Evaluation Summary	
Age range	Secondary (Key Stage 3, Years 7 and 8)
Number of students	11000 (based on an average of 100 students per school)
Number of schools	110 schools
Design	Two arm (2-year) cluster randomised control trial
Primary Outcome	Maths attainment gains as measured by a maths test at end of Year 8

MANCHESTER 1824
The University of Manchester

Independent Evaluation of Project ICCAMS Maths

A two-arm 2 year Cluster Randomised Control Trial

TRIAL PROTOCOL

Chief Evaluator: Dr Maria Pampaka (MP)

Protocol Version: 4

Protocol Date: 17 December 2015

ISRCTN

Funder: Education Endowment Foundation

Authorship: Maria Pampaka, Julian Williams, Graeme Hutcheson

Note/Acknowledgement: The **ICCAMS Maths** study was proposed by the Investigator Team (from Nottingham University, led by Professor Jeremy Hodgen, with the subcontracted Team from the Centre for Evaluation and Monitoring (CEM, Durham University, led by Victoria Menzies)). Many of the main study design elements were proposed by the Investigator Team. These and many other aspects of the study and the detail of the evaluation are determined collaboratively between the three teams and EEF. The independent randomisation, data collection and analysis of the ICCAMS Maths study are the responsibility of the independent Evaluation Team, led by Maria Pampaka. The purpose of this document (the Evaluation Protocol) is to describe that evaluation process and, in line with EEF policy, is formally authored by the Evaluation Team, but in order to do that well the study research proposal has been incorporated within this evaluation protocol.

Table of Contents

BACKGROUND: THE ICCAMS STUDY

The ICCAMS Maths intervention was developed in a recently completed 4.5 year ESRC-funded project[1], which focused on improving teaching and learning in Key Stage 3 mathematics. ICCAMS Maths is designed to teach two mathematical areas that are a key part of the Key Stage 3 curriculum, but which cause particular problems to students: algebra and multiplicative reasoning. The teaching development programme is comprised of 40 lesson plans, 20 associated assessment pre-tasks, and an extensive teacher professional development (PD) programme.

The lessons are designed to help teachers use formative assessment (and feedback) in mathematics, helping them to identify the problems students struggle with and how to address them. Activities are set in contexts that students can engage with, are collaborative, and use visual representations to help develop understanding. The lessons are designed to address the key conceptual ideas underlying the Key Stage 3 mathematics curriculum in algebra and multiplicative reasoning, and are aligned with the revised mathematics National Curriculum.

The current study aims to compare the effect of the ICCAMS Maths intervention (when delivered 'at distance' through the institutional network of Maths Hubs) to a 'business-as-usual' control group in a cluster randomised controlled trial, and with a particular focus on addressing the mathematical learning needs of low attaining students in deprived socio-economic contexts.

During the first development / pilot year of the project, the ICCAMS Maths intervention will be adapted (i) for teaching over two years (rather than over one academic year as in the original project) and (ii) to provide particular support for low attaining students and their classes. In addition, the University of Nottingham team will develop material that more explicitly describes the ICCAMS Maths PD programme so that it can be delivered independently and with a degree of fidelity (including materials to train and support the PD leads). As a result, it is envisaged that, in June 2016, the evaluation protocol is likely to be slightly amended to reflect these adaptations.

THE EXISTING RESEARCH EVIDENCE

The initial ICCAMS ESRC study (ICCAMS 1) aimed to investigate ways of raising students' attainment and engagement with the use of formative assessment in order to inform teaching and learning of mathematics in secondary school. The focus was on the mathematical areas of algebra and multiplicative reasoning for Key Stage 3, which are considered to be the cause of particular problems for students. The first phase of the study involved a longitudinal national survey of Year 9 students which used tests first developed in the 1970s under the framework of Concepts in Secondary Mathematics and Science (CSMS) in order to provide up-to-date empirical evidence on (i) current lower secondary students' understandings of and difficulties with algebra and multiplicative reasoning, (ii) rates of progression across

[1] The ICCAMS Maths intervention was developed in the ICCAMS 1 project (Increasing Competence and Confidence in Algebra and Multiplicative Structures), which was funded by the Economic and Social Research Council (ESRC), grant reference RES-179-25-009 (2008-2012). For more information, see: http://iccams-maths.org/

Key Stage 3 (KS3), and (iii) differential performance across the cohort. In addition, this survey enabled a comparison of students' understanding over time. This comparison provided evidence of a decline in students' understanding in algebra and ratio since the 1970s (Hodgen, Brown, Küchemann, & Coe, 2010; Hodgen, Coe, Brown, & Küchemann, Under review). In response to this, the 'ICCAMS 1 team' suggested the need for a more topic-focussed formative assessment approach to the teaching of these areas (Hodgen, Brown, Kutchemann, & Coe, 2010). Research suggests formative assessment is an effective approach to increasing attainment and engagement (e.g., Black & Wiliam, 1998; Black et al., 2011). Indeed, although Formative Assessment is not currently included amongst the topics in the EEF Toolkit, three key elements of FA are amongst the approaches with the highest impact: feedback, peer tutoring and metacognition/self-regulation.

