

Study Plan for Pupil Premium statement research

June 2025



PROJECT TITLE	Pupil Premium Statement Research (2024/25 Academic Year)
RESEARCH ORGANISATION	Behavioural Practice, Verian UK
NAMED RESEARCHER(S)	Natalie Gold, Luke Taylor, Adam Ozer, Alex Thornton, James Thom, Sarah Bowen
TYPES OF SETTING(S)	Primary, Secondary, Special schools
TYPE OF RESEARCH PARTICIPANTS	Schools

Study plan version history

VERSION	DATE	REASON FOR REVISION
1	June 2025	

Contents

Introduction	3
Aims and Objectives	3
Background	3
Research Questions	4
Approach and methodology	5
Research question rationale	5
Sampling and data collection	7
Sample population	7
Sample design	9
Data collection	10
Data processing	11
Coding framework	11
Data extraction	12
Merging with other datasets	13
Design weighting	14
Analysis	14
Reporting	16
Timeline	16
Ethical considerations and data protection	17
Risks and mitigations	18
Project team	22

Introduction

Aims and Objectives

The EEF has commissioned Verian to undertake this research to better understand how schools are spending the Pupil Premium. The results will help the EEF and policy makers to better support schools using Pupil Premium funding to improve outcomes for disadvantaged pupils.

Background

The Pupil Premium grant is additional government funding provided to publicly-funded schools serving reception to Year 11 in England to improve educational outcomes for disadvantaged pupils.¹ In FY2023–24, the total value of the Pupil Premium was over £2.9 billion, and was allocated to eligible schools based on the number of:

- pupils who are recorded as eligible for free school meals, or have been recorded as eligible in the past 6 years (referred to as Ever 6 FSM)
- children previously looked after by a local authority or other state care, including children adopted from state care or equivalent from outside England and Wales

An evaluation of the Pupil Premium was conducted in 2013,² and more recently in 2024.³ The 2024 evaluation found that schools view the premium as an important source of funding when planning support for disadvantaged pupils. Around four in five schools (82%) reported that they would have to cut some forms of support they currently offered to disadvantaged pupils if they did not have Pupil Premium funding.

To ensure that Pupil Premium is focused on effective approaches to raising the educational attainment of disadvantaged pupils, schools must spend their Pupil Premium grant on evidence-informed activities. These activities should be in line with the [‘menu of approaches’](#) set by the Department for Education (DfE) and the EEF’s 3-tiered approach, which helps schools allocate spending across the following areas:

- developing high-quality teaching, for example through professional development and recruitment and retention
- providing targeted academic support, such as one-to-one or small group tuition
- tackling non-academic barriers to academic success, such as difficulties with attendance, behaviour and social and emotional wellbeing

Each year, schools are required to publish a Pupil Premium strategy statement on their website setting out how they plan to spend the Pupil Premium allocation. Statements are expected to follow the DfE’s [template](#), and provide information on:

- the school

¹ Schools eligible for the Pupil Premium include local authority maintained schools and academies, free schools and non-maintained special schools (including alternative provision academies). See <https://www.gov.uk/government/publications/pupil-premium/pupil-premium#funding-criteria> for more information.

² <https://assets.publishing.service.gov.uk/media/5a7c8811ed915d48c2410633/DFE-RR282.pdf>

³ The evaluation also covered perceptions of Recovery Premium funding, which was a time-limited grant as part of the government’s package of funding to support pupils whose education was impacted by COVID-19

https://assets.publishing.service.gov.uk/media/67dd5b7f0114b0b86e59f42b/Pupil_premium_and_recovery_premium_evaluation.pdf

- the funding allocation
- challenges to achievement identified among the school's disadvantaged pupils
- intended outcomes and success criteria
- planned and costed approaches to spending
- the evidence used to support the choice of approaches

The EEF has also developed a range of resources to support schools in using Pupil Premium funding effectively, including:

- the [EEF guide to the Pupil Premium](#)
- an [evidence brief](#)
- [discussion prompts](#) for governors and trustees
- a [school's guide to implementation](#)

Previous research has examined how schools use Pupil Premium funding. The 2024 evaluation collected data from school staff using surveys and interviews. The results showed that schools tend to spend the Pupil Premium funding in line with DfE's published menu of approaches and distribute funding evenly between the EEF's 3-tiers.⁴

The Sutton Trust conducts an annual survey of teachers and school leaders to understand how they are spending budgets, including Pupil Premium.⁵ From the most recent figures from 2025, 46% of senior leaders reported using the funding to cover gaps in their school's budget, particularly for teaching assistant salaries (79%), staff salaries (46%), academic interventions (e.g. tutoring) for non-Pupil Premium pupils (41%), and extracurricular activities/trips for non-Pupil Premium pupils (33%).

Pupil Premium statements are a valuable source of information to understand how schools are using funding. However, while schools are expected to publish Pupil Premium statements for each academic year publicly, Pupil Premium statements are not centrally collected.⁶ Therefore, there is an incomplete picture (beyond the contributions of previous research outlined above) as to how schools allocate their Pupil Premium funding each year. This research will build on this evidence base by collating and analysing information from a sample of Pupil Premium statements.

This research will run twice, first for a sample of statements published for the 24/25 academic year and then again for a sample published for the 25/26 academic year.⁷ This study plan will be updated for the second research cycle.

Research Questions

This research will address the following research questions, see Table 1.

