had moderately or greatly decreased their school's ability to implement PSQM, 37% (18/48) thought their ability to deliver PSQM had increased either greatly or moderately. This points to the fact that there was no clear consensus on the impact of COVID-19 on the implementation of PSQM. Interestingly, differences at the extremes were more pronounced: 4% (2/48) said that COVID-19 had greatly decreased their ability to implement PSQM while around 15% (7/48) thought it had greatly increased their ability, suggesting that, overall, fewer respondents felt that PSQM had been significantly negatively impacted by COVID.

Teachers, instead, seemed to point more at the negative impact of COVID-19 in their survey responses. In fact, 39% (17/44) of respondents said that their ability to engage with PSQM had decreased after COVID-19 (compared with pre-COVID), while 50% (22/44) said that it had neither decreased nor increased. An equally small number of respondents (2/22) said that their ability to engage with PSQM greatly decreased, and greatly increased.

Some schools found the PSQM workload burdensome in the context of COVID-19

When asked to provide comments on their schools' involvement with PSQM in an open-ended survey question, some headteachers mentioned that they found it difficult to manage the PSQM-related workload under the unprecedented COVID-19 circumstances. For example, one respondent reported that teachers found it difficult to collate evidence during lockdown and another added that the number of action plans to complete were not very useful and burdensome. Another headteacher said that:

> *'The complication of COVID-19 and the extension to the scheme did put additional pressures on the school and staff, particularly with remote teaching and learning, bubble restrictions, remote meetings and general micro-management' (headteacher intervention survey, open-ended response).*

Similar comments were found in the hub leader survey in answer to an open-ended question about respondent's experience in supporting SSLs during COVID-19. One hub leader said that collating evidence proved more difficult for some schools due to COVID-19 restrictions but in their case, leaders adapted well to collecting different sources of evidence (for example, home-learning evidence). Also, one SSL reported in the equivalent open-ended question in the SSL survey that it was extremely difficult to generate the evidence required for the portfolio due to COVID-19 and school closures.

Hub leaders reported little differences in SSL engagement as a result of the pandemic, though there was a perceived loss of 'community feeling' due to moving activities online

Almost half of the hub leaders (46.2%, 6/13) reported in the survey that the engagement of SSLs neither increased nor decreased compared with before the pandemic. An equal number of respondents (23%, 3/13) said that SSLs' engagement moderately decreased, and moderately increased.

At the same time, however, a number of interviewees highlighted issues with moving the training online, including a decrease in the 'community feeling' of school hubs (INT-HL-05 and INT-HL-07), a loss of networking opportunities (INT-HT-06), and difficulties in building relationships (INT-HL-02 and INT-HL-03). However, one of the two interviewees (INT-HL-07) also added that the 'PSQM support' team was extremely responsive in providing support and additional resources during COVID-19.

The extremely positive role played by the PSQM team in supporting schools emerged also in interviews. Two SSLs mentioned the deadline extension to gather evidence, refresher training, additional spotlight sessions, and continued support post-PSQM submission as examples (INT-SSL-04 and INT-SSL-03). One hub leader (INT-HL-06) praised the PSQM team's responsiveness to school's needs more broadly.

Conclusion

Table 19: Key conclusions

Key conclusions
1. Due to COVID-19, the primary outcome for this evaluation was not collected and so no measure of impact on science attainment is reported and there is no security rating for this trial. Key conclusions are based on qualitative data from the implementation and process evaluation.
2. Overall, headteachers and SSLs attributed several changes that were in line with PSQM's theory of change. For example, they reported that taking part in PSQM had raised the profile of science in their schools, with more time being dedicated to science. They also reported that those in science leadership in their school had become more confident and credible.
3. The training was well received by SSLs. There was over 90% attendance at all five sessions and SSLs reported finding all sessions highly relevant. Hub leaders reported that SSL engagement was high throughout the programme.
4. Headteachers and SSLs in PSQM schools were more likely to report a shared understanding of the importance and value of science across the school compared to other schools. However, headteachers and SSLs in PSQM schools reported similar levels of engagement with continuing professional development (CPD) and with science enrichment activities compared to other schools.
5. Interviews and surveys suggest that willing and enthusiastic staff are a key facilitator for the implementation of the PSQM process. Time pressures on staff were identified as a key barrier. Collating evidence for submissions and the administrative work associated with the programme were considered time-consuming by headteachers and SSLs.

PSQM is aimed at improving school-wide science teaching and raising the profile of science in primary schools. This efficacy trial was planned as a two-arm, cluster-randomised controlled trial (cRCT) and sought to help determine whether PSQM leads to observable improved outcomes in children's science outcomes. However, as a result of the COVID-19-related lockdowns and social distancing measures introduced in England from March to June 2020 and from January to March 2021, activities for the evaluation of PSQM were revised. Consequently, all planned impact evaluation activities were foregone and the revised evaluation draws on IPE data only. While past research followed a design comparable to this rescoped trial, the current evaluation is substantially more extensive, covering over 70 PSQM schools—compared to eight schools in research by White et al. (2015, 2016)—in addition to having a comparison control group with staff from over 70 schools. Hence, this evaluation builds on past research and provides insights based on a notably larger pool of experiences. Overall, IPE data suggests that PSQM was well received and provides evidence that respondents implementing PSQM perceive that it can help improve teacher and pupil outcomes in line with what is outlined in the theory of change.

PSQM training and implementation

Training was well attended by SSLs with the majority of respondents rating it as 'highly relevant' and 'extremely useful'. In contrast, teacher training was less well attended after the first session, but this is probably linked to the fact that training was no longer compulsory for teachers after March 2021.

SSLs were also quite engaged with all compulsory sessions (Sessions 1, 2, 3a, and 4). Session 3b had much lower engagement but it should be noted that this session was optional given it was a repeat of training previously provided to support schools once they reopened after COVID-19 closures. Furthermore, interviews with hub leaders suggest that most SSLs were very engaged in many aspects of the PSQM programme, evidenced by use of the chat function during online training and the trainings attendance rate. Hub leaders reported that SSL engagement was hindered by increased workload levels caused by the COVID-19 disruption, which led to some schools being worse off than others. This suggests that SSL engagement was high overall, with some variation likely linked to pressures caused by COVID-19.

Looking across to other metrics of engagement with the programme, the majority of hub leaders reported that SSLs were ‘greatly engaged’ with the VLE and all schools submitted tasks for the PSQM award. This supports the conclusion that schools were engaged with the programme and keen to comply with requirements.

Changes driven by PSQM

Overall, school staff attributed a number of changes to PSQM that were in line with PSQM’s theory of change, including increases in the confidence and credibility of SSLs, an improved vision for science teaching and learning, and a general agreement that science teaching had improved as a result of participation. Furthermore, the majority of school staff noted an improvement in pupil outcomes following their participation in PSQM.

Interestingly, while control schools reported similar numbers of CPD and science enrichment activities, PSQM schools perceived better outcomes for science teaching and learning. Compared to control schools, a higher proportion of PSQM school staff reported having a shared understanding of the values and principles of science, were more likely to say that they had developed a range of effective strategies for teaching and learning science, and were more likely to report positive outcomes for pupils. This suggests that while control schools were engaged in a similar number of science activities, PSQM was more successful at improving outcomes at the teacher and pupil level. However, it should be noted that the number of science activities reported by schools varies from school to school.

Barriers and facilitators for PSQM implementation

Despite COVID-19 being identified as a barrier for delivering PSQM, one key facilitator that appeared to aid implementation was the support offered by the PSQM team. The flexibility and support of the PSQM team was noted by a majority of SSLs and hub leaders, including the team’s proactiveness in adjusting timelines and the provision of a document to support schools to identify evidence of impact during COVID-19. In short, the PSQM team was perceived to offer a wide range of support and resources. Willing and enthusiastic staff was also identified as a key implementation facilitator: interviews and surveys both point to the importance of staff support at all levels. Participants highlighted the importance of having a shared understanding of the value of science in the school and noted that having this rest on only one individual could potentially hinder continuity for improvement.

While mixed, some feedback suggests that providing a more manageable workload in relation to the PSQM programme would facilitate its implementation. However, reported experiences that collecting documentation for PSQM submissions was an onerous task may have been exacerbated by the additional pressures COVID-19 disruptions brought about.

Overall, most school staff perceived the impact of COVID-19 on science teaching as being negative, mostly impacting their ability to provide practical science learning opportunities. However, some staff noted benefits resulting from the pandemic such as new opportunities provided by online learning and an increased focus on pupils’ wellbeing.

Limitations and lessons learned

While this evaluation provides support to suggest that implementation was, by-and-large, successful and was well received by school staff, evidence on impact is limited by the rescoped design. Without a counterfactual impact estimate it is difficult to conclude whether reported changes in practice led to the hypothesised impact on student attitudes and attainment. While there are initial indications that Year 5 teachers and SSLs changed their behaviour in line with PSQM’s theory of change, because the majority of the information comes from Year 5 teachers and SSLs involved in the delivery of PSQM there is a limit to how independent and objective they may be.

Finally, COVID-19 created unique circumstances that may have influenced how the programme was implemented and perceived. While the evaluation team attempted to document this, it is important to acknowledge that this was a unique period and may not be an accurate representation of the programme or people's views.

Future research and publications

Given the relatively successful implementation of PSQM during the evaluation and the fact that PSQM was well received, the team suggest that there is merit in conducting a robust trial of PSQM. While it must be acknowledge that some elements of PSQM were adapted as a result of COVID-19, the programme was found to be acceptable to participants.

Given the success of engagement with PSQM, despite the challenges of COVID, it would be interesting to see if any elements of implementation, such as online meetings or ad-hoc support, may be beneficial for future roll out.

References

Allison, P. D. (2012) 'Handling Missing Data by Maximum Likelihood', Haverford, PA: Statistical Horizons: https://statisticalhorizons.com/wp-content/uploads/MissingDataByML.pdf

Blouin, D., Teikan, A., Kamin, C. and Harris, I. B. (2018) 'The Impact of Accreditation on Medical Schools' Processes', *Medical Education*, 52 (2), pp. 182–191. DOI: 10.1111/medu.13461

Blumenfeld, P., Modell, J., Bartko, W. T., Secada, W., Fredricks, J. and Friedel, J. et al. (2005) 'School Engagement of Inner City Students During Middle Childhood', in C. R. Cooper, C. Garcia Coll, W. T. Bartko, H. M. Davis and C. Chatman (eds), *Developmental Pathways Through Middle Childhood: Rethinking Diversity and Contexts As Resources,* Mahwah, NJ: Lawrence Erlbaum (pp. 145–170).

Boyle, A., Taylor, A., Giacomantonio, C. and Sutherland, A. (2015) 'Using Ambulance Data to Reduce Community Violence: Critical Literature Review', *European Journal of Emergency Medicine*, 23 (4), pp. 248–252.

