Visible Classroom

Behavioural Insights Team

Education Endowment Foundation

Amendments

February 2017
Jessica Heal

Following the addition of a second treatment year to the intervention which follows the year 5 student cohort in the second year, several changes have therefore been made to the original protocol:

- Page 15 – additional note on the implementation and process evaluation now taking place over two years, as opposed to one. The method will remain the same and merely be duplicated over the second year
- Page 16 – Table of engagement, cut off figure has been updated to reflect that teachers who submit fewer than 3 recording in one sitting will not receive a report
- Page 18 – personnel changes to the evaluation team, Daniel Carr and Jessica Heal have taken over from staff who have left BIT or transferred abroad
- Page 21 – updated timeline to reflect the intervention being implemented over two years and the additional process evaluation year

Evaluation Summary	
Age range	Years 5 and 6
Number of pupils	11,760 (42 students per year group x 2 year groups x 140 schools)
Number of schools	140 schools
Design	School-level randomised controlled trial
Primary Outcome	Combined KS2 English and Maths for years 5 and 6

Background

The intervention aims to improve student achievement outcomes by supporting teachers' professional practice development. The proposed mechanism by which this improvement occurs is via the provision of personalised feedback on teaching. This feedback is designed to encourage teachers to reflect critically on their teaching and develop their classroom practice, which would be expected to have flow on effects for their pupils' learning and attainment.

The Visible Classroom (VC) intervention involves teachers' audio recording lessons and receiving detailed feedback on their teaching practices. Teachers use a smartphone or tablet app to record lessons and upload them to their own personal profile. Once uploaded, teachers receive a transcript of their lesson along with some high-level descriptive statistics of their lesson. Having uploaded a comprehensive amount of recordings (typically five hours per week), the recordings are then analysed by the University of Melbourne using their teaching rubric. Teachers then work with mentors to reflect and develop their practice. While the original intervention involved real-time captioning in the classroom, the revised design streamlines the process so that teachers can easily record and upload their lessons to the VC team. The purpose of the intervention is for teachers to be reflective about their practices and to improve their teaching through personalised feedback. This document outlines the detailed protocol of the evaluation of this intervention.

Trial objective

The trial's objective is to test whether providing feedback to teachers on their audio recorded lessons has an impact on student outcomes for KS2 results in maths and English. Further details on the VC intervention are outlined in the intervention section, below.

Research questions

The primary research question this evaluation seeks to answer:

- Does the VC intervention, increase the educational attainment of Y5 and Y6 students (age 9-11) in combined KS2 maths and English scores?

Secondary research questions will consider whether the intervention has an impact on:

- The maths and English results for each year group independently;
- The maths and English results of Free School Meal (FSM) students; and
- The impact of students whose teachers score in the bottom third of the VC rubric, as measured at baseline.

Further details are outlined in the outcome measure section, below.

Design

The trial is a two-arm, school-level cluster randomised trial. The two arms are (1) 'business as usual' control arm, and (2) the Visible Classroom intervention. Approximately 140 schools will be recruited to this trial.

Randomisation

Following recruitment of schools by the Schools, Students and Teacher Network (SSAT) and baseline data collection, schools will be randomised to either the control or treatment arm. This will be conducted using data analysis and statistical software Stata. The randomisation will follow a two-stage process:

1. The schools will be stratified on the basis of the proportion of FSM students (split across the median sample proportion), 2010-11 KS1 Average Point Score (split across the median sample score), and whether schools were offered entry into the trial on a capped or uncapped basis[1].
2. A random number will be generated within each block to ensure that the proportion of FSM students, KS1 results, and number of and capped and uncapped schools are balanced across trial arms. We will use data from DfE's Performance Tables to determine the blocking characteristics, and SSAT will note which schools are offered capped entry during recruitment.

The randomisation will occur following recruitment and after baseline data collection. The Stata code used for the randomisation will be recorded in the final report.