4

https://assets.publishing.service.gov.uk/media/67dd5b7f0114b0b86e59f42b/Pupil_premium_and_recovery_premium_evaluation.pdf

⁵ <https://www.suttontrust.com/wp-content/uploads/2025/04/School-Funding-and-Pupil-Premium-2025.pdf>

⁶ It is also not currently known how actual spend aligns with planned spend in the Pupil Premium statements. However, addressing this knowledge gap is out of scope for this project.

⁷ The project is scheduled to run for two academic years, with the option to extend it in future years.

Table 1. Research questions

Research questions	
1	What challenges do schools identify for pupils eligible for the Pupil Premium? How does this vary by school characteristic?
2	Which approaches are schools funding using the Pupil Premium? How does this vary by school characteristic?
3	How do schools' approaches to spending align with their identified challenges?
4	How does school spending align with the EEF's tiered approach? How does this vary by school characteristic?
5	What types of evidence are being cited in Pupil Premium statements? How does this vary by school characteristic?
6	Which external programmes are schools funding with their Pupil Premium grant? How does this align with the EEF's list of promising programmes? How does this vary by school characteristic?

Approach and methodology

Research question rationale

The rationale for each research question, and the approaches taken to address them are outlined below:

1. What challenges do schools identify for pupils eligible for the Pupil Premium? How does this vary by school characteristic?

The first step in The EEF Guide to the Pupil Premium is to identify challenges faced by the school's disadvantaged pupils. Schools are encouraged to draw on a range of data sources to identify these challenges and consider the underlying academic and non-academic causes of them. We will explore which challenges schools identify to understand which challenges are the most and least prevalent. See the **coding framework** section (page [11](#)) for more detail about how we will create a typology of challenges.

Challenges facing Pupil Premium eligible cohorts may vary between schools. For example, schools may face different challenges depending on the:

- school's local context, including their geographic location (region)
- demographics of their student population (percentage of cohort eligible for the Pupil Premium)
- how many pupils are eligible for the Pupil Premium (number of pupils eligible for the Pupil Premium)
- age and needs of pupils (phase)
- school's staffing capacity, resources, and funding structures (school size, school type)
- attainment of disadvantaged pupils (attainment of disadvantaged pupils in KS2 in the previous academic year, attainment of disadvantaged pupils in KS4 in the previous academic year)

To explore this, we will examine the number and percentage of schools facing challenges by region, percentage of cohort eligible for the Pupil Premium, number of pupils eligible for the Pupil Premium/allocation amount, phase, school size, school type, and the attainment of disadvantaged pupils in the previous academic year at KS2 and KS4.

2. Which approaches are schools funding using the Pupil Premium? How does this vary by school characteristic?

A key objective of this research is to understand which approaches schools are funding using their Pupil Premium funding. To explore this, we will examine how many schools plan to use each approach, based on the DfE's menu of approaches.

As with challenges, the approaches funded by the Pupil Premium may vary between schools. To explore this, we will examine the number and percentage of schools citing approaches by region, percentage of cohort eligible for the Pupil Premium, number of pupils eligible for the Pupil Premium/allocation amount, phase, school size, school type, and the attainment of disadvantaged pupils in the previous academic year at KS2 and KS4.

3. How do schools' approaches to spending align with their identified challenges?

In the Pupil Premium statement template, schools can indicate which challenges are addressed by each approach. This will allow us to explore which approaches schools use to target their challenges and improve outcomes.

4. How does school spending align with the EEF's tiered approach? How does this vary by school characteristic?

The DfE Pupil Premium strategy statement template organises approaches and spend to address challenges into three tiers.⁸ The tiered model aligns with DfE's menu of approaches and helps schools balance spend across each tier to create a balanced Pupil Premium plan. To explore how different schools allocate their Pupil Premium funds across the tiers, we will investigate what proportion of spend is being allocated to: 1) developing high-quality teaching, 2) providing targeted academic support, and 3) tackling non-academic barriers to academic success.

As with challenges and approaches, we will look at the proportion of spend across tiers by region, percentage of cohort eligible for the Pupil Premium, number of pupils eligible for the Pupil Premium, phase, school size, school type and the attainment of disadvantaged pupils in the previous academic year at KS2 and KS4.

5. What types of evidence are being cited in Pupil Premium statements? How does this vary by school characteristic?

We are interested in whether schools cite any research evidence in their Pupil Premium statements, and if they do cite evidence, from which sources. In particular, what types of DfE and EEF evidence sources schools are citing to support their Pupil Premium statements. There is also an interest in understanding what other types of evidence schools are citing (e.g., books, anecdotal evidence, government reports, blog posts).

It is possible that different types of schools cite different types of research evidence to support their Pupil Premium statements. Although we have no strong prior hypotheses as to why this may differ and for which types of schools, we will investigate whether the types of evidence cited varies by region, percentage of cohort eligible for the Pupil Premium, number of pupils eligible for the Pupil Premium/allocation amount, phase, school size, school type, and the attainment of disadvantaged

⁸ The three tiers as set out in the EEF's guide to the Pupil Premium (available here: <https://d2tic4wvo1iusb.cloudfront.net/production/documents/guidance-for-teachers/pupil-premium/guide-to-the-pupil-premium-2024.pdf?v=1727884053>) are 1) developing high-quality teaching, for example through professional development and recruitment and retention, 2) providing targeted academic support, such as one-to-one or small group tuition, and 3) tackling non-academic barriers to academic success, such as difficulties with attendance, behaviour and social and emotional wellbeing

pupils in the previous academic year at KS2 and KS4. We will report on all analysis run in the final report even if there are no trends of interest in the data.

6. Which external programmes are schools funding with their Pupil Premium grant? How does this align with the EEF's list of promising programmes? How does this vary by school characteristic?

