STATISTICAL ANALYSIS PLAN

Impact of Attendance Mentors in secondary schools: a two-armed school-level cluster randomised controlled trial and implementation and process evaluation

Centre for Evidence and Implementation

Principal investigators: Katherine Young; Jane Lewis

YOUTH ENDOWMENT FUND

Impact of Attendance Mentors in secondary schools: a two-armed school-level cluster randomised controlled trial and implementation and process evaluation

Statistical analysis plan

Evaluating institution: Centre for Evidence and Implementation

Principal investigator(s): Jane Lewis; Katherine Young

Project title[1]	Impact of Attendance Mentors in secondary schools: a two-armed school-level cluster randomised controlled trial and Implementation and Process Evaluation
Developer (Institution)	Etio (with delivery support from Thrive and ImpactEd)
Evaluator (Institution)	Centre for Evidence and Implementation
Principal investigator(s)	Katherine Young; Jane Lewis
SAP author(s)	Katherine Young, Amy Hall, Jane Lewis, Shania Rankin, Ariel Lindorff, Sharon Lee
Trial design	A two-armed cluster randomised controlled trial with random allocation at the school level and parallel implementation and process evaluation
Trial type	Efficacy

[1] Please make sure the title matches that in the header and that it is identified as a randomised trial as per the CONSORT requirements (CONSORT 1a).

Evaluation setting	School
Target group	Secondary-age pupils in Years 7-11 who are persistently (50-90% attendance) or severely (<50% attendance) absent from school, in mainstream secondary schools in England
Number of participants	63 schools; 17,325 pupils across control and intervention arms
Primary outcome and data source	Attendance: school Management Information System (MIS), data gathered via the ImpactEd platform; average attendance rate of pupils (2 terms before and after mentoring) who were persistently (50-90% attendance) and severely (<50% attendance) absent at baseline per school
Secondary outcome and data source	Social and behavioural difficulties: Strengths and difficulties questionnaire Total Difficulties Score and subscales (Emotional Symptoms, Conduct Problems, Hyperactivity/ Inattention, Peer Relationship Problems, and Prosocial Behaviour). Young Person self-report, measured annually in Year 8 pupils only Attainment (Pupil-level Attainment 8 score): ImpactEd platform and National Pupil Database (NPD) Exclusions and suspensions: ImpactEd platform and NPD

SAP version history

Version	Date	Changes made and reason for revision
1.2 [*latest*]		
1.1		
1.0 [*original*]	15.04.2025	*[leave blank for the original version]*

Any changes to the design or methods need to be discussed with the YEF Evaluation Manager and the developer team prior to any change(s) being finalised. Describe in the table above any agreed changes made to the evaluation design. Please ensure that these changes are also reflected in the SAP (CONSORT 3b, 6b).

Table of Contents

Introduction

The Attendance Mentors programme is based on the theory that building trust and a supportive relationship with a positive role model, alongside practical assistance to address attendance barriers, can lead to improvements in pupils' attendance and behavioural, emotional, academic, and overall developmental outcomes. The programme was previously evaluated in a pilot study in Middlesborough (York Consulting, 2024), with findings indicating an average increase in attendance of 11 percentage points during the intervention (from 53% at the start of the interventions to 64% during the intervention). Further information on the intervention and its Theory of Change are presented in the study protocol (https://youthendowmentfund.org.uk/wp-content/uploads/2025/04/Attendance-Mentors-Evaluation-Protocol-Efficacy-1.pdf).

The current efficacy trial is a hybrid effectiveness-implementation study (Curran et al. 2012), exploring both the impacts of Attendance Mentoring (impact evaluation) and what it takes to deliver it effectively (Implementation and Process Evaluation; IPE). This Statistical Analysis Plan provides details of the impact evaluation, for more information on the IPE, see the study protocol. The two-armed cluster randomised controlled trial will compare outcomes between schools randomised to receive the Attendance Mentors programme, and schools delivering business as usual.

