education sciences

MDPI

Article

The Mathematics Teacher Exchange and 'Mastery' in England: The Evidence for the Efficacy of Component Practices

Mark Boylan *, Bronwen Maxwell, Claire Wolstenholme, Tim Jay and Sean Demack

Sheffield Institute of Education, Sheffield Hallam University, Sheffield S1 1WB, UK; b.maxwell@shu.ac.uk (B.M.); c.e.wolstenholme@shu.ac.uk (C.W.); T.Jay@shu.ac.uk (T.J.); s.demack@shu.ac.uk (S.D.)

* Correspondence: m.s.boylan@shu.ac.uk

Received: 15 October 2018; Accepted: 11 November 2018; Published: 15 November 2018

Abstract: 'Mastery' is central to current policy in mathematics education in England, influenced by East Asian success in transnational assessments. We scrutinise the prospects for mastery pedagogies to improve pupil attainment in English primary schools. The Mathematics Teacher Exchange (MTE)—an element of the mastery innovation—involves teachers visiting Shanghai and then hosting Shanghai teachers in their schools. Informed by programme evaluation, core component practices are analysed, which were implemented by schools belonging to the first cohort of MTE schools. These consist of: varied and interactive teaching; meaningful and coherent mathematical activity; and full curriculum access for all. These elements are supported, optimally, by collaborative, embedded, and mathematically focused professional development. Details of the implemented pedagogy and forms of professional development are reported. Differences from prevailing practice in primary mathematics in England are highlighted. Evidence is reviewed from quasi-experimental trials, reviews and meta-analyses, and rigorous observational studies of the efficacy of practices similar to the MTE mastery pedagogy components in order to assess the prospects for increases in pupil attainment. The analysis suggests that many of the specific practices, if considered individually, have the potential to improve attainment, though overall policy ambitions may not be realised. Based on the review, component practices are identified for which existing evidence justifies immediate implementation by schools and teachers. In addition, practices that would benefit from further testing and evaluation are highlighted.

Keywords: mastery; mathematics; primary; innovation; impact; professional development; transnational; comparative; Shanghai

1. Introduction

'Mastery' became central to current policy in mathematics education in England in 2015. Although the 2014 curriculum [1] does not refer to mastery, government efforts to influence mathematics teaching and funding for professional development promote what is described as mastery or teaching for mastery. Thus, mastery is what the government is choosing to do [2] in mathematics education. In the past, in the U.S. and other Western countries, mastery has been mostly associated with a teaching and curriculum approach that was formulated by Benjamin Bloom, and the regular use of formative assessment [3–5]. However, mastery in current English policy discourse is informed by Western characterisations of East Asian education systems and practices, and particularly those in Shanghai and Singapore, due to their current success in transnational assessments, including the Programme for International Student Assessment (PISA).

Here, we assess the prospects for mastery in the English system: whether it will be a seasonal vogue or lead to lasting changes in practices and, if changes are effected, whether they will improve learner outcomes. In order to make this assessment, we adapt theory-based evaluation principles to develop a model of core components of the pedagogy that is promoted within the loosely defined policy. We then consider the extent to which the components are different to previous practices in English schools (and ones that continue in many schools) and assess the existing research evidence that the implemented practices might improve pupil outcomes. The approach is to compare enacted practices, as identified via an evaluation of a sample of Mathematics Teacher Exchange (MTE) schools, to evidence for the efficacy of similar practices.

In education policymaking in many jurisdictions, there is a notable trend of promotion of evidence-based and evidence-informed teaching and policy [6–8]. However, this does not necessarily mean that innovative educational policies are developed or informed by rigorous consideration of evidence, given that evidence in educational policy is sometimes used selectively [6]. Thus, it is important that evidence for the efficacy of policy innovations is scrutinised.

What has come to be called the 'teaching for mastery' programme, and its genesis, is described in detail later in the paper, as are other innovations under the mastery banner. In summary, it consists of a range of interconnected policy initiatives that have at their centre a promoted pedagogy that is formulated as 'teaching for mastery' [9,10]. This pedagogy is styled as a framework for the implementation of East Asian practices (particularly influenced by Shanghai and Singaporean mathematics education). Prior to the full articulation of teaching for mastery and the teaching for mastery programme, implemented by the National Centre for Excellence in the teaching of Mathematics (NCETM), one of the first government-funded parts of the programme, launched in 2014–2015, was the Mathematics Teacher Exchange (MTE). The MTE's first cohort involved teachers and school leaders from 48 schools in England visiting Shanghai for a week-long visit and hosting Shanghai teachers in England for two weeks. Although policy borrowing in mathematics education has long been a feature across systems internationally [11,12], including, in England, for example, the National Numeracy Strategy [13], the MTE is an innovative and arguably internationally unique professional development programme in that it involves an exchange of teachers with the aim of leading to system-wide change and wide-scale adoption of another education system's practices.

Currently, the evidential basis for transnational policies—such as mastery—often centres on the fact that one education system is more successful on international tests than the adopting system. In adopting East Asian practices, it is hoped that England's relative performance in these comparative assessments will improve [14]. However, concerns about the feasibility of policy importation aside, there are longstanding concerns about whether reasons for differences in PISA (Programme for International Student Assessment) outcomes, for example, can be ascribed to teaching approaches, given that systemic factors and cultural contexts also seem important. Thus, we contend that the comparative performance of national systems is not a sufficient basis to justify policies. It is important that more comprehensive and rigorous evidence is considered for possible effects of transnational policies and practices to enable an assessment to be made about the potential impact in a different context.

The methodological challenges of evaluating transnational policy innovations are heightened in relation to the mastery innovation. For example, assessing the extent to which a policy has been implemented as intended is challenging if the 'as intended' is not specified in advance. For practitioners, dilemmas arise about whether or not to adopt the promoted practices (in cases where they are not prescribed) if the evidence base is not clear at least in relation to when the practices may be appropriate and for whom and for what purpose [15].

Given the significance of this practice in government policy and also the interest in East Asian pedagogies internationally, we believe it is important to consider: (i) the prospects for improvements in pupil attainment; (ii) which of the various practices that are promoted are likely to be more effective; and (iii) where there is need for further research to establish efficacy. These are the aims of the paper.

There are two parts to this process: firstly, to identify the extent to which the practices that schools that participated in MTE cohort 1 are implementing under the banner of mastery are different to previously prevailing methods in England; and, secondly, to consider what the evidence base is that these practices, if adopted, might lead to improved pupil outcomes.

To address this aim and the two subsidiary objectives, we draw on another strand in that we make a critique from within the evidence-based or 'what works' paradigm, though we do this with caution given concerns about unintended consequences of such a stance [16,17].

Our methodological approach in the paper is to adapt the 'theory of change' methodology [18,19]. We identify components [20] of the enacted practices of schools engaged in the MTE. To define these components, we draw on our empirical study of the implementation of the mastery practices by the MTE cohort 1 schools, which is referred to in this paper as the MTE mastery pedagogy [21–23]. During the first two years of the MTE evaluation, 134 interviews were conducted with teachers and leaders from the first cohort of MTE primary schools. The outcomes of the analysis of the interview data are supported by a more extensive data corpus and described below. By doing this, we consider the enacted practices rather than the policy intentions or descriptions. Although it is beyond the scope of this paper to provide a full comparative analysis of English and East Asian practices (see [21] for more extended discussion of this), by describing the changes that have been enacted by schools, we point to ways that these are informed by East Asian practices. We then consider the evidence for each component or sub-component, drawing on a synthesis of previous research, including the outcomes of randomised controlled trials. In the subsequent section, the mastery innovation and the MTE are discussed in more detail. Although this paper is primarily intended as a theoretical contribution, the MTE evaluation methodology is outlined so that the basis for the model of critical components that follows is made transparent. Taking each component in turn, we then consider the evidence from reviews and a meta-analysis of prior interventions, and so the second question—the prospect for changes in outcomes—is addressed.

This analysis suggests that many of the specific practices, if considered individually, have potential to improve attainment. We conclude by identifying practices for which this evidence for efficacy supports immediate implementation by schools and teachers. In addition, practices that would benefit from further testing and evaluation are highlighted.

2. Mastery as Policy Enactment

In this section, we describe the mastery innovation, including the Mathematics Teacher Exchange initiative. First, we consider the different ways the term 'mastery' is used. Second, brief descriptions of East Asian practices are provided, with a focus on the Shanghai approach. An important issue here is that East Asian pedagogy is embedded in systems and cultures with specific characteristics.

2.1. Meanings of Mastery

Mastery is a flexible term that is employed in a variety of ways [24–27] and is used to refer to various related concepts:

- beliefs about learners' potential and, related to this, a type of personal construct about learning: a 'mindset' [28];
- a form of curriculum [9];
- a particular quality of learning [24,29];
- pedagogy and, in particular, two different approaches to teaching and learning:
    - learning for mastery [3–5] (which confusingly is a description of an approach to teaching as much as to learning);
    - East-Asian-informed mastery pedagogies, including 'teaching for mastery' [10].

2.2. East Asian Mathematics Teaching and Mastery

Referring to the way mathematics is taught in East Asia, the term 'mastery' entered into use in England relatively recently. The term is not, it appears, common parlance among teachers in either Shanghai or Singapore, the two principal education systems currently of interest in England. Indeed, the timing of its emergence as a label for East Asian practices appears to coincide with its adoption by a curriculum and professional development programme that was promoted to schools by a social enterprise, which is now known as the Maths Mastery programme, around 2012. The term was then taken up by the National Centre for Excellence in Teaching Mathematics (NCETM) in 2014 with the formulation 'teaching for mastery' [10].

Before describing pedagogical differences, it is important to note cultural and systemic factors. It appears that such factors are at least some of the reasons for educational success [30,31] in countries that have become PISA 'reference societies' [32]. Relevant issues include the high cultural expectations and beliefs that ability is malleable rather than fixed, and the later age at which formal schooling begins. Teacher expertise and deployment are also important factors. In Shanghai, primary mathematics is taught by specialist teachers who only teach mathematics compared to English teachers who teach across subjects, and Shanghai teachers typically teach only for 60–80 min daily and with greater continuity from year to year [33,34]. Primary mathematics teachers in Shanghai typically undertake 340–560 h of subject-specific professional development in the first five years post-qualification [33,34]; professional development is collaborative, research-oriented, and linked to university expertise. This contrasts sharply with Continuing Professional Development (CPD) for English primary teachers, in which professional development may be generic, course- or skill-based, and ad hoc (Cordingly et al. 2018). What we posit below as 'critical components' of the MTE mastery pedagogy do not include the above factors, which may also be important in explaining differences in outcomes.