Schools may choose to allocate Pupil Premium funding to external programmes. To explore this we will examine the proportion of schools who use their Pupil Premium to fund programmes on the EEF's [list of promising programmes](#). We will also explore which external programmes on the EEF's list of promising programmes are being funded by different schools. The list of EEF's promising programmes for raising attainment differs depending on phase of education. Therefore, we will examine how the funding of EEF promising programmes varies by phase.

We will also investigate which external programmes are being funded by region, percentage of cohort eligible for the Pupil Premium, number of pupils eligible for the Pupil Premium/allocation amount, school size, school type, and the attainment of disadvantaged pupils in the previous academic year at KS2 and KS4. We have no strong prior hypotheses as to why funding external programmes may differ across these school characteristics. However, we will report on all analysis run in the final report even if there are no trends of interest in the data.

Sampling and data collection

Sample population

We will begin by compiling a **database that contains all schools that are in scope for the research**. We propose to use a similar approach to the one that we successfully implemented for the recent DfE Pupil Premium and Recovery Premium Evaluation.⁹

We will first identify the schools that have received Pupil Premium funding for the 2024/25 academic year using the published **DfE Pupil Premium allocation files**. DfE allocations are made on the basis of financial years rather than academic years, and payments are made quarterly. As a result, to identify all schools that obtained funding during the 2024/25 academic year we will combine four DfE Pupil Premium allocation files from two different financial years:

- **2024-25 financial year:** [Pupil premium: allocations and conditions of grant 2024 to 2025 - GOV.UK](#), three files:
 1. Pupil premium allocations 2024 to 2025 financial year (September 2024)
 2. Pupil premium allocations 2024 to 2025 financial year (December 2024)
 3. Pupil premium allocations 2024 to 2025 financial year (March 2025)
- **2025-26 financial year:** [Pupil premium: allocations and conditions of grant 2025 to 2026 - GOV.UK](#), one file:
 4. Pupil premium allocations 2025 to 2026 financial year (March 2025)
This file consists of indicative allocations for each school at the start of the 2025/26 financial year. Actual allocation data will not be published until later this year (too late to be used when defining and drawing the sample).

⁹

https://assets.publishing.service.gov.uk/media/67dd5b7f0114b0b86e59f42b/Pupil_premium_and_recovery_premium_evaluation.pdf

Schools (based on their DfE Unique Reference Number (URN)) that appear in **any of these four files** will then be included in our master list of schools in scope for the study – as they will have received funding for at least some of the 2024/25 academic year.

Using the school URN we will then augment the sample frame by merging in additional variables from the DfE [Get Information About Schools](#) (GIAS) database (an extract downloaded on 1/5/2025). This includes variables which will be useful to define the population (e.g. school type, whether schools are closed), to structure the sample design (e.g., region), to find Pupil Premium statements (the URL of each school's website), and for analysis (e.g., total number of pupils at each school).

Next, we will put in place a number of **exclusions to remove schools that are not in scope** for the study:

- **De-duplication:** Some schools converted to an academy during the 2024/25 academic year, and DfE assigned them a new URN after the conversion. As a result, schools that changed status during the year can appear multiple times in the DfE Pupil Premium allocation files (under the old URN pre conversion, and under the new URN after the conversion). To avoid double counting these schools in the population (and potentially sampling the same school twice) we will exclude the old URN (where “Academy Converter” is given as the reason why an establishment is closed in GIAS) from the sample frame and only retain a record for the new URN.
- **Closed schools:** The GIAS establishment status field has been used to identify schools that are now closed. We have decided to exclude these from the study, as they form a very small part of the population (<0.1%), and as they are now closed their website may now be offline and Pupil Premium statements therefore unavailable.
- **All through schools:** For practical reasons, we will exclude “all through” schools (<1% of all schools in scope). The rationale for this is as follows:
 - They are a very distinctive group and cannot be simply allocated to a phase (primary or secondary)¹⁰ for the analysis.
 - They are too few of these schools to enable standalone analysis, unless we disproportionately sampled a large proportion of them (which would affect the precision of other estimates e.g. by region and phase).
- Please note that middle schools are in scope for the research, as DfE formally allocate all middle schools to a phase in GIAS (either “Middle deemed primary”, or “Middle deemed secondary”).
- **Pupil Referral Units (PRUs) and other Alternative Provision (AP):** We will also exclude these establishments. These only make up a small minority of schools that receive Pupil Premium funding (AP colleges c.0.5%, and PRUs a further c.0.5%). As a result, they are both too few in number for standalone analysis to be possible without substantial boosts (which would negatively affect all other analysis).
- **Schools with 5 or fewer FSM-eligible pupils:** Under DfE's conditions of grant for Pupil Premium, schools with five or fewer eligible pupils do not have to publish a Pupil Premium statement. As a result, schools that have fewer than six eligible pupils in **all four** of the DfE allocation files used to compile the sample frame have been excluded from the study.

¹⁰ Allocation to primary or secondary could theoretically be made on the basis of the year groups covered by each school. Although in some instances they may cover an equal number of primary and secondary year groups – in which case it would not be clear where they should be assigned.

Schools are in scope if they had at least six eligible pupils in at least one of the allocation files.

Once these exclusions have been made, we will be left with a master list of schools that received Pupil Premium funding for the 2024/2025 academic year and that are in scope for the study.

It should be noted that Special Schools will be covered by the study. Special schools will be reported on separately from other schools (e.g., when reporting by phase), ensuring that consistent comparisons can be made between the new 2024/25 data and data from previous years.