The primary aim of this study is to determine whether delivery of the Attendance Mentors programme increases the rates of attendance among persistently (50-90% attendance) and severely (less than 50% attendance) absent pupils. The secondary aims are to investigate the change in attendance among pupils who receive mentoring, and at a whole school level, as well as whether changes in attendance differ by age, ethnicity, gender, baseline attendance, free school meals (FSM) eligibility, special education needs and disabilities (SEND) status, and those with previous suspensions. We will also explore the time course of change in attendance, as well as impacts on secondary outcomes: social and behavioural difficulties, exclusion and suspension rates (and reasons), and academic attainment.

Design overview

Table 1 provides a summary of the trial design, as presented in the study protocol for the impact evaluation.

The primary impact evaluation question is:

RQ1: What is the difference in attendance rates (measured over 2 terms before/after mentoring; as measured by school administrative data, collected via the ImpactEd platform[2]) of pupils who were persistently or severely absent at baseline **from before to after implementation of the Attendance Mentors programme** in comparison with those in control schools receiving business-as-usual provision?

The secondary (exploratory) impact evaluation questions are:

RQ1(a) What is the difference in attendance rates of persistently and severely absent pupils **who take part in Attendance Mentoring**, in comparison with a matched control group of persistently and severely absent pupils in control schools?

RQ1(b): What is the difference in attendance rates **at the whole school level** for schools with Attendance Mentors in comparisons with those in control schools?

RQ1(c): Does the impact of Attendance Mentors differ by age, ethnicity, gender, baseline attendance, FSM eligibility, SEND status (including pupils with an education health and care plan (EHCP)), or for those who have previously been suspended?

RQ1(d): Among pupils who have received Attendance Mentoring, what is the time course of impact on a termly basis (i.e., when are changes in attendance observed, and how long do they last)?

RQ2: What is the difference in social and behavioural difficulties (as measured by the Strengths and Difficulties Questionnaire, SDQ) of persistently and severely absent pupils in Year 8 in schools with Attendance Mentors in comparison with those in control schools?

RQ3: What is the difference in exclusion and suspension rates and reasons (as measured by school administrative data and NPD data) of persistently and severely

[2] Data from the National Pupil Database (NPD) will be used as a back-up data source should issues with school uptake of the ImpactEd platform arise.

absent pupils in schools with Attendance Mentors in comparison with those in control schools?

RQ4: What is the difference in attainment (attainment 8 score) of persistently and severely absent pupils in schools with Attendance Mentors in comparison with those in control schools?

Table 1. Summary of trial design

Trial design, including number of arms		A school-level cluster randomised controlled efficacy trial with 1:1 allocation into two arms (treatment and control)
Unit of randomisation		School
Stratification variables (if applicable)		Priority Education Investment Area (PEIA)
Primary outcome	variable	Attendance
	measure (instrument, scale, source)	School attendance (percentage, including both authorised and unauthorised absences) obtained from school administrative data, collected via the ImpactEd platform/NPD[3]. Baseline is attendance for

[3] As stated in the protocol, we aim to use ImpactEd data wherever possible, if issues are encountered, or if data is unavailable, NPD data will be used as an alternate data source

		the two terms prior to January 2025[4], with follow-ups assessing attendance for the three terms prior to January 2026 and January 2027
Secondary outcome(s)	variable(s)	Social and behavioural difficulties, continuous variable Academic attainment , continuous variable Rates of exclusions, binary variable, and suspensions, categorical variable Reasons of exclusions and suspensions, categorical variables
	measure(s) (instrument, scale, source)	Strengths and Difficulties Questionnaire (SDQ) (Goodman, Meltzer & Bailey 1998) - total difficulties score self-reported by pupils at pre-randomisation and at 1 and 2-years post-intervention roll out, for Year 8 pupils only. Attainment 8 scores for pupils who complete their GCSEs during the study period; school administrative data (ImpactEd platform) and NPD Pupil level count of, and reasons for, suspensions and exclusions; school administrative data (ImpactEd platform) and NPD
	variable	Pupil level percentage attendance for the two terms prior to January 2025

[4] Please note that in the protocol it was stated 3 terms would be used, but this was revised to two terms as this is a more relevant time period more aligned with the eligibility criteria for mentoring

Baseline for primary outcome	measure (instrument, scale, source)	School administrative data, collected via the ImpactEd platform
Baseline for secondary outcome	**variable**	As above
	measure (instrument, scale, source)	As above

Sample size calculations overview

Table 2 provides details of sample size calculations conducted prior to recruitment and randomisation (as presented in the study protocol, where rationale for values selected is described in detail) and updated figures informed by the rates of recruitment achieved (i.e., numbers of eligible pupils in schools that are participating in the study, minus those who opted out).