The main pedagogical differences between England and Shanghai are now considered. Chinese lessons, in general, routinely start from a problem rather than a lesson objective as in England and include frequent metacognitive and reflective discussion of key concepts and metacognitive processes [35–37] with learning objectives introduced at an appropriate time [37].

In Shanghai, pupil talk is an instructional priority [38] as is the precise use of mathematical language [39]. Teaching integrates the development of conceptual understanding and problem-solving with a proficiency in routine skills [40,41] through conceptual and procedural variation [39,42]. Both of forms of variation rely on a careful choice of examples or tasks. Conceptual variation focuses on varying representations or exemplification to allow for mathematical structures and meaning to be highlighted. Procedural variation is concerned with the application of procedures or algorithms in a purposeful way so that the choice of tasks similarly allows for an understanding of the relationships between procedures and mathematical meaning. In-depth learning is emphasised before moving on to new topics (hence the adoption of the term 'mastery' to describe East Asian approaches). This contrasts with England, where showing progress in a single lesson is prioritised [12].

In Shanghai practice, homework is assessed daily with immediate intervention by qualified teachers, as needed, so that all pupils progress through the curriculum together [33,34]. Arguably, this approach of regular formative assessment supports teaching in heterogeneous groups in primary schools in Shanghai. In England, pupils of similar prior attainment are frequently grouped in classes (sets) or sit together in class or both [43].

Whilst the MTE involves engagement with Shanghai education specifically, in the current mastery innovation, Singaporean practice is also influential and has informed change in a number of schools in the project either directly through the use of Singaporean-derived resources, such as textbooks, or indirectly through the diffusion of ideas from those engaged in Singaporean approaches. Many of the features of Shanghai mathematics teaching are also found in Singapore. However, a notable difference is a more systematic and theorised use of representations in Singapore, informed by [44] Bruner's forms of representation, and described here as concrete/pictorial/abstract [45], whereas in Shanghai visual models and representations are important and there is less use of concrete materials [41].

2.3. The Mastery Programme

Under the banner of mastery, a variety of projects, programmes, and innovations are promoted, some pre-dating the widespread use of the term 'mastery'. Linked to a revised mathematics curriculum [1,29], mastery as a policy innovation consists of a number of overlapping initiatives that are promoted and enacted by a variety of not-for-profit and for-profit private enterprises and state-funded actors. The most notable government-funded organisation in this context is the National Centre for Excellence in Teaching Mathematics (NCETM), which operates with and through the Maths Hub network in England. The 35 Maths Hubs form a regional network that organises and promotes professional development and improvement in mathematics. Thus, mastery policy is implemented through network governance [46–48].

The mastery innovation is the principal government policy currently guiding mathematics education reform and improvement in England; it combines a number of textbook resource and professional learning programmes that are intended to support the development of mathematics mastery teaching in English schools. These include the Singaporean-influenced Maths No Problem textbook scheme, which is supported by a CPD programme. The textbook scheme and related CPD is independent of the government programme. However, since the first MTE, schools engaging with mastery Maths Hub activity can access a subsidy if using Maths No Problem. Similarly, since 2018, Power Maths, which is based on Shanghai textbooks, has been approved for subsidy. Maths No Problem was launched in 2007, and is based on a translation of a Singaporean textbook. Mathematics Mastery, which is similarly informed by Singaporean practices, was developed initially from 2009 by the Ark multi-academy trust. Further development was then funded by an Education Endowment Foundation grant and subject to independent evaluation (see [49,50]). Mathematics Mastery provides sets of lesson activities, teacher support materials, and a professional development programme. More recently, an alternative translation of a Singaporean textbook—Inspire—has been evaluated [51] and gives indications of possible positive effects on attainment.

Interest in East Asian mathematics pedagogies led to two study visits to Shanghai by the National Centre for School Leadership [33,34] and promotion of mastery by the NCETM [9], leading to mastery being a central focus for the Maths Hub network initiative. The NCETM has established a CPD course for teachers—the Primary Mathematics Teaching for Mastery Specialists (PMTMS) programme—with a cohort of 140 teachers per year.

2.4. The Mathematics Teacher Exchange

The MTE began with primary school teachers in May 2014. The declared aim of the programme at its launch was to fundamentally change the teaching of mathematics in English schools through implementing the mastery approach that teachers adopt in Shanghai [21]. In the first exchange, 64 teachers and school leaders drawn from 48 primary schools visited Shanghai for one week, and then hosted Shanghai teachers within their primary schools for two weeks later in the same school year. Teachers from England observed Shanghai teachers delivering mathematics lessons in Shanghai, and later observed them teaching pupils in the English schools during the reciprocal exchange visit. English teachers also observed and engaged in Shanghai teachers' planning and professional development. The English schools that participated in the exchange began to make changes to their mathematics teaching practices, and most shared their learning with other primary schools. Since this first exchange, there has been substantial government investment in what has become known as 'teaching for mastery'. The MTE continues to be important to the mastery innovation with 70 schools per year directly involved in the MTE, drawn from alumni of the PMTMS. Further commitments to the programme were announced in 2018 [52].

Schools participating in the exchange are influenced by the different discourses and meanings of mastery, but increasingly by the concept and principles of 'teaching for mastery', the designation the NCETM uses to promote its East-Asian-informed pedagogy. However, our focus is MTE cohort 1 schools' innovation in response to their exchange experiences, and other influences prior to, and in parallel with, the development of the 'teaching for mastery' programme, rather than assessing the extent to which a specific meaning of mastery—teaching for mastery—has been taken up by those schools.

3. Methodology

In this section, we describe the methodology of the research that is reported in the paper. There are two aspects to this. The first is the evaluation methodology of the MTE that was used to generate the data on changes in the practices of cohort 1 MTE schools. The second is the analysis of these data to develop the model of critical components. A more complete description of the evaluation methodology has been reported elsewhere, including the approach to evaluating the impact of the exchange on attainment, which is not a specific concern in this paper [21,23].

3.1. The MTE Evaluation Methodology

Describing the methodology of the MTE evaluation provides context and also serves to make clear the empirical and analytical basis for the model. Using a longitudinal mixed-methods design [53], the evaluation aimed to determine the potential of learning from Shanghai mathematics education to impact on teaching in England. Although not a focus of this paper, as part of the evaluation, the effects on pupils' mathematics attainment over a three-year period was investigated through a quasi-experimental study using propensity score matching [21,23]. A model of conjectured core components was derived principally from an analysis of data that were collected from participants in 48 schools, as well as a textual analysis of documents that were produced by the NCETM and others. The data corpus principally drawn on in this paper, which is a subset of the full evaluation dataset, is summarised in Table 1.

Table 1. The subset of the Mathematics Teacher Exchange (MTE) evaluation dataset that informs the paper.

Year	Data
2015	Interviews: 107 teachers based in England interviewed at MTE cohort 1 schools comprising 60 exchange participants and 47 other teachers and school leaders. Of these interviews, 88 were face-to-face interviews and 19 were telephone interviews; 12 Maths Hub leaders by telephone; 4 National NCETM or DFE by telephone; and 6 Shanghai teachers interviewed face-to-face. A survey of mathematics coordinators in 48 MTE cohort 1 schools undertaken by the evaluation team.
2016	A review of 28 school reports prepared for the NCETM in August 2015 and the NCETM's analytical summary.Telephone interviews with a teacher or school leader from 43 of 48 MTE cohort 1 schools. An analysis of reports on the activity of cohort 1 MTE schools administered by the NCETM.

The research was conducted with institutional ethical approval and in accordance with the British Educational Research Association guidelines [54], with further scrutiny at each data collection stage by the Department for Education (the funder of the research) to ensure that the burden on schools and participants was minimised. Participation in interviews and surveys involved voluntary consent in all cases. For telephone interviews, information was sent to interviewees in advance, and consent was confirmed verbally and audio recorded. In face-to-face interviews, written consent was obtained.

To exemplify the approach to data collection and analysis, we provide detail here of the procedures used in the second round of interviews. We choose these interviews for exemplification of the approach because the analysis presented of MTE schools' practices is derived mainly from the analysis of the second round of interviews [23]. These were key to respondents checking [55] the emerging findings from the analysis of data that were collected in the first year, and so to the findings reported here.

Returning to data collection and analysis, interviews consisted of closed and open questions, with closed questions entered immediately by fieldworkers into a database and checked later by a second researcher using the transcripts to ensure reliability. Interviews were transcribed and analysed using NVivo 10 qualitative data analysis software. A variety of protocols and processes were used to ensure that the analysis was reliable, including multiple coding by different researchers, with discussion and additional coding by other researchers as needed to arrive at a consensus of interpretation. In addition, the rigour of the research was supported by scrutiny from a Department for Education project steering group consisting of policy and other stakeholders as well as academic experts.

An adaptive theory methodology was used [56] that combines inductive and deductive analysis and draws on prior research and scholarship about East Asian and other effective mathematics teaching as well as the first year findings (analysed in similar ways). Emergent themes and categories were identified [57] across the different interviews. Coding was organised into six broad thematic categories, as well as contextualising information and case attributes. These six themes comprised 16 sub-themes, many of which had further sub-categories. A total of 224 distinct codes were identified. This set was then used to inform the model that is presented in the next section, which includes relevant themes and sub-themes. Other themes, for example relating to work with other schools and enabling and constraining factors, are not pertinent to our main concerns here. Similarly, the model of change as presented focuses on mathematical outcomes; however, interviewees also reported a variety of affective outcomes for learners, such as increased resilience or more positive attitudes to mathematics.

3.2. The MTE Core Components and Theory of Change

In this section, we present (what we posit to be) the core components of MTE cohort 1 schools' mastery implementation. Core components "refer to the essential functions or principles, and associated elements and intervention activities" [58], and may comprise theory-based and empirically derived principles, contextual factors, structural elements, and/or a specific intervention practice (ibid, p. 4).