Sample design

Based on the profile of schools that are in scope for the study (using the process outlined above), we recommend a disproportionate design, boosting low-incidence population sub-groups of analytical interest:

- There is substantial variation in the number of in-scope schools in each region (ITL1 – former Government Office Regions). Most notably, schools in the North East only comprise 5% of the total. We therefore recommend implementing a disproportionate design with an equal number of schools sampled from each region.
- Schools in receipt of Pupil Premium funding are heavily skewed in terms of phase (reflecting the wider profile of schools – where primary schools are more numerous as they tend to have fewer pupils). Over three-quarters (78.8%) are primary, 15.9% are secondary, and 5.3% are not allocated to a phase (Special Schools).¹¹ We therefore recommend boosting the representation of secondary schools and Special Schools.

We do not feel that there is a need to vary sampling probabilities by **school type** or by **Free School Meals eligibility**:

- For analysis by school type, the population distribution is such that oversampling will not be required to allow for standalone analysis of mainstream academies and mainstream LA maintained schools.
- We recommend clustering schools into quartiles or quintiles based on FSM eligibility, as this will generate categories with sufficiently large sample sizes for robust analysis.

Based on our analysis of the population profile, **we recommend an overall sample size of 1,395 schools**. Within this overall total, we propose to construct 27 strata based on region (nine - ITL1 regions) and phase (three - primary, secondary, N/A (Special Schools)). **We will sample 155 schools from each region: 50 secondary schools, 90 primary schools and 15 Special Schools.**

We propose to draw a **stratified random sample**. Within each stratum we propose to sort the sampling frame by a range of relevant variables: school size (bands based on number of pupils), FSM eligibility (banded), school type, local authority. We will then draw a systematic random sample of schools (varying the sampling probability in each stratum based on the agreed design). The implicit stratification will help ensure that the selected sample is representative of the population in respect to the factors included in the sorting.

This design allows for robust estimates nationally and for key sub-groups. Table 2 shows our estimates for effective sample size (neff) and the precision of estimates (the maximum 95% Confidence Intervals) overall and for key analytical sub-groups (after applying design weighting to compensate for the disproportionate sample design). We have also included figures to show indicative sample sizes allowing for 8% of schools to not have published a statement (based on EEF insights on this topic).

¹¹ DfE do not allocate special schools to a phase, and they will therefore be treated as a standalone category.

Table 2. Estimates for effective sample size (*neff*) and the precision of estimates overall and for key analytical sub-groups.

	Analysis based on all schools			Analysis based on those that publish a statement (92%)		
	n	Neff	Max 95% CIs	n	neff	Max 95% CIs
Overall	1395	c.1110	±2.9%pts	c.1283	c.1020	±3.1%pts
Phase						
Mainstream Primary	810	c.760	±3.5%pts	c.745	c.700	±3.7%pts
Mainstream Secondary	450	c.420	±4.8%pts	c.414	c.390	±5.0%pts
Special schools (not allocated to a phase by DfE)	135	c.120	±8.9%pts	c.124	c.110	±9.2%pts
Region						
East Midlands	155	c.130	±8.7%pts	c.143	c.120	±9.1%pts
East of England	155	c.130	±8.6%pts	c.143	c.120	±9.0%pts
London	155	c.140	±8.3%pts	c.143	c.130	±8.6%pts
North East	155	c.130	±8.5%pts	c.143	c.120	±8.9%pts
North West	155	c.130	±8.5%pts	c.143	c.120	±8.9%pts
South East	155	c.130	±8.5%pts	c.143	c.120	±8.9%pts
South West	155	c.130	±8.7%pts	c.143	c.120	±9.0%pts
West Midlands	155	c.140	±8.4%pts	c.143	c.130	±8.7%pts
Yorkshire and the Humber	155	c.130	±8.7%pts	c.143	c.120	±9.1%pts
Type						
Mainstream academies & free schools	c.768	c.600	±4.0%pts	c.704	c.550	±4.2%pts
Mainstream LA maintained	c.492	c.440	±4.7%pts	c.453	c.400	±4.9%pts
Special schools	135	c.120	±8.9%pts	c.124	c.110	±9.2%pts
FSM % (quintiles)						
1 (lowest proportion)	c.225	c.200	±7.0%pts	c.207	c.180	±7.3%pts
2	c.257	c.210	±6.7%pts	c.236	c.200	±7.0%pts
3	c.283	c.220	±6.6%pts	c.260	c.200	±6.9%pts
4	c.300	c.220	±6.6%pts	c.276	c.200	±6.9%pts
5 (highest proportion)	c.330	c.200	±6.9%pts	c.304	c.190	±7.2%pts

Data collection

When the sample will be drawn

We will draw the sample once this study plan has been fully approved. Based on the current timetable, we anticipate that we will begin drawing the sample starting the week of June 2nd. We will use the latest published [DfE Pupil Premium allocation file](#) (currently March 2025) as the base for the sample frame, as this is limited to schools that are in receipt of Pupil Premium funding. Using the school URN, we will then augment the sample frame. First, we will merge in relevant fields from the DfE Get [Information About Schools \(GIAS\)](#) database. This includes variables useful for structuring the sample design (e.g., region, detailed school type), to find Pupil Premium statements (the URL of each school's website) and for analysis (e.g., total number of pupils at each school). Next, we will merge in DfE [school performance data](#) pertaining to the academic performance of disadvantaged pupils at KS2 and KS4 in the previous academic year.