However, despite widespread take-up of FA nationally and internationally, there is a developing body of evidence that teachers have considerable difficulties implementing these ideas (e.g. Smith & Gorard, 2005). This may be because formative assessment has been described vaguely and is thus difficult to implement, or at least it is difficult to be sure when it is being authentically implemented (Bennett, 2011). It may also be – particularly for secondary contexts where pedagogy is formulated as Pedagogical Content Knowledge - because formative assessment has largely been described generically rather than in subject-specific terms (Watson, 2006). There is evidence that teachers find it more straightforward to implement the generic and more procedural techniques of formative assessment such as traffic lights, and 'no hands up and lolly sticks' practices (Marshall & Drummond, 2006). Whilst these techniques can be a useful means to an end, they do not embody the key aspects of formative assessment that have been shown to be effective. For example, the evidence on effective feedback indicates that it should be specific to the task, yet metacognitive in purpose (e.g., see the EEF Toolkit, and Wiliam, 2007). Teachers' ability to use formative assessment in mathematics is therefore limited by their knowledge about key mathematical ideas, how they develop in context, and the likely progression of student learning in them. Thus if teachers focus on teaching mathematical procedures they may find it difficult to see what is causing problems for students in mastering and applying these, and may thus have difficulty responding to the students' difficulties (Hodgen, 2007; Watson, 2006). What is needed – in conclusion – is mathematically and task situated formative assessment practices: this is what ICCAMS Maths provides.

In developing this mathematically oriented approach to formative assessment, the ICCAMS approach was guided by the following principles, drawn from the research literature on mathematics teaching and learning:

- To set activities in *realistic* contexts (e.g., Streefland, 1991). By realistic, we do not mean that all the activities are set in real life contexts that students may have encountered, but rather contexts that the learners can imagine and engage with.
- To make *connections* between mathematical ideas (e.g., Askew et al, 1997).
- To encourage students to *collaborate* and *talk* (e.g., Slavin et al, 2009).
- To use *multiple representations*, such as the Cartesian graph or the double number line, to help students better communicate, understand and connect mathematical ideas and to help teachers appreciate learners' difficulties (e.g., Gravemeijer, 1999).

The ICCAMS approach to PD is informed by the literature on teacher professional development (Adey, 2006; Adey, Hewitt, Hewitt, & Landau, 2004; Cordingley, Bell, Evans, & Firth, 2005) and by the same principles that guide the ICCAMS lessons.

In the original ESRC-funded study, the ICCAMS approach was evaluated in the third phase through an intervention study with a group of Year 8 students in a wider group of schools. This intervention study showed that ICCAMS students made greater progress than a matched control group, a gain equivalent to typical growth in scores over a year (Hodgen, Coe, Brown, & Kuchemann, 2014, p. 171).

THE PROPOSED ICCAMS MATHS INTERVENTION

The initial project team, (ICCAMS-1) in collaboration with a group of teachers, developed a series of 40 research-informed lesson plans, associated assessment pre-tasks and professional development activities aimed to help teachers to integrate formative assessment within the secondary mathematics curriculum. These 40 lessons and 20 assessment starter activities were split equally between algebra and multiplicative reasoning. The lessons are organised in pairs: two closely linked lessons together with a related assessment pre-activity; thus FA is contextually linked to the situated task and mathematics pedagogy. The lessons are organised in five extended blocks themed as follows:

1. Symbols, functions and the Cartesian graph (Algebra focused)
2. Models *of* and *for* multiplicative reasoning (Multiplicative Reasoning focused)
3. Manipulation and equivalence of expressions (Algebra focused)
4. Fractions, division and ratio (Multiplicative Reasoning focused)
5. Linking algebraic and multiplicative ideas

The intervention aims to provide opportunities for teachers to collaborate in school as well as in professional development (PD) sessions which are centred around the assessment starters and lessons, as described below:

- In the PD session ***before*** teaching the lesson, teachers do (and extend) the mathematical tasks that the students will do, consider the possible difficulties students will have (including errors and misconceptions) and collaboratively plan how they will teach the lessons in their own classes, given the guidance of the lesson plans etc.
- In the PD session ***after*** teaching the lesson, teachers reflect on the students' learning, the potential for its development in 'generalisation', the mathematics involved, the formative assessment strategies used, and how the lesson fits within the sequence of lessons. They also consider possible adaptations of the lesson.
- Some lessons are dealt with in more depth than others, but sessions are designed to ensure that during the project the teachers consider the key ideas underlying Key Stage 3 algebra and multiplicative reasoning, and how they are developmental in practice in their classrooms.
- Additionally, teachers read and discuss some research relating to ICCAMS (usually presented in a short professionally focused publication), learn to use the GeoGebra graphing package, and consider ways to use similar approaches in mathematics lessons more generally.
- In each PD session, teachers observe (implying structured observation) and discuss a video either of a lesson or of a small group of students grappling with a key idea in algebra or multiplicative reasoning.

The ICCAMS Maths intervention will be adapted for teaching the 40 prototypical lessons over the first two years of secondary school: Year 7 and Year 8. The CPD-pedagogic model of this intervention involves a minimum of 2 teachers from each school deemed to have potential and recommended as such by the school, (with one of them being ideally the Maths Head of Department or someone of clear influence for the school 'cascade') attending the formal PD with these teachers responsible for cascading the training to their colleagues in school. It should be noted that materials for cascade training will be provided for the core teachers. Over the two years of the project there are nine PD sessions for teachers which will be organised and led by the PD consultant in each area. Teachers will be provided with a handbook containing details of the intervention and lessons plans for the 40 lessons involved in the intervention. It is likely that the order of lessons and structure of the PD sessions will change during the development and pilot stage to reflect the particular multiplicative needs of Y7. The protocol will be revised accordingly by June 2016.