Participants

The target number of schools for this intervention is 140. In order to be eligible, schools must have at least two teachers across Y5 and Y6. Prior to recruitment, interested schools will have a short conversation with SSAT, the team supporting the recruitment of schools. This will include completion of a short telephone survey where their eligibility will be assessed, their motivation for

[1] This stratification covariate was added in mid-August as a result of a decision during school recruitment to begin capping the number of teachers per school who could be involved in the trial. This decision was made to avoid a cost overrun in the delivery team's budget. As capped schools may select which teachers can be involved in this trial on the basis of merit or some other element correlated with student attainment, we made the decision to include this covariate in our stratification.

joining the intervention and their availability to attend a training day in early October 2016. Schools will be required to sign an MOU before enrolment into the study.

The eligibility criteria for schools includes:

- Upfront transfer of data to the evaluators including eligible student UPNs;
- A completed MOU;
- Agreement that teachers in both the VC intervention and control arm complete a survey at the end of the trial period;
- Inclusion of both Y5 & Y6 in the intervention, with a minimum of two teachers providing baseline data and attending the training session;
- Must have access to an iPad/smartphone and sufficient internet connection to upload recording (internet connection to be verified at baseline data gathering);
- The schools must not be using Visible Learning plus, a similar intervention designed by the University of Melbourne Project Team.

The sample is expected to include a higher proportion of disadvantaged schools, otherwise we expect it to be nationally representative. There will not be specific eligibility criteria for students.

Outcome Measures

The primary outcome measure will be combined maths and English KS 2 results for Y5 & Y6. If results are consistent, the analysis will be done on the combined Y5 & Y6 data. Taking into account potential issues of implementation for Y6, as it is a 'high-stakes' year, we will alternatively report two headline findings recorded one year apart for Y5 and Y6 individually.

Secondary outcomes will include:

- English and Maths separately for each year group (4 outcomes)
    1. KS2 English results for Y5
    2. KS2 English results for Y6
    3. KS2 maths results for Y5
    4. KS2 maths results for Y6
- FSM subgroup analysis for (2 outcomes)[2]
    5. Combined English and maths results for Y5
    6. Combined English and maths results for Y6

[2] If results are consistent the analysis will be done on the combined Y5/6 data. If not, then it will be done on just Y5 as the reason we added in Y5 was because of concerns that Y6 is a high stakes year and this may affect intervention delivery, as well as longer-term effects being of interest.

- Teacher subgroup analysis for (2 outcomes)[3] – This is a subgroup of potentially 'high impact' teachers who are predicted to be less competent according to the rubric for whom you might expect more change to occur.

    7. Combined English and maths results for Y5
    8. Combined English and maths results for Y6

Additional exploratory analysis will include:

- Dosage analysis (IV approach) of the combined time the teacher spends (1) recording lessons and (2) using the online feedback which is available for teachers on their online dashboard.
- Analysis using the survey of teacher practice to understand how self-reported changes in teaching practice moderate any effect on attainment.

As in the primary analysis, this exploratory analysis will involve combined English and maths results for Y5 & Y6 pooled, where results are consistent, otherwise they will be reported separately.

Intervention

Inadequate description of interventions is a perennial problem in a range of disciplines. A recently published framework – the Template for Intervention Description and Replication (TIDieR; Hoffman et al, 2014) – offers a standardised approach that we have adapted to suit school-based interventions. The intervention is described using this adapted template below, followed by the participant flow diagram.