How the data will be collected

To minimise manual data collection, we propose to begin with a stage of automated searches using the Google search API and web scraping. Based on previous efforts to collect this data for DfE, we anticipate that the proportion of statements found automatically will be relatively low. This step is still worth including as it will reduce the manual burden. We propose to first use the Google search API to look for a pdf including the school's name, the school website (where available in GIAS), and the term ("Pupil Premium"). Where the Google API returns a pdf, we will download the document directly. Where the Google API returns a web link (to a page on the school website, but not a pdf), we will then scrape the URL for any links that include "pupil", "premium", and "pdf", and download any relevant files which are found. Any pdf files downloaded as part of this automated process will then be manually reviewed by a junior researcher to ensure that the file is correct.

Following this, manual searches will be conducted for all schools where we did not find the correct Pupil Premium statement during the automated phase.

Data processing

Coding framework

We will work closely with the EEF to develop and refine the coding framework to address the research questions. The coding framework variables will include challenges identified, approaches, evidence and guidance cited, and external programmes. When available, we will use DfE and the EEF guidance as a source for codes. For example, for approaches we will use the menu of approaches cited in the [guidance for school leaders](#) and for external programmes we will use the EEF's [list of promising programmes](#).

To refine the coding framework further, two researchers will extract from 20 statements (10 each) selected purposively at random, to include a range of schools representing all key variables.¹² We will agree a list of relevant quotas with the EEF, but expect they would include variables like region, phase, and Pupil Premium cohort %. We will then refine the codes for all variables. We can also update the variables themselves if feedback from the data extractors suggests that additional or amended variables would be helpful.

The approach to developing coding categories for each variable is summarised in Table 3.

Table 3. Approach to developing the coding framework for each variable

Variable	Approach to developing coding categories	Research Question(s)
Challenges identified	We will agree an initial list of codes with the EEF. Then, we will conduct a thematic analysis of the data extracted from 20 statements to refine the codes. We anticipate that the total number of codes will range between 15 to 25.	1, 3

¹² In our experience, 20 statements should be enough to reach data saturation (when no new themes are emerging). However, if after having coded 20 statements, we are still finding new themes coming up, or if the experts on our team think the 20 statements have not covered all the types of content that schools write in Pupil Premium statements, then we will draw another sample of 10 statements at a time, and continue extracting until we reach data saturation.

Variable	Approach to developing coding categories	Research Question(s)
Approaches	We will begin with the ‘menu of approaches’ cited in the guidance for school leaders , that sits within the 3-tiers, to code approaches. We will add additional codes based on our experience of approaches named in Pupil Premium statements. We will then work with the EEF to add additional codes based on a thematic analysis of 20 statements.	2, 3
Evidence and guidance cited	We will work with the EEF to develop an initial list of evidence and guidance codes, based on school Pupil Premium resources. We will then add any other codes based on a thematic analysis of 20 statements.	5
External programmes	We will begin with the EEF’s list of promising programmes , and add any other programmes cited in the 20 statements used to refine the coding framework.	6

Data extraction

Our in-house Data Science team will oversee data extraction and coding. We will complete the data extraction and coding processes using a bespoke Python script and a GPT-4o model deployed from our Microsoft Azure data platform.

Once the sampling process is complete and all statements have been collected, we will use a Python script to upload the text from each school’s Pupil Premium Statement into a single data frame. Each row of the data frame will represent a single Pupil Premium statement, with columns containing the text of the statement, the school’s name, and the school’s unique ID number.

Afterwards, we will use an additional Python script to clean and process the text from each statement to prepare it for use in quantitative and qualitative analyses. Leveraging a GPT-4o that has been deployed from our secure data platform, researchers will sift through the text of each Pupil Premium statement and identify text that is relevant to specific queries. Using the agreed-upon variable coding framework as a guide, we will carefully design prompts (instructions for the AI model), that will allow researchers to create binary coded columns based on specific queries (e.g. mark as “1” if it mentions increased high quality teaching for students from disadvantaged backgrounds, “0” if it does not). The researchers will also use the AI model to extract text relevant to the query, allowing one to capture and check the text that the AI model has flagged. The resulting dataset contains both the relevant text and the coding of the text (or quantitative data). In our experience, this requires minimal data cleaning.

We propose two methods of assessing intercoder reliability to ensure the reliability of AI text extraction and coding.

1. Researchers will manually identify texts that mention specific topics using basic keyword searches. We will then feed those texts to the AI model and examine the output to ensure that the AI is not misidentifying or excluding excerpts we are certain are relevant to the topic. This is useful for conducting AI prompt engineering and identifying potential false negative results.
2. We will have the AI code the same 20 statements as our researchers and compare the results using Cohen’s Kappa. We suggest a success threshold for intercoder reliability at 90% (as we use for researchers coding the Crime Survey for England and Wales).

If intercoder reliability for either human-AI pair falls below 90%, we will take the following steps to identify issues and improve performance:

1. We will review data for potential errors in text formatting that could prevent the AI model from reading the text correctly.
2. We will review our AI prompt engineering to identify, adjust, and retest specific instructions or queries that may be producing suboptimal results.
3. We will feed the original intercoder reliability test texts to the newly modified AI to determine if its performance has improved.
4. We will select a new sample of texts for both the human coders and AI to process to ensure that the improvements are robust, as measured by Cohen's Kappa > 90%.

If necessary, we will repeat these steps until we have sufficient interrater reliability. We will directly compare discrepancies in the text flagged by the human researcher and the AI model. If a large portion of discrepancies appear to be driven by the AI model's output and not human error, our Data Science team will conduct prompt engineering tweaks to improve the AI model's accuracy before repeating the intercoder reliability process.

Merging with other datasets

We will merge variables from a range of other datasets into the data frame compiled from the Pupil Premium statements to maximise analytical potential. This information will be merged on using the school URN.