One aim of this study is to describe the ICCAMS Maths PD intervention in such a way that it can be delivered without direct support from the original developers, the University of Nottingham team. Hence, the Durham team will be responsible for leading the trial and the training for PD leads, although the University of Nottingham team will provide expert advice and support. In addition, the Durham team will be responsible for collecting data on implementation fidelity and for maintaining contact with participating schools (in conjunction with the PD leads).

This CPD model has implications for the implementation of the intervention (i.e. it is designed to be relatively feasible for schools nationwide) and it also creates opportunities for the evaluation to consider whether different implementations of PD led to differential effects (e.g. greater effects for those directly involved in the PD than cascade teachers).

METHODS

The ICCAMS-2 study will be implemented in two main phases. Phase 1, the Development and Piloting will be led by Nottingham, whilst Phase 2, the Main Trial, will be led by Durham and the Evaluation will be led by Manchester.

ICCAMS 2 DEVELOPMENT AND PILOT PHASE

Phase 1 is planned to take place between June 2015 till July 2016 with a primary objective to design and pilot the research methodology. The particular aims of this stage stem from the teams' interest in better understanding (a) the specific difficulties that low attaining students face in mathematics classrooms, (b) how the ICCAMS Maths intervention can be better adapted to suit these students' learning needs, and (c) how to support teachers to more fully adopt the ICCAMS Maths approach and lessons with low attaining students. During this stage a replicable model of PD for the intervention needs to be developed and described (Durham). Finally, and more relevant to the evaluation, we aim to develop and evaluate the fidelity measures, as well as pilot the other relevant instruments for the secondary outcomes. The main objectives to be accomplished within this phase along with roles and actions are listed below:

- Development and adaptation of approaches to specifically support the implementation of the ICCAMS Maths pedagogic methods with low attaining groups (Nottingham).
- More explicit description of the ICCAMS Maths PD programme so that it can be delivered independently and with some sort of metric of fidelity. (Nottingham with support from Durham.)
- Development of a cascade model of the delivery of PD to other teachers in school (Nottingham with advice from Durham).
- Adapt the ICCAMS intervention for delivery over two years rather than one as in the original project (Nottingham).
- Develop and pilot fidelity evaluation and measures for the intervention (Durham and Nottingham, with support from Manchester).
- Pilot MALT tests for calibrating the sub-measure of multiplicative reasoning and algebra (Manchester).
- Pilot student attitudinal measures and teachers' attitudes and practice (Manchester).

The newly developed training model for PD Leads, the teacher PD and the new resources will be piloted in a piloting phase with eight to ten schools in one area (close to Nottingham) between January 2016 and July 2016. Data will be collected through (a) interviews with students in small groups about mathematics, (b) evidence from students work (via examples), (c) interviews with teachers and students about the ICCAMS-2 intervention, (d) observations of lessons and PD sessions, and (e) piloting of fidelity and other measures to be used for the main trial with students and students.

It is expected that a key outcome of Phase 1 will be a revision to the intervention which will lead to a revision of the outline of the ICCAMS-2 intervention and consequently the evaluation protocol.

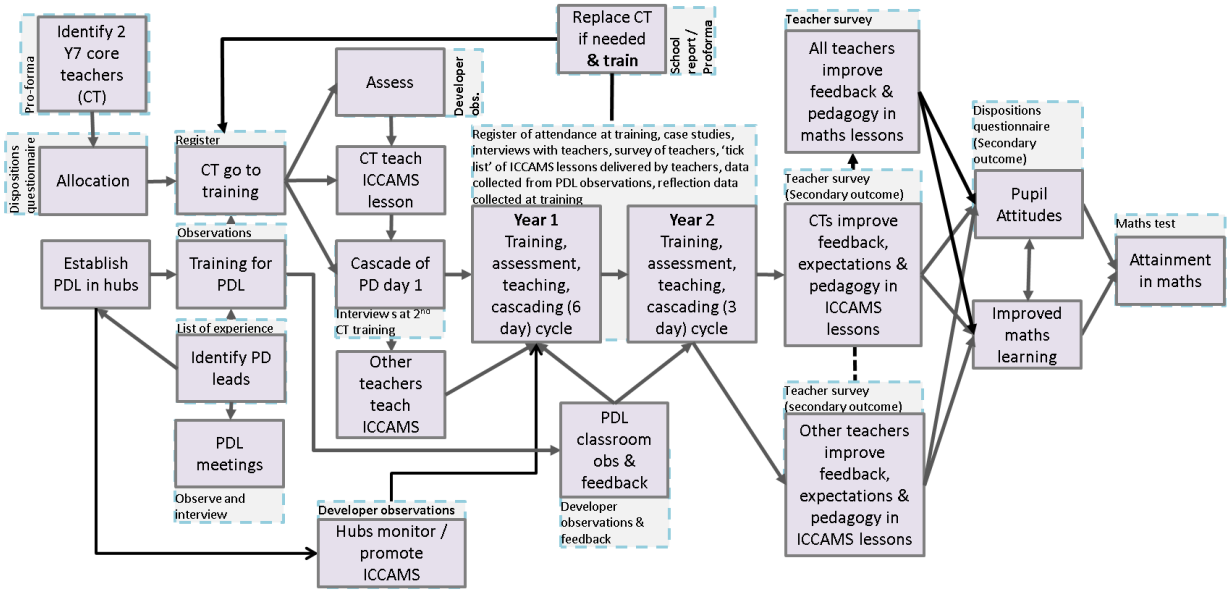

Figure 1: Logic Model for ICCAMS2

THE ICCAMS 2 MAIN TRIAL DESIGN

The main trial will be a cluster randomised controlled trial which will then run for two academic years with the intervention schools running the intervention for all students

in Year 7 initially and for the same students again when they are in Year 8 (September 2016 to July 2018). The roles in the trial will be as follows:

- Durham will recruit schools from five areas focused around Maths Hubs and PD Leads (with support from Manchester and Nottingham).
- Manchester will randomise schools within each hub to intervention or 'business as usual' control in July 2016.
- Durham will be responsible for the delivery of the intervention and maintaining contact with control schools during the project (with guidance from Nottingham).
- Manchester will be responsible for conducting the primary and secondary outcome assessments which will be collected at the end of the second academic year. They will also conduct an implementation and process evaluation during the year (with some support and data from Durham).