Table 1 TIDieR Intervention Details

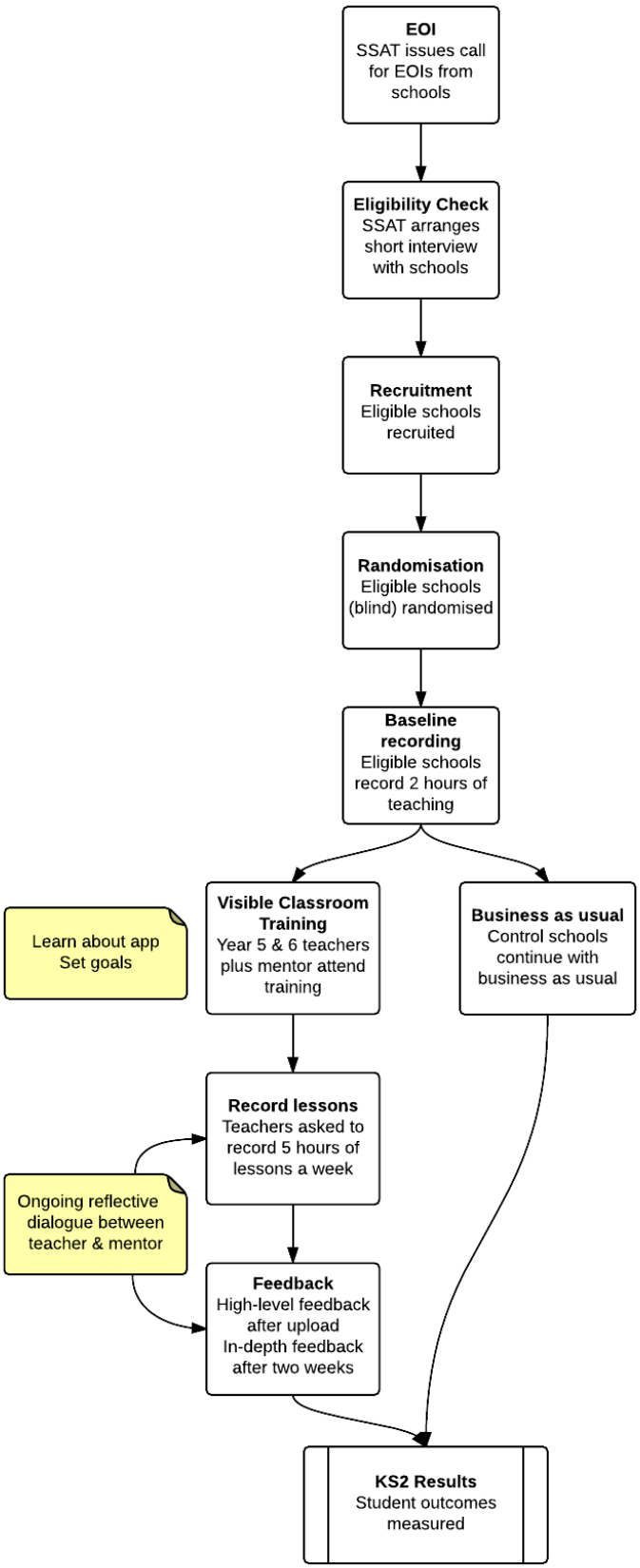
1	Brief name	The Visible Classroom (VC)
2	Why: Rationale, theory and/or goal of essential elements of the intervention	The intervention aims to improve student achievement outcomes by supporting teachers' professional practice development. The proposed mechanism by which this improvement occurs is via the provision of real-time, personalised feedback on teaching. This feedback is designed to encourage teachers to reflect critically on their teaching and develop their classroom practice, which would be expected to have flow on effects for their pupils' learning and attainment.
3	Who: Recipients of the intervention	Literacy-based subject teachers in regions XXX in Y5 & Y6

[3] If results are consistent the analysis will be done on the combined Y5/6 data. If not, then it will be done on just Y5 as the reason we added in Y5 was because of concerns that Y6 is a high stakes year and this may affect intervention delivery.

4	What: Physical or informational materials used in the intervention	Teachers use a smartphone or tablet app to record lessons and upload them to their own personal profile. Once uploaded, teachers receive a transcript of their lesson along with some high-level descriptive statistics of their lesson. Having uploaded a comprehensive amount of recordings (typically five hours per week), the recordings are then analysed by the University of Melbourne using their teaching rubric. Nominated mentor teachers will be encouraged to participate in an online webinar. In this interactive session they will be introduced to the principles of effective mentoring in education, and provided with practical strategies for how to use the feedback generated through the VC system to establish a meaningful and constructive professional dialogue with their mentee teacher. Participating teachers will be provided with a training manual on their attendance at the face-to-face training which includes information about the intervention aims and pedagogical model, how to use the app, and how to reflect on their teaching using the intervention materials (the verbatim lesson transcripts, the data dashboard and the feedback reports). The verbatim lesson transcripts are provided for download from the VC app in a word or pdf file. The data dashboard is also accessible through the app, and provides the teacher with key metrics about their teaching (e.g., number of pupil questions, pace of instruction) that is calculated from an analysis of the captioned lesson. This information is presented visually in the form of annotated graphs. The feedback report provides the teacher with a breakdown of their teaching over the 5 captioned lessons. Their lessons are coded against an evidence-based rubric of effective teaching practice. This feedback will be provided within a maximum of two weeks after the teaching takes place. Through this process, areas of strength and areas for improvement are signalled to the teacher both in written and visual form. Trends over time are also provided and their practice is compared against 'teachers like them' to provide a context/benchmark for their performance. Finally, they are provided with tailored suggestions for improvement, and some prompts for discussion with their mentor.