CFE (2019) 'Understanding the "State of the Nation" Report of UK Primary Science Education: A Baseline Report for the Wellcome Trust', CFE Research: https://wellcome.ac.uk/sites/default/files/understanding-state-of-the-nation-report-of-uk-primary-science-education.pdf

Cundill, B. and Alexander, N. D. (2015) 'Sample Size Calculations for Skewed Distributions', *BMC Medical Research Methodology*, 15 (1), p. 28. DOI: 10.1186/s12874-015-0023-0

Darling-Hammond, L., Hyler, M. E. and Gardner, M. (2017) 'Effective Teacher Professional Development', Palo Alto, CA: Learning Policy Institute: https://learningpolicyinstitute.org/product/teacher-prof/dev

Dong, N. and Maynard, R. (2013) 'PowerUp!: A Tool for Calculating Minimum Detectable Effect Sizes and Minimum Required Sample Sizes for Experimental and Quasi-Experimental Design Studies', *Research on Educational Effectiveness*, 6 (1), pp. 24–67. DOI: 10.1080/19345747.2012.673143

Dong, N., Kelcey, B., Spybrook, J. and Maynard, R. A. (2017) 'PowerUp!-Moderator: A Tool for Calculating Statistical Power and Minimum Detectable Effect Size of the Moderator Effects in Cluster Randomized Trials' (Version 1.08) [Software]: http://www.causalevaluation.org/

EEF (2015) 'EEF Guidance on Cost Evaluations', London: Education Endowment Foundation: https://v1.educationendowmentfoundation.org.uk/uploads/pdf/EEF_guidance_to_evaluators_on_cost_evaluation.pdf

EEF (2018) 'Statistical Guidance for EEF Evaluations', London: Education Endowment Foundation: https://educationendowmentfoundation.org.uk/public/files/Evaluation/Writing_a_Protocol_or_SAP/EEF_statistical_analysis_guidance_2018.pdf

Fredricks, J. A. and McColskey, W. (2012) 'The Measurement of Student Engagement: A Comparative Analysis of Various Methods and Student Self-Report Instruments', in *Handbook of Research on Student Engagement*, Boston, MA: Springer (pp. 763–782).

Garet, M. S., Heppen, J. B., Walters, K., Parkinson, J., Smith, T. M., Song, M., Garrett, R., Yang, R., Borman, G. D. and Wel, T. E. (2016) 'Focusing on Mathematical Knowledge: The Impact of Content-Intensive Teacher Professional Development' (NCEE 2016-4010), Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Garet, M., Wayne, A., Stancavage, F., Taylor, J., Eaton, M., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., Doolittle, F. and Warner, E. (2011) 'Middle School Mathematics Professional Development Impact Study: Findings After the Second Year of Implementation' (NCEE 2011-4024), Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gore, J., Lloyd, A., Smith, M., Bowe, J., Ellis, H. and Lubans, D. (2017) 'Effects of Professional Development on the Quality of Teaching: Results from a Randomised Controlled Trial of Quality Teaching Rounds', *Teaching and Teacher Education*, 68, pp. 99–113. DOI: 10.1016/j.tate.2017.08.007

Hanbury, N., Prosser, M. and Rickinson, M. (2008) 'The Differential Impact of UK Accredited Teaching Development Programmes on Academics' Approaches to Teaching', *Studies in Higher Education*, 33 (4), pp. 469–483. DOI: 10.1080/03075070802211844

Hanley, P., Slavin, R. and Elliott, L. (2015) 'Thinking, Doing, Talking Science: Evaluation Report and Executive Summary', London: Education Endowment Foundation.

Hedges, L. V. (2007) 'Effect Sizes in Cluster-Randomized Designs', *Educational and Behavioral Statistics*, 32 (4), pp. 341–370.

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L. and Sterne, J. A. C. (2011) 'The Cochrane Collaboration's Tool for Assessing Risk of Bias in Randomised Trials', *British Medical Journal*, 343, d5928.

Hoffmann, T., Glasziou, P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D., Barbour, V., Macdonald, H., Johnston, M., Lamb, S., Dixon-Woods, M., McCulloch, P., Wyatt, J., Chan, A. and Michie, S. (2014) 'Better Reporting of Interventions: Template for Intervention Description and Replication (TIDieR) Checklist and Guide', *BMJ.*, 348, g1687.

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R. and Kerr, K., (2016) 'Implementation and Process Evaluation (IPE) for Interventions in Education Settings: An Introductory Handbook', London: Education Endowment Foundation.

Kind, P., Jones, K. and Barmby, P. (2007) 'Developing Attitudes Towards Science Measures', *International Journal of Science Education*, 29 (7), pp. 871–893.

Kitmitto, S, González, R., Mezzanote, J. and Che, Y. (2018) 'Thinking, Doing, Talking Science: Evaluation Report and Executive Summary', London: Education Endowment Foundation: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/TDTS.pdf

Oakes, J. M. (2013) 'Effect Identification in Comparative Effectiveness Research', *Electronic Health Data and Methods*, 1 (1), 1004. DOI: 10.13063/2327-9214.1004

Ofsted (2013) 'Maintaining Curiosity: A Survey Into Science Education in Schools', Ofsted: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/379164/Maintaining_20curiosity_20a_20survey_20into_20science_20education_20in_20schools.pdf

Ofsted (2021) 'Research Review Series: Science': https://www.gov.uk/government/publications/research-review-series-science/research-review-series-science#fn:15

Ofsted (2019) 'Intention and Substance: Further Findings on Primary School Science from Phase 3 of Ofsted's Curriculum Research': https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/777992/Intention_and_substance_findings_paper_on_primary_school_science_110219.pdf

Opfer, D. (2016) '*Conditions and Practices Associated with Teacher Professional Development and Its Impact on Instruction in TALIS 2013'*, OECD Education Working Paper, No. 138, OECD Publishing.

Pintrich, P. R., Smith, D. A. F., Garcia, T. and McKeachie, W. (1993) 'Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (MSLQ)', *Educational and Psychological Measurement*, 53 (3), pp. 801–813.

Senn, S. (2004) 'Controversies concerning randomization and additivity in clinical trials', *Statistics in Medicine,* 23 (24), pp. 3729 – 3753.

Singh, K., Granville, M. and Dika, S. (2002) 'Mathematics and Science Achievement: Effects of Motivation, Interest, and Academic Engagement', *Educational Research*, 95 (6), pp. 323–332.

Sims, S. and Fletcher-Wood, H. (2018) 'Characteristics of Effective Teacher Professional Development: What We Know, What We Don't, How We Can Find Out', UCL Institute of Education: https://improvingteaching.co.uk/wp-content/uploads/2018/09/Characteristics-of-Effective-Teacher-Professional-Development.pdf

Skinner, E. A., Marchand, G., Furrer, C. and Kindermann, T. (2008) 'Engagement and Disaffection in the Classroom: Part of a Larger Motivational Dynamic', *Journal of Educational Psychology*, 100 (4), pp. 765–781. DOI:10.1037/a0012840

Snijders, T. A. B. and Bosker, R. J. (2012) *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2nd edn), London: Sage.

Speckesser, S., Runge, J., Foliano, F., Bursnall, M., Hudson-Sharp, N., Rolfe, H. and Anders, J. (2018) 'Embedding Formative Assessment: Evaluation Report and Executive Summary', London: Education Endowment Foundation: https://educationendowmentfoundation.org.uk/public/files/EFA_evaluation_report.pdf

Spybrook, J., Kelcey, B. and Dong, N. (2016) 'Power for Detecting Treatment by Moderator Effects in Two- and Three-Level Cluster Randomized Trials', *Educational and Behavioral Statistics*, 41 (6), pp. 605–627. DOI: 10.3102/1076998616655442

Sutherland, A., Strang, L., Stepanek, M., Giocmantonio, C. and Boyle, A. (2017) 'Using Ambulance Data for Violence Prevention: Technical Report', Santa Monica, CA: RAND Corporation.

Volkwein, J. F., Rattuca, L. R., Harper, B. J. and Domingo, R. J. (2006), 'Measuring the Impact of Professional Accreditation on Student Experiences and Learning Outcomes', *Research in Higher Education*, 48 (2), pp. 251–282. DOI: 10.1007/s11162-006-9039-y

White, E., Dickerson, C. and Mackintosh, J. (2015) 'Impact of Royal Society of Chemistry Bursary-Funded Primary Science Quality Mark on Primary Science Teaching: Final Report', University of Hertfordshire: https://uhra.herts.ac.uk/bitstream/handle/2299/16537/Impact_of_RSC_bursary_funded_PSQM_on_primary_science_teaching_final_110915.pdf?sequence=2&isAllowed=y.

White, E., Dickerson, C., Mackintosh, J. and Levy, R. (2016) 'Evaluation of the Primary Science Quality Mark Programme 2013-15', DOI: 10.13140/RG.2.1.1312.9209 https://www.researchgate.net/publication/305033454_Evaluation_of_the_Primary_Science_Quality_Mark_programme_2013-15

Wellcome Trust (2014) 'Primary Science: Is It Missing Out? Recommendations For Reviving Primary Science': https://wellcome.ac.uk/sites/default/files/primary-science-is-it-missing-out-wellcome-sep14.pdf

Appendix A: EEF cost rating

Table 20: Cost Rating

Cost rating	Description
£ £ £ £ £	*Very low:* less than £80 per pupil per year.
£ £ £ £ £	*Low:* up to about £200 per pupil per year.
£ £ £ £ £	*Moderate:* up to about £700 per pupil per year.
£ £ £ £ £	*High:* up to £1,200 per pupil per year.
£ £ £ £ £	*Very high:* over £1,200 per pupil per year.

Appendix B: Security classification of trial findings

This project and its evaluation were affected by the 2020 and 2021 partial school closures and social distancing regulations caused by the coronavirus (COVID-19) pandemic: the trial and PSQM submission date was extended and intervention schools were invited to receive additional delivery intended to partly compensate for lost time. Implementation and process evaluation data collection occurred throughout the evaluation until July 2021. Ultimately, the collection of pupil outcome data was not possible. The impact evaluation element was foregone and, therefore, we cannot estimate the impact of the project on science subject attainment.

Appendix C: Memorandum of Understanding

Evaluation of Primary Science Quality Mark

Memorandum of Understanding between trial schools, the University of Hertfordshire and RAND Europe

The University of Hertfordshire
Department of Education

Hatfield

Herts AL10 9AB

01707 281036

Associate Professor Jane Turner

PSQM Delivery Team lead

Associate Professor

j.turner@herts.ac.uk

RAND Europe

Westbrook Centre

Milton Road

Cambridge, CB4 1YG

01223 353329

Dr Alex Sutherland

RAND Evaluation Team lead

Senior Research Leader

alex_sutherland@rand.org

Memorandum of Understanding

Agreement to participate in the Evaluation of the Primary Science Quality Mark

Please complete the information in all of the boxes below, then sign and return this page to:

Post to:
Claire Harman: PSQM
Room R010A – School of Education
The University of Hertfordshire
Hatfield
Herts AL10 9AB

Scan and email to:

Claire Harman
C.harman2@herts.ac.uk

BOX 1: SCHOOL AND HEADTEACHER CONTACT DETAILS

School Name:

……

School Address:

……

Postcode:

BOX 2: INFORMATION ABOUT SCIENCE SUBJECT LEADER, YEAR 5 CLASSES AND NOMINATED TEACHER

Science subject leader ……………………………………………………………………………………………………

Email address of science subject leader

………

TRIAL BACKGROUND, AIMS AND DESCRIPTION

The information is for schools wishing to take part in the PSQM trial, a research and evaluation project looking at the impact of the Primary Science Quality Mark (PSQM) [39] in the 2018/19, 2019/20 and 2020/21 academic years. PSQM is an accreditation programme aimed at improving school-wide science teaching and raising the profile of science in primary schools through: (i) effective science leadership and (ii) supported school self-evaluation. Existing research suggests that PSQM can benefit schools in multiple ways, such as contributing to raising the profile of science in primary schools and providing schools with a framework and professional support for developing science leadership, teaching and learning.[40]

BOX 3: SCHOOL COMMITMENT

- We commit the school to the implementation and evaluation of the Evaluation of the Primary Science Quality Mark (PSQM) as described below.

- In doing so we understand and accept that the current PSQM evaluation is a randomised controlled trial, which means our school may not receive PSQM, but we will still engage in research activities.