RESEARCH QUESTIONS

The evaluation is designed considering the following principal question:

"Does the ICCAMS-trained teaching practice improve students' learning outcomes in Year 8, as compared to 'business as usual' teaching practice?"

In addition, the evaluation will investigate and where possible measure the effects of ICCAMS Maths intervention on (a) changes in teaching / pedagogy, (b) changes in learners' algebra and multiplicative strategies and reasoning, and c) changes in students/learners' disposition towards mathematics.

The following research questions (RQ) will guide the impact evaluation:

Primary Question:
RQ1: Do students in schools implementing ICCAMS Maths over a two year period demonstrate improvements in overall mathematical attainment compared to students attending control schools?

Secondary Questions:
RQ2: Do students in schools implementing ICCAMS Maths over a two year period demonstrate improvements in attainment in algebra (2a) and multiplicative reasoning (2b), compared to students attending control schools?

RQ3: Are effects on attainment different for students eligible for FSM? If so, how?

RQ4: Do students in schools implementing ICCAMS Maths over a two year period change their attitudes/dispositions to learn mathematics compared to students attending control schools?

RQ5: Is there an interaction between fidelity and attainment for the treatment schools?

Implementation and Process Evaluation Questions:

The main questions to be addressed through this process evaluation are listed below:

RQ6: How, and to what extent compared to control teachers, do teachers offered the ICCAMS PD change their perceptions of teaching practice (regarding feedback, expectations, and related pedagogy)?

- RQ6a: Are any changes in perceptions of practice moderated by the way PD is offered (i.e. cascade via core teachers)?
- RQ6b: Are these perceptions of teaching practice related to students' learning outcomes (both attainment and attitudes)?

RQ7: How, and to what extent do the involved stakeholders (e.g. PD Leads, schools and teachers) practise and adhere to the principles, guidance and materials?

- RQ7a: Are there differences in fidelity between Core and Cascade teachers?
- RQ7b: How frequently do teachers claim they implement the materials and for how long?
- RQ7c: What are the contextual factors that afford or constrain the quality of implementation and the cascading in school?

RQ8: How do learners engage with the key metacognitive, formative practices on which ICCAMS Maths is predicated, and how do contextual classroom factors and pedagogies mediate these?

RQ9: What is the nature of the changes to pedagogy required to explain effectiveness in terms of Learning Outcomes, and how are these mediated by use of the ICCAMS materials and PD? What are the implications for improvement of these materials and the PD intervention?

RQ10: What relevant mathematics and PD systems and practices are in place in schools randomly allocated to continue 'business-as-usual' control group?' And how do these associate with successful learning outcomes identified at classroom or school level?

IMPACT EVALUATION DESIGN

A 2-year cluster-randomised trial will be utilized. Participating schools will be the unit of randomisation.

RANDOMISATION

Random allocation will be at the school level and will take place in July 2016. As randomisation should be done within each regional hub with a maximum of 30 schools to randomise we will be using block stratified randomisation (Torgerson & Torgerson, 2008). In order to ensure balance in regards to previous attainment and proportion of FSM, blocks will be defined by the proportion of students in each school to achieve 5 A*-C in the 2015 GCSE examinations (above median and below median) and the proportion of students in each school to be eligible for Free School means (above and below median). This means that there will be up to 4 blocking variables (or strata) made up of the combinations of these two variables. Randomisation of schools within regional Maths Hubs (to achieve a 50:50 allocation) will be performed with the following steps: (a) Each school will be assigned a random generated number; (b) schools will be sorted by blocking variable and within each block, by the random number, (c) the first school will be randomised to treatment or control; (d) subsequent schools will be assigned in alternate outcomes from previous schools. This procedure will be repeated for each of the 5 hubs. The details of the randomisation process will be recorded (both school lists and employed tool for the

algorithm) and the outcome will be shared with the delivery team. The schools will then be informed of their random allocation at the beginning of the academic year and after the students have completed the first 'Disposition' Questionnaire (i.e. pre-survey. Information will be circulated via email (Durham, copying in Nottingham and Manchester – for transparency).

Schools allocated to the intervention arm of the trial will be trained and supported to implement the ICCAMS over a two-year period. Schools allocated to the control arm of the trial will be encouraged to continue practice as usual during the same period. £1500 will be offered to control schools as compensation for time and to avoid attrition: £500 at the beginning of the trial (upon provision of UPNs and completion of pre-survey) and £1000 following the completion of post-test measures. Schools in both arms of the trial will be required to sign a memorandum of understanding before randomisation, committing them to comply with the evaluation protocol whichever arm they are allocated to.