Customarily, in theory-based evaluation [18,19], the initial identification of core components might occur via an interview or a discussion with those designing or delivering the intervention (see [59]). Given the diffuse nature of the implementation and adaptation of Shanghai pedagogy at the local level in England, a proxy for this has been to consider: firstly, the analysis of Shanghai practices (discussed above); secondly, a review of NCETM and other mastery proponents' literature; but, most importantly, an analysis of MTE teachers' and school leaders' views of what have been the most important changes made. Thus, the model is empirically rooted in an implementation of mastery pedagogies. The derived model of critical components is shown in Table 2, below, and structured previous reporting on the MTE [21,23].

Table 2. The core components of the MTE mastery pedagogy.

Varied Interactive Teaching	Mathematically Meaningful and Coherent Activity	Engaging and Challenging the Whole Class	Knowledge of Mathematical Facts and Language
Substantial whole-class teaching in multiple part lessons with varying forms of activity	Depth and meaning for procedural fluency and conceptual understanding	Curriculum pace for whole-class access	Precise use of mathematical language by teachers and students
Interactive dialogue	Representations	Differentiation by deepening and support	Memorising facts, relationships, and structures
	Mathematically coherent resources	Responsive teaching: formative assessment informing differentiation	
		Flexible and responsive intervention	

An additional supporting core component was changes to the forms of professional learning to be:

- collaborative,
- embedded and close to daily practice, and
- mathematically focused.

This type of professional learning is a means of realising changes in pedagogy, as well as potentially an outcome of such changes. An example from the MTE is a professional development experience leading to teachers making greater use of representations; interviews reported instances where this enhanced pedagogical subject knowledge. Such increased pedagogical subject knowledge then potentially may directly influence outcomes separate from the identified core components. To put this another way, applying the core components may lead to a more general improvement in the quality of mathematics teaching as teacher knowledge increases.

As noted in Section 2.2, subject-specific collaborative professional development (PD) of the type undertaken by MTE schools contrasts with what is common in England in many primary schools (see [60]). That these forms of professional development represented a significant change to previous practice was attested to by MTE participants [21].

Before continuing, we identify a number of limitations of the model of core components. The theory of change, as presented, does not attempt to model outcomes of professional learning that might lead to enhanced pupil learning, unless they were identified through analysis in this study as core components of the MTE cohort 1 mastery model. These might include, for example, greater teacher enthusiasm for teaching mathematics, which in turn might lead to changes in pupil motivation or increased time spent planning mathematics lessons, and so improving their quality. As noted earlier, potential influences on pupil affect or dispositions are not included.

Furthermore, a variety of other changes are also excluded from the presented change model where the evidence is less clear that such changes are due to the exchange, or they may be a by-product of other changes. For example, it appears that, in some schools, the total curriculum time for mathematics has increased. However, it is not clear if this is a general phenomenon in English primary schools due to the revised curriculum [29]. Also, this change was not reported by all schools implementing other components. Thus, increased curriculum time was not included in the model.

However, changes in practice were included in the model if interviewees linked them more directly to the exchange experience as the reason for the change, even if anecdotally it appears that other schools not involved in the MTE may be making changes of these sorts. An example of this is an increased emphasis on the recall of factual knowledge in the MTE schools.

Finally, the model represents practices as commonly implemented by schools that have reported that a change in their practice was influenced by the MTE experience, and, specifically, the influence of Shanghai pedagogies. This was not the case for all of the original 48 schools.

4. Reviewing the Evidence for Efficacy: Sources, Relevance, and Effects

In the four sections of the paper that follow this one, we discuss each of the components and their sub-components in turn and what these mean in practice in MTE cohort 1 schools. We also consider the wider evidence base from prior research for whether or not adopting such practices may lead to changes in pupil outcomes. A systematic review was not appropriate to our purpose, given (1) the varied nature of the practices being promoted and implemented under the mastery banner and (2) the difference between the promoted practices and the previous research on mastery learning. To conduct the narrative review, relevant reviews and meta-analyses were identified, and these form the main source of evidence, as well as consideration of detailed findings from quasi-experimental trials that were identified from these reviews. However, where such data are lacking, we also considered rigorous observational studies. Our criteria of rigour for the observational form of studies are the ones that are commonly used in peer review, such as transparency of reporting, appropriateness of methodology, and recognised procedures followed in relation to research quality.

We do not, in general, use comparative evidence from different countries, or large transnational studies, such as TALIS (Teaching and Learning International Survey), PISA, or TIMMS (Trends in International Mathematics and Science Study). To do so would invite the charge of circularity, in light of Shanghai's success in PISA, given that the starting point is to consider the potential of Shanghai practices to lead to improved outcomes in England. However, in relation to textbook use, we do draw on such studies, given the relative lack of experimental studies outside of those that are focused on specific textbook schemes.

In discussing whether or not practices are effective in increasing attainment, where available we report numerical effect sizes; where this information is not available, we describe them as small, medium, or large (following [61]). A 'small' (low) effect size is about 0.2, a 'medium' effect size is about 0.5, and a 'large' (high) effect size is about 0.8 or higher [62]. Hattie [61] suggests that programmes with effect sizes of 0.4 or higher are worth considering for implementation.

In discussing this evidence, we consider how closely aligned the practices evaluated or researched in the considered sources are to the practices implemented in MTE schools. Four categories are used to summarise these relationships: closely aligned (the evidenced practices are closely related or very similar to the MTE practices), weakly aligned (the evidenced practices may be related to MTE mastery practice but are distinct from them), and somewhat aligned (neither closely or weakly aligned). In addition, in some cases, the alignment of the MTE cohort 1 mastery practices to evidenced practices varied; for example, being closely aligned to some previous studies but not to practices in others. In this case, the category mixed was used (the evidenced practices vary in how related they are to MTE practices). One possible criticism of our approach in relation to practices that are weakly aligned is that the relevance of the evidence that does exist is debatable. However, including all core components in the analysis does serve to underline that some of the proposed practices under the mastery banner do not, yet, have a substantial evidence base.

In addition, we consider the amount of evidence there is with respect to practices that are aligned to the MTE cohort 1 mastery practices. In relation to some practices, there is a large body of previous research on the efficacy of the practices. However, in other cases, there is less previous research. Here, we use the categories 'extensive', 'limited', and 'some'. Where the amount of evidence varies with regard to specific practices within a category or context, we use the term 'variable'.

In England, the Sutton Trust has popularised the concept of 'months' progress as a measure [63]. However, putting aside the debates about whether this measure is generally appropriate, in the case of a complex programme, such as the MTE, where multiple practices are being considered, referring to "X months' progress" might give the impression that the effects of practices could be added together in a simple way.

At the end of each subsection, we summarise: the evidence for potential impact using high/medium/low descriptors; the extent to which the practices that were evaluated in the relevant research are similar or not to those implemented in schools; and the amount of relevant evidence.

5. Varied Interactive Teaching

We now turn to the first of the model components. The development of mathematical understanding was supported in MTE schools by practices that together can be viewed as 'varied interactive teaching', which in turn can be considered as consisting of two components: substantial whole-class teaching in multiple part lessons with varying forms of activity, and interactive dialogue. In East Asia, whole-class teaching is emphasised [36,37], as is mathematical talk [38,39]. Differences with English practices were noted by MTE cohort 1 participants, and this was an aspect of practice that was one of the first to be focused on for change and development [21].

5.1. Substantial Whole-Class Teaching in Multiple Part Lessons with Varying Forms of Activity

The development of mathematical understanding was supported in schools by varied lesson activities and structures. There are two interrelated features to this component: the first is the episodic nature of lesson structure with multiple parts of the lesson. The second is potentially an increase in teacher-directed activity as well, specifically an increase in whole-class teaching.

'Varied' here refers to a range of different types of learning activity and frequent changes between these. This contrasts with adopting a more compartmentalised lesson structure, such as the National Numeracy Strategy 'three part lesson' [64]. MTE Schools implementing a Shanghai-informed mastery approach reported that lessons followed a pattern of multiple short periods of questioning and dialogue between teacher and pupils interspersed with short periods of pupils working on one or two problems or tasks. This form of lesson structure was described variously as 'to-ing and fro-ing', 'back and forth', 'short bursts of teacher-led activity', and 'I do, you do, I do, you do' [23].

Data were collected regarding time spent on different forms of interaction during mathematics teaching. Schools reported spending more than half (54%) of an average lesson on whole-class interaction; however, as indicated above, this was typically spread out over the lesson, alternating with other types of activity. Schools reported an increased emphasis on classroom talk that was focused on process and understanding and an emphasis on precision in the use of formal mathematical language.

There is little prior evidence on the relation between lesson structure or distinct lesson parts and attainment. However, such an approach potentially supports other features of the pedagogies implemented by schools; for example, the 'to-ing and fro-ing' could scaffold assessment for learning.

With regard to the second feature—teacher-directed activity including whole-class teaching—Reynolds and Muijs [13] reviewed research on effective teaching of mathematics. They observe that children learn more when they experience more time being actively taught or their activity is supervised by the teacher, rather than working independently for extended periods of time. Hattie [65] reports an effect size for direct instruction of 0.59 in a meta-analysis, and a more recent review [66] similarly reports medium to high effects in general and 0.25 as a minimum found across a large number of studies. However, it has been argued that the research underlying these contentions derives largely from direct instruction in U.S. schools [67] and studies validated largely in relation to the teaching of basic skills [68]. Also, the approaches adopted in MTE classrooms do not accord with some of the descriptions of direct instruction in the literature. For example, most MTE classes adopted all-attainment grouping or differentiation strategies that were appropriate to all-attainment teaching (see below).

In an English-based study of a primary education innovation, it was found that the time spent teaching the whole class was not associated directly with pupils' progress. However, whole-class teaching was associated with teacher behaviours or outcomes of behaviours that have been deemed to be effective, such as time on task and direct teaching, and negatively associated with less-effective practices, such as a high percentage of seat-work [13].

In summary, there is evidence that whole-class teaching can have a positive medium-strength effect on attainment. However, the relationship between the type of practices previously studied and those adopted in MTE schools is mixed given that, in some cases, the MTE mastery practices and the prior research appear to be closely aligned but in other cases more weakly aligned. In addition, for practices where there appears to be a closer alignment, the evidence is limited.