5	What: Procedures, activities and/or processes used in the intervention	During the face-to-face training, teachers will be introduced to the pedagogical model, the technological platform, and the ways in which they can use the feedback delivered to inform their practice. They will also take part in a goal-setting exercise. This training will be a mixture of presentations by the project team, interactive activities, and free time to explore the VC app and practice recording themselves speaking. During a one-hour webinar, mentor teachers will be engaged in an interactive conversation about effective mentorship processes, and provided with hypothetical scenarios, resources and ideas to support them to work with their mentee teachers in this project. To commence a captioned lesson, teachers will establish an audio link to the Ai Media captioning centre via the VC app, over a broadband internet connection via smartphone or tablet device. They will select record when they commence the lesson, and end the recording at the conclusion of the lesson. Trained staff will then generate a verbatim transcript of the lesson, from which the data dashboard is populated automatically through a series of algorithms. Teachers can access these 24 hours after the lesson through the VC app. A team of trained research assistants from the University of Melbourne then code the transcripts against an evidence-based rubric of effective teaching practice. Following the intervention, teachers in the intervention and control conditions will be sent an online survey focused on their own current teaching practices and experiences with VC.
6	Who: Intervention providers/implementers	The intervention is delivered in collaboration by three partners. *Ai Media* provide the technological platform for the captioning of lessons, provide the verbatim lesson transcripts, and populate the data dashboard with teaching analytics. *The University of Melbourne* deliver the training package to participating teachers and mentors, conduct the in-depth coding of lesson transcripts, and generate the tailored feedback reports. SSAT are responsible for recruiting schools, checking that schools are using the technology, and supporting them to do so effectively.
7	How: Mode of delivery	As detailed above, much of the intervention is delivered via technological platforms. Participating teachers attend a face-to-face training day, whilst their nominated mentors are encouraged to participate in an online webinar. Teachers record their lessons using the VC app, which can be accessed using a smartphone or tablet device. Their lesson transcript and data dashboard is also delivered through the app. After 5 captured lessons, the teachers receive their detailed feedback report via email, which they are encouraged to review along with their transcripts and dashboard with their nominated mentor face-to-face.

8	Where: Location of the intervention	The intervention is conducted in schools, with the exception of the face-to-face training which will be conducted off-site in a community venue.
9	When and how much: Duration and dosage of the intervention	Each participating teacher will receive 3 blocks of captioning and feedback, with 5 hours of lessons captured within each block. This will be delivered over the course of 3 school terms, within one academic year. For each hour they receive a verbatim lesson transcript, and a data dashboard. At the end of each block, they also receive an in-depth report on their teaching with tailored suggestions for improvement. They are also encouraged to spend a minimum of 1 hour with their nominated mentor at the end of each block. Thus in total, each teacher receives 15 transcripts and data dashboards, and 3 feedback reports, as well as 3 hours of collaborative professional dialogue with their mentor.
10	Tailoring: Adaptation of the intervention	The intervention involves tailored feedback to teachers and therefore it is adapted to the needs of each individual teacher. Additionally, the intervention will be piloted in a small number of schools, where feedback will be incorporated into the design of the intervention.
11	How well (planned): Strategies to maximise effective implementation	Teachers' adherence to the intervention will be tracked via the Visible Classroom app. Web analytics can provide evidence of teachers engagement with the intervention, including number of captioned lessons, reads of transcripts and access rates of data dashboards. Teachers will also be encouraged to record (via the app) when they meet with their mentor teacher. Teachers who are detected as having low levels of engagement with the intervention will be contacted by SSAT for additional support and trouble-shooting.
12	How well (actual): Evidence of implementation variability	Online activity on the dashboard and website will be measured and will act as a robust measure of implementation. This will be complemented by a process evaluation aimed at exploring the barriers to implementation and the mechanisms that were helpful.