PARTICIPANTS (SCHOOL AND STUDENTS SAMPLE)

Eligible **schools** should be mainstream English state secondary schools with more than two class intake for Year 7 (ideally not in special measures) and with, ideally, higher than average levels of FSM eligibility. Eligible schools should not be involved with the Maths Hub's KS3 Multiplicative Reasoning project[2]. Schools will only be eligible to take part in the study if they agree to all of the study requirements outlined in the Memorandum of Understanding (MoU) with Schools (Appendix to be added with the June revision). The trial schools will be recruited (by Durham, supported by Nottingham and Manchester) in five regionally-based groups to facilitate the hub-based PD (The number of schools also taking part in the Schools, Students and Teachers Network SSAT trial, "Whole school Embedding Formative Assessment Project" or any other special program deemed related should be minimised). It was suggested to target three areas that are currently underrepresented in EEF trials: the North West (excluding Manchester), the North East (including Newcastle) and the East Midlands. All five areas will be chosen based on relationships and networks with the Maths Hubs in the region, location of potential PD leads and in coordination with the National Centre for Excellence in the Teaching of Mathematics NCETM. It is anticipated that the protocol will be revised by June 2016 with the final regions and Hubs to be involved.

110 schools (to ensure that the trial is sufficiently powered with a low level of school-level attrition) will take part in this trial (55 in each arm). The schools would be spread across the five areas with up to 30 schools in each area. All schools within these areas will be invited (in writing) to the recruitment events in each area. Durham and Nottingham will work with the maths hubs in each area to increase prominence and reach of the project to support recruitment. PD leads will also support with the recruitment of schools to the project.

All **students** in Year 7 at the beginning of the 2016/17 school year are the target cohort (excluding those without parental consent). We estimate a sample of 11,000 students, based on an estimated average of 100 per school. Since the target year is Year 7, it will not be possible to collect UPNs from the schools or obtain opt-out consent until September, which is after randomisation. Therefore, even though

[2] See, for example: http://www.george-spencer.com/images/itl_demo/doc/Maths%20Hub/Events/KS3%20Multiplicative%20Reasoning%20Project.pdf

randomisation will take place in July, schools will be informed after the completion of the pre-survey and provision of necessary information, as per MoU.

Children eligible for FSM will be a sub-group for this trial – the effect of the intervention will be analysed within this sub-group in view of EEF's primary remit of narrowing the attainment gap for such students and in line with differential gains established for children from poorer socio-economic backgrounds for related universal programmes.

SAMPLE SIZE CALCULATIONS

The sample size, i.e. the number of cluster (schools) needed for each of the two arms of this study, has been determined based on the following assumptions:

- a minimum detectable effect size of 0.15 . This was deemed a worthwhile effect given the estimated cost of the intervention and the cascade delivery of PD within schools. The previous evaluation of ICCAMS (Hodgen et al., 2014) also suggests an effect size of this order is a reasonable target.
- 80% power and alpha of 0.05,
- ICC of 0.12, based on a combined consideration of suggestions/assumptions in relevant literature (Hedges & Hedberg, 2007; Spybrook & Raudenbush, 2009)
- A pre-test post-test design with 0.65 correlation

With these assumptions it was estimated that 50 schools will be required per trial (assuming number of students in Year7/8 in these schools ranging from 75 to 150 based on the eligibility criteria set earlier).

OUTCOME MEASURES

Baseline Outcome: Schools will be asked to provide name, unique pupil number (UPN), gender, EverFSM 6 status and date of birth (DOB) for all eligible students at baseline. This will allow us to collect all Key Stage 2 (KS2) results from the National Pupil Database (NPD).

The **primary outcome** will be the age standardised score on the Mathematics Assessment for Learning and Teaching (MALT) test. This assessment is a test of general maths but also includes some conceptual elements of maths. Attainment will be measured at the end of Y8, with the MALT test (paper – 40 minutes). Administration of the tests will be directly invigilated by the evaluator team and implemented under exam conditions in schools. The timing of the testing may take advantage of the infrastructure in place for GCSE.

Three **secondary outcomes** will also be measured:

(1) An attainment sub-scale on MALT of "multiplicative reasoning"
(2) An attainment sub-scale on MALT of "algebra"
Both of these are likely to be more sensitive to the intervention, and will be validated during the pilot stage.
(3) Student attitudes will be measured at post-test using a survey of dispositions towards mathematics. Such measures had been previously validated with

similar populations (Pampaka et al., 2013; Pampaka & Wo, 2014) and will be piloted to ensure validity and reliability for this study during the pilot stage.

These secondary outcomes will be developed and validated by the evaluator team and shared with the developer teams for input. As mentioned, these will be based on previously validated instruments designed for similar populations of students and teachers (www.teleprism.com; www.transmaths.org). The student attitudinal instrument will be delivered at the start of the academic year 2016-17 (pre-survey) and along with the post-test.

Outcome measures, when possible (e.g. MALT subscales and attitudinal measures for students) will be calibrated using the Rasch modelling framework (Bond & Fox, 2001; Wolfe & Smith Jr., 2007a, 2007b) which allows for objective measurement: the outcomes will thus be similar to standardised scores (which will also be preferred regarding the KS2 outcomes).

IMPLEMENTATION AND PROCESS EVALUATION METHODS

We propose to collect further evidence through a process evaluation with an aim to isolate the causal explanations of the intervention impact under usual conditions. Our implementation and process evaluation focus around fidelity and integrity (RQ7), dosage and exposure (RQ7b), participant responsiveness (RQ8), possible adaptations (RQ9) and some evidence for the 'business as usual' practices in control schools (RQ10). The process evaluation will also provide opportunities to explore teachers' attitudes and practices, as well as the effect of cascading (by comparing cascade vs core teachers). Finally, more intensive case studies including observations of key lessons will identify learning and teaching practices mediated by ICCAMS materials (and other contextual factors, eg institutional norms) that can explain targeted learning outcomes whose improvement is being measured statistically.