5.2. Interactive Dialogue

There were two aspects to interactive dialogue reported by MTE teachers: teacher–pupil talk and pupil–pupil talk. The degree of interaction between teacher and pupils, often through questioning, potentially supports the development of understanding and allows the teacher to assess understanding in order to guide further teacher elaboration or explanation. Teacher elaboration [69] and teacher clarity are positively associated with improved outcomes, with Hattie [61] reporting in general that teacher clarity has an effect size of 0.75.

The value of rich classroom talk is supported by research on dialogic teaching [70,71]. Interaction in dialogic teaching encourages pupils to engage in meaningful talk that is cumulative, supportive, reciprocal, collective, and purposeful [70]. So, questions are structured to provoke thoughtful answers and these answers prompt further questions. Answers are seen as the building blocks of dialogue rather than its end, and exchanges among teacher and pupils are chained into coherent lines of discussion rather than left disconnected [70]. This is similar to the patterns of talk that are reported by those implementing changes informed by the MTE.

Mercer and Sams [72] report an evaluation of the 'Thinking Together' intervention for mathematics learning in Year 5 using dialogic principles. This intervention led to substantial changes in classroom practice (seven teachers undertook the intervention with 196 pupils) and significant gains in mathematics scores when compared with a control group, with an effect size of +0.59. The Thinking Together intervention consisted of 12 lessons, focusing on data handling, properties of numbers, and number sequences. Although positive effects were observed, the sample size was small, and the intervention focused on a particular section of one curriculum subject. More recently, Jay et al., [71] reported an efficacy trial of dialogic teaching that involved Year 5 classes (9–10 year-old children) in 78 schools over two school terms (approximately seven months). A small positive effect was seen in intervention schools, compared to control schools, for attainment in English, mathematics, and science. In mathematics, an effect size of 0.09 was reported: a minimal effect. The smaller effect that is reported here may be due to the larger scale of the study and the attempt to introduce a complex intervention across the whole curriculum.

Beyond research on dialogic teaching, aspects of the type of classroom communication that are promoted as part of a mastery approach have been found to have moderate to high utility in a variety of studies. These aspects include elaborative interrogation (generating explanations of why a fact or concept is true) and self-explanation (explaining how new information is related to known information or explaining steps taken in problem-solving) [69].

To summarise the above discussion, a recent trial in England reported low but positive effects of practices similar to those that were identified in MTE schools. However, other studies have reported higher effects. Previously evaluated practices appear to be closely aligned to those that are implemented in MTE schools. However, overall, the amount of evidence is variable, given few impact studies of more closely aligned practices but more extensive evidence of related practices.

6. Mathematically Meaningful and Coherent Activity

An analysis of Shanghai teaching emphasises the importance of developing conceptual understanding and fluency through mathematically meaningful and coherent activities. In the implementation in England, three aspects are prominent:

- depth and meaning: for example, by focusing on a mathematical structure often by aiming to use conceptual and procedural variation [39];
- a much increased and more sophisticated use of mathematical models, including visual and concrete representations; and
- the use of resources that are designed to communicate mathematics coherently to pupils, and by implication to teachers; for example, the use of textbooks.

The evidence for the efficacy of each of these approaches is considered in turn.

6.1. Depth and Meaning

Here, the promotion of mathematical depth and meaning is considered in relation to practices that are separate from the use of models and representation. A limitation of the approach to review we have taken is a discontinuity between practices advocated for as mastery in this regard and prior evidence. 'Depth' is an important concept in the mastery discourse in England with a recommendation to seek to go deeper rather than introduce new content. However, depth in itself has not been a subject for evaluation of effect, possibly because it is difficult to define. However, depth is associated with emphasising mathematical thinking, conceptual understanding, and mathematical structures.

In English mastery approaches, this is supported by a variety of approaches, including the careful selection of tasks and series of questions, often described in terms of the principles of variation theory [39]. There is a lack of meta-analyses for studies exploring the use of tasks and task construction. However, specific programmes have provided some evidence for efficacy, particularly where teacher subject knowledge supports effective use [73,74].

The type of approach that is adopted by many MTE schools accords also with evidence for the effective use of problem-solving, as well as the teaching of problem-solving as a pedagogical strategy [75]. One aspect of this is assisting pupils to reflect on solving mathematical problems (nine studies with effect sizes of 0.31–2.54 for immediate post-test, op cit). A second aspect is instruction in multiple problem-solving strategies (four studies, effect sizes 0.20–0.33 for immediate post-test). Evidence for other approaches to multiple problem-solving strategies, such as comparing strategies or generating multiple strategies, is more mixed.

Thus, a range of effect sizes are reported in interventions that promote teaching for depth and meaning, although in different ways. In relation to task use or the careful selection of tasks, while attention to these aspects is recommended, there is a lack of secure evidence. The evidence base for teaching problem-solving strategies is stronger; however, the relationship to MTE practices is mixed rather than clear. Thus, effect sizes are low to high, the relationship between prior evidence and MTE practices is mixed, and the extent of the evidence base is variable.

6.2. Models and Representations

Shanghai schools make greater use of visual representations than is often found in English schools. MTE schools have adopted these approaches and, influenced by Singaporean pedagogy, they have also increased the use of concrete materials and manipulatives, particularly extending this across all age ranges (and not just younger children as previously) and attainment levels (rather than focusing use on low-attaining pupils) [21].

The importance of pupils' understanding and the use of different representations is widely noted in the mathematics education literature [76]. However, there is mixed evidence for efficacy [77]. Sowell found no positive effect when comparing pictorial and abstract/symbolic representations, although a meta-analysis of studies with pupils with learning disabilities found positive results [78]. Woodward et al. [75] drew on six randomised controlled trials to recommend the use of visual representations; effect sizes varied from 0.21 to 1.87 on immediate post-test. The use of visual representations is specifically recommended to support low-attaining pupils [79].

A weakness in the research base and, consequently, the meta-analyses is that they focus on experimental designs where pictorial representations (and this also relates to the concrete materials that are discussed below) were the central or main feature of the intervention or programme. However, in the MTE cohort 1 schools adopting a mastery pedagogy, the use of models and representations is embedded in other practices. Nevertheless, there are a number of programmes, such as the 'Increasing Competence and Confidence in Algebra and Multiplicative Structures' (ICCAMS) project, which make considerable use of multiple representations embedded in a pedagogy that has some features similar to the way in which representations feature in mastery practices [80]. These programmes have shown positive effects, but are not interventions specifically to promote pictorial or visual use. Similarly, the value of modelling is also supported indirectly by evidence of the impact of approaches such as Realistic Mathematics Education on problem-solving skills and conceptual understanding [81].

There is also positive evidence for the effectiveness of the use of concrete materials or manipulatives [82]. A meta-analysis [77] indicated positive effects when compared with a symbolic abstract-only condition (for example, pencil and paper work from a textbook) across all elementary grade levels. More recent studies, albeit involving small numbers of students, have shown similar outcomes (e.g., [83–85]).

There is also potentially an effect on teacher knowledge from adopting these practices. However, the influence of an increased use of representations on teachers' knowledge and practice is complex. For example, greater use may or may not lead to changes in how representations are used in terms of classroom activity, but may change classroom talk. The increased use is also sensitive to forms of professional development [76].

In summary, the evidence for the effectiveness of the extensive use of models and representations in mathematics varies across different studies, from low to high efficacy. This is based on multiple randomised controlled trials (RCTs). However, such RCTs often examine representation or model use in isolation rather than as practices that are embedded in more complex pedagogies. Nevertheless, there appears to be relatively close alignment between MTE school practices and the available evidence in terms of the type of models used and the evidence is extensive.

6.3. Mathematically Coherent Resources

Many schools involved in the MTE have changed the resources that they use either to inform planning or directly as learning activities with pupils. For many of them, this has meant adopting textbooks that are translations or based on those found in East Asia. As yet, no systematic research has been undertaken on how similar or different the translated versions are to the East Asian originals.

Research on mathematics textbooks is generally focused on their role, the analysis and comparison of content, and how they are used [86]. In a review of research in this area, no efficacy trial is reported (op cit), indicating that where trials are undertaken (such as those reported in [87]), theorising about textbook use in itself is not a priority. Evidence on the relationship between textbooks and pupils' learning is often based on a comparison of selected textbooks rather than comparing use of a textbook with no textbook use at all. Slavin and Lake [87] reviewed a number of studies involving textbooks and reported mixed outcomes, including when looking at repeated trials of the same textbook.

However, generalising from these findings is problematic in relation to the materials used by schools in the MTE, not least because schools employ them in a variety of ways: to inform planning, use directly with children, or a mixture of both. Thirty of the original 48 schools were either directly using materials with students or informing lesson planning with materials that were either translations of, or designed and informed by, Shanghai or (in most cases) Singapore materials [23]. With regard to the Singapore materials, as noted earlier, there is evidence of positive gains from engagement in such programmes from previous trials [50,51].

Nevertheless, although textbooks can promote "deep and connected knowledge" [88] (p. 186), it is important to recognise that how teachers use textbooks and how they interpret them varies and they cannot be analysed only in terms of content [88] (ibid). An additional and indirect potential effect of the use of textbooks is improving teachers' mathematical knowledge. The issue of subject knowledge and primary teachers has been recognised as important for some time in England (see, for example, [89]); teachers' mathematical knowledge contributes to gains in pupils' mathematical achievement [90].

Evidence for the effect in general for the use of textbooks is limited, and there have not been trials of the Shanghai- or Singapore-informed materials that most MTE schools were using (if they were using such materials at all). Based on the limited available data, the alignment between evaluations of historic interventions involving textbooks (e.g., [87]) and MTE practices is weak; that is, practices are not closely aligned with previously reported interventions. For example, with regard to textbooks, one recent Singapore-informed textbook subject to a quasi-experimental evaluation [51] was used by only a few of those MTE cohort 1 schools adopting textbooks [23]. Most schools that used textbooks to inform planning or with pupils were using materials that had not been subject to an impact evaluation at the time of the research. Given the limited evidence on the impact of the use of textbooks in general and the lack of specific evidenced in relation to textbooks actually being used, it is not possible to categorise the potential effect of using textbooks, or indeed other mathematically coherent resources, separately from specific programmes.

