School incentive payments for mentor engagement in training
Statistical Analysis Plan

Evaluator (institution): NFER
Principal investigator(s): Dr Ben Styles

Education Endowment Foundation

PROJECT TITLE[1]	School incentive payments for mentor engagement in training: A randomised controlled trial
DEVELOPER (INSTITUTION)	Department for Education (commissioner), Ambition Institute and Teach First (providers)
EVALUATOR (INSTITUTION)	National Foundation for Educational Research
PRINCIPAL INVESTIGATOR(S)	Dr Ben Styles.
PROTOCOL AUTHOR(S)	Dr Joana Andrade, Dr Ben Styles, and Sarah Tang
TRIAL DESIGN	Two-arm cluster randomised controlled trial with random allocation at school level
PUPIL AGE RANGE AND KEY STAGE	n/a
NUMBER OF SCHOOLS	225
NUMBER OF PUPILS	n/a
PRIMARY OUTCOME MEASURE AND SOURCE	Time spent engaged in mentorship training activities (source: providers' management information (MI) systems)
SECONDARY OUTCOME MEASURES AND SOURCES	Mentorship dropout(1), Mentorship retention(2), Mentor retention(3), and Mentor replacement(4) (source: providers' management information (MI) systems)

(1) Mentorship dropout being defined as all the mentorships established in a school at the beginning of the programme are terminated.
(2) Mentorship retention being defined as all the mentorships established in a school at the beginning of the programme are still ongoing.
(3) Mentor retention being defined as all the mentors enrolled in a school at the beginning of the programme are still in the programme.
(4) Mentor replacement being defined as a mentor that enrolled in a school but dropped out of the programme having been replaced by an additional mentor.

SAP version history

VERSION	DATE	REASON FOR REVISION
1.0	March 2021	*N/A*

[1] Make sure that the project title here matches the title of the document and the protocol. Please ensure that there is an identification as a randomised trial in the title as per CONSORT requirements.

- The primary outcome has been changed since the protocol was written. While we are still looking at mentor engagement time in training, we are planning on considering this per mentorship as opposed to per mentor. Under the revised approach, if a mentor drops out of the programme but is replaced then the new mentor's training time will be included in the analysis as this is part of the same 'mentorship' (and the programme for the ECT is continuing with a mentor).
- More detail has been provided on the two secondary outcome measures in the original protocol and a further two have been added.
- The first dropout measure is now defined in terms of mentorship dropout rather than mentor dropout, acknowledging that an ECT's mentorship can carry on even if the mentor originally assigned to the ECT left the programme.
- There are three further secondary outcomes – *mentorship retention*, *mentor retention,* and *mentor replacement* ((a) and (b)). The retention measures are school-level binary measures that signals if all the original mentors in a school are still enrolled in the training programme (*mentor retention*) and if all the mentorships established in a school at the beginning of the programme are still ongoing (*mentorship retention*). *Mentor replacement* is a school-level binary measure that signals whether there has been any replacement in mentors in a school with an additional mentor. Replacement refers to an additional mentor joining the programme to replace a mentor who has dropped out. *Mentor replacement (a)* will consider schools that only had one mentor-ECT pairing whereas mentor replacement (b) will consider schools that had more than one mentor-ECT pairing.
- The payments were delayed from autumn to spring term and this has been updated.

School incentive payments for mentor engagement in training
Statistical Analysis Plan

Evaluator (institution): NFER
Principal investigator(s): Dr Ben Styles

Education Endowment Foundation

Table of contents

Introduction

The school incentive payments for mentor engagement in training trial has been commissioned by EEF as part of the evaluation of the early roll-out of the core induction programme[2]. The Framework is part of the Government's teacher recruitment and retention strategy, which aims to increase the resources and opportunities open to teaching professionals, alongside the number of those in the profession.