Table 2. Overview of sample size calculations

		Protocol	**Randomisation**
Minimum Detectable Effect Size (MDES)		0.23 95% CI [0.069, 0.392]	0.23 95% CI [0.068, 0.392]
Pre-test/ post-test correlations	level 1 (participant)	0.5	0.5
	level 2 (cluster)	0.5	0.5

		Protocol	Randomisation
Intracluster correlations (ICCs)	level 1 (participant)	n/a	n/a
	level 2 (cluster)	0.2	0.2
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		Two-sided	Two-sided
Average cluster size		275	362
Number of clusters	intervention	31	32
	control	31	31
	total	62	63
Number of participants	intervention	8,550	12,096
	control	8,550	10,326[5]
	total	17,050	22,422

[5] Data from one control school is currently unavailable due to technical issues, this figure will be updated when available.

Assumptions for sample size calculations

A priori power calculations were conducted in PowerUpR (Ataneka et al., 2023) using a multilevel model of a cluster RCT, 1:1 allocation, a two-tailed alpha = 0.05, and power of 0.8. As detailed in the protocol, the final number of schools randomised was 63, with 32 allocated to the intervention condition and 31 to the control condition. For the purposes of power calculations, sample sizes were based on an equal number of schools in the intervention and control groups (31 in each). A priori power calculations were based on the anticipated number of pupils who could receive the mentoring intervention (38 mentors, each with 25 mentees for each of 9 waves = 8550 pupils), as a conservative (i.e. lowest) estimate of the number of pupils *eligible* to receive mentoring.

Randomisation took place on 17th February 2025 with a total of 63 recruited schools, stratified within each of 7 PEIAs (6 in Blackpool, 10 in Ipswich, 8 in Norwich, 10 in Nottingham, 9 in Portsmouth, 10 in Rochdale, 10 in Walsall). Randomisation was completed using the rarPar function of the randomizeR package (version 3.0.2; R version 4.4.2, 2024-10-31). Indicative code for randomisation is presented in the Appendix. Randomisation was stratified by PEIA to ensure 1:1 allocation of intervention and control schools within each region. In Portsmouth, where an uneven number of 9 schools were recruited and Etio's delivery requirements indicated 5 schools should receive the intervention, the list of schools was randomly re-ordered, then randomisation conducted with the first 8 schools on the list, and the 9th school was allocated to receive the intervention.

Post-randomisation power calculations were based on the numbers of pupils in intervention and control schools who were *eligible* to receive mentoring at beginning of the trial, and who did not withdraw consent for participating in the trial, plus a 20% increase for each new intake of pupils who would reach Year 8 within the study (Year 8 is reported by delivery partners as the 'cliff edge' of attendance where attendance problems begin to emerge). Therefore, the final sample size was estimated at 140% of the current total eligible pupils.

Secondary research questions will use different samples of pupils, restricting to only those pupils who *receive* mentoring (compared to a matched group from control schools, RQ1a), comparing effects at a whole school level (RQ1b), and comparing impact on the SDQ among Year 8 pupils only (as the SDQ was administered only to this year group, as outlined in the protocol, RQ2).

As detailed in the study protocol, a priori power calculations for the minimum detectable effect size (MDES) were based on ranges of estimates for pre- and post-test correlations and

intra-cluster correlations (ICC), drawing both on previous literature, and preliminary analysis of attendance data held by DfE. The final power calculation was based on pupil and school-level correlations of r = 0.5, and school-level ICC of 0.2, resulting in an MDES of 0.23 (note this was the most conservative estimate of MDES with ranged from 0.047-0.23). Simplified approximation of this effect size in terms of percentage attendance indicates sensitivity to 6% change (e.g., from 80% to 86% attendance, or an increase from 152 to 163 days attendance within a school year).