7. Full Curriculum Access for All

In Shanghai primary schools, there is an emphasis on the whole class learning curriculum content together. English schools have mirrored this and supported pupils to access content by:

- slowing the pace through the curriculum to ensure greater access by more learners;
- decreasing the use of in-class or across-class grouping by prior attainment ('ability' grouping) and instead adopting differentiation strategies that are focused on deepening and support, rather than through tasks pre-allocated to particular groups of pupils, to support pupils to access the *full* curriculum; and
- responsive intervention, including the use of a daily intervention, which is in some cases linked to timetabling changes; for example, splitting lessons or introducing short additional periods typically to promote factual knowledge.

The evidence for the efficacy of these sets of practices is now considered.

7.1. Curriculum Pace for Whole-Class Access

In prevailing English practice, at least prior to the mastery innovation, classes in primary schools are often divided either into sets or within-class attainment groups. Different work is allocated to children who are perceived to have different levels of 'ability' as measured in attainment tests. A consequence of this is that not all pupils access the full curriculum. The increase in whole-class teaching outlined above potentially reverses this. Interestingly, at the time of the implementation of the National Numeracy Strategy, the arguments made for adopting interactive whole-class teaching included enabling low-attaining pupils to access the whole curriculum, although the case was made more explicitly about the literacy strategy than for mathematics [91]. One consequence of this practice was to slow the curriculum pace down. By curriculum 'pace' we mean the rate at which curriculum content is covered, and this related to how quickly material is taught [92] as opposed to the pace of changes of type of activities or practices within the lesson, which also relates to the frequency of interaction and varying types of activity. To distinguish between these two forms of pace, we refer here to this second form of pace as the rhythm of the lesson, and previously we have also referred to this as 'tempo' (see [21]). The rhythm of a lesson metaphorically is formed by regarding each change in activity as a beat. An activity here can mean on the scale of a part of a lesson or at a finer grain in terms of who is speaking or leading in an interaction. Thus, it is possible for the curriculum pace to slow whilst the rhythm increases.

However, the rationale for slowing the curriculum pace was related not only to issues of engaging all learners, according to MTE interviewees. Those involved in the exchange also noted the 'tiny steps' that Shanghai teachers would take in breaking down procedures or concepts to ensure full understanding by all [21,23].

There is little specific research on interventions concerning the effect of slowing the curriculum pace. Arguably, 'learning for mastery' [3–5] involves slowing the curriculum pace to ensure understanding before new content is introduced. Similarly, this is a feature of some forms of direct instruction [66], in which mastery of previous material is required before new material is introduced. Essential to learning for mastery is a view that everyone can succeed mathematically, except those with specific cognitive disabilities, if appropriate resources, support, time, and teaching are provided. A similar sensibility is promoted in current East-Asian-inspired conceptions of mastery teaching. Features of learning for mastery approaches are [4]:

- diagnostic pre-assessment;
- high-quality group-based instruction;
- monitoring of progress through regular formative assessment; and
- high-quality corrective instruction for individuals or groups.

Thus, it may be that evidence for learning for mastery is relevant to current English mastery initiatives; however, the relationship is not clear. Learning for mastery has been extensively evaluated and is considered to be effective [62].

Qualitative studies indicate that pupils' experience of a fast curriculum pace in secondary school mathematics classes can lead to alienation of learners and a sense of them not being able to keep up [92]. Thus, it is plausible that there may be a mediating change in affect when the pace is slowed. However, whether primary school pupils also experience a fast curriculum pace, and if so whether this has a negative effect on affect, is not known.

Watson and De Geest [93] report that teachers working with lower-attaining pupils who were found to be successful in comparison with other teachers with similar groups, specifically in developing pupils' mathematical thinking, worked to a different timescale than others. These teachers spent longer on each topic than was recommended or usual, and spent longer on specific tasks to support participation, reasoning, understanding, and mathematical connectedness. Thus, if curriculum pace is important then the mechanism may be, at least in part, that a slower pace enables teachers to employ the types of teaching approaches that are discussed above rather than because a slow pace leads directly to greater learning.

To summarise, there is a lack of specific evidence on the relationship between curriculum pace and attainment or on interventions to slow curriculum pace. Evidence from mastery learning may be relevant, as may be qualitative research concerning learner engagement and disaffection. However, this does not allow for determining a potential size of effect. The relationship between the evidence that does exist and the MTE practices is weakly aligned given the different contexts of the research.

7.2. Differentiation by Deepening and Support

Differentiation by deepening and support contrasts with differentiation by 'ability'. As discussed above, constructs of ability are often central to how pupils are organised in English classrooms either by setting (grouping pupils by perceived ability into different classes) or by in-class grouping or both. This generally means that pupils have different access to the curriculum with different tasks allocated to them depending on teacher judgement. Many schools adopting Shanghai-informed pedagogies adopted all-attainment grouping. In the schools that embraced all-attainment teaching, efforts were made to change discourses of ability and to refer to children in different ways. Further, even in schools where pupils sat with others of similar perceived ability, schools had commonly changed practice to give the same tasks to all children and then provide additional support to those that need it (see below in the section on 'responsive intervention') or deepening tasks for others. It was not predetermined who would receive additional support or the deepening tasks. Schools adopting such practices assented to a label of this approach as 'differentiation by deepening and support'.

There is limited evidence, in primary education, about differences in outcomes between attainment grouping and all-attainment grouping (colloquially 'mixed ability'). Overall, no positive difference is found either way [43]. Thus, the encouragement by government and its agencies, such as OFSTED (The Office for Standards in Education—the English schools' inspection service), during the 1990s to increase grouping by attainment [94] is not evidence-based.

A meta-analysis that does not distinguish between secondary and primary schooling suggests a small overall negative effect of grouping by attainment [62]. Thus, it might be expected that moving to all-attainment grouping in the MTE schools could, if anything, have a small positive effect. Moreover, schools reported making considerable changes to their approach to differentiation. However, the new primary mathematics curriculum discourages 'acceleration'; that is, providing pupils with material or content drawn from future years' programmes of study. Rather, it encourages deepening of understanding as the main differentiation approach. Thus, it is unclear to what extent the changes to differentiation practices were due to the participation in the exchange or a general response to curriculum change.

In summary, the evidence shows that adopting all-attainment teaching practices is unlikely to greatly affect attainment, so any positive effect in MTE mathematics classes is likely to be low; however, a negative effect is also unlikely. MTE school practices appear to be closely aligned to the evidence base, save for a lack of evidence in the primary education context. Thus, overall, the relationship between evaluated practices and the MTE is mixed and the extent of evidence is variable.

7.3. Responsive Teaching: Formative Assessment

Although not an explicit mastery practice, formative assessment practice has changed in MTE cohort 1 schools as part of their approaches to changing differentiation practices and engaging the whole class. Formative assessment is enacted on two different timescales. One is in the 'back and forth' process described above, in which pupil responses inform teacher response and importantly when to take the next step. In the second timescale, each lesson is a time unit to inform the responsive intervention discussed above.

Formative assessment is carried out in order to identify gaps in pupils' knowledge and so enable teachers to focus on those areas where pupils need further input to build on their existing knowledge and therefore make progress [95]. Evidence shows that a number of strategies contribute to effective formative assessment; for example, clarifying learning goals with learners and empowering pupils to be a learning resource for one another [96]. Using more informal assessment and formative evaluation is considered to be effective practice with medium to high effect sizes reported in the meta-analyses [65].

There is particularly strong evidence for the use of effective feedback in assessment [62]. For example, looking across the 12 meta-analyses assessing the influence of feedback, Hattie and Timperley [97] found an average effect size of 0.79.

In order to carry out effective formative assessment in mathematics, teachers need to know the mathematical misconceptions that pupils may develop or have already developed [98]; these can then be identified and explained to pupils as part of assessment feedback [98,99]. In Shanghai classrooms, pupils' explanation and reasoning skills are embedded further through the routine presentation and discussion of common mathematical misconceptions. Teachers in the second year of MTE interviews reported how their lesson preparation had started to include detailed planning of questions that related to anticipated misconceptions.

Although formative assessment practices with medium to high effects have been identified in the literature based on extensive evidence, the relationship between MTE practices and those evaluated in the meta-analyses on formative assessment is not clear; at best, the evidence is somewhat aligned with MTE mastery practices.

7.4. Responsive Intervention

Informal interventions and formal interventions in mathematics are both established features of practice in English primary schools [100]. Many MTE schools reported to have changed their approach to identifying pupils for intervention. Previously, schools often used teaching assistants in class to support the same groups of pupils in every lesson or to work with small groups on a weekly basis. Teaching assistants support qualified teachers and work under their direction. Deployment of teaching assistants in this way has been found to be ineffective in England [101].

Instead, MTE schools increasingly identified pupils daily who needed additional tuition or support with particular lesson content; this also informed and supported the greater use of assessment for learning and formative evaluation. Pupils were supported on a one-to-one or small-group basis by the teacher.

One-to-one tuition in mathematics can be an effective intervention [63,102,103]. Effect sizes reported in a recent trial where tuition was undertaken by a qualified teacher in England were reported as 0.34 for general mathematics attainment and 1.11 for a test of early number knowledge and skills [103]. A trial of an intervention where tuition was undertaken by teaching assistants, who were provided with specific training and materials, found effect sizes of 0.21–0.24 depending on the details of the training and materials used. However, some caution is needed here, as the evidence cited is for specific programmes that were aimed at remediation. Further, such programmes have been found to produce high initial gains, but these can then tail off or drop once the programme ends; for example, the Maths Recovery programme [104]. On the other hand, it may be that an intervention that is more closely aligned with the regular programme of study could be more effective than a supplementary programme, given that it may enable students to access regular teaching more fully. However, this is untested, although there is a trial of an MTE-informed daily intervention approach now underway [105].

Overall, previous research suggests there is potential for low to medium effects of changes in intervention practices in MTE schools and this is based on multiple trials. However, given the specificity of the evaluated programmes, the relationship of MTE practices to the evidence base is weak in some cases and somewhat aligned in others, and so mixed overall.