Three main approaches will be followed to collect evidence in order to answer these questions, in addition to the evaluation team attending recruitment and training events.

TEACHER SURVEYS

Mathematics teachers in these schools will be a participating group of interest. The subgrouping here for the intervention group will consider their level of participation to PD: Direct or via school cascading. Teacher knowledge, beliefs and perceptions of practices to be collected through teacher questionnaires at the start (September/October 2016) and end of the project (May/June 2018) in all schools: it seems sensible and crucial for the evaluation of this intervention to have a measure of what is happening during the mathematics teaching in order to objectively 'monitor' practice. This is even more crucial since inevitably, the PD offered to some teachers in each school/department will spread to others in the department- a 'cascade' effect – but the extent of this cascade is likely to vary dramatically from school to school. Such measures (i.e. practices) would help shed more light on the effect of PD on changing the teachers' beliefs and practices and will provide useful aggregate variates for modelling the school level effects on attainment. For these measures we will also build on previous work (Pampaka et al., 2012). It is expected that the teacher instruments will be provided both online and in hard copies. This

information will be explored also in relation to teachers' attendance at PD sessions and the fidelity categories to be devised.

In order to **monitor the control conditions** we also intend to collect information from control schools/teachers through these surveys about their experience and any potential PD going on in their schools.

CAPTURING FIDELITY

We intend to identify through classroom/school case studies (during the pilot phase) what are the practical forces limiting 'fidelity' and then working with Durham (and Nottingham) we will support the design of an instrument for the PD Leads visiting schools to 'measure' this in the main trial. Its use and validity will also be triangulated by researchers from the evaluation team during the case studies and through a light touch invigilation in a small number of schools (about 10). The purpose is to satisfactorily take account of the currently anticipated variations in implementation, and in its perceived fidelity. After the pilot phase it is also expected that an agreement should be reached between the involved teams as to what a high, medium and low fidelity ICCAMS school is based on these measures. Such a categorisation will also be used as part of secondary analysis (see later). Logs of attendance will also be shared (Durham will share with Manchester) in order to provide another indication of fidelity.

IN-DEPTH LONGITUDINAL CASE STUDIES

These will involve interviews with involved stakeholders (PD Leads, teachers, students) and lesson observations with both directly trained teachers, as well as those via the cascade model. In addition we intend to collect survey data from students in these case studies in regards to their perceived engagement with the material. We also intend to observe 5 PD meetings and events in the first year and 2 in the second. For better coverage of the intervention we propose to choose our case study schools by selecting at least 2 from each consultant (10 schools). The selection will also take into account the initial responses of teachers in the surveys (i.e. with an aim to observe teachers in the spectrum of a practice scale). Based on early visits to these schools we will progressively focus on a selected group of about 5 'telling' schools and classrooms that include a variety of practices likely to offer explanations of variation in teaching and associated learning outcomes. These more intensive case studies will seek explanatory frameworks including the ways institutional and cultural norms, pedagogy, and ICCAMS materials and training mediate learning practices in the classroom. Despite the structured agenda for data to be collected from these case study sites, we aim to keep these case studies as open ended in nature as possible including the unexpected in schools that seem different/extreme or otherwise interesting. We anticipate that in order to achieve this we will need between at least 5 to 10 visits at each of these schools for the duration of the project, and in some cases involving a team of at least two researchers. The purpose of these case studies is to test the ICCAMS 'theory' to its limits and develop new hypotheses and explanations for phenomena that are unanticipated at this moment.

In addition we intend to choose a few 'control' schools including a diversity of telling and interesting responses to teacher surveys for close follow up interviews by telephone, exploring the way that 'business as usual' has developed in the perception of the key school contacts (at least 10 control schools).

The overall aim of this process evaluation is thus to gather more evidence about the mechanisms that support or inhibit the implementation of the ICCAMS approach, its effects in classroom practices, and learners' outcomes, thus explaining the findings of the impact evaluation.

ANALYSIS PLAN

Analysis will need to account for the fact that schools were randomised into groups, while the outcome measures are collected from the students. Therefore multilevel models will be employed to estimate a school-level and a student-level (and preferably class-level if teachers are matched to classes and students) variance, in order to allow for schools to differ regarding their average outcome. Each model will include the outcome of interest as dependent variable (i.e. students' maths overall score, algebra, multiplicative reasoning, students' dispositions/attitudes, and teachers' perceptions of practice, etc) and the following covariates will be included as independent variables: an indicator of group membership (ICCAMS Maths Intervention vs Control), student's KS2 (average/maths) score (and maths attitudes at start, when appropriate), gender, age (at post-test), and FSM.

Analysis will be conducted using the principles of intention to treat, including all schools and students in the groups to which they were randomised irrespective of whether or not they actually received the intervention.

A standardised effect size will be calculated and reported using Hedge's *g* (Cohen's *d* bias corrected), when needed (e.g. in FSM analysis), and accompanied by 95% confidence intervals as per EEF specifications. In particular, statistical significance will be assessed using two-sided tests at the 5% level unless otherwise stated. Regression based methods of analysis will be used. Estimates of effect with 95% confidence intervals (CIs) and p-values will be provided as appropriate. Methods for handling missing data and further detail on analyses will be provided within a statistical analysis plan (which will be appended with the revised protocol after the pilot stage – June 2016).