- We understand and accept that the school is not officially part of the trial until the requested pupil data has been sent to the Delivery Team.

Headteacher name

……

Quality_Mark_programme_2013-15

The aim of this trial is to evaluate the effects of the PSQM on pupil outcomes in science. The results of this research will make an important contribution to understanding what works in primary science teaching and learning. By participating in the trial, schools contribute to furthering our knowledge on strategies for whole-school improvement in science teaching and learning.

The PSQM programme will be led by The University of Hertfordshire (the Delivery Team) and will be independently evaluated by RAND Europe (the Evaluation Team) [41]. RAND Europe is an independent not-for-profit research institute whose mission is to help improve policy and decision making through research and analysis. Pupil assessment will be conducted by the National Foundation for Educational Research (NFER), a centre for educational research and development in the UK (the Assessment Implementers), sub-contracted by RAND Europe. The PSQM team is based at the University of Hertfordshire and is supported by the Primary Science Teaching Trust[42]. The trial is funded by the Education Endowment Foundation (EEF)[43] and the Wellcome Trust[44]. The EEF is an independent charity that funds research to test various ways of raising attainment in English schools, and Wellcome is an independent charity supporting science education in the UK.

The evaluation of PSQM is a 'randomised controlled trial' (RCT). This means that half the schools, chosen at random within each PSQM hub, will take part in the intervention (the PSQM programme). These are the Intervention Schools. The other half will not receive the intervention. These are the Control Schools. Once the PSQM programme is completed, outcomes from children in the Intervention Schools will be compared to those in the Control Schools to find out whether the intervention has made a measurable difference. As such, both Intervention and Control Schools are vital to the evaluation. The random allocation of schools by RAND Europe is integral to the evaluation design, as this is the best and most rigorous way to measure the effects of the PSQM programme on children's science outcomes. It is therefore important that schools understand, and consent to, the random allocation process. This means that each school may or may not be allocated to the PSQM programme, as determined by the randomisation process.

As whole-school changes take time, we will evaluate the effect of the intervention on pupil science attainment and attitudes following the 2019/20 implementation year with the two successive Year 5 cohorts in Intervention and Control Schools. Cohort A will include children in year 5 during the academic year 2019/20 and cohort B will include children in year 5 during the academic year 2020/21.

Overview of the two pupil groups involved in the trial

	Academic year 2018-19	Academic year 2019-20	Academic year 2020-21
		School works towards a PSQM in the Intervention Schools	
Pupil Cohort A	Year 4	Year 5 (tested at the end of the year)	
Pupil Cohort B	Year 3	Year 4	Year 5 (tested at the end of the year)

We will also survey teachers, science subject leaders and headteachers in all trial schools. In addition, four schools among those assigned to the Intervention (PSQM), representing a diverse set of characteristics, will be approached for in-depth case studies. These will involve interviews with school teaching staff and headteachers, as well as parents, and – where possible - school governors, as well as visits to PSQM training sessions attended by the year 5 teacher and science subject leader from those schools. Interviews and visits to the training sessions

[41] https://www.rand.org/randeurope/research/projects/primary-science-quality-mark.html
[42] https://pstt.org.uk/
[43] https://educationendowmentfoundation.org.uk/
[44] https://wellcome.ac.uk/

allow the Evaluation Team to gain a more in-depth understanding of what PSQM involves in practice for participating teachers, subject leaders, and schools. This information is important because it helps us to understand what other schools need to do if they chose to participate in PSQM later.

Intervention Schools will not be charged to take part in the PSQM programme and will receive a payment of £1,500 towards teaching cover and £120 towards travel costs. Control Schools will receive a payment of £1,500 on completion of the trial. They may spend this on whatever they like, including the Primary Science Quality Mark (PSQM), once the evaluation is completed at the end of 2020/21 academic year. Note that Control Schools will not be allowed to participate in or purchase PSQM while the trial is running, however they could sign up to participate in the academic year 2021/22.

WHAT IS INVOLVED IN PSQM?

During the 2019/20 academic year, **Intervention Schools** (i.e. those randomly allocated to participate in PSQM), will undertake the following:

- Science subject leaders and Year 5 teachers from the Intervention Schools will attend the equivalent of a minimum of 2 days' PSQM training between September 2019 and June 2020.
- Please note that attendance at the PSQM training is **mandatory** for all Intervention Schools.
- The training – which will cover an introduction to PSQM, creating and executing an action plan, and writing a reflective submission with evidence on the completed work – can be completed over *either:*
    - two full days *or* four half days
- Members of the Evaluation Team may visit some of the training sessions as part of the independent evaluation
- Online mentoring via the PSQM Virtual Learning Environment.
- The science subject leader and the Y5 teacher will collate and submit the evidence for the relevant PSQM award.

Please note that if the school has more than one Year 5 teacher, the school needs to nominate <u>one</u> teacher and class to participate in training.

The following table presents an overview of PSQM participation for an Intervention School.

Half term 1	Half term 2	Half term 3	Half term 4	Half term 5	Half term 6
CPD: Introduction to PSQM	CPD: Action planning		CPD: Reflective writing	CPD: Compiling a submission	
Y5 Teacher and subject leader complete self-assessment Agree award, Carry out principles activity and pupil voice survey , Appraise School Development Plan for science	Y5 Teacher and subject leader write action plan, Begin reflective logs	Y5 Teacher and subject leader implement action plan	Y5 Teacher and subject leader continue to implement action plan	Y5 Teacher and subject leader draft reflections, Begin to compile portfolio	Y5 Teacher and subject leader complete submission
On line mentoring - throughout					

OVERVIEW OF TRIAL TIMELINE (for summary table see p7)

2018/19

- In 2018/19, the Delivery Team will contact ALL schools to gather data about ALL current Year 3 and 4 pupils (names, sex, date of birth, Free School Meal status, and Unique Pupil Numbers). Current Year 4 pupils will be in Cohort A and Year 3 pupils will be in Cohort B.
- Prior to pupil data being sent to the Delivery Team, parents must be sent information and withdrawal forms by the school *and* have the opportunity to return these. Parents can withdraw their children at any time from the trial, but the initial withdrawal forms should be returned by parents within 2 weeks.
- If parents choose to withdraw their children from the trial later on, their data will not be collected or will be deleted, as appropriate (see Privacy Notice).
- Data for all Year 3 and 4 pupils is required so that:
    i. we ensure we have pupil data for all the pupils that will eventually be in the Year 5 trial class in academic years 2019-20 and 2020-21 – as larger schools typically cannot be certain which pupils will be in which class in Year 5 until quite late in the year, we collect all this data first then delete data that is not needed once we know which classes pupils are in;
    ii. we can plan and undertake outcome assessment; and
    iii. track pupil attainment longitudinally using the National Pupil Database.
- In schools with multiple Year 5 classes, the Evaluation Team will delete all data for children who are not in the trial Year 5 class once class lists have been finalised at the beginning of academic years 2019-20, and 2020-21.
- To minimise burden on pupils and schools, the Evaluation Team will use the National Pupil Database to collate 'baseline' (pre-Intervention) data about participating children's KS1 attainment in English and mathematics and/or Early Years Foundation Stage.

2019/20

- Year 5 children (Cohort A) in both Intervention and Control Schools from the nominated class will sit a brief science knowledge and science attitudes assessment.
- The nominated year 5 teacher and science subject leader will be asked to complete a survey about practices and attitudes related to teaching science.
- Surveys will also be sent to headteachers regarding science teaching, and will (for Intervention Schools) also include questions regarding the intervention implementation and additional costs associated with PSQM.
- Interviews on programme implementation will be conducted in several Intervention Schools.

2020/21

- Year 5 children (Cohort B) in both Intervention and Control Schools from the nominated class (Year 5 class taught by the nominated teacher) will sit a brief science knowledge and science attitudes assessment.
- The nominated year 5 teacher and science subject leader will be asked to complete a survey about practices and attitudes related to teaching science.

Trial timeline

Academic year 2018-19		
ALL SCHOOLS	**By end April 2019**	Sign MOU Send parents project information sheet with withdrawal form and privacy notice Send teachers privacy notice Return pupil information to PSQM: names, dates of birth and Unique Pupil Numbers for current Y3 (Cohort B) and Y4 (Cohort A) children
	July 2019	Schools notified of allocation by the Delivery team
Academic year 2019-20		
INTERVENTION SCHOOLS	**September 2019 to end June 2020, with training in:** **- Sept/Oct 2019** **- Nov/December 2019** **- April 2020** **- May 2020** **PSQM submission deadline - June 2020** **PSQM award notification - Sept 2020** **PSQM award event - Nov 2020**	PSQM intervention: Subject leader and Y5 teacher attend training and complete all PSQM processes in school
CONTROL SCHOOLS	**September 2019 to end June 2020**	Business as usual
ALL SCHOOLS	**June 2020**	Cohort A children (current Year 5 in selected class) sit a brief science knowledge and science attitudes assessment
Y5 teachers, science subject leaders in ALL SCHOOLS	**June 2020**	Complete a questionnaire about practices and attitudes related to teaching science
Head teachers in ALL SCHOOLS	**June 2020**	Complete survey about science teaching
Head teachers, teachers, subject leaders, parents and governors in SELECTED INTERVENTION SCHOOLS	**June 2020**	Take part in interviews about PSQM processes and science teaching
Academic year 2020-21		
ALL SCHOOLS	**September 2020 to end June 2021**	Business as usual
ALL SCHOOLS	**June 2021**	Cohort B children (current year 5 in selected class) sit a brief science knowledge and science attitudes assessment
Y5 teachers, science subject leaders in ALL SCHOOLS	**June 2021**	Complete a questionnaire about practices and attitudes related to teaching science

Head teachers, teachers, subject leaders, parents and governors in SELECTED INTERVENTION SCHOOLS	**June 2021**	Take part in interviews about science teaching
Academic year 2021-22		
	Sept 2022	Publication of evaluation report

RESPONSIBILITIES

The Delivery Team will:

- Be responsible for the collection of completed school Memorandum of Understanding (MOUs).
- Collect and collate Year 3 (Cohort B) and 4 (Cohort A) (in 2018/19) pupil data from all schools, consisting of: names, sex, date of birth, Free School Meal status, and Unique Pupil Numbers, sending this data to the Evaluation Team as soon as it is available from the school.
- Confirm which parents have withdrawn the children and send this to the Evaluation Team (so that the datasets can be checked by the Evaluation Team and any withdrawn pupils removed).
- Provide PSQM training (equivalent to two days) to the Year 5 teacher and the science subject leader.
- Provide feedback and support to the science subject leader and Y5 teacher via the PSQM Virtual Learning Environment
- Coordinate the assessment and awarding of Primary Science Quality Marks
- Communicate with the schools about all aspects of PSQM training, submission and assessment.
- Facilitate visits to PSQM training by the Evaluation Team.
- Be the first point of contact for any questions about the PSQM programme.

The Evaluation Team will:

- Design and conduct the independent evaluation of PSQM, which consists of the following activities:
  - Randomise participating school to Intervention/Control condition.
  - Coordinate the assessment of Year 5 children and surveying of teachers and subject leaders, which will be carried out by NFER at end of Year One (Cohort A) and Two (Cohort B) (2020 and 2021).
  - Develop and implement data collection tools.
  - Using the Unique Pupil Numbers of Year 3 and 4 pupils received from schools to link to National Pupil database to compile baseline data.
  - Analyse the outcome and process data from the trial.
  - Ensure all members of the Evaluation Team and Assessment implementers are appropriately trained and have full Disclosure and Barring Service (DBS) clearance where needed.
  - Publish a public report on the findings from the trial on the EEF website.
- Be the first point of contact for any evaluation-related questions about the trial.