8. Knowledge of Mathematical Facts and Discourse

The fourth component of the implemented mastery pedagogy focused on knowledge of mathematical facts and discourse. There were two aspects to this:

- an increased emphasis on factual recall often leading to specific episodes in the day to support this; and
- a focus on the precise use of mathematical language.

Both of these sub-components supported pupils' capacity to engage mathematically, including in interactive dialogue.

8.1. Precise Mathematical Language

The precise use of mathematical language was an aspect of Shanghai pedagogy that was often remarked on by the English MTE teachers who visited Shanghai and one that was readily adopted. There is some evidence for the effectiveness of teaching approaches that include this. A review of mathematical vocabulary was included as a technique in a programme of meta-cognitive instruction for low-attaining students [106]. The programme as a whole had a strong effect (2.18), although the role of language precision in relation to other aspects of the intervention is not identifiable separately. Given that this specific evidence is lacking, the relationship between MTE practices (which also vary across schools) and the one identified study is weak, thus meaning that identifying a likely effect size is not possible. Further, evidence in this area is limited to a small number of trials rather than based on a meta-analysis.

8.2. Memorising Mathematical Facts, Relationships, and Structure

As with curriculum pace, MTE teachers visiting Shanghai remarked on the level of factual knowledge of the children. They considered this to be important in two ways: directly as a feature of mathematical proficiency, but also in supporting pupils to access the curriculum and participate in the pedagogical approaches, such as interactive dialogue. The importance of factual knowledge was often translated in MTE cohort 1 schools with a particular focus on knowledge of times tables. This concern may reflect not only the lessons drawn from engagement in the MTE but also the current policy and societal discourses about this issue and the increased emphasis in the new curriculum and assessment arrangements. In MTE schools, the most common way to address factual knowledge was via the introduction of an additional, short period explicitly for learning and practising number facts, often focused on times tables but not exclusively so. These periods might be built into lessons—thus echoing one approach to the National Numeracy 'starter'—or as an additional 5–10 min episode at another time in the day.

There is some evidence for the effectiveness of these kinds of approaches, and they are recommended as part of interventions to improve mathematical outcomes specifically for "pupils struggling with mathematics", with a recommendation of 10 minutes of regular practice to build "fluent retrieval of basic arithmetical facts" [79] (p. 37). Some studies compare different approaches to learning facts and so do not give insight into the effect of introducing such practices. However, three studies comparing regular practice with a 'no instruction' condition report effect sizes of 0.19–0.60 (op cit, p. 83). A weakness in the evidence base is a focus only on younger primary pupils in these studies (U.S. grades 1–3). Supplemental teaching/episodes of this sort have been found in other studies and reviews to have positive effects in general, particularly for low-attaining students [107].

Thus, there is low to medium efficacy reported for programmes that target factual knowledge in mathematics. Given both the variability of how learning of facts was implemented in MTE schools, and variation across the evaluated programmes, the alignment of practices is mixed with some alignment to reported practices in some schools and weak alignment in relation to others. The evidence base is limited with only three RCTs reported, and these focused on the early primary years.

9. Professional Development

In each of the previous sections, we have included a discussion of ways in which MTE schools have implemented or adapted Shanghai practices, and we have considered the evidence in relation to each of the core components of the approach. However, the evidence on effective professional development tends to report and describe professional development holistically. So, in this section, we firstly describe the research evidence on effective professional development and then discuss Shanghai practice and the MTE schools' interpretation and here consider the three sub-components or features highlighted earlier in proposing core components: collaborative; embedded and close to daily practice; and mathematically focused.

However, before doing so, we expand on an earlier note that professional development in Shanghai contrasts with that found in England. In Shanghai, teachers engage in intensive, ongoing school-based professional development [108], with mathematics teachers undertaking 340–560 h in the first five years post-qualification [33,34] (NSCL, 2013, 2014). MTE cohort 1 schools use a variety of sustained professional development activities to support the implementation of mastery practices, which can be categorised as formal, embedded, or informal activities. Formal activities include in-school CPD sessions and staff meetings, Teacher Research Groups, attendance at external courses, conferences and workshops, in-class support and modelling, and observation. Embedded activities relate to engagement in teaching practices, such as collaborative lesson planning, which in turn gives rise to new learning. Informal professional learning activities relate to unplanned informal teacher discussion that takes place, for example, in the staff room during break times.

9.1. Effective Teacher Professional Development

The forms of professional development found in Shanghai accord with research evidence on effective professional development. Reviews identify that it should have the following features:

- sustained over time [109];
- localised, adaptive, and varied [60,109,110];
- tailored to teacher needs [110,111] including being shaped by the context in which teachers practice [60,112];
- teachers are agentic [111] and encouraged to adapt practices rather than reproduce them [109];
- professional development involves enquiry as a key tool [110,111]; and
- collaboration is central [110–113].

Turning to evidence of the effect of professional development on pupil outcomes, Lomos et al.'s [114] meta-analysis of the impact of five studies of professional learning communities found a small effect (0.25) on pupil achievement. However, the data are primarily drawn from secondary schools, and the lack of a conceptual and empirical definition of professional learning communities is noted by the authors. Specifically, in relation to primary mathematics, programmes that focus on mathematics content have been found to be effective [87].

In synthesising nine studies where the features of professional development corresponded to the characteristics listed above, Yoon et al. [109] identified overall moderate effect sizes of 0.54 on average, and Hattie [65] conducted two syntheses of larger studies reporting effect sizes of 0.51 (2011) and 0.62 (2009).

With regard to these findings, professional development may produce a direct effect on pupils that is not mediated by the sorts of practices already discussed as being core components of MTE pedagogy. For example, professional development may lead to improved teacher self-esteem and so influence the classroom climate, which is itself linked to a positive impact [61,65]. However, professional development is mainly effective because it leads to changes in practice, so the possible professional development impact may already be factored into the potential effects discussed in previous sections.

9.2. MTE Teacher Professional Development

Three features of teacher professional development in MTE schools were identified from the analysis, and each of these is discussed in turn below. Taken together, they accord with the features that were identified as effective and considered above.

9.2.1. Collaborative Professional Development

MTE cohort 1 schools implemented Teacher Research Groups (TRGs) as an adaptation, at a smaller scale, of the ones teachers observed in Shanghai. In Shanghai, extensive networks of vertically and horizontally organised TRGs are found [108,115]. Immediately following the exchange year, mastery-focused TRGs were formed of teachers within the MTE schools or also including teachers drawn from nearby schools. Subsequently, TRGs or working groups were formed and led by mastery specialists within the Teaching for Mastery Programme, and involved schools from across the 35 regional Maths Hub areas.

The characteristics of TRGs in the MTE schools differed from those found in Shanghai. One difference is that, in Shanghai, vertical organisation of TRGs across districts and with University-based mathematics educators provides a defined structure at the system level that leads to a relatively close coupling [116] between TRGs. This differs from the looser and more ad-hoc coupling (if any) between MTE cohort 1 TRGs through Maths Hubs. In addition, Some MTE cohort 1 schools were influenced by the Japanese lesson study approach (e.g., [117]) involving joint planning, observing the lesson (in person or on video), and post-lesson discussion.

The lesson study influence meant that the ‘research’ aspect of a teacher research group was preserved. However, in other schools, the format, even if referred to as a TRG, was different to the Shanghai model and more similar to cascade practices similar to those found as legacies of the National Numeracy Strategy. In these cases, joint planning was omitted, and instead the delivery of the lesson by the teacher who had participated in the MTE was the focus of a post-lesson discussion. Post-lesson discussions in these cases included the rationale for planning, as well as the group’s detailed exploration of what had happened during the lesson. There was variation between MTE cohort schools as to whether TRGs were sustained over time or were one-off activities and as to whether they were restricted to the MTE participant’s school or included teachers from other schools. MTE cohort 1 teachers considered TRGs to be effective in supporting professional learning.

In addition to formal TRGs, increases in collaborative lesson planning were reported by (most) MTE schools, and interviewees also pointed to the importance of modelling mastery teaching outside of the more formal framework of TRGs, either to a group of teachers or on a one-to-one basis.

9.2.2. Embedded and Close to Daily Practice

Not all MTE schools reported sustaining TRGs beyond hosting Shanghai teachers. An alternative to TRGs (or in some cases supplementing them) was the strategy of embedding professional development in regular activities. As a result, professional development was linked to daily practice and this was also reported as being effective. Examples of these embedded activities were teachers' use of textbook exercises or other mastery-oriented resources as well as novel artefacts for planning that supported the focus on key mathematical ideas or structures. A specific example here was the use of 'concrete-pictorial-abstract' as a planning rubric. So, for example, schools reported using the rubric as a means to structure lessons and to include different representations or models, following a sequence from the use of manipulatives or objects, visual representation—often some form of bar model—before the introduction of symbolic representation. As noted above, planning would often have a collaborative aspect and this was highlighted as a change of practice in many schools [21,23].

9.2.3. Mathematically Focused

A feature of both the collaborative and embedded professional development was a focus on mathematics and mathematical meaning. For example, the detailed focus on developing conceptual and procedural fluency in lesson planning was perceived to enhance subject knowledge and support teachers in developing questions. One approach to this was to focus on identifying the connections between different mathematical concepts in planning to identify how mathematical meaning progressively developed [21]. Textbooks were perceived to improve subject knowledge and confidence; for example, in how to use resources effectively. Some schools chose to select questions from a variety of East-Asian-informed textbooks, and the process of deciding which resources to use could entail discussion of mathematical structure [23].

9.3. Potential Effects of MTE Professional Development

The forms of professional development that were enacted by MTE schools accord with the general features of effective professional development. In relation to studies of effectiveness that report medium effect sizes, MTE professional development has the potential to realise this. However, the extent to which the more fine-grained details of practices in MTE schools are similar to practices in the reported studies is mixed. Whilst there is extensive evidence of the effects of PD on teacher practice, the amount of evidence on pupil outcomes is variable, at least in relation to the type of professional development engaged in by MTE cohort 1 schools.

10. Prospects for Improvement in Attainment Outcomes

The central motivation for the mastery policy in England is the potential to improve pupil attainment [14,21]. In this section, we summarise the review of literature on the efficacy of the identified core components. However, the extent to which the evidence is applicable to components of mastery pedagogy varies. Table 3 summarises this evidence (note that Q-E denotes quasi-experimental).