Analysis

The trial overview will be reported in accordance with the CONSORT statement for cluster randomised trials (Campbell et al., 2012). Multilevel models will be used (pupils nested within schools, with random intercepts for school), to analyse both the primary and secondary outcomes (R package lme4). Post-intervention attendance (i.e., attendance in the 2 terms after the end of programme delivery) will be the dependent variable in the primary outcome model, with predictors at the school level (intervention condition), and the pupil level (pre-intervention attendance). As mentoring is implemented at a school level, and not within specific classes or year-groups, it was decided that modelling of additional nesting variables of year group or class, was not appropriate.

Our primary model will be an intention-to-treat (ITT) analysis, examining attendance for all pupils eligible for attendance mentoring in intervention and control schools (independent of whether the mentoring was offered or taken up in intervention schools). Eligibility for mentoring is defined as pupils in Years 7 to 11 whose attendance is below 90% in the two academic terms prior to school randomisation (one previous term for those in Year 7). This. Includes all pupils who are either classed as persistently absent (50-90% attendance), or severely absent (less than 50% attendance). In the primary model, we do not differentiate between these categories of absence (exploratory subgroup analyses will examine differences by attendance band, detailed below). Variables that were initially considered for stratification of randomisation (rates of free school meals, special educational needs, attendance bands and school size) but were unfeasible to implement will be considered for inclusion as covariates, based on the distribution across groups (as described in detail in the protocol). Noting findings from the pilot study demonstrating 30% non-engagement for those referred into attendance mentoring (York Consulting, 2024), we will also conduct exploratory compliance analyses, examining outcomes only for pupils who engaged to varying degrees (see compliance section below).

All other analyses are exploratory. To correct for multiple comparisons across a large number of tests, we will use the Romano-Wolf procedure for familywise error correction (using R package, crctStepDown; Romano and Wolf, 2005). For all analyses, outcomes data will be examined to determine the most appropriate modelling approach, including considerations to transform the data to a normal distribution or models based on alternative distributions (e.g., zero-inflated models for count data or fractional models for data constrained between 0-1). We prefer to use linear models where possible, providing key assumptions are met, to aid with ease of interpretability. To examine subgroups of interest (e.g., pupils who have received a first suspension, pupils from minority ethnic backgrounds – see section on subgroup analyses below), interaction effects will be included in multilevel models to assess the statistical significance and effect size of differential impact. Analyses will be conducted in R (R version 4.4.2, 2024-10-31).

Primary outcome analysis

The primary outcome variable is percentage attendance in the two terms before and after receiving mentoring (where unavailable, one term of attendance will be used), with matched periods in control schools.

The model for the primary analysis is as follows (analysis code shown in Appendix):

$$y_{ij} = \beta_0 + \beta_1 Group_j + \beta_2 Baseline_{ij} + u_j + e_{ij}$$

where

y_{ij}	is the post-mentoring attendance for pupil *i* in school *j*
β_0	is the intercept
$\beta_1 Group_j$	is the group (intervention vs. control) for school *j*
$\beta_2 Baseline_{ij}$	is the baseline attendance for pupil *i* in school *j*
u_j	is the random intercept for school *j*

e_{ij} is the residual error term for pupil *i* in school *j*

The primary model is simple, evaluating the main effect of the intervention, and modelling school-level variance using a random intercept for transparency and ease of interpretation. Exploratory subgroup analyses (described below) will include additional covariates for rates of free school meals, special educational needs, attendance bands and school size following exploration of variance and distribution across intervention and control schools. We will assesss the relationship between each covariate and the primary outcome using univariate analyses, and include only those demonstrating a statistically significant association in our covariate analyses.

Secondary outcome analysis

RQ2: What is the difference in social and behavioural difficulties (as measured by the Strengths and Difficulties Questionnaire, SDQ) of persistently and severely absent pupils in Year 8 in schools with Attendance Mentors in comparison with those in control schools?

This analysis will use the model specified for RQ1, but with the following differences: i) substituting rates of attendance before and after mentoring with SDQ Total Difficulties scores, ii) including a dummy covariate, and interaction term, for the term in which mentoring was received. The dummy covariate is required as the SDQ is completed on an annual basis, whereas mentoring is delivered on a termly basis, with 9 delivery cycles across the trial. Analyses will be repeated for the Total Difficulties score, as well as each of the 5 subscales (Emotional Symptoms, Conduct Problems, Hyperactivity/ Inattention, Peer Relationship Problems, and Prosocial Behaviour). Romano-Wolf Family-Wise Error correction will be used to correct for multiple comparisons.