Primary Analysis: For the primary outcome (RQ1) we will conduct intention-to-treat analysis (Gupta, 2011), operationalized as two-level hierarchical linear models (student, school), with intervention group (e.g. ICCAMS vs. control) and blocking stratifiers (e.g. % FSM) entered at the school level.

Secondary Analysis: The effect of the intervention on attainment (both overall and subscales) will also be analysed by repeating the primary analysis in the sub-group of students who are eligible for FSM. Secondary outcomes (i.e. attitudinal student measures) will be analysed in a similar manner to the primary outcome. We will also perform secondary analysis within the treatment group to explore the interaction between fidelity and attainment.

Note: A detailed analysis plan along with a data management plan will be appended with the revised protocol after the pilot in June 2016.

COSTS

As part of the evaluation we will collect data regarding the following costs for schools implementing their responsibilities in the intervention, as well as outside school costs of implementation. We note:

- Direct, marginal costs (i.e. costs directly attributable to the school's participation in the intervention): it is envisaged that this will include the financial costs for providing the training, for school visit (if necessary) and the cost of implementing the intervention and other necessary resources for the successful implementation. In particular, there will be costs related to salary costs (for teachers time to take part in training), purchasing/printing resources and fees for services.
- In addition it is expected that such intervention will entail 'Hub' costs of administration and implementation monitoring.
- Pre-requisites, especially in relation to the delivery of the intervention (and regarding the equipment needed or available at schools).

Data will be collected from the developer, as well as directly from schools as part of the process evaluation (e.g. interviewing via case studies, school-level surveys and/or use of pro-forma spreadsheets) to uncover the expected and any unexpected costs of this intervention.

ETHICS AND REGISTRATION

Each of the participating institutions will ensure ethics clearance within their institution.

Parental opt-out consent will be sought for collecting and using data for the trial. Opt-in consent will be sought from students and teachers for observations, interviews and surveys.

The Evaluation team will register the trial with ISRCTN (www.controlled-trials.com) once the protocol is agreed.

PERSONNEL

Roles and responsibilities will be updated with the revised version (June 2016) and after the evaluation team is established.

RISKS

The main anticipated risks along with the mitigation considerations are presented in the following table.

Risks	**Assessment**		**Mitigation**
	Likelihood	**Impact**	
School and student recruitment	Low/Medium	High	Recruitment will be a collaborative approach between the implementation team and the evaluation team. All parties will work with networks of schools.
Possibility of attrition for those schools that agree to enter	Low/Medium	Medium	Control schools will be provided with a financial contribution of £1500 to acknowledge the time and resources required to take part in the study. Schools

			will sign up to MoU informing them of all aspects of the trial. Contact will be made with schools during the trials via social media and newsletters.
Refused Access to NPD	Low	Low	Data will be collected directly from schools.
Missing outcome data (MALT tests)	Medium	High	We can offer schools automated reports for students based on MALT test, subject to ethically-approved data management protocols. Expertise in the evaluation team can alleviate partly missing data issues via imputation at the analysis stage if such a problem arises (Pampaka, Hutcheson, & Williams, 2014).
Other aspects of implementation relating to the PD, the PD leads and the hubs	Low/Medium	Medium	PD leads would be contracted to provide the training and would agree to different aspects. If a PD lead were not able to carry on a member of the Durham (or Nottingham) team could step in for an interim period until another PD lead could be recruited.
Researcher Loss (illness etc)	Medium	Medium	Each institution has a large department with numerous researchers to take on. Senior staff can stand in if necessary
Administration of data	Low	High	All three institutions will agree to follow a data management plan registered online which details the acceptable processes for transferring and storing data as well as which data will be stored and used by each party. All processes specified will comply with the policies of the three institutions.
Management of the trial & the evaluation/test data			Clear protocols in place for communication between all three teams so that all teams are informed about any communication with schools. Regular and frequent communication between teams on milestones.
Maintaining and monitoring fidelity (intervention and control) [It is essential that as many schools as possible maintain a high level of implementation fidelity]	Medium	Medium	The evaluation team will provide support to the Implementation and Delivery teams, including a robust recording mechanism to allow for triangulation. Three forms of triangulation, supporting our observations and reports from a number of viewpoints will be utilised: (a) Comparing and contrasting evidence about the same actions and activities, from different stakeholders, (b) Scrutinising events from different perspectives by making use of a variety of methods for collecting information, and (c) Using 'outsiders' as observers.

Data Protection Statement for Manchester:

http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=14914

TIMELINE

Date	Team*	Activity
1st August 2015	All	Project starts
Aug/15 to Jan/16	N, D	Intervention development work
Oct/15 to Jan/16	All	Development of new instruments and institutional ethic clearance
Jan/16 to Jun/16	N, (All)	Piloting of intervention, training and materials
June to Aug 16	E	Validating new instruments and sub-scales
Jan/16 to Jun/16	D, (E)	Recruitment of schools to trial
Oct/15 to Mar/16	D	Recruitment of PD Leads
April 2016	D	Training of PD Leads
July 2016	E	Randomisation of schools
September 2016		Pre-survey First year of intervention begins
September 2016 to July 2018	E	Process Evaluation
September 2017		Second year of intervention begins
June/July 2018	E	Final outcome assessment
December 2018	E	Draft report
March 2019	E	Final report
31st March 2019	All	Project end date

*Nottingham=N, Durham=D, Manchester=E(valuation)

REFERENCES

Adey, P. (2006). A model for the professional development of teachers of thinking. *Thinking Skills and Creativity, 1*(1), 49-56.