Participant flow diagram

Power Calculation Specifications and Assumptions

These power calculations are based on the analysis of Y5 & Y6 separately.

- Sample size= 2,940 per arm. This is estimated on the assumption that 140 schools participate in the trial, with an average of 1.5 classes per year group and an average class size of 28 students (42 students in each school cluster).
- Randomisation is at the school-level.
- Intracluster correlation coefficient set at 0.15[4]
- As tracking will be possible, through Unique Pupil Numbers (UPNs), we expect attrition to be minimal, however below we do set out a process in the instance that significant data is missing. Additionally, participants' KS1 results will be incorporated to increase power along with the variables used for blocking during randomisation. We assume a correlation coefficient of 0.7 between KS1 and KS2 results[5].
- Hypotheses:
  - Null hypothesis: There will be no difference in combined maths and English KS2 scores between students whose teachers used Visible Classroom and those whose teachers did not.
  - Alternative hypothesis: There will be a difference in combined maths and English KS2 scores between students whose teachers used Visible Classroom and those whose teachers did not (i.e. a two-sided alternative hypothesis).
- Power: 80% ; Significance level: 5%. These figures are set to the standardised values denoted to power and statistical significance.
- We will have 2 trial arms (a treatment and control)

Below is the output from our power calculations to determine the minimum detectable effect size (MDES). As described above, we expect attrition not to be a significant risk due to UPN data being an eligibility requirement of schools, and therefore rely on an MDES of 0.139.

Level of attrition	**No. of schools**	**MDES**
No attrition; with KS1 regression	**140**	**0.139**

[4] This is conservatively based on estimates from Hedges & Hedberg (2007), denoting an ICC of 0.13-0.21 for schools. Hedges, L. & Hedberg, E. (2007). Intraclass correlation values for planning group-randomized trials in education. Educational Evaluation and Policy Analysis, 29 (1), 60-87.

[5] Source: https://educationendowmentfoundation.org.uk/public/files/Evaluation/EEF_Evaluation_Pre-testing_in_EEF_evaluations.pdf

FSM power calculations

Based on the same assumptions as above, we calculated the following MDES for FSM students. The calculation primarily varied by numbers per cluster (16.7% of students received FSM as determined by the average across Primary Schools on the EduBase dataset).

Level of attrition	**No. of schools** **(7 FSM students per cluster)**	**MDES**
No attrition; with KS1 regression	140	0.177

Analysis Strategy

Our primary analysis will concern our combined maths and English KS2 results. This continuous outcome measure will be estimated using a linear prediction model of the form:

$$Y_{ij} = \alpha + B_1 T_j + \boldsymbol{B_2 X_j} + B_3 S_{ij} + u_{ij}$$

i: student

j: school

where, Y_j is our outcome variable. This is a continuous outcome measure representing the combined maths and English KS2 score of the student. These scores are standardised by design as Key Stage outcomes.

T_j is a binary treatment indicator, set to one if the student's teacher received the Visible Classroom intervention (treatment), or zero otherwise.

α is a constant term. It can be interpreted as the response level of participants in the omitted category.

X_j is a vector of blocking variables.

S_{ij} is the student-level measure of KS1 results.

u_j is the individual residual clustered at the school level.

Secondary Analysis

Three groups of secondary outcome measures will be considered:

- English and Maths separately for each year group (4 outcomes)
    9. KS2 English results for Y5
    10. KS2 English results for Y6
    11. KS2 maths results for Y5
    12. KS2 maths results for Y6
- FSM subgroup analysis for (2 outcomes)[6]
    13. Combined English and maths results for Y5
    14. Combined English and maths results for Y6
- Teacher subgroup analysis for (2 outcomes)[7] – This is a subgroup of potentially 'high impact' teachers who are predicted to be less competent according to the rubric for whom you might expect more change to occur.
    15. Combined English and maths results for Y5

[6] If results are consistent the analysis will be done on the combined Y5/6 data. If not, then it will be done on just Y5 as the reason we added in Y5 was because of concerns that Y6 is a high stakes year and this may affect intervention delivery, as well as longer-term effects being of interest.