The Schools will:

- Inform and consult school staff about the programme.
- Consent to randomised allocation.
- Distribute the Parental Information and Privacy Notice to parents.
- Distribute and collect withdrawal forms from parents.
- Distribute the Staff Privacy Notices.
- Collate names, sex, date of birth, Free School Meal status, and Unique Pupil Numbers of Year 3 and 4 pupils where parents have not withdrawn the children and send this to the Delivery Team.
- Share the names of the Year 5 teacher and science subject leader nominated for PSQM with the Delivery Team, and notify them of any staff changes.
- If the school is allocated to the intervention group: release the Y5 teacher and science subject leader to attend (the equivalent of) two days of PSQM training between September 2019 and June 2020.
- If the school is allocated to the intervention group: release the subject leader to complete all aspects of PSQM submission, and to lead monitoring and development activities – minimum three days.
- Accommodate Year 5 pupil assessment to be carried out by the Assessment Implementers (NFER) at the end of Year One (Cohort A) and Year Two (Cohort B).
- Accommodate teacher and science subject leader questionnaires be carried out by the Assessment Implementers at the end of Year One and Year Two.
- Facilitate data collection by the Evaluation Team as required - interviews, visits to PSQM training, document analysis.
  - Headteachers and other staff members in the Intervention and Control Schools will receive short online surveys.

- Interviews – in the four selected schools, several staff members and parents as well as school governors will be invited to take part in brief semi-structured telephone interviews (lasting no more than 20 mins) to discuss their experience of PSQM. Participation will be voluntary.
- Visits to PSQM training held on premises

To conduct the evaluation, schools are asked to share the specified data with the Delivery Team (PSQM) and the Evaluation Team (RAND Europe). Schools will notify the Delivery Team immediately if a school wishes to withdraw from the PSQM intervention or the trial data collection. *Note that even if schools withdraw from PSQM or the other trial activities, it is vital to the project that all pupils are assessed at the end of the first and second year of the trial, respectively.*

USE OF DATA

All data—including children's assessment responses—will be treated with total confidentiality. Data will be collected by the Evaluation Team and Assessment Implementers. Data analysis will be conducted by the Evaluation Team. After anonymisation, data will be shared with the EEF for archiving. No individual child, teacher, staff member, or school will be identified in any reports arising from the research. The trial is undergoing ethical review from RAND and the University of Hertfordshire. The intervention will not commence until approval is received.

Please see the Privacy Notices for further detail on how pupil and staff personal data will be collected and used.

Updated MoU sent to schools (June 2020)

Summary of amendments to Memorandum of Understanding between trial schools, the University of Hertfordshire and RAND Europe
June 2020- for information only, no signature required

Page 1.

- New project lead at Rand: Elena Rosa Brown

Page 3-4.

- There is now only one group of children involved in the trial, Cohort B, currently in year 4. Cohort B will be tested in June 2021 (when they are in Year 5), and no testing will take place for Cohort A (current year 5) which was planned for this June.
- Work on the PSQM programme will continue in school into academic year 20-21 (also page 4, 5 and 9).

Page 4:

- There will be 5 case study schools.
- The case studies will include documentary review and school site visits by the evaluation team in June 2021, but not observations of PSQM training (also page 6).
- Science subject leaders and Year 5 teachers from the Intervention Schools will attend the equivalent of a minimum of 2 days' PSQM training between September 2019 and February 2021 (also page 5 and 9).

Page 5:

- Training and submission deadlines moved to academic year 2020-21 (also page 9).

Page 7-8

- Revised timeline.

Follow-up communication with schools on further changes (March 2021)

<u>EMAIL TO PSQM INTERVENTION SCHOOLS</u>

To: Headteacher

CC: Y5 teacher and science subject leader

Subject: EEF PSQM TRIAL – CHANGES TO TESTING REQUIREMENTS MARCH 2021

Dear headteacher

Thank you for your continued participation in the PSQM evaluation trial.
We understand that this continues to be an incredibly challenging time for schools and as a result the plans for the testing phase of the trial have been altered, to reduce the requirements for participating schools. These changes are summarised for you below. There is no need to reply to this email, unless there have been any changes to the relevant staff contact details since your last update to us in December 2020/January 2021.

Summary of changes to testing phase of PSQM Evaluation Trial

1) The planned testing and survey of pupils in June 2021 is <u>cancelled</u> and there will not be any requirement for pupils to participate in any testing or survey in the future.

2) RAND Europe will send a short online survey about science in your school for you, your science subject leader and year 5 teacher to complete and return in May 2021. These surveys should take no more than 15 minutes to complete.

3) <u>10</u> schools only will be selected as case study schools and RAND Europe will be in touch with the selected schools from May 2021. The case studies will involve telephone interviews with relevant school staff and a review of documents pertaining to science in your school.
The selected schools will receive an additional £100.00 in compensation from PSQM for their participation.

Additional training sessions

We are pleased to confirm that the 2 additional training sessions, to provide extended PSQM hub leader support for your science subject leader, will go ahead as planned in the summer term. By continuing the relationship, the hub leader will be able to provide some of the support and encouragement that was inevitably missed this year and ensure that your school benefits from the full PSQM experience. These sessions are fully funded and there will be no cost to the school.

If you have any queries, please don't hesitate to get in touch.

With our best wishes

Jane Turner

EMAIL TO PSQM CONTROL SCHOOLS

To: Headteacher

CC: Y5 teacher and science subject leader

Subject: EEF PSQM TRIAL – CHANGES TO TESTING REQUIREMENTS MARCH 2021

Dear headteacher

Thank you for your continued participation in the PSQM evaluation trial.
We understand that this continues to be an incredibly challenging time for schools and as a result the plans for the testing phase of the trial have been altered, to reduce the requirements for participating schools. These changes are summarised for you below. There is no need to reply to this email, unless there have been any changes to the relevant staff contact details since your last update to us in December 2020/January 2021.

Summary of changes to testing phase of PSQM Evaluation Trial

1) The planned testing and survey of pupils in June 2021 is cancelled and there will not be any requirement for pupils to participate in any testing or survey in the future.

2) RAND Europe will send a short online survey about science in your school for you, your science subject leader and year 5 teacher to complete and return in May 2021. These surveys should take no more than 15 minutes to complete.

3) 10 schools only will be selected as case study schools and RAND Europe will be in touch with the selected schools from May 2021. The case studies will involve telephone interviews with relevant school staff and a review of documents pertaining to science in your school.

Payments
Your school has already received a payment of £300 for participation in the trial as a control group school. A further payment of £1,200 was due to be made once all evaluation activities were complete. Even though outcome testing will not go ahead, we will pay the remaining £1,200 before the end of the summer term, in recognition of your support for all other data collection and to thank you for your collaboration in this trial. If your school is selected as one of the 10 case study schools then you will need to provide the information required for the case study before the payment is made.

If you have any queries, please don't hesitate to get in touch.

With our best wishes

Jane Turner

Appendix D: Participant information sheet

Independent Evaluation of the Primary Quality Science Mark

Information sheet for participants asked to take part in interviews

Background to the Independent Evaluation of the Primary Science Quality Mark

The Primary Science Quality Mark (PSQM) was initiated in 2008 at the University of Hertfordshire to raise the profile of science in primary schools in England and promote professional development in science teaching and leadership. The PSQM programme is led by The University of Hertfordshire and is being independently evaluated by RAND Europe. The study is funded by the Education Endowment Foundation and Wellcome Trust.

RAND Europe was commissioned to look at the effectiveness of the intervention. The evaluation team will provide real-time feedback to allow continuous improvement in the design and delivery of the intervention, and robust evidence to allow the Education Endowment Foundation to make decisions about its future funding.

As part of the evaluation activities, we are conducting case studies with a number of randomly selected schools. The case studies involve interviews with a number of individuals and will help us better understand their perception of the programme and practical aspects of delivering the programme (e.g. what makes it easy/hard to deliver the programme and what resources are needed).

What will taking part in the interviews involve?

We would like to invite you to take part in an interview with a member of the evaluation team.

The interview would be arranged at a time that is convenient for you and would last between 30 and 45 minutes.

The interview would be conducted over Microsoft Teams.

There are no wrong or right answers to questions asked during the interview – we're just interested in hearing about your experiences and views. You don't have to answer any questions you don't want to, and you can stop and take a break at any point.

You can also change your mind about taking part at any point during the interview.

With your permission, we will make a digital audio-recording of the interview. Recordings will be kept securely, not accessible to anyone outside the evaluation team, and will be deleted at the end of the evaluation.

How will the information from the interviews be used?

The information you provide will be used only for the purposes of the Independent Evaluation of the PSQM. Findings from the interviews will help inform the design and implementation of the future editions of PSQM. Interview findings may be included in reports which will be published and made available to the public. The results of the evaluation may also be published in academic journals and other short summaries available to the public.

All these reports will bring together the data for all interviewees and it will not be possible to identify any individual. Findings (including any direct quotes) used in the reports will be anonymised. We will take care to remove information that could identify an individual as the source of the quote. It will not be possible to identify you in any publication from the evaluation.

What about the security of my personal information?

Strict arrangements will be in place to make sure that information collected from you is stored securely in line with the General Data Protection Regulation (GDPR) 2018 and the Data Protection Act (2018). These guidelines mean that your information will be kept safe and secure.

One year after the evaluation is completed all data collected through the interview will be deleted, along with any information that could be used to identify you.

All data will be kept in confidence. Only members of the evaluation team based at RAND Europe will have access to your data, which will be processed in accordance with all legal obligations. Only aggregated data and summaries will be reported.

One year after the evaluation is completed all audio recordings will be deleted, along with any information that can be used to identify you, such as your name. The evaluation team will only keep written notes of the interview.

Voluntary participation

Participation in the evaluation of the intervention is entirely voluntary. You may refuse to participate or refuse to answer specific questions at any time without penalty. Your contribution will provide valuable information that will help us to evaluate PSQM and we do hope that you decide to take part. However, participation is voluntary, and you can change your mind about participating in the evaluation at any stage without penalty.

Consent forms

Before participating in an interview, we will ask you to confirm your voluntary agreement to take part. If you are happy to proceed with the interview, please sign and return a copy of the consent form to the evaluation team.

Your agreement confirms that you have received and understood the information sheet (this document) and have had time to decide if you want to take part. You can withdraw from participation at any time, without giving a reason, and there will not be a penalty of any kind.

Consenting in the evaluation

If you agree to take part in the interview but change your mind at a later stage, you have the right to ask for the information collected from you in the interview to be deleted and not used in the evaluation. You can ask for your data to be deleted by emailing **PSQM@randeurope.org** or RAND Europe's Data Protection Officer at **redpo@randeurope.org.**

Do you have any questions?

If you have any questions about the interview or the Independent Evaluation of PSQM, please get in touch with a member of the evaluation team using the following email address: **PSQM@randeurope.org.**

Appendix E: Privacy notice

Privacy Notice
PSQM Case Study Interviews

About the project, who we are and what data we collect

The accompanying information sheets outline information about the project. This privacy notice outlines how your data will be used as part of the project.

RAND Europe Community Interest Company is a not-for-profit research organisation registered in the UK conducting independent research to inform policy.

In this project we will collect your name, email and telephone number.

Why are we collecting it?

We are collecting your data in order to arrange your participation in a telephone interview. These interviews make up part of the research activity being undertaken to evaluate the PSQM programme.

What is the legal basis for processing your data?