The trial is designed as a randomised controlled trial (RCT) to evaluate the impact of an incentive payment of £775 to schools per mentor. The payment is intended to cover the costs of staff training time – explicitly the costs associated with mentor training. Mentor training takes place over the two-year programme but is frontloaded in the first year. The trial will therefore measure mentor engagement in training over the first year of the programme.

To support the ECF, the Department for Education (DfE) has commissioned four providers to develop and deliver a Full Induction Programme (FIP) for early career teachers and their in-school mentors. It was acknowledged during the design stage of the RCT that working with as few providers as possible would be more practical and efficient, but none of the four providers would be able to deliver a large enough sample to detect a small enough effect size. As such, the evaluation team agreed to restrict the trial to the two providers that, by autumn 2020, were closer to reaching their recruitment targets, and were considered to have comprehensive online platforms likely to provide useful management information (MI) for the trial. Those were Ambition Institute and Teach First.

Mentors will share their expertise with early career teachers (ECTs) in order to help them to develop and feel prepared for their future careers, with most mentors supporting one early career teacher. As part of the core induction programme, training and resources will be developed by the providers and the participating mentors will receive:

- 36 hours of training over the two-year induction period based on the ECF
- high-quality resources to support their mentoring
- funding to cover mentors' time with the mentee in the second year of teaching.

Ambition Institute and Teach First will offer a series of tailored activities and resources like webinars and online training, induction and coaching sessions, and seminars.

The incentive payment is separate from any funding received as part of the early roll-out of the ECF and will be paid by the DfE directly to schools during the spring term.

For the trial we will consider the primary research question:

RQ1 What is the effect of financial payments to schools participating in the early roll-out of the ECF on the level of mentor engagement in training as measured by the number of minutes spent engaged in training?

And the secondary research questions:

RQ2 What is the effect of financial payments to schools on the retention of mentors in the programme as measured at the end of each term?

[2] The core induction programme is a new two-year professional development entitlement for early career teachers and their mentors. It is underpinned by the Early Career Framework: https://www.gov.uk/government/collections/early-career-framework-reforms.

RQ3 What is the effect of financial payments to schools on the retention and dropout of mentorships as measured at the end of each term?

If the early career teacher (ECT) has a mentor (whether this is their original mentor, an additional mentor to replace the original mentor or another ECT's mentor) then the mentorship is considered to be continuing.

RQ4: What is the effect of financial payments to schools on the replacement of mentors who drop out in (a) schools where there is only one mentor and (b) schools where there is more than one mentor?

Design overview

Trial design, including number of arms		Two-armed randomised control trial
Unit of randomisation		School
Stratification variables (if applicable)		Provider
Primary outcome	variable	Time spent engaged in mentor training activities in a mentorship.
	measure (instrument, scale, source)	School-level bespoke measure. Total time of attendance at seminars, coaching sessions and time spent accessing materials on the providers' online platforms divided by the number of ECTs in the school. (source: providers' MI systems)
Secondary outcome(s)	variable(s)	Mentorship dropout, mentorship retention, mentor retention, and mentor replacement
	measure(s) (instrument, scale, source)	(1) Mentorship dropout: **All** the mentorships established in the school at the beginning of the programme are terminated. (2) Mentorship retention: All of the mentorships established in the school at the beginning of the programme are still ongoing. (3) Mentor retention: All of the mentors in the school at the beginning of the programme are still enrolled in the programme. (4) Mentor replacement: The mentors in a school who left the programme during the first year are replaced with an additional mentor. (source: providers' MI systems)
Baseline for primary outcome	**variable**	n\a
	measure (instrument, scale, source)	n\a
	variable	n\a

Baseline for secondary outcome	measure (instrument, scale, source)	n\a

Randomisation

The providers were aiming to recruit approximately 600 early career teachers and their mentors to take part in the programme, with more than one mentor based in the same school being a likely scenario. Considering this, and also that financial incentives will be paid to schools and not to individuals, the evaluation team chose to randomise at school level rather than at mentor level.