This is detailed in the equation below:

$$y_{ij} = \beta_0 + \beta_1 Group_j + \beta_2 Baseline_{ij} + \beta_3 \mathrm{X}_i + \beta_4 Group_j * \mathrm{X}_i + u_j + e_{ij}$$

where

y_{ij} is post-mentoring 'SDQ Total Difficulties' for pupil *i* in school *j*

β_0 is the intercept

$Group_j$	is group (intervention vs. control) for school *j*
X_i	is a dummy variable for a covariate of term of delivery
$Baseline_{ij}$	is baseline 'SDQ Total Difficulties' for pupil *i* in school *j*
u_j	is the random intercept for school *j*
e_{ij}	is the residual error term for pupil *i* in school *j*

RQ3: What is the difference in exclusion and suspension rates and reasons (as measured by school administrative data and NPD data) of persistently and severely absent pupils in schools with Attendance Mentors in comparison with those in control schools?

This question seeks to explore whether the Attendance Mentors programme has an impact on exclusion and suspension in intervention compared to control schools. The variance and distribution of exclusion and suspension rates will be examined for suitability for multilevel modelling, categorising permanent exclusion as a binary variable (yes/no) and suspension (aka fixed-term exclusion) as a categorical variable (0 suspensions, 1 suspension, >1 suspension). We anticipate using zero-inflated or Poisson distribution models given the relative low incidence of these outcomes (e.g., 4.13 suspensions, 0.05 permanent exclusions per 100 pupils in Autumn 2023/24, DfE 2024). The equation from RQ1 will be used, substituting attendance outcomes for exclusion and suspension outcomes, using the glmTMB R package to implement alternative distributions (Brooks et al., 2017). The exclusion model will also remove the 'baseline' term for exclusion, as this will be zero for all pupils. Where available, the number of days or sessions suspended will also be descriptively analysed for exploratory purposes but will not be included as a primary outcome due to expected non-normality and data sparsity. Reasons for exclusion and suspension will be explored using descriptive analyses.

RQ4: What is the difference in attainment rates (attainment 8 score) of persistently and severely absent pupils in schools with Attendance Mentors in comparison with those in control schools?

Attainment 8 scores will be accessed via the ImpactEd platform and the NPD for two cohorts of Year 11 pupils who complete their GCSEs during the evaluation period. A continuous

score, ranging from 0-90 where higher scores indicate stronger academic performance, will be obtained for each pupil. The outcome will not be averaged at the school level; instead, individual pupil-level scores will be used in the analysis. As with RQ3, distributions of attainment rates will be explored for suitability for use in a multilevel model, with alternative models considered as necessary. The equation for RQ1 will be used, substituting outcomes for attendance with outcomes for attainment (pupil level attainment 8 score), and removing the baseline term for this variable as it is only available as a one-time measure.

Subgroup analyses

The following sets of analyses aim to examine outcomes in different groups of pupils than specified in the primary outcome analysis. All of these analyses are exploratory.

RQ1(a) What is the difference in attendance rates of persistently and severely absent pupils **who take part in Attendance Mentoring**, in comparison with a matched control group of persistently and severely absent pupils in control schools?

This analysis will examine differences in attendance rates among those who participated in mentoring (as opposed to those *eligible* for mentoring in the primary research question), compared to a matched sample of pupils from control schools. Pupil matching in control schools will be conducted by asking schools to complete a form in which they identify which pupils would have been put forward to *receive* mentoring from within their pool of pupils who would have been *eligible* for mentoring, had they been randomised to the intervention arm. This approach aims to approximate the subjective element of the selection process conducted within intervention schools, albeit with the acknowledged limitation that the process will not be an exact match, particularly as we are unable to control for participants who would be offered mentoring, but choose not to engage. Interpretation of this model will include consideration of these limitations, and will acknowledge the extent to which inferences can be reasonably drawn. Since randomisation and the beginning of delivery, we now agree that the priority is to minimise the burden on control schools to avoid drop-out from the trial. Statistical power is also more affected by a reduction in the number of schools providing data, than the number of individuals per school (for example, reducing from 62 schools to 52 schools would change our MDES from 0.23 to 0.26). As such, we are proposing to conduct the matching only once during delivery. To match procedures in interventions schools as closely as possible, we will request that control schools identify 30 pupils from any year who they would have put forward for mentoring, aligning with recruitment for Wave 3 (Nov/Dec 2025). Power calculations indicate that this would provide a MDES = 0.24 [CI = 0.072, 0.415]. Analyses will be conducted using the same model as specified in RQ1.