We are using your data on the basis of our legitimate interests. Your data is collected and processed solely to facilitate your voluntary contribution to the project. The data is not excessive and will be used for the purposes of contacting you to arrange your participation in the interview, as required to meet the project goals. These project goals have been explained to you and will lead to a wider public benefit through our work. The data is necessary for the purpose of the project, as without we would be unable to undertake the interview or recognise your contribution. We also judge that there is very limited scope for harm to you as appropriate data handling safeguards have been put in place. As such the approach to processing balances our legitimate interests against your interests, rights and freedoms.

What do we use the data for?

We will use your data to contact you. We will use a random unique ID to attribute any contribution of yours that is used in our report, not your name.

How do we share the data, and how do we keep your data secure?

We will keep all data safe on our secure servers. We will not share your data with any third parties.

How long do we keep your data?

Your data will be deleted within 12 months of the end of the project (end of project estimated April 2022).

What choices do you have in our use of your data?

You may contact us to request the deletion of your personal data.

What are your rights?

RAND Europe operates in accordance with the Data Protection Act 2018 and EU law including GDPR. You are provided with certain rights that you may have the right to exercise through us. In summary those rights are:

- To access, correct or erase your data. Your right to erase your name in relation to any attribution shall expire after it has been submitted for publication.

- To object to the processing of your data. Your right to object to processing of your name in relation to any attribution shall expire after it has been submitted for publication.
- To request that our processing or your data is restricted. Your right to restrict processing of your name in relation to any attribution shall expire after it has been submitted for publication.

If you wish to exercise any of these rights please contact the RAND Europe Data Protection Officer by email at **REDPO@randeurope.org** or in writing to Data Protection Officer, RAND Europe, Westbrook Centre, Milton Road, Cambridge, CB4 1YG, UK.

How do you contact us?

You can contact us by email at **PSQM@randeurope.org**

Appendix F: Withdrawal form

Evaluation of Primary Science Quality Mark

Please only complete if you do NOT want your child's data to be used in this trial

If you are happy for your child's data to contribute to this trial, then you do not have to do anything.

If you do NOT want your child's data to be used for this research, you can withdraw your child via returning this form to the school or completing a form online.

I, the undersigned, hereby do NOT give permission for my child's data to be used for the purposes of the trial of the Primary Science Quality Mark.

Child's full name: ____

School:______________________________________Class __________________

Parent/guardian name: ____________________________________

Parent/guardian signature: ____________________________________

Date: ____________________________________

Appendix G: Documentary review of PSQM submissions

The documents analysed for the submission review[45] provided evidence on 26 key dimensions that define good practices in terms of science teaching that ought to be enhanced by PSQM. The 26 dimensions are outlined in Table 21. They pertain to key PSQM aspects such as enriching the science learning experience, improving the shared understanding of the value of science across the school, a focus on CPD, the centrality of science enquiry, guiding principles for science teaching, a better understanding of science assessment and monitoring of teaching practices, a commitment to expand children's science capital, among other aspects. On average, each school provided evidence on 22.5 out of 26 dimensions in at least one document, which shows a high degree of fidelity to the programme implementation.

In conducting the submission review we looked at 30 schools. Each school submitted 16 number of documents. After an initial screening to identify the most informative documents (based on usefulness of information contained and ease of comparability with other schools' documentation), four documents were selected (these were standard across all 30 schools). This led to a total of 120 documents selected for an in-depth review. We reviewed all 120 documents checking whether each of them provided evidence on each of the 26 dimensions. We then used the two criteria below to identify the 26 dimensions with most and least evidence:

a) How many schools provide evidence on [dimensions 1 to 26] in at least one of the four documents? This criterion focuses on schools.
b) How many of the 120 documents reviewed provide evidence on [dimensions 1 to 26] (in %)? This criterion focuses on the overall number of documents across all schools.

Both criteria should be considered together, as they complement each other. For example, two dimensions where the same number of schools provided evidence in at least one document (criterion a), may vary in the overall number of documents evidencing it at aggregate level (criterion b).

To classify the dimensions from most to least amount of documentation supporting it, we calculated the first, second, third and fourth quartile of the number of documents evidencing the dimension. Each dimension was assigned into one of four quarters of the distribution (see Table 21 below).

[45] Four documents were analysed, for 30 schools (120 documents in total):

1. *Calendar*: illustrates schools' programme of science events, specifying which age group they are aimed at (EYFS, KS1, LKS2, UKS2, whole school)
2. *CPD Log*: schools' record of CPD activities for subject leaders and teachers
3. *Portfolio*: summarises approach to science teaching and how they implemented PSQM's programme
4. *Reflections form*: we selected one out of the four 'PSQM Action to Reflection Plan' forms that submitted

The Theory of Change details when they are drafted (at which stage of the training) and submitted.

Table 21: Submission review analysis overview

Dimension	Criteria		
	a. Number of schools providing evidence on this dimension in at least one document	b. Percentage (%) of the 120 documents reviewed providing evidence on this dimension	Quarters based on the number of schools
There are appropriate links with families, other schools, communities and outside organisations to enrich science learning. (Evidence on science visits, visitors and/or science weeks for my class (these can all be virtual or in person))	30	79.20%	fourth quarter
A shared understanding of the importance and value of science	30	68.30%	fourth quarter
Engagement with professional development to improve science teaching and learning- regular CPD for all teachers led by SL who is proactive in seeking out new pedagogies/resources. Evidence on teachers or SSL being able to identify needs for CPD.	30	68.30%	fourth quarter
A shared understanding of the purpose and process of science enquiry. Evidence that teacher encourages: children use full range of enquiry types, asking and answer own question, investigate independently, enquiry part of almost every lesson	30	66.70%	fourth quarter
An agreed set of principles for good science teaching and learning	30	61.70%	fourth quarter
Shared understanding of the purposes of science assessment and current best practice- formative assessment used regularly to improve lesson, and summarised periodically for reporting, progression is tracked. Evidence on SSL ability to evaluate science teaching and improve it, SSL identifying high quality science teaching resources, disseminating to teacher colleagues	30	61.70%	fourth quarter
Appropriate and active goals for developing science	30	58.30%	fourth quarter third quarter
Evidence on school/classes/pupils participating in science enrichment activities (i.e. any structured activities focused on science enrichment on top of what a school would normally provide to pupils to improve their science learning).	29	79.20%	third quarter
Evidence around pupil engagement, confidence and attainment in science. Is there evidence that this is targeted? That this has improved?	29	73.30%	third quarter
A clear development cycle linked to an established monitoring programme (learning walks, lesson observations, book scrutiny, pupil voice, data analysis)	29	60.80%	third quarter
There are appropriate links between science and other learning.	29	60.80%	third quarter
A clear vision for the teaching and learning of science	29	50.80%	third quarter
There is a commitment to developing all children's science capital- every class has science visits and visitors; teacher know the importance of developing positive attitudes towards science; parent are involved; science weeks take place; children engaged in global scientific issues, children care for their environment, children are engaged in national scientific initiatives.	28	76.70%	third quarter
Regular outdoor science, children grow fruit and vegetables, get to know their local environment well	28	58.30%	second quarter

Regular science staff meetings (science discussed at staff meetings)	28	47.50%	second quarter
A range of effective strategies for teaching and learning science which challenge and support the learning needs of all children - strategies are identified in accordance with development needs, not ad hoc and effectiveness evaluated - use of up top date resources e.g. Explorify, BAES	27	56.70%	second quarter
SL actively engaged in external networks to develop science	27	55.00%	second quarter
Evidence on teachers reflecting regularly on their science teaching practice to inform future teaching and learning	27	52.50%	second quarter
Range of up-to-date, quality resources for teaching and learning science which are used regularly and safely – including IT e.g. data loggers, IPADs etc	26	48.30%	second quarter
A science display in every classroom	26	44.20%	second quarter first quarter
Science visible on the website and in communal areas	26	42.50%	first quarter
Links with other schools to share and develop science	25	48.30%	first quarter
School membership of ASE	16	26.70%	first quarter
Science featuring in school newsletter	15	15.00%	first quarter
Reports to governors include science	14	22.50%	first quarter
A designated budget for science resources and CPD	8	10.00%	first quarter

The same criteria were used to assess the documents provided by control schools. However, given the small number of documents received a similar summary as above is not provided.

Appendix H: PSQM evaluative criteria

These criteria were effective at the time of the trial, but a more recent version was updated in 2020.

<table>
<tr><th colspan="3"></th><th>PSQM award</th><th>PSQM GILT award</th></tr>
<tr><td rowspan="34">ESSENTIAL</td><td rowspan="10">SUBJECT LEADERSHIP</td><td rowspan="2">SL1 There is a clear vision for the teaching and learning of science</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• establishing a clear vision for science;
• the development of school principles for science teaching and learning by teachers and children?</td><td>• a clear vision for science, which is embedded and regularly reviewed;
• school principles for science teaching and learning, which have been developed by the whole school community and are reviewed regularly?</td></tr>
<tr><td rowspan="2">SL2 There is a shared understanding of the importance and value of science</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• the school community's developing understanding of the importance and value of science?</td><td>• the school community's shared understanding of the importance and value of science?</td></tr>
<tr><td rowspan="2">SL3 There are appropriate and active goals for developing science</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• effective support of the subject leader's development of science through school strategic planning processes?</td><td>• effective support for the development of science through embedded school strategic planning processes;
• the subject leader's contribution to whole school leadership and strategy?</td></tr>
<tr><td rowspan="2">SL4 There is a commitment to the professional development of subject leadership in science</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• the subject leader's engagement with professional development and learning;
• the subject leader's interest in science, which is communicated to others?</td><td>• the subject leader's engagement with sustained professional development and learning;
• the subject leader's engagement with the primary science education community, which is shared with others?</td></tr>
<tr><td rowspan="2">SL5 There are monitoring processes to inform the development of science teaching and learning</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• the subject leader using a range of processes to monitor science teaching and learning;
• the subject leader ensuring that pupil voice is valued and responded to;
• the subject leader sharing outcomes with colleagues and implementing appropriate actions?</td><td>• school wide processes to monitor, evaluate and develop science teaching and learning, which are rigorous and embedded;
• monitoring and evaluation processes which use evidence and views from all stakeholders and sources;
• school development priorities which are informed by monitoring and evaluation processes?</td></tr>
<tr><td rowspan="6">SCIENCE TEACHING</td><td rowspan="2">T1 There is engagement with professional development to improve science teaching and learning.</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• staff engagement with relevant internal or external professional development;
• the subject leader's provision of professional support for colleagues in response to development needs?</td><td>• staff engagement with a sustained programme of relevant internal or external professional development;
• the subject leader's provision of regular, sustained support for colleagues in response to development needs?</td></tr>
<tr><td rowspan="2">T2 There is a range of effective strategies for teaching and learning science which challenge and support the learning needs of all children.</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• teachers using a range of effective strategies for teaching science which challenge and support the learning needs of all children;
• the subject leader introducing new strategies for teaching science in response to development needs?</td><td>• teachers using and evaluating a range of evidenced based strategies for teaching science which challenge and support the learning needs of all children;
• the subject leader developing existing strategies and introducing new ideas for teaching science in response to development needs?</td></tr>
<tr><td rowspan="2">T3 There is range of up-to-date, quality resources for teaching and learning science which are used regularly and safely.</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• resources that are audited annually, well-organised and accessible;
• children regularly and safely using appropriate practical and digital resources, information texts and the outdoor environment?</td><td>• resources that are audited in line with development planning for science, are well-organised and accessible;
• children regularly and safely using a wide range of appropriate practical and digital resources, information texts and the outdoor environment;
• links teachers make with outside agencies to borrow or source additional equipment where necessary?</td></tr>
<tr><td rowspan="6">SCIENCE LEARNING</td><td rowspan="2">L1 There is a shared understanding of the purpose and process of science enquiry.</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• children using different enquiry types to answer scientific questions about the world around them;
• children developing independence in:
  o asking scientific questions
  o planning how to investigate them,
  o carrying out and evaluating investigations?</td><td>• children using a range of enquiry types to answer scientific questions about the world around them;
• children independently:
  o asking scientific questions,
  o planning how to investigate them
  o carrying out and evaluating investigations?</td></tr>
<tr><td rowspan="2">L2 There is a shared understanding of the purposes of science assessment and current best practice.</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• teachers using a range of strategies and processes for formative, summative and statutory assessment, which reflect the school understanding of the purposes of assessment in science and current best practice;
• the subject leader developing assessment practice?</td><td>• regular evaluation of practice and processes for formative, summative and statutory assessment which ensure that they reflect the shared understanding of the purposes of assessment in science and current best practice;
• a school wide commitment to continually improving assessment practice?</td></tr>
<tr><td rowspan="2">L3 There is a commitment to developing all children's science capital</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• the subject leader promoting initiatives that encourage all children to think that science is relevant and important to their lives, now and in the future?</td><td>• the school community supporting and promoting initiatives that encourage all children to think that science is relevant and important to their lives, now and in the future?</td></tr>
<tr><td rowspan="4">WIDER OPPORTUNITIES</td><td rowspan="2">WO1 There are appropriate links between science and other learning.</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• curriculum planning that links science to other areas of learning?</td><td>• whole school planning that links science to other areas of learning including specific links with other core subjects;
• science being part of whole school initiatives?</td></tr>
<tr><td rowspan="2">WO2 There are appropriate links with families, other schools, communities and outside organisations to enrich science learning.</td><td colspan="2">What is the impact on science teaching and learning of:</td></tr>
<tr><td>• children taking part in some initiatives supported by other organisations to enrich science learning;
• children's science learning including topical science events;
• children carrying out science activities with their families?</td><td>• children taking part in a range of initiatives supported by other organisations to enrich science learning;
• children's science learning including topical science events;
• children regularly carrying out science activities with their families?</td></tr>
</table>