An initial school randomisation took place in October 2020. As specified in the protocol, the project statistician was not blinded to group allocation. The randomisation, stratified by provider in order to guarantee a balanced distribution of the incentives, was based on two lists of schools provided by Ambition Institute and Teach First.

After the initial randomisation, the evaluation team realized that one of the lists contained both schools and multi-academy trusts (MATs). A total of five MATS were randomised. In four instances several schools and the MAT they belonged to had been randomised separately, and, in a separate instance, a MAT was randomized but not its associated schools.

After conferring, the evaluation team decided to exclude the randomised MATs from the analysis as invalid cases while inviting the schools in the five MATs that were not randomised to join the trial. The MAT schools that were eligible and willing to take part in the trial could then be randomised at a later stage, and considered to have joined the trial on a second stage of recruitment. The second stage randomisation would proceed in the same manner as the initial one, stratified by provider. In practice, no further schools signed up to the trial so a second randomisation was not necessary.

The R code used to perform the stratified randomisation is included in this statistical analysis plan as an appendix. All the calculations were performed in R 4.0.3.

Sample size calculations overview

		Protocol	Randomisation
Minimum Detectable Effect Size (MDES)*		0.33	0.38
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		2	2
Average cluster size		1	1
Number of schools	Intervention	150	113
	Control	150	112
	Total	300	225

*There is no baseline for the trial so pre-post correlation is assumed to be zero.

The two providers selected to take part in the trial suggested that they would be able to recruit 200 schools each. Any changes in implementation must have a larger effect size than that desired in the more important downstream outcome of interest (in this case ECT retention) and therefore require a smaller sample size to detect. Four hundred schools, randomised equally, would be powered to detect an effect size of 0.28.

While writing the protocol a success rate of 75 per cent in the recruitment of schools to take part in the trial was taken as a reasonable estimate, and out of the 365 approached schools we estimated that 300 would take part. This yields an effect size of 0.33. In fact, at the time of recruitment to the trial the providers had recruited slightly fewer schools which meant that around 360 schools were approached to take part in the trial. Subsequently only 225 signed up which increases the MDES to 0.38 as shown in the table.

Power calculations were completed using a bespoke Excel spreadsheet assuming a two-group independent sample t-test design.

Analysis

The primary and secondary analyses will follow EEF guidelines, with intention-to-treat being assumed in both cases.

Although the data possesses a nested structure, with mentors clustered into schools, it is not amenable to multi-level modelling. A quick exploratory inquiry revealed that the majority of the schools in the sample have a small number of mentors in their staff, typically one or two, and multilevel models would not be able to estimate mentor-level variance.

Given the clustered nature of the data and the unsuitability of multi-level models, we have decided to consider instead outcomes that are aggregated or summarized at school level and run the analyses resorting to single level OLS and logistic regression models.

Primary outcome analysis

A bespoken measure of engagement with training will be a calculated using MI from providers' customer relationship management (CRM) systems. The termly collections of data on participation in training across the different types of provision[3] (live seminars, coaching and online self-study) will be combined in order to estimate how long, measured in minutes, each mentor engaged in training. In the pilot study for the programme the authors noted that they suspect a 'dormant time effect' suggesting not all time logged to the material was spent actively working on it. If the information collected by the providers allows for an estimation of dormant time effects, these will be accounted for.

The school-level primary outcome measure will be computed as the total engagement time measured across the mentors of each school divided by the number of ECTs being mentored, which corresponds to an estimate of the average length of time spent in mentor training per mentorship. If a mentor drops out of the programme but a replacement is found for the ECT, then the new mentor's training time will be included in the measure. We are thus defining the mentorship by the original ECT, i.e. the original list of ECTs enrolled at the beginning of the programme determines the mentorships in the analysis[4]. We feel this is

[3] The provisions are specific to each provider and differ across the programme.
[4] We have included all the ECTs that enrolled, although some of them were only assigned a mentor at a later stage. The training time of the mentor was accounted for, even if the training occurred before the assignment of mentor to ECT.

important as the aim of the programme is to train ECTs and provided the paired ECT still has a mentor (whether a new or existing mentor) the programme can continue. Following the same rationale, mentors that are not paired with ECTs will be excluded from the analysis as not meeting the eligibility criteria for the trial. For the primary outcome measure, if an ECT drops out of the programme then the ECT, as well as their associated mentor training time, will not be included in the analysis (unless the mentor is mentoring another ECT still in the programme).