[7] If results are consistent the analysis will be done on the combined Y5/6 data. If not, then it will be done on just Y5 as the reason we added in Y5 was because of concerns that Y6 is a high stakes year and this may affect intervention delivery.

16. Combined English and maths results for Y6

In relation to the analysis of English and maths scores separately, the empirical strategy will be the same as the above, substituting the outcome variable, Y_{ij}, with each result separately (i.e. Y5 English, Y5 maths, Y6 English, Y6 maths).

For FSM students and for the teacher subgroup analysis, the empirical strategy will remain the same with the addition of interaction analysis considering the interaction between the intervention and the FSM students and teachers who are deemed high on adherence to the rubric at baseline.

Additional exploratory analysis will include:

- Dosage analysis (IV approach) of the combined time the teacher spends (1) recording lessons and (2) using the online feedback which is available for teachers on their online dashboard.
- Analysis using the survey of teacher practice to understand how self-reported changes in teaching practice moderate any effect on attainment.

Missing data analysis

Due to the nature of the outcome variable (i.e. national standardized tests), we do not foresee missing data to be problematic. However, if this is the case we present the procedure for sensitivity analysis based on the missing data. Missing data presents a problem for analysis, whether a pupil is missing a value for an outcome variable (post-test score) or for covariates (e.g. pre-test score). If outcome data is 'missing at random' given a set of covariates then the analysis has reduced power to detect an effect; if data is 'missing not at random' (for example, differential dropout in the intervention and control groups for unobserved reasons) then omitting these pupils (as in the primary 'completers' analysis) could bias the results. Conducting sensitivity analysis through imputing missing data could improve the robustness of the analysis and examine how sensitive the results are to alternative assumptions.

Every school that will be randomised will complete a post-test in the form of KS2. However, some individual pupils within those schools may not be present due to absence or because they left the school. Randomisation will be stratified to account for differences across some observable characteristics (proportion of FSM students and school size) with the objective of avoiding differential attrition. Post-trial and pre-analysis balance checks will be conducted using regression analysis to determine whether randomisation was successful across the groups and the observable. Each observable will be taken in turn and analysed against the groups. If statistically significant differences are observed based on the stratification variables (i.e. proportion FSM, school size and KS1 results) these characteristics will be controlled for during the analysis.

Where >5% of the sample for each school is missing or where imbalance is found between the groups, we will perform sensitivity analysis and impute the missing scores for those pupils using imputation, conditional on KS1 National Pupil Database (NPD) data. The correlation between KS1 and KS2 score is high, so the imputed values will be estimated with a high level of precision.

Standardised effect size

This will be calculated through standardization of the outcome, such that the mean will be 0 and the standard deviation will be 1. The purpose of accounting for any clustering effect is to minimise a possible false conclusion by using robust standard error, which is not based only on residual variance, but on total variance (Verbeke and Molenberghs, 2005). Effect sizes for cluster randomised and multi-site trials, as in the current instance, should incorporate this important statistical principle

by using total variance to avoid inadvertent over-estimation of effect sizes when there are substantial differences between schools.

Confidence intervals

We will estimate a 95% confidence interval alongside the standardised effect size to give the precision with which the effect size has been estimated. The upper and lower bounds of the confidence interval will be calculated as the effect size plus/ minus the product of the critical value of the normal distribution (≈ 1.96) and the standard error of the group indicator coefficient estimated from the multilevel model.

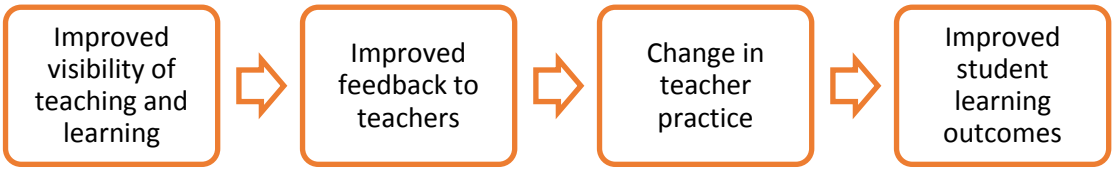
Implementation and process evaluation methods

The implementation and process evaluation will explore (1) the theory of change related to the intervention, (2) the barriers to, and conditions for, effective delivery of the intervention, and (3) the related costs of implementation incurred by schools and teachers. Each of these sections will cover, the specific research question and the methods and tools for measurement. The process evaluation will capture both year 1 and year 2 of the evaluation and follow the same structure each year.