The data sources we propose to use are:

- DfE Pupil Premium Allocation files¹³ (the same files as used to compile the sample frame) – the total number of pupils, the number of pupils eligible for the Pupil Premium, and the Pupil Premium allocation amount (£).
- DfE Get Information About Schools database – phase, school type, region, school size (number of pupils)¹⁴, etc.
- DfE 2023/24¹⁵ performance data pertaining to the academic performance of disadvantaged pupils at KS2 and KS4:
 - **Key Stage 2** – for example:
 - Percentage of disadvantaged pupils reaching the expected standard in reading, writing and maths
 - Percentage of disadvantaged pupils achieving a high score in reading and maths and working at greater depth in writing
 - **Key Stage 4** – for example:
 - Average Attainment 8 score per disadvantaged pupil

¹³ It is preferable to use the allocation files for %FSM and number of pupils – as it is up to date and relates to the 2024/25 academic year. The information currently available in GIAS relates to the January 2024 school census (i.e. the previous academic year). The GIAS FAQs indicate that GIAS will be updated with January 2025 school census data in July / August 2025 – if this is available in time we could use GIAS instead.

¹⁴ As it stands GIAS has data on: (i) number of pupils, and (ii) number of pupils eligible for FSM relate to the January 2024 school census (i.e. last academic year). DfE anticipate that this data will be updated in July or August 2025 to reflect the January 2025 school census.

¹⁵ 2024/25 performance data will not be available in time for the analysis – as this will only be published towards the end of 2025.

- Adjusted Progress 8 measure - disadvantaged pupils.
- % of disadvantaged pupils achieving standard 9-4 passes in GCSE English and maths
- % of disadvantaged pupils achieving strong 9-5 passes in GCSE English and maths

Design weighting

We will draw a disproportionate sample of schools to enable robust analysis by region and phase/type. That means schools in smaller regions, secondary schools and Special Schools will be over-represented in the full sample. To compensate, we will calculate a design weight for each school – the inverse of its sampling probability – and report weighted estimates.

The only information required to produce design weights is the sampling probability for each school, which should be available as soon as the sample is drawn. We will therefore produce the design weights once the study sample is agreed and drawn.

Analysis

We will produce weighted descriptive statistics for key outcomes with robust confidence intervals adjusted to take the weighting design into account. This will include the frequency (N and %) of schools who published a Pupil Premium statement for the academic year and for whom no published statement could be found. Additionally, we will present the frequency (N and %) of schools who used the DfE template and those who did not.

Below, we set out the relevant statistics for each of the study's research questions:

1. What challenges do schools identify for pupils eligible for the Pupil Premium? How does this vary by school characteristic?

Frequency (N and %) of schools reporting and not reporting each challenge identified in the code frame. This will provide us with descriptive statistics on what the most and least commonly cited challenges are and how many schools identified each challenge.

Frequency (N and %) of schools reporting and not reporting each challenge identified in the code frame, broken down by:

- **region** using ITL1 – former Government Office Regions (GOR) – from DfE's Get Information About Schools database to define schools in London, the South East, Yorkshire and the Humber, North West, East of England, West Midlands, North East, South West, and the East Midlands.
- **percentage of cohort eligible for the Pupil Premium** using total number of pupils eligible for the Deprivation Pupil Premium divided by the number of pupils on roll from DfE's Pupil Premium Allocation files (DfE PP allocation 24/25 FY September 2024).
- **number of pupils eligible for the Pupil Premium** using total number of pupils eligible for the Deprivation Pupil Premium (DfE PP allocation 24/25 FY September 2024) from DfE's Pupil Premium Allocation files.
- **phase** using PhaseOfEducation from DfE's Get Information About Schools database to define primary (including middle deemed primary) and secondary (including middle deemed secondary) schools.
- **school size** using NumberOfPupils from DfE's Get Information About Schools database.

- **school type** using EstablishmentTypeGroup from DfE's Get Information About Schools database to define mainstream academies & free schools, mainstream local authority maintained schools, and special schools.
- **attainment of disadvantaged pupils** using percentage of disadvantaged pupils reaching the expected standard in reading, writing and math (PTRWM_EXP_FSM6CLA1A) as a measure of KS2 attainment and average Attainment 8 score per disadvantaged pupil (ATT8SCR_FSM6CLA1A) as a measure of KS4 attainment from DfE 2023/24's performance data.

To avoid sub-groups that are too small for robust analysis, we will generally treat each characteristic independently rather than crossing them. For continuous characteristic variables, we will break the sample into quintiles or quartiles (depending on the subgroup sizes) and report the frequencies in the top and bottom quintile/quartile for comparison.

2. Which approaches are schools funding using the Pupil Premium? How does this vary by school characteristic?

Frequency (N and %) of schools reporting and not reporting approaches taken (of those specified by DfE) in the code frame. We will present these statistics overall and broken down by the characteristics listed in Research Question 1.

3. How do schools' approaches to spending align with their identified challenges?

Frequency (N and %) of approaches funded by schools to address each of the top 10 most commonly cited challenges.

4. How does school spending align with the EEF's tiered approach? How does this vary by school characteristic?

Proportion (%) of total Pupil Premium spend in each school allocated to 1) developing high-quality teaching, 2) providing targeted academic support, and 3) tackling non-academic barriers to academic success. We will present these statistics overall and broken down by the characteristics listed in Research Question 1.