Due to different numbers of mentorships in each school, we need to divide our total mentor training time attended in each school by either the number of mentors per school or the number of ECTs. If we divide by mentors then we risk understating the training time per mentorship if there is replacement of mentors with new mentors. We feel that by calibrating the total time spent on mentor training by dividing by the number of ECTs we include all the training that has taken place associated with that mentorship. One consequence of this is that schools where more replacement has taken place have the opportunity to receive more total hours mentor training per mentorship. We have chosen to divide by the number of ECTs as the programme (including the mentor training) is ultimately for the benefit of the ECT. It is also important to bear in mind that there is likely to be replacement across treatment and control. One of the secondary measures considers whether replacement when a mentor drops out is more likely in schools receiving the incentive than those not.

The primary analysis will determine if giving the schools in the programme a financial incentive (in addition to funding received as part of the existing programme) had an effect on how long their mentors engaged in training activities. For this purpose we will fit a single-level regression model with the dependent variable as school level estimated engagement time per mentorship, measured in minutes. In addition to group allocation, a categorical variable that reflects the provider stratifier of the randomisation will also be included in the model.

The full equation for the OLS model is:

$$Y = \beta_0 + \beta_1 intervention + \beta_2 provider + \epsilon$$

Where Y is the school-level average length of engagement in mentor training activities per mentorship, $intervention$ is the intervention/control dummy, and $provider$ is the indicator for provider.

The OLS model will be run in R (version 4.0.3).

Secondary outcomes analysis

The secondary analyses will assess the impact of giving schools an extra financial incentive on *mentorship dropout, mentorship retention, mentor retention* and *mentor replacement* at the end of each term of the first year of the programme, whenever applicable, and also at the end of the first year.

The first and second variables are related. *Mentorship dropout* indicates if **none** of the mentorships in a school are ongoing and *mentorship retention* indicates if **all** of the mentorships in a school are still ongoing. If there is only one mentorship in a school then the two variables will be identical but if not, the two variables measure different ends of the scale – one where all mentors drop out and the other where no mentors drop out. It is useful to include both variables in the analysis as *mentorship dropout* could be considered to reflect implementation failure where the programme has failed in a school as no mentorships have

continued. The *mentorship retention* variable captures **any** termination of mentorship as it signals when just a single mentorship in a schools has ended. Mentorship can end due to ECT dropout, for personal reasons, for reasons associated with the school as well as factors relating to the programme itself.

Mentor retention asks whether all the mentors originally enrolled in the programme remain in the programme in a particular school. Mentors could leave the programme for a range of reasons including if their ECT drops out.

The last variable – *mentor replacement* – measures whether any mentors that dropped out of the programme have been replaced. In schools where only one mentor is taking part in the programme *mentor replacement* is unambiguously defined as a school-level binary measure that signals whether the mentor that dropped out has been replaced with an additional mentor or the mentorship terminated.

In schools where more than one mentor is taking part in the programme, a mentor dropping out can either lead to:

- the mentor be replaced with an additional mentor (replacement),
- the ECT mentorship being taken over by another mentor already engaged in the programme (resassignment), or
- the mentorship being terminated (termination).