Adey, P., Hewitt, G., Hewitt, J., & Landau, N. (2004). *The professional development of teachers: Practice and theory*. Dordrecht: Kluwer.

Askew, M., Brown, M., Rhodes, V., Johnson, D., & Wiliam, D. (1997). *Effective Teachers of Numeracy* (Final Report). London: King's College.

Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5-25.

Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2011). Can teachers' summative assessments produce dependable results and also enhance classroom learning? *Assessment in Education: Principles, Policy & Practice, 18*(4), 451-469. doi: 10.1080/0969594X.2011.557020

Black, P., & Wiliam, D. (1998). *Inside the black box: raising standards through classroom assessment*. London: King's College.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. NJ: Lawrence Erlbaum Associates Inc.

Cordingley, P., Bell, M., Evans, D., & Firth, A. (2005). *The impact of collaborative CPD on classroom teaching and learning. Review: What do teacher impact data tell us about collaborative CPD?* . Research Evidence in Education Library, London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London

Gravemeijer, K.: 1999, 'How emergent models may foster the constitution of formal mathematics', *Mathematical Thinking and Learning, 1* (2), 155–177.

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass Correlation Values for Planning Group-Randomized Trials in Education. *Educational Evaluation and Policy Analysis, 29*(1), 60-87.

Hodgen, J. (2007). *Formative assessment: tools for transforming school mathematics towards dialogic practice?* . Paper presented at the Proceedings of CERME 5: Fifth Congress of the European Society for Research in Mathematics Education (pp. 1886-1895), Larnaca, Cyprus.

Hodgen, J., Brown, M., Kutchemann, D., & Coe, R. (2010). Mathematical attainment of English secondary school students: a 30-year comparison. *Paper presented at the British Educational Research Association (BERA) Annual Conference*, University of Warwick.

Hodgen, J., Coe, R., Brown, M., & Kuchemann, D. (2014). Improving students' understanding of algebra and multiplicative reasoning: Did the ICCAMS intervention work? In S. Pope (Ed.), *Proceedings of the 8th British Congress of Mathematics Education* (pp. 167-174). UK: http://www.bsrlm.org.uk.

Hodgen, J., Coe, R., Brown, M., & Küchemann, D. E. (Under review). Educational standards over time: changes in mathematical understanding between 1976 and 2009 in England.

Marshall, B., & Drummond, M. J. (2006). How teachers engage with Assessment for Learning: lessons from the classroom. *Research Papers in Education, 21*(2), 133-149.

Pampaka, M., Hutcheson, G., & Williams, J. (2014). Handling missing data: analysis of a challenging data set using multiple imputation. *International Journal of Research & Method in Education*, 1-19. doi: 10.1080/1743727X.2014.979146

Pampaka, M., Williams, J. S., Hutchenson, G., Black, L., Davis , P., Hernandez-Martines, P., & Wake, G. (2013). Measuring Alternative Learning Outcomes: Dispositions to Study in Higher Education. *Journal of Applied Measurement, 14*(2), 197-218.

Pampaka, M., Williams, J. S., Hutcheson, G., Wake, G., Black, L., Davis, P., & Hernandez - Martinez, P. (2012). The association between mathematics pedagogy and learners' dispositions for university study. *British Educational Research Journal, 38*(3), 473-496.

Pampaka, M., & Wo, L. (2014). *Revisiting Mathematical Attitudes of students in Secondary Education. In Liljedahl, P., Oesterle, S., Nicol, C., & Allan, D. (Eds.) Proceedings of the Joint Meeting of PME 38 and PME-NA 36, Vol. 4, pp. 385-392. .* Paper presented at the The Joint Meeting of PME 38 and PME-NA 36, Canada, Vancouver.

Smith, E., & Gorard, S. (2005). 'They don't give us our marks': the role of formative feedback in student progress. *Assessment in Education: Principles, Policy and Practice, 12*(1), 21-38.

Spybrook, J., & Raudenbush, S. W. (2009). An Examination of the Precision and Technical Accuracy of the First Wave of Group-Randomized Trials Funded by the Institute of Education Sciences. *Educational Evaluation and Policy Analysis, 31*(3), 298-318.

Slavin, R.E., Lake, C., & Groff, C. (2009), Effective programs in middle and high school mathematics: A best evidence synthesis. *Review of Educational Research, 79* (2), 839 – 911.

Streefland, L. (1991). *Fractions in realistic mathematics education: a paradigm of developmental research*. Dordrecht: Kluwer Academic Publishers.

Torgerson, D. J., & Torgerson, C. J. (2008). *Designing Randomised Trials in Health, Education and the Social Sciences: An Introduction*: Palgrave Macmillan.

Watson, A. (2006). Some difficulties in informal assessment in mathematics. *Assessment in Education, 13*(3), 289-303.

Wolfe, E. W., & Smith Jr., E. V. (2007a). Instrument Development Tools and Activities for Measure Validation Using Rasch Models: Part I - Instrument Development Tools. *Journal of Applied Measurement, 8*(1), 97-123.

Wolfe, E. W., & Smith Jr., E. V. (2007b). Instrument Development Tools and Activities for Measure Validation Using Rasch Models: Part II - Validation Activities. *Journal of Applied Measurement, 8*(2), 204-234.

Wiliam, D. (2007). Keeping learning on track: Formative assessment and the regulation of learning. In J. F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1053–1098). Greenwich, CT: Information Age Publishing.