Theory of Change

This section of the process evaluation considers the specific mechanism at play in the intervention to initiate change in our outcome measure - student attainment. It is imperative is to determine the theory underpinning why exactly a change would occur. The intention here is to generate an initial theory to test, while also considering an open-ended approach so as to not overlook any key insights from the process evaluation.

The theory of change, as described in NatCen's Evaluation Report[8], asserts a cyclical four-step process stating that "Changes in teacher practice and mindset, such that teachers become evaluators of their own teaching … leading to increased attainment, particularly amongst disadvantaged students"[9]:



This evaluation will consider the same theory of change as defined by NatCen and the University of Melbourne. Data to answer this question will be collected through an online survey of the teachers in the intervention condition post-intervention as well as in the control group. The survey will be administered to teachers and mentors in both the treatment and control arms to help establish whether any change in children's outcomes occurs through changes in teacher practice. The University of Melbourne will provide teachers in control schools with a low-cost incentive to complete the survey (e.g. invitation to a celebration event after Y5 complete KS2). The online survey will be sent to teachers in both groups at the same time to minimise the costs of surveying control group teachers as well as intervention group teachers. This will include measures such as a locus of control survey, a survey related to the teaching rubric set out by the University of Melbourne and a self-reflection survey (this survey is attached as an addendum to this document in PDF format).

Engagement in the intervention

The intervention has many stages of engagement, in order to ascertain barriers and conditions to implementation. Critical features of the intervention, with a sliding scale of engagement, are outlined below. It considers the rounds of lessons and recordings. The app

[8] https://educationendowmentfoundation.org.uk/uploads/pdf/Visible_Classroom_(Final).pdf

[9] Reference to student access to VC has been removed since the intervention design has been amended following the pilot evaluation.

and website analytics will allow us to monitor fidelity and dosage of the intervention and subsequently identify low and high adherence schools.
In order to ensure the integrity of the data is maintained, it is important to outline the minimum standards a school meets for us to ascertain they are in receipt of the intervention. In this case, we will not consider a teacher to be on-treatment if they do not receive 2 detailed feedback reports from UoM. This would mean that to be on treatment teachers have to record and upload at least 6 lessons throughout the year. This is based on the premise that fewer than 3 reports will not generate a report.

Table of Engagement

Submit lessons	Round 1	Round 2	Round 3	Total submissions
High	5	4-5	4-5	13-15
Medium	3-4	3-4	3-4	7 - 12
Low	0-2	0-2	0-2	0 - 6

Barriers to and Conditions for Effective Delivery

Here, the evaluation seeks to determine what the barriers to effective delivery are, and what the conditions for effective delivery are. In order to answer these questions, the evaluation will employ a multi-pronged approach, using pre-existing channels in order to reduce burden on respondents and ensure high quality data. This will involve:

- Semi-structured interview with SSAT re: recruitment of schools, delivery of intervention, and support requirements;
- Semi-structured interview with school mentors re: delivery of intervention and support requirements;
- Semi-structured interview with high and low adherence teachers re: delivery of intervention, satisfaction with VC and support requirements;
- Web analytics to monitor usage of VC app and website dashboard;
- Brief survey to all teachers' post-intervention, to capture satisfaction, perceptions of usefulness and feedback on the intervention.

The interviews with school mentors and teachers will cover 6-8 schools in total. They will be chosen based on a mix of low and high adherence to the instructions. The app and website analytics will allow us to monitor fidelity and dosage of the intervention and subsequently identify low and high adherence schools. Additionally, the evaluators will attend a training session and review the materials provided to teachers.

Costs of Implementation

Description of how cost data will be collected, and a breakdown of the costing scope (i.e. whether or not teacher time, administration etc. are costed).