If all the schools with more than one mentorship follow a consistent replacement/reassignment pattern, i.e. all the mentors that drop out in a school are replaced or all the ECTs are reassigned, we can define for the purposes of the analyses mentor replacement as a school-level measure. However, if one or more schools simultaneously replace and reassign mentors, the measure will have to be defined in terms of both replacement and reassignment, and we will consider two binary measures:

1. At least one mentor was replaced by an additional mentor.
2. At least one ECT was paired with a mentor already taking part in the programme.

All the analyses of secondary outcomes can be performed considering regression models (logit) with the outcomes of interest as dependent variables. As is the case with the primary outcome model, a dummy that accounts for the randomisation stratifier (provider) will also be included as a covariate.

The equation for the secondary outcomes logit models is:

$$logit(p) = log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 intervention + \beta_2 provider + \epsilon$$

Where p is the probability of the outcome of interest: mentor retention, mentorship retention mentorship dropout, or mentor replacement, and the dependant variables as defined in the primary outcome analysis.

The logit models will be run in R (version 4.0.3).

Subgroup analyses

No subgroup analyses are planned for this trial.

Additional analyses

Taking advantage of having three points of data collection, corresponding to the three terms of the academic year, we will run additional analyses in order to better understand the profile of mentor engagement overtime. We will repeat the primary outcome analysis considering the engagement with training of mentors during the first term (i.e. considering just the data collected during the first term) and the engagement with training during the two first terms (i.e. considering just the data collected during the first and the second terms). This corresponds to running OLS models identical to the primary outcome OLS model, except for the dependent variable that is, in these cases, the school-level average length of mentorship training activities during the first academic term, and the average length of engagement during the first two terms.

Longitudinal follow-up analyses

No longitudinal follow-ups were planned for this trial.

Imbalance at baseline

To assess imbalance between intervention and control groups at baseline we will produce cross-tabulations of background characteristics of the schools in the sample. We will examine the following background characteristics: proportion of FSM eligible pupils within the school, if the school is rural or urban, type of school governance, and latest Ofsted rating. To run this analysis, we will link the schools taking part in the trial to the relevant information contained on the most up to date edition of NFER's registry of schools.

Missing data

The trial data is administrative data collected from the providers' MI systems. We expect missingness levels to be much lower than five per cent, in terms of data collection.

However, the dropout of ECTs from the trial can have an impact in terms of the computation of the primary outcome. The dropout of an ECT corresponds to the termination of the corresponding mentorship and the removal of the data corresponding to the mentorship from the analyses. We will evaluate if ECT dropout is potentially biasing the results of our primary outcome analysis by running logistic regression models to determine if ECT dropout is conditional on treatment allocation and also by means of a sensitivity analysis. The sensitivity analysis will build on a multi-level multiple imputation that can be implemented in R (version 4.0.3) using the packages MICE and smcfcs.

For each school that is affected by ECT dropout we will impute the value of the primary analysis outcome based on the school characteristics described on the imbalance at baseline section of this analysis plan, proportion of FSM eligible pupils within the school, if the school is rural or urban, type of school governance, and latest Ofsted rating, as well as provider and number of mentorships established at the onset of the trial. The results of the OLS models run with imputed values will then be compared with the results of the primary outcome analysis.

Compliance

No compliance analyses were planned for this trial.

Effect Size Calculation

As advised by the EEF 2018 guidelines, for the primary outcome analysis we will be reporting the effect size as Hedges' g. The effect size will be calculated according to the formula

$$\boldsymbol{g = \frac{\bar{o}i - \bar{o}c}{s^*}}$$

The numerator $\boldsymbol{\bar{oi} - \bar{oc}}$ is the difference between the intervention and control group in terms of the mean value of the outcome being assessed, and corresponds to the coefficient of the randomisation group dummy variable on the OLS regression model. The denominator $\boldsymbol{s^*}$ is the pooled standard deviation of the outcome and is given by the formula

$$s^* = \sqrt{\frac{(N_i - 1)s_{i+}^2 (N_c - 1)s_c^2}{N_i + N_c - 2}}$$

with N_i and N_C being the number of elements in the intervention and control groups, and s_i and s_c the standard deviations of the outcome measured in the intervention and control groups.