PSQM OUTREACH award

ESSENTIAL	SUBJECT LEADERSHIP	**SL1 There is a clear vision for the teaching and learning of science**	**What is the impact on science teaching and learning of:** • a clear vision for science, which is embedded and regularly reviewed; • school's principles for science teaching and learning, which have been developed by the whole school community and are reviewed regularly?
		SL2 There is a shared understanding of the importance and value of science	**What is the impact on science teaching and learning of:** • the school community's shared understanding of the importance and value of science?
		SL3 There are appropriate and active goals for developing science	**What is the impact on science teaching and learning of:** • • effective support for the development of science through embedded school strategic planning processes; • the subject leader's contribution to whole school leadership and strategy?
		SL4 There is a commitment to the professional development of subject leadership in science	**What is the impact on science teaching and learning of:** • the subject leader's engagement with sustained professional development and learning; • the subject leader's engagement with the primary science education community, which is shared with others?
		SL5 There are monitoring processes to inform the development of science teaching and learning	**What is the impact on science teaching and learning of:** • school wide processes to monitor, evaluate and develop science teaching and learning, which are rigorous and embedded; • monitoring and evaluation processes which use evidence and views from all stakeholders and sources; • school development priorities which are informed by monitoring and evaluation processes?
	SCIENCE TEACHING	**T1 There is engagement with professional development to improve science teaching and learning.**	**What is the impact on science teaching and learning of:** • staff engagement with a sustained programme of relevant internal or external professional development; • the subject leader's provision of regular, sustained support for colleagues in response to development needs.
		T2 There is a range of effective strategies for teaching and learning science which challenge and support the learning needs of all children.	**What is the impact on science teaching and learning of:** • teachers using and evaluating a range of evidenced based strategies for teaching science which challenge and support the learning needs of all children; • the subject leader developing existing strategies and introducing new ideas for teaching science in response to development needs?
		T3 There is range of up-to-date, quality resources for teaching and learning science which are used regularly and safely.	**What is the impact on science teaching and learning of:** • resources that are audited in line with development planning for science, are well-organised and accessible; • children regularly and safely using a wide range of appropriate practical and digital resources, information texts and the outdoor environment; • teachers making links with outside agencies to borrow or source additional equipment where necessary?
	SCIENCE LEARNING	**L1 There is a shared understanding of the purpose and process of science enquiry.**	**What is the impact on science teaching and learning of:** • children using a range of enquiry types to answer scientific questions about the world around them; • children independently: o asking questions, o planning how to investigate them o carrying out and evaluating investigations?
		L2 There is a shared understanding of the purposes of science assessment and current best practice.	**What is the impact on science teaching and learning of:** • regular evaluation of practice and processes for formative, summative and statutory assessment which ensure that they reflect the shared understanding of the purposes of assessment in science and current best practice; • a school wide commitment to continually improving assessment practice?
		L3 There is a commitment to developing all children's science capital	**What is the impact on science teaching and learning of:** • the school community supporting and promoting initiatives that encourage all children to think that science is relevant and important to their lives, now and in the future?
	WIDER OPPORTUNITIES	**WO1 There are appropriate links between science and other learning.**	**What is the impact on science teaching and learning of:** • whole school planning that links science to other areas of learning including specific links with other core subjects; • science being part of whole school initiatives?
		WO2 There are appropriate links with families, other schools, communities and outside organisations to enrich science learning.	**What is the impact on science teaching and learning of:** • children taking part in a range of initiatives supported by other organisations to enrich science learning; • children's science learning including topical science events; • children regularly carrying out science activities with their families?
	PROFESSIONAL DEVELOPMENT	**O1. There is a commitment to leading professional development and learning in science in other schools**	**What is the impact on science teaching and learning in other schools of:** • the subject leader regularly sharing good practice beyond their own school; • the planning and evaluation of science outreach initiatives; • effective cross phase links within and/or between schools?
EITHER OR/AND	LOCAL ACTIVITY	**O2. There is a commitment to working with other community groups and organisations to develop their science teaching and learning**	**What is the impact on community engagement in science or one or more of:** • developing a resource for others e.g. pond at public nature reserve? • collaboration with local science activity e.g. industry, environment, astronomy? • science shared beyond the school community (beyond informing parents and governors)? • projects to develop science capital?
	WIDER ACTIVITY	**O3. There is a commitment to sharing expertise in science teaching and learning beyond the immediate community**	**What is the impact on science teaching and learning of one or more of:** • training people in industry to work in schools? • reciprocal global links? • working with colleagues in Initial Teacher Training? • sharing expertise through : o writing for journals e.g. ASE, Phizzi News TES? o regular online blogging? o contributing to published resources? o presenting at conferences? o contributing to policy level activity? In the case of publications, conferences and consultations, evidence of impact may be limited so evidence of activity will be acceptable.

Appendix I: Pilot for measures of pupil attitudes to science

A pilot for measures to capture pupil attitudes to science and scientific inquiry in the PSQM trial

Prepared by RAND Europe

July 2020

RAND EUROPE

Background

The current randomised controlled trial of PSQM assesses the PSQM's impact on science attainment (primary outcome) and science attitudes (secondary outcome).

The protocol for the PSQM trial currently states that:

> *To assess changes in pupils' attitudes towards science, we suggest carrying a post-test survey at the same time as the primary outcome assessment. The attitudinal measure at post-test will also be administered by NFER using paper forms. The attitudinal measures will be compiled in machine-readable forms, to allow scanning, data entry and scoring by RAND Europe.*
>
> *Enjoyment of science, confidence in science and engaging teaching in science will be measured using the 'enjoyment of science' subscale adapted from the Trends in International Mathematics and Science Study (TIMMS) Grade 4 surveys from TIMMS 2015.13 As described above and illustrated in Table 1, the attainment measure from 2019/20 will also be used as a secondary outcome.*
>
> *If possible, science enquiry skills will be captured from relevant items in the science attainment test by Hanley and colleagues.*

Since the protocol was drafted, and pending *revisions* relating to the measure for the primary outcome (science attainment), more work has been carried out to specify and strengthen the quality of the measure of science attitudes in the trial.

This can be achieved by piloting a higher number of measures than planned and selecting those that are most appropriate to include in the final secondary outcome assessment. The pilot will allow for decisions about this final measurement to be based on robust and relevant evidence collected from a sample of pupils similar to those participating in the PSQM trial.

Aim of pilot

The aim of this pilot is therefore to test a number of these proposed measures, to understand:

- The most robust measures to be included, from a validity and reliability perspective
- The best combination of practical measures to capture pupils' science attitudes

Methodological approach

Sample

A total of 126 Year 5 pupils from 3 schools responded to the questionnaire.

Administration

The questionnaire was distributed by class teachers, with pupils asked to mark their responses on paper. The questionnaire distributed to pupils is included in Appendix 1.

Data was then manually entered into a spreadsheet for the purposes of analysis.

Measurement scales

The guiding principles behind attitudinal measurement as part of PSQM are:

- The measure should capture attitudes to science validly and reliably
- The measure should be easy to administer to pupils in the trial (e.g. appropriate reading age, length, content)
- The measure should be comparable to previous measures used, either in evaluations of science interventions or international studies, in as much as possible

The following scales were included in the pupil questionnaire:

Scale 1: Enjoyment of science	
1	I enjoy learning science
2	I wish I did not have to study science
3	Science is boring
4	I learn many interesting things in science
5	I like science
6	I look forward to learning science in school
7	Science teaches me how things in the world work
8	I like to do science investigations
9	Science is one of my favourite subjects
Response options	disagree a lot - disagree a little - agree a little - agree a lot
Source	TIMSS 2015

Scale 2: Interest in science	
1	Science lessons make me think
2	I look forward to my science lessons
3	Science lessons are interesting
4	I would like to do more science at school
5	Science is fun
7	I enjoy discussions in science lessons
8	Science lessons are boring
9	I think science lessons are more for boys
10	We often have discussions in science lessons
11	We do a lot of writing in science lessons
12	It is important that we learn science
13	I like thinking about scientific ideas
Response options	disagree a lot - disagree a little - agree a little - agree a lot
Source	Amended from: Hanley et al., 2018, from TDTS, from Kind, Jones & Barmby, 2007

Scale 3: Science self-efficacy	
1	I find science difficulty to understand
2	I am just not good at science
3	I understand everything in my science lessons
Response options	disagree a lot - disagree a little - agree a little - agree a lot
Source	Hanley et al., 2018, from TDTS, from Kind, Jones & Barmby, 2007

Scale 4: Inquiry-based science learning	
1	In science lessons I am given opportunities to find the answers to science questions myself
2	In science lessons I am given opportunities to find out how things work
3	The science we do in school has links to my everyday life
4	In science lessons I am given opportunities to ask science questions which I can investigate myself
5	In science lessons I am given opportunities to design my own science investigations
6	In science lessons I am given opportunities to talk about what might happen before doing a science investigation and give reasons
7	In science lessons I am given opportunities to explain what I notice in science investigations

8	In science lessons I am given opportunities to use scientific vocabulary to explain what I have learned
9	In science lessons I am given opportunities to work with others to investigate science questions
10	In science lessons I am given opportunities to do investigations to learn science facts
11	I do science investigations outside of school
Response options	disagree a lot - disagree a little - agree a little - agree a lot
Source	New, to reflect key concept in PSQM's theory of change

Feedback on these question (attached to each of the measurement scales above)	
1	Were these questions easy to answer?
2	Were these questions easy to understand?
Response options	Very easy – easy – difficult – very difficult
Source	New, for purposes of piloting of measure

Analysis

The data analysis was conducted in four steps:

Step 1: descriptive analysis: to understand how pupils respond and what the distribution of data is, for each respective question

Step 2: exploration of pupils' feedback on the specific questions

Step 3: scale-specific measurement focus: to understand how each measurement scale operates independently; if there are any items that could be removed from each scale (to shorten administration time, and in line with feedback from pupils) while maintaining the integrity of the measurement; how internally consistent each scale is (this will use statistical techniques including exploratory factor analyses; reliability analyses).