Table 3. A summary of evidence for the efficacy of the core components.

Component Group	Component	Effect	Alignment of MTE Practices to Evidence Base	Extent of the Evidence Base	Notes
Varied interactive teaching	Substantial whole-class varied teaching in multiple part lessons	Medium	Mixed	Limited	Lack of extensive randomised controlled trial (RCT)/quasi-experimental (Q-E) evidence on some aspects of lesson structures. Issue of relationship between direct instruction and MTE practices.
Varied interactive teaching	Interactive dialogue	Low to high	Closely aligned	Variable	A recent trial in England reports a low but positive effect size; other studies relevant to MTE practices report high effects. Trials of very similar practices more limited but other relevant Q-E evidence.
Mathematically meaningful and coherent activity	Depth and meaning	Low to high	Mixed	Variable	Lack of RCT/Q-E on task use. Multiple RCTs on problem-solving strategies.
Mathematically meaningful and coherent activity	Models and representations	Low to high	Closely aligned	Extensive	Multiple studies, RCT, and other interventions.
Mathematically meaningful and coherent activity	Mathematically coherent resources	N/A	Weakly aligned	Limited	A lack of RCT/Q-E on task use. RCT on textbook use focused on specific schemes so not applicable.
Full curriculum access for all	Curriculum pace for whole-class access	NA	Weakly aligned	Limited	Evidence from mastery learning trials may be applicable but this is not clear. Features of MTE practice have some similarity with reports of practices.
Full curriculum access for all	Teaching to attainment by deepening and support	Low	Mixed	Variable	Reported reviews are of overall negative effects of 'ability' grouping. However, relative lack of research in primary years.
Full curriculum access for all	Responsive teaching: formative assessment	Medium to high	Somewhat aligned	Extensive	Reported practices appear similar to those reported in the literature and the meta-analyses are based on a large number of interventions with different features.
Full curriculum access for all	Responsive intervention	Low to medium	Mixed	Extensive	Trials of one-to-one intervention are specific to programmes and may not be comparable to MTE embedded practices.
Knowledge of mathematical facts and language	Precise mathematical language	N/A	Weakly aligned	Limited	Single study and use of language embedded in another study.
Knowledge of mathematical facts and language	Memorising facts, relationships, and structures	Low to medium	Mixed	Limited	Three RCTs, however in the early primary years.
MTE Professional development	Collaborative, embedded and close to daily practice, and mathematically focused	Medium	Mixed	Variable	Evidence for impact of continuing professional development (CPD) on pupils is limited. More extensive evidence for an impact on teacher practices.

Some caution is needed when interpreting this summary. In some cases, effect sizes may be based on specific tests that are related to the practice. For example, the effect size reported of the efficacy of instruction periods to enhance factual knowledge is not a measure of general mathematical attainment but the effect on recall of factual knowledge. Further, these different possible effects cannot be simply added together. For example, an increased use of multiple representations can support greater classroom talk particularly of the sort requiring pupil explanation. However, it is also possible that the combination of these practices may have a multiplying effect, with some practices amplifying others. It may be that the effect of professional development itself would not be additional to the effect of potential changes of practice that stem from the professional development learning.

Overall, although this is a complex intervention and one that is evolving, the evidence for specific components suggests that it has the potential to lead to positive effects on pupil attainment. However, a key driver for the policy, as pointed to in the introduction to this paper, is the comparative difference between English outcomes on international tests and those in East Asian education systems. Although the review of the efficacy of components indicates the potential for improvements, these appear to be modest relative to differences on PISA tests.

For comparison, in 2012, the difference on PISA maths tests between Shanghai and the U.K. was 119 points [118]; this is on a test that was originally standardised to give a standard deviation of 100 and a mean of 500. If an equivalent of an effect size was to be calculated, then the standard deviation of both Shanghai and U.K./England test scores would need to be considered. However, an indication for comparison with the evidence discussed below is given by the estimated effect size equivalent of 1.19 standard deviations. Notwithstanding earlier cautions about translation of effect sizes into units of time, the Organisation for Economic Cooperation and Development OECD suggests that this difference is equivalent to approximately three years of schooling according to the OECD [119], and more than one year if using Education Endowment Foundation (EEF) guidance on converting standard deviation into months' progress [120]. This suggests that policy-makers' hopes may not be realised if the aim is to match outcomes in East Asia given the type of increases in outcomes as a result of an application of practices similar to those of the core components.

11. Priorities for Implementation and Evidence Gathering

In this section, we consider the implications of the analysis of evidence summarised in the previous section. We discuss priorities for implementation of a mastery pedagogy and professional experimentation and for further evidence gathering. Arguably, different variations of mastery pedagogies—indigenous East Asian pedagogy and the current teaching for mastery formulation—need to be considered in holistic ways as different practices that are mutually supportive. Nevertheless, the analysis of components of MTE mastery pedagogy presented here, and the review of the relevant evidence, allows us to identify specific practices that schools and teachers may choose to prioritise for implementation. Although mastery pedagogy is promoted as a coherent whole, in practice, schools and teachers make choices about what aspects to implement first [22,23]. What then does the review of evidence suggest about priorities for implementation if the main purpose is to change attainment outcomes?

We contend that evidence-informed priorities should be those practices for which there are:

- examples of evidence of medium or high efficacy for related practices; and
- a moderate or close relationship between this evidence and the mastery practices; and
- the evidence is robust.

Applying these criteria to the summary in Table 3 suggests that the following practices should be considered as priorities for implementation:

- Interactive dialogue and mathematical problems/tasks that promote this;
- A greater use of models and representations with multiple models used in relation to the same mathematical concepts; and

- Responsive teaching using formative assessment.

In addition, collaborative professional development that is mathematically focused, embedded, and close to practice is recommended to support change. The evidence for the impact of such professional development on attainment is limited. However, given (1) that it has been demonstrated to be effective in changing practice, and (2) it is a route to professional experimentation in relation to the above priorities, then it should be adopted by schools that are attempting to implement mastery approaches and/or the recommended components. Whilst the analysis presented has focused on MTE cohort 1 mastery pedagogy, these components or practices are recognisable within other mastery pedagogy formulations.

The practices that the analysis has pointed to as priorities for implementation align with general recommendations and guidance for mathematics teachers based on review (e.g., [73,74]). This is not surprising given that the same evidence base informs both the specific mastery-oriented review presented in this paper and more general reviews of effective practice in mathematics teaching. However, in addition to these recommendations for implementation at the school and classroom level, the analysis of the evidence also indicates areas for further specific evidence gathering in relation to mastery pedagogies, alongside evaluation of the implementation of mastery approaches as a whole. These specific practices require evaluation given that they are currently promoted in a range of formulations of mastery pedagogy. The analysis in this paper has focused on the mastery pedagogy implemented by MTE cohort 1 schools; given subsequent developments, formulations in the policy promoted 'teaching for mastery' would be the most appropriate foci.

Evaluation here would entail both considering the potential effect of the practices but also a more in-depth consideration of how practices are implemented. As noted above, there is a relationship between whole-class teaching as implemented in MTE cohort 1 schools and apparently similar practices for which there is evidence. An example is whole-class teaching in the MTE mastery pedagogy and direct instruction, for which there is an established evidence base [66]. Thus, well-designed trials in mastery contexts are recommended. Currently, the use of mathematically coherent resources—in particular textbooks—is being promoted, including through government subsidy for purchase. However, the evidence for the effectiveness of different sets of materials is not clear, and, as noted, textbooks, for example, are used in a variety of different ways [23]. This suggests the need for multiple arm trials with in-depth mixed-method case studies to support implementation and process evaluation.

A quickly implementable change that MTE cohort 1 schools made following visiting Shanghai [22] was to slow the pace at which material was covered, whilst at the same time maintaining or quickening the rhythm of interactions and movement between types of activity within lessons. This is advocated with teaching for mastery as a way of promoting depth and keeping the class together. However, whether this is an effective practice, and if so, for whom, and whether it has any negative consequences for some learners, is an open question. It is also beyond the scope of this paper to explore possible tensions between the content of the national curriculum with age-related expectations and a longer-term view of learning [93] involved in a slower curriculum pace. However, such tensions were identified by some of the MTE participants [21].

The emphasis on memorising mathematical facts appears to be supported by research on cognitive load theory, as well as having an a priori appeal: the more mathematical knowledge one has, the more likely one is going to be able to apply this. Similarly, the precise use of mathematical languages appears to be potentially positive. However, what is not clear is whether the practice in itself is effective or whether it is effective because it encourages dialogue, interaction, and thinking. Fostering precision in language can be embedded in different pedagogies.

Finally, whilst collaborative, embedded, and mathematically focused CPD was recommended above, given its potential effect on teacher practice, the evidence for its effect on pupils remains limited. Thus, further research is required and specifically on the forms of collaborative PD being developed and enacted through the mastery innovation.

In summary, the following recommendations are made for prioritising the implementation of practices that are promoted under the banner of mastery by teachers and schools based on the extant evidence for effect on attainment:

- Interactive dialogue and mathematical activity that promotes this;
- Greater use of models and representations;
- Responsive teaching using formative assessment; and
- Mathematically focused, embedded, and close to practice collaborative professional development to change practice.

Further, the analysis presented herein has identified the following practices that require additional research and evaluation to establish their effectiveness in relation to attainment.

- Whole class mastery teaching;
- Use of textbooks (and specifically those recommended to schools);
- Curriculum pace;
- Memorising factual knowledge;
- Precise use of language; and
- mastery-focused CPD.

That priorities are identified does not mean that other components of mastery pedagogies are not also potentially of value and worth implementing, but that, as yet, the evidence for their efficacy is not established.

12. Conclusions

In the introduction to the paper, we identified a number of concerns that arise when evaluating complex policy innovations with multiple components and actors. Such concerns are heightened when policy innovation is informed or justified by transnational comparison fuelled by PISA and the influence of the OECD as well as other transnational bodies, and the innovation involves transnational flows or policy mobilities [13].