5. What types of evidence are being cited in Pupil Premium statements? How does this vary by school characteristic?

Frequency (N and %) of schools citing each type of evidence identified in the code frame. 'Type' will span multiple dimensions, which will be agreed with EEF in the coding framework. We expect it to include frequency of citations to different types of evidence, frequency of references to the EEF, and frequency with which different EEF resources are cited. We will present these statistics overall and broken down by the characteristics listed in Research Question 1.

6. Which external programmes are schools funding with their Pupil Premium grant? How does this align with the EEF's list of promising programmes? How does this vary by school characteristic?

Frequency (N and %) of schools who report using their Pupil Premium to buy an external programme on the EEF's list of promising programmes, and the frequency (N and %) of schools who report using their Pupil Premium to buy each promising programme. We will present these statistics overall and broken down by the characteristics listed in Research Question 1.

Reporting

The results will be shared with the EEF at the end of each research cycle. The first research cycle will cover Pupil Premium statements published for the 24/25 academic year, and the second for the 25/26 academic year. The results will include a research report of 20-30 pages, summarising the approach, results, and lessons learned and a presentation to the EEF. The raw data from the coded Pupil Premium statements will also be shared with the EEF at the end of each research cycle. The project will run between April 2025 and September 2026. Table 4 shows the key dates and activities across the duration of the project.

Table 4 sets out the project timeline including reporting dates.

Timeline

The project will run between April 2025 and September 2026. Table 4 shows the key dates and activities across the duration of the project.

Table 4. Timeline

Dates	Activity	Staff responsible/leading
April 2025	Set up meeting	Verian and the EEF
May 2025	Draft and review the study plan	Verian and the EEF
5th June 2025	Sign off on the study plan	Verian and the EEF
June 2025	Data collection	Verian
	Coding frame development	Verian
16th June 2025	Sign off on the coding frame	Verian and the EEF
June 2025	Intercoder reliability checks (using 20 statements) to refine coding frame and AI model	Verian and the EEF
June – July 2025	Data extraction	Verian
July 2025	Analysis	Verian
July – August 2025	Drafting the written report (AY 24/25)	Verian
August 2025	Verian delivers findings presentation from (AY 24/25) to the EEF	Verian
August – September 2025	Written report (AY 24/25) review and peer review	Verian and the EEF
12th September 2025	Sign off on written report (AY 24/25)	Verian and the EEF
September 2025	Verian delivers raw data extracts (AY 24/25) to the EEF	Verian
	Review and lessons learned from AY 24/25	Verian and the EEF
April 2026	AY 25/6 kick off meeting	Verian and the EEF

Dates	Activity	Staff responsible/leading
	Update study plan (specifying changes from AY 24/25 if required)	Verian
April – May 2026	Collect sample	Verian
	Update coding frame for AY 25/26	
May – June 2026	Analysis	Verian
June – July 2026	Drafting the written report (AY 25/26)	Verian
July 2026	Verian delivers findings presentation from (AY 25/26) to the EEF	Verian
July – August 2026	Written report (AY 25/26) review and peer review	Verian and the EEF
28th August 2026	Sign off on written report (AY 25/26)	Verian and the EEF
August 2026	Verian delivers raw data extracts (AY 25/26) to the EEF	Verian

Ethical considerations and data protection

Verian puts all projects through our internal ethics process, assigning the project a RAG rating which determines what ethics procedures must be followed. This project was coded as Green 'low risk' as we will not be working directly with children and the Pupil Premium statements will be publicly available. Therefore, no formal ethics procedures are needed for this project.

Verian holds ISO27001 and Cyber Essentials Plus. We abide by the Data Protection Act 2018 and embed Data Protection by Design in all our work. We have a GDPR champion, and a Quality and Information Security team consult on all data privacy issues. Project data will be held securely on Verian's servers. We do not anticipate processing personal data for this project, as all data will come from publicly available information. However, in the case that sample schools upload personal data in their Pupil Premium statements we will take steps to identify and remove this before sharing the raw data extracts with the EEF (see the risk and mitigation 12 in Table 5 for more information).

Risks and mitigations

Table 5. Risks and mitigations

	Risk	Likelihood	Impact	Mitigation
1	Some schools may not have a Pupil Premium statement listed, or not have one updated for the right year(s)	Medium	Low	We assume that there will be a proportion of schools that do not publish their Pupil Premium statements on the website. We will include these schools in the sample and report the frequency (N and %) of schools for whom we could not find a published Pupil Premium statement for the academic year. Assuming that 8% of the sample will not have published a Pupil Premium statement based on our estimates, this should not affect the precision of quantitative estimates too much (+- 3%pts overall, and +-9%pts for region subgroups).
2	Schools included/that have Pupil Premium statements available are not representative of the wider school sample/their student bodies are not representative of the general student populations	Low	High	<p>Key demographics we will consider in designing our final sampling frame include (school-level) location, type of school, representation from Education Investment Areas (EIAs), and teacher-level characteristics including gender, tenure as teacher, ethnicity, and subject. We propose a few further measures to ensure representation in the chosen sample:</p> <ul style="list-style-type: none"> • Implementing a disproportionate design with an equal number of schools sampled from each region. • Boosting the representation of secondary and AP establishments to allow for robust analysis by school phase (i.e. primary, secondary, EP) • Clustering schools into quartiles or quintiles based on FSM eligibility – as this would generate categories with sufficiently large sample sizes. <p>Once we draw the sample of 1,395 schools, we will check the design weighted profile against the profile of the entire eligible population to ensure the sample is broadly representative of the wider population across a range of variables.</p>