RQ1(b): What is the difference in attendance rates **at the whole school level** for schools with Attendance Mentors in comparisons with those in control schools?

This analysis aims to examine the impact of attendance mentors on whole school populations and will be conducted using the same model as specified in RQ1 but for an analytical sample including all pupils in each school.

RQ1(c): Does the impact of Attendance Mentors differ by age, ethnic group, gender, baseline attendance, FSM eligibility, SEND status (including pupils with an education health and care plan (EHCP)), or for those who have previously been suspended?

This research question will examine whether there is differential impact of the attendance mentors programme on primary and secondary outcomes based on the demographic variables listed. Descriptive breakdown of eligible pupils by each demographic variable will be provided in accordance with the YEF Demographic Data Policy. Each demographic variable will be inspected for appropriateness of inclusion within analytical models (exploring the variance – where relevant – and distribution across intervention and control schools). For subgroups in which there is insufficient sample size (i.e., <5%) to examine within regression models (e.g., minority ethnic groups), we will explore recategorizing data to create larger subgroups for statistical analysis and will use cross-tabulation to report differences in original categorisations descriptively. Where subgroup sizes are ≥5%, we wil conduct exploratory analyses with justification based on their representation in the sample. Such analyses will be explicitly noted as exploratory, particularly where sample sizes are small and underpowered.

For ease of interpretation of regression models, each variable will be included in a separate model, covariates will be included as dummy variables where appropriate (i.e., for ethnic group, gender, FSM eligibility, SEND status, and previous suspension) and interaction terms will be included to explore the impact of each variable on attendance. This will follow the same equation as listed for RQ2, but substituting terms of mentoring delivery for dummy variables and interaction terms related to demographic variables (except when modelling the SDQ outcome, in which both demographic variables and the dummy variable for term of delivery will be included).

Further analyses

RQ1(d): Among pupils who have received Attendance Mentoring, what is the time course of impact on a termly basis (i.e., when are changes in attendance observed, and how long do they last)?

This exploratory analysis was specifically requested by a funder of the evaluation and seeks to provide practical insight into the timing of impact. Longitudinal (termly) data will be extracted for pupils receiving mentoring, and a matched control group, and descriptive analyses will explore mean rates of attendance, and clustered standard error (calculated using R package sandwich) to visually display the timing and duration of impact (plotted using R package ggplot2). Termly data will be aligned across groups receiving the intervention in different waves (i.e., 1 term pre-mentoring, 1 term post-mentoring etc.).

Interim analyses and stopping rules

There are no planned interim analyses or stopping rules based on statistical analyses.

Longitudinal follow-up analyses

No follow-up analyses will be conducted beyond the time period specified in the primary analysis.

Imbalance at baseline

We will report on the following attendance variables obtained from intervention and control schools at baseline for the two terms preceding the beginning of mentoring (Autumn 2024 and Winter 2025):

1. Overall school attendance
2. Rates of persistent absence
3. Rates of severe absence
4. Proportion of students across attendance bands

The following additional variables from the ImpactEd database will be reported at the school level:

5. Proportion of pupils with free school meal eligibility (FSM6_p; categorised as: 1= True, 0= False)
6. Proportion of pupils with special educational needs (categorised as: 1= True (any), 0= False)
7. School size (continuous measure)
8. Proportion of pupils who were excluded (categorised as: 1= True, 0= False)
9. Proportion of pupils who received suspensions (0, 1 or >1 suspension)
10. Mean SDQ Total Difficulties baseline scores (continuous score between 0-40), and mean subscale scores (Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, Peer Problems, and Prosocial Behaviour)
11. Mean age (in years)
12. Proportion of pupils in each gender category (categorised based on NPD standard: male, female)
13. Proportion of pupils in each ethnic group (categorised based on NPD standard: Any Other Ethnic Group, Asian, Black, Chinese, Mixed, Unclassified, White)

Attendance measures (1-4) will be reported to examine the success of randomisation. Additional variables (5-13) will be reported to demonstrate sample characteristics, and to inform suitability for subgroup/covariate analyses. Continuous variables will be reported with means and standard deviations, categorical variables will be reported as counts/percentages.