Step 4: convergent validity: undertaken by assessing the relationships between the concepts measured with the respective measurement scales, this will provide an understanding of whether, and how, the four measurement scales overlap (this will use statistical techniques including correlations and further factor analyses).

Results

Descriptive statistics

Pupil responses to each question are illustrated in the table below. Overall, between 110 and 118 pupils responded to each respective question, of the 126 pupils that engaged with the questionnaire.

Scale		**Response options**				**Respondents**
	Question	**Disagree a lot**	**Disagree a little**	**Agree a little**	**Agree a lot**	
1						
1	I enjoy learning science	4	21	48	44	117
2	I wish I did not have to study science	50	38	21	9	118
3	Science is boring	49	30	29	7	115
4	I learn many interesting things in science	1	11	39	67	118
5	I like science	8	24	37	47	116
6	I look forward to learning science in school	10	32	38	36	116
7	Science teaches me how things in the world work	5	7	42	64	118

8	I like to do science investigations	7	6	31	73	117
9	Science is one of my favourite subjects	32	29	31	26	118
2						
1	Science lessons make me think	3	13	52	49	117
2	I look forward to my science lessons	14	31	36	35	116
3	Science lessons are interesting	6	14	41	55	116
4	I would like to do more science at school	22	28	32	35	117
5	Science is fun	6	24	38	45	113
6	I enjoy discussions in science lessons	15	28	35	36	114
7	Science lessons are boring	65	28	14	9	116
8	I think science lessons are more for boys	88	10	10	6	114
9	We often have discussions in science lessons	3	15	47	51	116
10	We do a lot of writing in science lessons	7	29	45	33	114
11	It is important that we learn science	5	5	31	75	116
12	I like thinking about scientific ideas	12	26	39	39	116
3						
1	I find science difficult to understand	28	40	38	8	114
2	I am just not good at science	40	36	27	10	113
3	I understand everything in my science lessons	21	37	33	22	113
4						
1	In science lessons I am given opportunities to find the answers to science questions myself	4	17	56	36	113
2	In science lessons I am given opportunities to find out how things work	2	7	30	74	113
3	The science we do in school has links to my everyday life	14	29	42	28	113
4	In science lessons I am given opportunities to ask science questions which I can investigate myself	4	21	50	35	110
5	In science lessons I am given opportunities to design my own science investigations	16	27	42	28	113
6	In science lessons I am given opportunities to talk about what might happen before doing a science investigation and give reasons	4	15	36	57	112
7	In science lessons I am given opportunities to explain what I notice in science investigations	6	19	35	53	113
8	In science lessons I am given opportunities to use scientific vocabulary to explain what I have learned	4	17	46	50	117
9	In science lessons I am given opportunities to work with others to investigate science questions	3	7	45	62	117
10	In science lessons I am given opportunities to do investigations to learn science facts	1	8	59	49	117
11	I do science investigations outside of school	31	27	26	33	117

Scale feedback

Pupils were asked to rate if each scale was easy to understand, and respectively easy to answer. Pupils were presented with the pictorial response options only, which we interpret as “very difficult” – “difficult” – “easy” – “very easy". Regardless of the exact meaning of the pictorial response options understood by the pupils, the symbols are progressively more positive.

Overall, between 111 and 115 pupils provided answers to these rating questions, with 86.9% of pupils providing positive answers to these rating questions.

In terms of ease of understanding, Scale 1 was rated most positively, with 91.3% of pupils providing positive answers: 69.6% of pupils chose the symbol indicating scale questions were very easy to understand; and an additional 21.7% indicated the questions were easy to answer. These were followed by Scale 3, and Scale 2. Of these two scales, Scale 2 had a higher overall positive response (89.5%, compared to 86.5% for Scale 3), but a lower proportion of pupils indicating that the questions were very easy to understand (61.4% compared to 64% for Scale 3).

For Scale 4, overall positive responses stood at 83.3%: 54.4% of pupils chose the symbol indicating that the questions were very easy to understand, and an additional 28.9% indicated the questions were easy to understand.

These differences were only statistically significant in relation to Scale 4, which was significantly less positively rated by pupils compared to Scale 1.

Scale	"These questions were easy to understand"				Respondents
					
1	2	8	25	80	115
2	3	9	32	70	114
3	5	10	25	71	111
4	4	15	33	62	114

In terms of ease of answering, Scale 3 was rated most positively overall (89.4% of pupils chose options indicating the questions were very easy or easy to answer). This is not surprising given that it was the shortest, at only 3 questions. Scale 2 and 4 followed, and performed similarly to each other, with 86.6% and 86% respectively providing positive responses: 56.3% (Scale 2) and 55.3% (Scale 4) of pupils indicated that the questions were very easy to answer.

Scale 1 attracted a higher proportion of responses indicating that the questions were very easy to answer compared to Scale 3 (58.8%, compared to 57.5% for Scale 3), and an additional 23.7% of pupils who suggested it was easy to answer. However, the data distribution also suggested that Scale 1 attracted the highest proportion of negative responses, with 17.5% of pupils indicating questions were very difficult or difficult to answer.

None of these differences, however, were statistically significant.

Scale	"These questions were easy to answer"				Respondents
					
1	7	13	27	67	114
2	2	13	34	63	112
3	5	7	36	65	113
4	7	9	35	63	114

Overall, on their own, these results would suggest that Scales 1, 2, and 3 are contenders for later use, and Scale 4 performed the worst overall.

Scale reliabilities and underlying factor structures

To understand if each of the four scales are reliable and perform as expected from a conceptual perspective (i.e. if they each address one single underlying factor that captures most of the variance in pupils' outcomes), a reliability analysis followed by a factor analysis was undertaken.

To explore reliability, or internal consistency, Cronbach's alpha was calculated for each scale. Values above 0.8 indicate high reliability, and suggest the measure is appropriately internally consistent for later use. We also report if modifying the scale by removing questions (indicating which) would improve the reliability of the scale

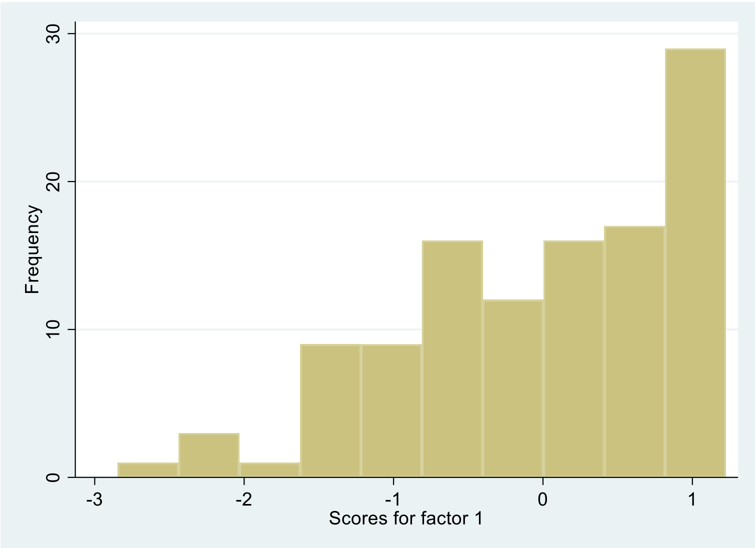
The factor analysis looked at the underlying structure of each scale. The analysis allows for the extraction of factors from the variance structure of responses to the questions making up each scale. When choosing or developing these scales, we hypothesised that each would represent a single factor, i.e. the concept the scale is meant to measure. We undertook an exploratory factor analysis to explore what the data structure would suggest, and report the number of factors suggested by the analysis, together with an indication of whether each question substantially contributes to the extracted factors (and which questions do not).

Scale	Reliability coefficient (>0.8 is desirable)	Would removing questions improve reliability?	Number of extracted factors (1 is expected)	Do all questions substantially contribute to the scale?
1	0.89	No	1	Yes
2	0.83	Yes (9 & 10)	1	No (9 & 10)
3	0.66	No	1	Yes
4	0.80	Yes (5)	1	No (5 & 9)

Results suggest that Scale 1 offered the best reliability, and the best result in relation to its underlying factor structure. Scale 3 resulted in the lowest reliability coefficient[46] and on the basis of that alone would not be recommended for later use.

Distributions

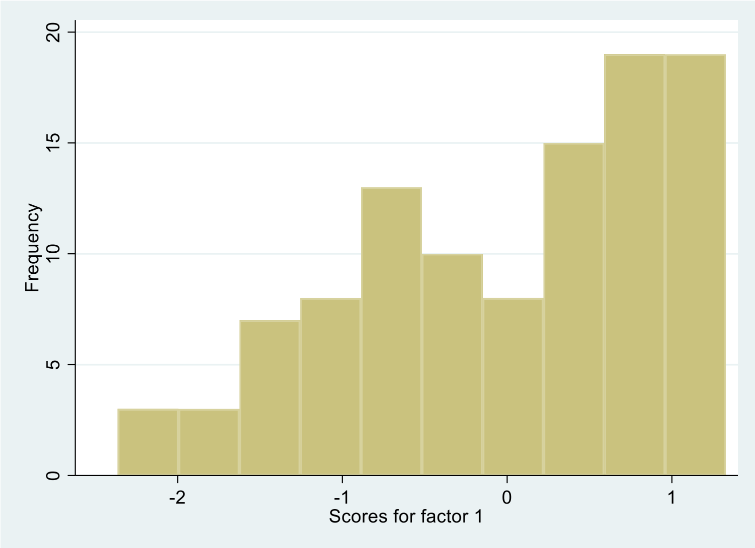
The factor score predicted for Scale 1 has the following distribution:



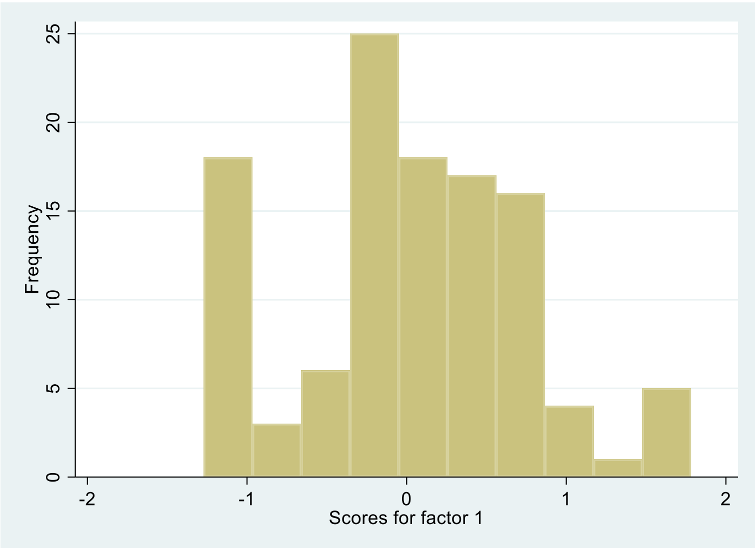
This is not ideally normal but there is sufficient variation in the scores.