The associated confidence interval will be computed by multiplying the standard errors of the intervention group by the left-tailed inverse of the Student's t-distribution with a probability of 2.5% and the number of degrees of freedom associated to the intervention group. The confidence intervals for the standard errors will be converted to effect size confidence intervals using the same formula as the effect sizes themselves.

The effect sizes for the secondary outcomes analysis will be reported as odds ratios (OR). The odds ratios will be calculated according to the formula

$$\boldsymbol{OR = \exp(\beta_1)}$$

Where β_1 is the coefficient of the intervention/control dummy variable of the logit models.

Appendix

Code for randomisation

```
## School Randomisation-Stratified by Provider (Ambition Institute or Teach First)


###Packages

library(openxlsx)



#1. Set work directory

setwd("…")



#2.identify project

project<-"…"


#3.identify classification: c, r or p

classification<-"R"


#4. Number of the randomisation: 1st, 2nd, 3rd ...

randomisation<-1

randomisation<-as.character(as.roman(randomisation))


###5. Load data

Experiment=read.xlsx("…")

colnames(Experiment)


###Check for duplicates: no duplicates

sum(duplicated(Experiment$NFER.No))==0

###Check for missing unique identifiers: no missings

sum(is.na(Experiment$NFER.No))==0
```

```
###Identify stratification and unique identifier variables

#6.list the stratification variables
stratification<-list("Organisation")
n_strats<-length(stratification)

#4.unique identifier variable
ui<-"NFER.No"

###5. What time is now? (hh.mm)
time_now<-10.02

aux<-100*trunc(time_now)+100*(time_now-trunc(time_now))
set.seed(aux)
seeds<-sample(1:9999,size=(n_strats+2))


#Keep the original order of the columns
originalColOrder<-colnames(Experiment)

###Adding a variable that will allow for the recovery
##of the original order of the data frame rows later on
Experiment$originalRowOrd<-1:nrow(Experiment)

### Ordering Experiment by unique identifier
Experiment<-Experiment[order(Experiment[ui]),]

### Assigning a random order to the stratification
rands<-paste("rand",as.character(1:n_strats),sep="_")

for (i in 1:n_strats){
```

```
aux<-as.data.frame(sort(unique(Experiment[,stratification[[i]]])))
set.seed(seeds[1])
seeds<-seeds[-1]

aux[rands[i]]<-sample(1:nrow(aux))

Experiment<-merge(Experiment,aux,by.x=stratification[[i]],by.y=colnames(aux)[1])
}

###Randomise by unique identifier
set.seed(seeds[1])
seeds<-seeds[-1]
Experiment["rand_ui"]<-sample(nrow(Experiment))

###Reorder the rows of Experiment by rands and rancluster
rands<-c(rands,"rand_ui")
aux<-do.call(order,Experiment[rands])
Experiment<-Experiment[aux,]

###Assigning Control or Intervention Group
aux<-rep(1:2,times=1+round(nrow(Experiment)/2))
Experiment$grp<-aux[1:nrow(Experiment)]

rands<-c(rands,"grp")

aux<-data.frame(group=c("control","intervention"))
set.seed(seeds[1])
aux$randgroup<-sample(1:2)

Experiment<-merge(Experiment,aux,by.x="grp",by.y="randgroup")
```

```
##Returning the data frame to its original order

Experiment<-Experiment[order(Experiment$originalRowOrd),]


###Removing the variables that are no longer necessary

originalColOrder<-c(originalColOrder,"group")

Experiment<-Experiment[,originalColOrder]


###Put it out: type of document to be decided later

csvname<-
paste(paste(paste(paste(project,"Randomisation",sep="_"),randomisation,sep=""),classification,sep="
_"),

      "csv",sep=".")

write.csv(Experiment,csvname,row.names =F)
```