Missing data

We will quantify and report the extent of missing data for all primary and secondary outcomes, reporting the number and percentage of available complete data at school and pupil level, as well as exploring patterns of missing data.

Our primary outcome is attendance gathered via the ImpactEd platform, which pulls data directly from school Management Information Systems. As such, we anticipate a low level of missing endline data (<5%), mainly reflecting pupils who left or moved schools during the trial period. We will explore patterns of missing data using descriptive statistics and a multi-level logistic regression model to assess whether missingness (binary variable) is predicted by variables in our primary analyses (as has been conducted in other cluster RCTs funded by YEF; e.g., Riordan et al., 2024). If no predictor of missingness is identified (consistent with MCAR assumptions), no further adjustment will be applied, and analyses will be completed with full cases only. If missingness is related to observed variables (consistent with MAR assumptions), we will apply multiple imputation to account for the missing data where feasible (R package mice). Multiple imputation will include all variables in the substantive model, predictors of

missingness, and auxiliary variables as appropriate. The number of imputations will be guided by the proportion of incomplete cases (e.g. at least equal to % missing; with a minimum of 10 imputations and increasing to 20–50 with higher missingness levels). Sensitivity analyses will compare imputed results with complete case analyses to evaluate robustness. Imputation will not be attempted if the level of missing data exceeds 40%, instead, results of the complete case analysis will be reported with interpretative limitations.

For secondary outcomes, we anticipate more missing data, primarily in the SDQ which requires administration by schools. The level and pattern of missing data will be reported descriptively, however, because secondary outcomes are exploratory, we will not conduct logistic regression or multiple imputation for these but focus our efforts on developing and assessing a robust model for the primary outcome (as recommended in YEF statistical analysis guidance). If any secondary outcome has 40% or more missing data, no inferential analysis on that variable will be conducted.

Compliance

Compliance with mentoring is defined at the individual pupil level, based on pupils attending a minimum 7 mentoring sessions, as agreed with the delivery partner during the set-up phase. The primary analysis will examine outcomes across all pupils *eligible* for mentoring, irrespective of whether they were *offered* mentoring, whether they *received* it, or whether they *complied* with it. In this regard, the primary analysis is considered an 'intention-to-treat' (ITT) model, with the intervention conceptualised as being targeted to this group of eligible pupils, even though not all will receive mentoring. This decision was taken to allow for an accurate and matched control condition in schools delivering BAU, where there would be no hand selection of pupils suitable for mentoring by mentors. Therefore, the most robust comparison group is pupils who would be *eligible* for mentoring in both intervention and control schools.

We will explore two alternate approaches to modelling compliance. First, we will operationalise as a binary variable, 0 = not offered mentoring, or offered mentoring, but did not comply to 7 or more sessions, 1 = complied to 7 or more sessions. Secondly, we will operationalise as a continuous variable reflecting the number of sessions attended. We will compare outcomes across these two alternative modelling approaches.

To evaluate the impact of compliance on outcomes, an Instrumental Variables (IV) approach will be taken, using a two stage least squares (2SLS) model. Replicating the approach taken

by other cluster RCTs funded by YEF (e.g., Riordan et al., 2024), we will estimate a first stage model as follows:

$$Compliance_{ij} = \beta_0 + \beta_1 Group_j + \beta_2 Baseline_{ij} + u_j + e_{ij}$$

$$u_j \sim N(0, \sigma_u^2)$$

$$e_{ij} \sim N(0, \sigma_e^2)$$

The outcomes of this model will be reported, along with correlations between the outcome and compliance, and an F-test (following YEF guidance). The predicted values of $Compliance_{ij}$ (i.e., $\widehat{Compliance}_{ij}$) from the first stage model will be used to estimate the second stage model of the outcome measure:

$$y_{ij} = \beta_0 + \beta_1 \widehat{Compliance}_{ij} + \beta_2 Baseline_{ij} + u_j + e_{ij}$$

The primary outcome of interest is β_1 which will estimate the effect of the intervention among pupils who completed 7 or more mentoring sessions.