Risk	Likelihood	Impact	Mitigation
3	Thematic analysis on our sample of 20 statements is not enough to generate all the codes for the coding frame	Low	Medium Sampling 20 statements allows us to sample purposively from schools with the whole range of key characteristics. In our experience, 20 statements should be enough to reach data saturation (when no new themes are emerging). However, if after having coded 20 statements, we are still finding new themes coming up, or if the experts on our team think the 20 statements have not covered all the types of content that schools write in Pupil Premium statements, then we will draw another sample of 10 statements at a time, and continue extracting until we reach data saturation.
4	AI misidentifies or excludes text excerpts that we are certain are relevant to the topic (false negative)	Low	High Researchers will manually identify texts that mention specific topics using basic keyword searches. We will then feed those texts to the AI model and examine the output to ensure that the AI is not misidentifying or excluding text excerpts that we are certain are relevant to the topic.
5	The AI hallucinates (false positives)	Low	High We are not using generative AI for complex tasks, but only to identify information in texts and create binary outputs (with links to the actual text). This reduces the risk of hallucinations, which is greater in generative processes. We will pilot the model we use and carefully measure any hallucinations and errors in the first outputs to decide if targeted mitigation is necessary (e.g., increasing human research support). We will also give the AI model very clear instructions based on the coding framework agreed with the EEF (which will be based on the coding frame from the EEF's previous Pupil Premium statement coding exercise). We will check interrater reliability with two humans and implement mitigations if Cohen's Kappa falls below 90% (see risk below).
6	Interrater reliability between humans and AI is low (Cohen's Kappa is below 90%)	Low	Medium If Cohen's Kappa is below 90%, we will take the following steps: (1) We will review data for potential errors in text formatting that could prevent the AI model from reading the text correctly. (2) We will review our AI prompt engineering to identify, adjust, and retest specific instructions or queries that may be producing suboptimal results. (3) We will feed the original intercoder reliability test texts to the newly modified AI to determine if its performance has improved. (4) We will select a new sample of texts for both the human coders and AI to process to ensure that the improvements are robust, as measured by Cohen's Kappa > 90%. If necessary, we will repeat these steps until we have sufficient interrater reliability. However, we deem the risk of interrater unreliability low, as a similar methodology has yielded intercoder reliability as high as 99% in our previous work. If after several rounds of comparisons between human and AI coder and

Risk	Likelihood	Impact	Mitigation	
			subsequent adjustments yield results just shy of the 90% benchmark, we will seek to maximise Cohen's Kappa and continue with a disclaimer. However, if the human and AI are still providing very different outputs after several attempts to reconcile them, we will have to conduct a very thorough audit into the problem or potentially drop the variable code if it is of low importance.	
7	If we use different AI models in each academic year, then they might assign codes differently	Low	Low	We will use the same AI for AY24/5 and AY25/6. If the AI model is updated in between AY24/5 and AY25/6, we will continue to use the model we used in AY24/5; if that model has been retired, then we will re-run the AY24/5 data with the updated model for consistency. This will ensure we are comparing like with like.
8	Challenges meeting timeline	Low	Medium	We are used to completing reviews in tighter timelines (2-3 months). We propose a team who are experienced in extracting and analysing data. We will apply established management processes to ensure delivery goes to plan, including sharing details of leave with EEF over the summer and flagging anywhere that we need EEF's input to stay on track. We will also work with the EEF to update the timeline in AY 25/26 to reduce timeline pressure (e.g., starting work on the next phase earlier).
9	Challenges meeting timeline with the study plan review and sign off	High	Medium	To maximise the quality of the first draft, we will prepare a skeleton study plan to agree content, structure, template, and depth with the EEF. We will also prepare and share a clear timeline so both teams know when to expect to see and review the study plan. On each review, we will highlight which sections include content that impacts when the next stage of research can begin (i.e. agreement on the research questions means we can begin refining the coding framework). Our team is experienced at delivering high quality outputs within tight timeframes.
10	Key staff turnover or absences	Medium	Low	We have other staff members with the required skills across Behavioural Practice, Data Science and Methods Teams. Our project management systems ensure all project team members are fully briefed; quality procedures ensure complete documentation; key roles are deputised so that informed back-up is always available.
11	Misunderstanding your needs	Low	High	Our Project Director and Project Manager have delivered previous projects for the EEF and are familiar with its ways of working and requirements. We will have regular catch-ups and an extended lessons learned/ kick off meeting before starting the AY25/26 research.

	Risk	Likelihood	Impact	Mitigation
12	Personal data in Pupil Premium statements	Low	High	We will agree with the EEF the definition of personal data (e.g. email addresses, phone numbers, names of pupils or staff members), then screen the raw data extracts for personal data. We will take steps to remove this personal data and complete additional checks to make sure all personal data has been removed. We have also added an extra week to the timeline to ensure all checks have been completed. Our staff members have experience with these tasks on similar projects with large datasets.

Project team

Table 6. Project team

Name	Job title	Role on project
Prof. Natalie Gold	Head of Trials	Project director overseeing the design, delivery, and reporting of the whole project
Dr Sarah Bowen	Associate Director	Project manager
Luke Taylor	Senior Methods Director	Methods expert designing the sampling and weighting
Dr Adam Ozer	Senior Data Scientist	AI and data expert leading on the AI-powered text data processing
Alex Thornton	Head of Education Research	Subject area expert advising on the coding frame and interpretation of results
Dr James Thom	Senior Methods Director	Methods expert overseeing quantitative analysis
James Gurd	Project support	Providing project support on data collection, intercoder reliability checks, analysis and reporting
Dr Varvara Kuz	Project support	Providing project support on data collection, intercoder reliability checks, analysis and reporting