Scale 2 has the following distribution of the factor score, performing similarly to Scale 1:

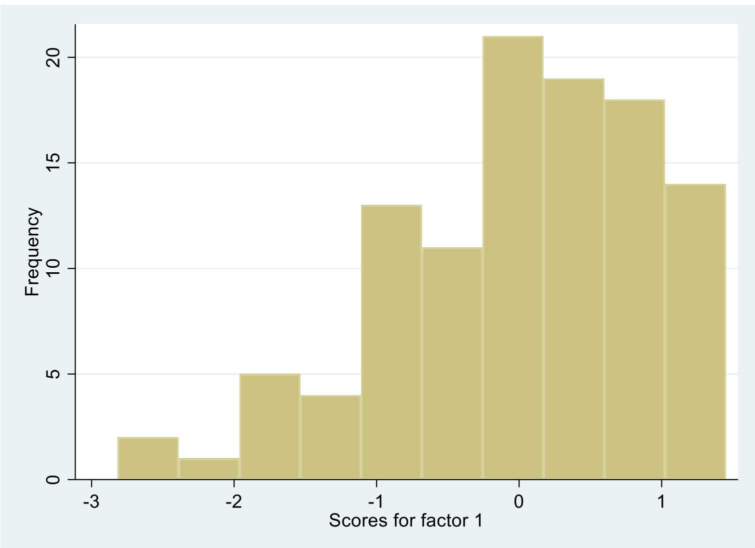
[46] We acknowledge, however, that reliability coefficients are influenced by the number of questions in a scale, so with 3 questions only, Scale 3 reliability may be downwardly biased.



Scale 3 has a better distribution, most closely approximating a normal distribution:



Finally, Scale 4 displays a distribution broadly similar, but slightly more normal, than Scales 1 and 2.



Convergent validity

To further understand the relative performance of these scales, we explored the manner in which they are related to each other.

To understand if the four scales relate to each other as expected, we computed factor scores following each factor analysis originally undertaken for each scale. These factor scores provide a continuous measure (in a single variable) for each scale that considers the relative contribution of each question to its corresponding scale. Scale 3 was generated from the factor analysis as a negative score, where higher values indicate a more negative attitude to science, in line with the fact that two of three questions for Scale 3 were negative (e.g. question 1: "I find science difficult to understand").

We then explored the pairwise linear correlations between these factor scores. Linear correlations above 0.9 are considered very high, and we also consider correlations above 0.6 as high. Correlations between 0.3 and below 0.6 are moderate, with anything below 0.3 considered low.

We found a high positive correlation (r=0.90) between Scale 1 and Scale 2, suggesting enjoyment and interest in science are highly linked to each other; and moderate correlations between Scale 1, and 3 (r=-0.55) and 4 (r=0.55). This suggests that high enjoyment of science does not fully map onto science self-efficacy, or pupils' experiences of inquiry-based science learning.

These latter associations were matched for Scale 2 as well, with moderate correlations between Scale 2, and 3 (-0.52) and 4 (0.56). A above, this suggests that interest in sciences only partially maps onto self-efficacy and inquiry-based science learning.

Finally, Scale 3 and Scale 4 were negatively and moderately correlated (r=-0.38).

Conclusions and recommendations

Drawing together information about scale reliabilities and underlying factor structures, pupils' reported ease of understanding and ease of answering questions, and the relationships between the different scales, combined with the guiding principles for selecting the scale for later use whereby we prioritise comparability with previous studies and retain the relevance to PSQM's theory of change, we do not believe that any of the scales in their original configuration are suitable for use as a secondary outcome measure.

However, we then explored the possibility of compiling scale from all existing questions, using factor analysis.

First, we explored the possibility that *all* above questions would provide better results than the scale structure above. The aim of this analysis was to understand if a different model of science "attitudes" was desirable over the four different scales. We undertook an exploratory factor analysis using all scale questions asked of pupils. This analysis suggested that there either 4 or 5 underlying factors to the whole set of questions. The 4-factor solution was preferred upon further investigation. These 4 factors only somewhat reflected the original 4-scale design of the questionnaire, but in a manner that did not suggest this was a better solution, for reasons that follow.

From this analysis, a majority of Scale 1 questions and a majority of Scale 2 questions contributed to a single factor, with the exception of questions 7 and 8 from Scale 1 (both 'real-life focused), and questions 8 and 9 from Scale 2 (which did not substantially contribute to any factor). This factor also saw substantial contributions from question 2 of Scale 3 ("I am just not good at science"), and question 11 of Scale 4 ("I do science investigations outside of school"). There does not appear to be a coherent conceptual underpinning to all these questions clustering together that would supersede the original scale design, as the questions are too diverse in focus.

The real-life questions from Scale 1 (7 & 8) contributed to a separate factor alongside questions 2, 3, 7 and 9 from Scale 4, also focused on "how things work", "everyday life", and "science investigations". This suggests that this new factor is essentially capturing pupils' responses to how science offers lessons about real world phenomena, separately from how they feel about science.

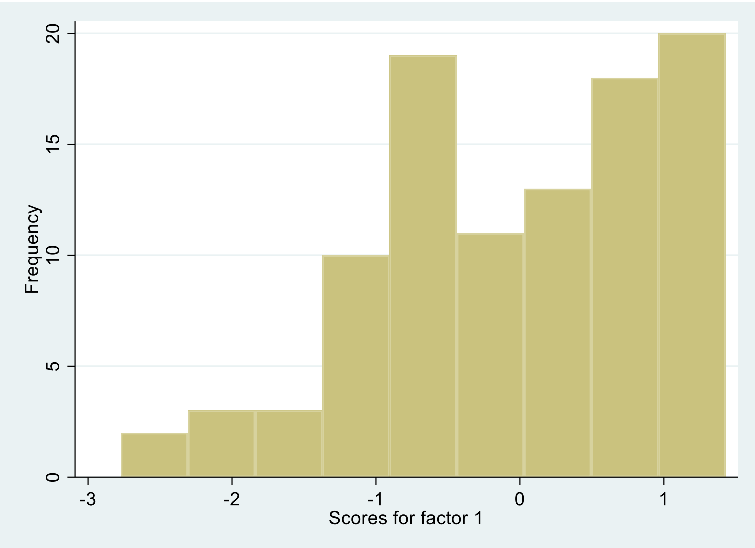
A third factor was made up of questions from across Scales 2 and 4 that referenced the opportunity for pupils to explain/work independently (e.g. question 1 Scale 4, “ In science lessons I am given opportunities to find the answers to science questions myself”), providing some support to an overall internally consistent factor.

The fourth factor resulting from this analysis encompassed only the first and third question of Scale 3. This would result in too short of scale for actual use.

We therefore took the step of only including in a separate factor analysis the questions that made substantial contributions to any of the above factors (i.e. that had high factor loadings). This led to the following questions being included: all Scale 1 questions; questions 1-7 and 11 and 12 from Scale 2; none of the questions from Scale 3; and questions 2,3,4,7,11, and 12 from Scale 4.

The factor analysis suggested that the underlying structure for these questions was as follows: one very strong factor, capturing over 70% of variance across all questions; and a second factor, much weaker than the first capturing an additional 8% of variance, to which only questions from Scale 4 contributed, but in a manner that reflected these questions’ contribution to the first factor. In other words, the contributions of these questions (2, 3, 4, 7, and 8 from Scale 4) were not substantially different between the two factors, with questions 3, 4, and 8 actually contributing more to the first, overall stronger, factor.

We therefore retained one single factor, and generated a factor score, as per the previous approach. The distribution of this overall score is:



This is not perfectly normally distributed, but the distribution better approximates a normal distribution than any of the Scales 1, 2, or 4 individually.

We then checked how this new scale related to Scales 1, 2, and 4, by computing linear correlations between the respective factor scores, and concluded that the new scale is strongly associated ($r>0.95$) to Scale 1 and 2 respectively, and moderately strongly ($r>0.6$) with Scale 4.

Finally, to understand the internal consistency of the scale, we computed Cronbach’s alpha as a reliability coefficient for the above questions, resulting in a reliability coefficient of 0.94. This is very high, suggesting that there may be some degree of overlap between questions and potentially that removing some questions would not negatively impact the reliability of the scale. However, since the aim of this exercise is not data reduction, we suggest retaining all questions in this scale. We would also be in a position to explore, outside of the remit of the evaluation, aspects of response consistency, with question 3 from Scale 1 and question 5 from Scale 2 opposites of each other.

Therefore, our recommendations is to use the new composite scale as a secondary measure in the PSQM evaluation, given the fact that: it has high reliability; it displays a factor structure that allows it to be analysed as one single outcome measure; it includes questions that would allow for question-by-question comparison with other surveys and international data sources; and importantly, it aligns with both theoretical and PSQM-related theory of change considerations by including questions on science enjoyment, science confidence, and inquiry-based science learning.

We believe that this scale would be answerable by pupils within a 15-minute time period.

Therefore, the scale we propose to use would be:

New scale	
1	I enjoy learning science
2	I wish I did not have to study science
3	Science is boring
4	I learn many interesting things in science
5	I like science
6	I look forward to learning science in school
7	Science teaches me how things in the world work
8	I like to do science investigations
9	Science is one of my favourite subjects
10	Science lessons make me think
11	I look forward to my science lessons
12	Science lessons are interesting
13	I would like to do more science at school
14	Science is fun
15	I enjoy discussions in science lessons
16	We do a lot of writing in science lessons
17	It is important that we learn science
18	In science lessons I am given opportunities to find out how things work
19	The science we do in school has links to my everyday life
20	In science lessons I am given opportunities to ask science questions which I can investigate myself
21	In science lessons I am given opportunities to explain what I notice in science investigations
22	In science lessons I am given opportunities to use scientific vocabulary to explain what I have learned
23	I do science investigations outside of school

Response options disagree a lot - disagree a little - agree a little - agree a lot

Appendix J: Thematic coding for interview analysis

Interviews with Hub Leaders

- Overall experience delivering PSQM
- Preparation to become Hub Leader
- Ongoing support
- Experience working with schools
- Experience working with SSLs
- Experience using PSQM evaluative criteria
- Key enablers and challenges
- Changes in science teaching and practice in schools
- Improving implementation of PSQM
- Sustaining PSQM changes
- COVID-19 impact

Interviews with Headteachers

Background
- Motivation of schools to take part in PSQM
- Previous participation in science programmes

Experience with PSQM
- Overall experience
- Experience of the training
- Support, activities and resources

Impact of PSQM at school
- Science teaching
- SSL practice in leading science
- Extent to which changes will be sustained

Implementation of PSQM
- Key enabler and challenges
- Costs
- Suggestions for improving implementation

Science and CPD activities
- Science CPD in last two years (apart from PSQM-related CPD)

Impact of COVID-19 on science teaching and implementation of PSQM

Interviews with SSLs

Background
- Qualifications of SSL

Experience with PSQM
- Overall experience
- Experience of the training
- Support, activities and resources

Impact of PSQM at school
- Science teaching
- SSL practice in leading science
- Changes in students
- Extent to which changes will be sustained

Implementation of PSQM
- Experience of implementing
- Key enablers and challenges
- Costs
- Suggestions for improving implementation

Science and CPD activities

- Science CPD in last two years (apart from PSQM-related CPD)

Impact of COVID-19 on science teaching and implementation of PSQM

Interviews with teachers

Experience with PSQM

- Overall experience

Impact of PSQM at school

- Science teaching
- Changes in students
- Extent to which changes will be sustained

Implementation of PSQM

- Key enablers and challenges

Impact of COVID-19 on science teaching and implementation of PSQM

Education Endowment Foundation

The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

https://educationendowmentfoundation.org.uk

@EducEndowFoundn

Facebook.com/EducEndowFoundn