Information will be requested from the University of Melbourne to capture the cost of teacher training, recruitment of and transcript analysis by RAs, any additional materials or support provided to schools.

For mentors and teachers, we will use usage data from the platform to determine the amount of time a teacher typically spent engaging with the intervention. We will additionally request for information on activities related to the intervention they completed outside of interacting with the platform and app.

From mentors, we will request information via survey on the amount of time spent in relation to the intervention (e.g. training, discussions with teachers, additional queries).

From teachers, we will request information via survey on the amount of time spent in relation to the intervention (e.g. training, discussions with mentors, time added to lessons due to recording equipment, if any, time spent on the website dashboard).

Ethics and registration

- Ethical approval will be requested through BIT's Academic Advisory Panel and the University of Melbourne Ethics Committee.
- Opt-out consent will be requested from parents on behalf of students, as recommended in the EEF guidance report.
- The trial will be registered at the International Standard Randomised Controlled Trial Number (ISRCTN) (register the trial at: www.controlled-trials.com)

Data management

Randomisation code and associated outputs will be securely saved.

Personnel

Project Team

- John Hattie, University of Melbourne
- Janet Clinton, University of Melbourne
- Anne-Marie Duguid, SSAT

Evaluation Team

- Michael Sanders, The Behavioural Insights Team
- Daniel Carr, The Behavioural Insights Team
- Jessica Heal, The Behavioural Insights Team

Risks

Risk/Issue	Detail	Resolution	Date/ Solver
Insufficient school recruitment	An insufficient number of schools may be recruitment to adequately power the study.	SSAT have extensive experience in recruiting schools for trials and additionally have a very large network of school to contact. The likelihood of this becoming an issue is minimal at present. Additionally, BIT has experience in school recruitment and retention and will support the project team, if needed.	March - September 2016/BIT.
School dropout post randomisation	Schools may initially sign-up but withdraw after assignment to condition.	Schools will be required to share student UPN to track to the NPD in advance of the intervention being rolled out. Additionally, schools will be randomised after recruitment but in advance of baseline data capture. This is in order to notify intervention schools to place a hold on a date for the training day. If there is drop out prior to the training day that leads to imbalance, schools will be re-randomised.	September 2016/BIT
Lack of fidelity to the protocol.	Non-adherence to the protocol and associated instructions for teachers may dilute the effect of the intervention.	Dosage and fidelity will be tracked via the submission of lesson recordings. Instrumental variable analysis can allow us to dissect the impact of dosage on the outcome variables. Although it may impact the external validity of the trial, there is also an option of monitoring use of the app and dashboard so that low-fidelity teachers can be prompted.	Post intervention/BIT

Missing outcome data.	Outcome data may not be collected for all participants.	As we will be using the NPD the likelihood of this will be minimal. To mitigate this issue, we will be requiring schools to send student UPNs to BIT in advance of randomisation, as a requirement to participate. Additionally above we describe the analysis strategy to overcome issues in this regard.	May - September 2016/BIT
Withheld consent to link to NPD	Using the NPD to create school-level covariates should not require parental consent. Using the NPD to identify FSM status and KS2 outcomes requires opt-out consent, as per the EEF's document on "Consent and the Data Protection Act: Advice for evaluators".	In line with Worth et al. (2015) and the EEF guidance document, we anticipate that opt-out consent will be sufficient. Given the low burden that this imposes on schools and teachers, we believe we should be able to provide the necessary information to parents in treatment and control schools and do not anticipate high or non-randomly varying levels of opt-out.	September 2016/BIT

Timeline

Date	Activity	Responsibility
April 16 – Sep 16	Recruitment of schools	SSAT
March 16 - April 16	Ethical clearance	UoM/BIT
April 16 – Sep 16	Eligibility interviews with schools	SSAT
Sep 16	Randomisation	BIT
Sep 16 - Oct 16	Training with schools	UoM
Sep 16 – June 18	Full rollout of intervention	UoM
May 17 - July 17	Year 1 Process Evaluation	BIT
May 18- July 18	Year 2 Process Evaluation	BIT
Nov 18	First draft of evaluation full report	BIT