Please note that compliance will also be explored via the subgroup analysis of RQ1(a) an ITT model examining only those pupils who were *offered* mentoring, compared to pupils in control schools who would have been offered mentoring (identified via a survey of control schools, asking them to select pupils who they think might benefit from mentoring, for a subsample of pupils).

Intra-cluster correlations (ICCs)

We will use ICCs to estimate the amount of variance in the primary outcome (attendance) is attributable to variance within and between schools. The ICC is defined as:

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

Where

σ_u^2 = variance between schools

σ_e^2= variance within schools (i.e., between pupils in a school)

These will be completed for an empty model (without covariates) and for the primary analysis model, as well as the main analytical model for each secondary outcome.

Presentation of outcomes

Hedge's g effect sizes will be calculated, in line with YEF's statistical analysis guidance (YEF, 2021):

$$ES = \frac{(\bar{Y}_t - \bar{Y}_c)_{adjusted}}{s^*}$$

Where

$(\bar{Y}_t - \bar{Y}_c)_{adjusted}$ = the difference in means between the intervention and control group, adjusting for the baseline predictor

s^*= the pooled unconditional variance of the two groups, defined as

$$s^* = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Where

n_1= control group sample size

n_2= intervention group sample size

s_1= control group standard deviation

s_2= intervention group standard deviation

95% confidence intervals of all effect sizes will be calculated by inputting the upper and lower confidence limits from the primary model into the formula for calculating effect sizes. We will provide a conversion from effect sizes into numerical/percentage change in attendance.

References

Ataneka, A., Kelcey, B., Dong, N., Bulus, M., & Bai, F. (2023). PowerUp R Shiny App (v. 0.9) manual. https://powerupr.shinyapps.io/index/

Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling.

Campbell, M. K., Piaggio, G., Elbourne, D. R., & Altman, D. G. (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ*, 345.

Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., & Stetler, C. (2012). Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Medical care*, *50*(3), 217-226. https://doi.org/10.1097/mlr.0b013e3182408812

Department for Education (2024). Suspensions and permanent exclusions in England statistics. https://explore-education-statistics.service.gov.uk/find-statistics/suspensions-and-permanent-exclusions-in-england/2023-24-autumn-term

Riordan, S., Hodgen, J., & Taylor, R. (2024). Grassroots efficacy trial: evaluation of an anti-conflict intervention by cluster randomised controlled trial: Statistical analysis plan. https://youthendowmentfund.org.uk/wp-content/uploads/2023/02/Grassroots-Statistical-Analysis-Plan-July-2024.pdf

Romano J. P., Wolf M. 2005. Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association* 100: 94–108. https://doi.org/10.1198/016214504000000539.

YEF (2021). Analysis guidance. Youth Endowment Fund. Last updated June 2021. https://res.cloudinary.com/yef/images/v1623145483/cdn/6.-YEF-Analysis-Guidance/6.-YEF-Analysis-Guidance.pdf

York Consulting (2024). Evaluation of the attendance mentors pilot: Year 1 findings. https://www.gov.uk/government/publications/evaluation-of-the-attendance-mentors-pilot

Appendix

Randomisation

```
# Indicative code for randomisation within a single PEIA:

# set randomisation parameters and generate randomisation sequence

params <- rarPar(N, K=2, ratio = rep(1,2), groups = c("Intervention","Control"))

(rs <- genSeq(params))

# output saved with randomisation seed for replicability

saveRand(rs, file = "Blackpool_randList.csv")
```

Primary analysis

```
# Model specification

attendance_model <- lmer(attendance_post ~ 1 + group + attendance_pre + (1|School),
data = attendance_data)


# Model specification - covariates

attendance_model_covariate <- lmer(attendance_post ~ 1 + group + attendance_pre +
covariate + covariate:group + (1|School), data = attendance_data)
```

YOUTH ENDOWMENT FUND

Impetus

Home Office

youthendowmentfund.org.uk

hello@youthendowmentfund.org.uk

@YouthEndowFund

The Youth Endowment Fund Charitable Trust

Registered Charity Number: 1185413