The approach we have taken in this paper is to analyse a complex, at-scale intervention and then interrogate the evidence base. This novel approach is recommended for consideration by researchers analysing similar innovations, whether the initiatives are informed by transnational practices or not. Further, if the components of an intervention are specified in advance—for example, after piloting—then the process we have outlined could be used to inform the decisions of policy-makers to consider before rather than during or after an innovation has taken place at scale.

The mastery innovation in England is not only an at-scale policy initiative but also concretises the growing influence of international comparisons and transnational flows in mathematics education. In the mastery ecology, there have been a variety of descriptions put forward of the meaning of mastery, including several formulations in relation to mastery as pedagogy. In this paper, we have presented a conceptual framework for a mastery based on a systematic study of schools involved in the Mathematics Teacher Exchange, which are also influenced by the other various initiatives and perspectives. Thus, the framework describes what schools are doing as a set of core components rather than an advocated-for pedagogy.

In discussing the mastery innovation, we pointed to ways in which Shanghai (and to an extent other East Asian) practices were different to English primary mathematics practices and we extended this discussion by considering the core components. Many schools, though not all, participating in the MTE report changes in practices influenced by the MTE experience [23].

Looking beyond the MTE schools, the ambition for mastery is to impact on primary mathematics education generally [14] and so to lead to systemic change; hence, the investment in further exchanges and the Primary Mathematics Teaching for Mastery (PMTMS) programme, a CPD programme to train 'mastery specialists'. However, although part of the original intention was for all schools involved in the MTE to share their learning with other schools, this has not happened in all cases. However, through additional support via the PMTMS programme, this is likely to occur with schools participating in future exchanges. Nevertheless, the experience of the National Numeracy Strategy is that there can be gaps between what schools say they are doing [71] and what they are actually doing, and interpretations of, in that case, 'whole-class interactive teaching' can vary considerably. Further, although the funding to support the mastery initiative appears to be substantial, this level of support for primary schools is still quite small given the size of the sector. Thus, whether policy ambitions will be realised in terms of system-wide changes of practices is yet to be determined.

Before concluding, we acknowledge a number of limitations of the paper. As noted above, the evaluation data largely comprise self-reports of lead teachers concerning practice; nevertheless, identifying core components was feasible, if the data is understood as giving insight at least into intended changes. Further, as discussed earlier, our focus is specifically on changes in schools engaged in the MTE and, within that group, those schools that continued to engage with the evaluation and are actively implementing their mastery approaches. It is not necessarily the case that other schools who espouse adoption of mastery are acting in the same way. Although the specific formulation of mastery, used in this paper, differs from the NCETM descriptions of 'teaching for mastery', or Singaporean-informed programmes, such as Maths No Problem or Mathematics Mastery, there is a recognisable kinship. Moreover, due weight was given in identifying components to the implementation of those lead teachers who have subsequently engaged in the teaching for mastery professional development. In the evidence review, we found that evaluated practices previously reported vary in the extent to which they are aligned with practices found in the mastery innovation. Important questions about the potential effects of the adoption of different practices on both pupil and teacher affect towards mathematics have also not been addressed. One instance of this is in relation to all-attainment teaching, and, more generally, supporting low-attaining students to think mathematically (see [93]).

Putting aside these cautions, the analysis presented here indicates that, when considered as a set of components, many of the practices that are currently being implemented under the banner of 'mastery' in those schools who participated in MTE cohort 1 do represent approaches to teaching and learning mathematics that have evidence of being effective. However, outcomes of the MTE evaluation will help to determine whether this is realised in practice. The evidence summarised herein is also potentially useful for practitioners who are considering adopting some but not all of the mastery practices or considering how to phase their introduction, as we have identified which practices have the clearest existing evidence of effectiveness. In addition, we have pointed to areas for focused research on potentially effective components where there is limited evidence thus far of efficacy.

The analysis presented of the core components and their relation to evidence from previous studies of interventions provides the first in-depth and research-informed analysis of how these practices are a significant departure from previous dominant practices in mathematics teaching in England, and that, taken as a whole, there is evidence that the new practices could lead to higher attainment by pupils.

As noted above, the impact of the MTE and the implementation of mastery practices will be evaluated through statistical analysis [24]. If this demonstrates positive effects, then some of the implications that follow are:

- the evaluation would provide evidence that it is the pedagogy and classroom practice, rather than only or as well as other factors, that are important to the East Asian success in mathematics;
- that pedagogical practices can be transplanted successfully; and
- that the mastery innovation as implemented in MTE schools should be taken up by teachers and schools in England and considered by policy-makers elsewhere.

Although the outcomes of the evaluation of the MTE will not address these issues definitively, if no or little impact is found, this would suggest that the policy or its enactment should be reviewed and further evaluation undertaken of components or of the mastery approach as a whole.

The mastery innovation in England was not designed by a systematic review of evidence on effective practices in mathematics, but from a desire to emulate the success of Singapore and Shanghai. This has led to an ambitious programme to change mathematics teaching in England. In a globalised world, the influence of international comparison is unlikely to diminish in the near future. This means that critical analysis is needed of similarities and differences between practices in other contexts as compared to those in PISA reference societies, of how such practices are enacted, and what evidence there is for their efficacy in different educational systems, settings, and cultures. This paper has provided a model for undertaking such analysis, and has identified the potential for the effectiveness of components of East-Asian-informed mastery practices in England. Moreover, it has identified those practices, currently being implemented by a sample of schools adopting mastery approaches, that have evidence for effectiveness and so can be recommended for immediate implementation. It has also identified practices that are promoted as part of the mastery agenda for which there is, as yet, not enough evidence to recommend and that require further research. This does not mean that the latter category of practices should not also be the focus of professional experimentation by schools and teachers, but that such experimentation followed by systematic research is needed to establish their efficacy. In general, the identification of mastery components and the analysis of evidence points to the need for a more circumspect and careful approach to innovation informed by East Asian mathematics education if the aim is to impact on mathematical attainment. Whilst this is true in relation to England, it also has implications for other policy-makers and educators influenced by PISA outcomes: enthusiasm for change needs tempering with careful consideration of evidence.

Author Contributions: Conceptualization, M.B., B.M., T.J., and S.D.; Data curation, M.B. and C.W.; Formal analysis, M.B., B.M., C.W., T.J., and S.D.; Funding acquisition, M.B., B.M., and T.J.; Investigation, M.B., B.M., C.W., T.J., and S.D.; Methodology, M.B., B.M., T.J., and S.D.; Project administration, M.B., B.M., C.W., and T.J.; Resources, M.B.; Software, M.B., B.M. and C.W.; Supervision, M.B.; Writing (original draft), M.B., B.M., C.W., T.J., and S.D.; Writing (review & editing), M.B., B.M., C.W., and T.J.

Funding: The empirical research was funded by the Department for Education, England under the contract for the Longitudinal Evaluation of the Mathematics Teacher Exchange: China-England. The review of evidence, further analysis and writing of the paper was funded by Sheffield Hallam University.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. DfE. Mathematics Programmes of Study: Key Stages 1 and 2. Available online: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335158/PRIMARY_national_curriculum_-_Mathematics_220714.pdf (accessed on 12 November 2018).
2. Dye, T. *Understanding Government Policy*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1992.
3. Guskey, T.R. *Implementing Mastery Learning*, 2nd ed.; Wandsworth: Belmont, CA, USA, 1997.
4. Guskey, T.R. Lessons of mastery learning. *Educ. Leadersh.* **2010**, *68*, 52–57.
5. Slavin, R.E. Mastery learning reconsidered. *Rev. Educ. Res.* **1987**, *57*, 175–213. [CrossRef]
6. Brown, C. *Making Evidence Matter: A New Perspective for Evidence-Informed Policy Making in Education*; IOE Press: London, UK, 2013.
7. Brown, C. *Evidence-Informed Policy and Practice in Education: A Sociological Grounding*; Bloomsbury Publishing: London, UK, 2015.
8. Greany, T.; Maxwell, B. Evidence-informed innovation in schools: Aligning collaborative research and development with high quality professional learning for teachers. *Int. J. Innov. Educ.* **2017**, *4*, 147–170. [CrossRef]
9. NCETM. Mastery Approaches to Mathematics and the New National Curriculum. Available online: https://www.ncetm.org.uk/public/files/19990433/Developing_mastery_in_mathematics_october_2014.pdf (accessed on 12 November 2012).

10. NCETM The Essence of Mathematics Teaching for Mastery. Available online: https://www.ncetm.org.uk/files/37086535/The+Essence+of+Maths+Teaching+for+Mastery+june+2016.pdf (accessed on 12 November 2012).
11. Ochs, K. Cross-national policy borrowing and educational innovation: Improving achievement in the London Borough of Barking and Dagenham. *Oxf. Rev. Educ.* **2006**, *32*, 599–618. [CrossRef]
12. Winstanley, C. Alluring ideas: Cherry picking policy from around the world. *J. Philos. Educ.* **2012**, *46*, 516–531. [CrossRef]
13. Reynolds, D.; Muijs, D. The effective teaching of mathematics: A review of research. *Sch. Leadersh. Manag.* **1999**, *19*, 273–288. [CrossRef]
14. Gibb, N. Building a Renaissance in Mathematics Teaching. Available online: https://www.gov.uk/government/speeches/nick-gibb-building-a-renaissance-in-mathematics-teaching (accessed on 12 November 2018).
15. Haynes, L.; Service, O.; Goldacre, B.; Torgerson, D. Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials. Available online: https://www.gov.uk/government/publications/test-learn-adapt-developing-public-policy-with-randomised-controlled-trials (accessed on 12 November 2018).
16. Biesta, G. Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educ. Theory* **2007**, *57*, 1–22. [CrossRef]
17. Biesta, G. Why 'what works' still won't work: From evidence-based education to value-based education. *Stud. Philos. Educ.* **2010**, *29*, 491–503. [CrossRef]
18. Weiss, C.H. Theory-based evaluation: Past, present, and future. *New Dir. Eval.* **1997**, *76*, 41–55. [CrossRef]
19. Rogers, P.J. Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation* **2008**, *14*, 29–48. [CrossRef]
20. Durlak, J.A.; DuPre, E.P. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* **2008**, *41*, 327–350. [CrossRef] [PubMed]
21. Boylan, M.; Wolstenholme, C.; Maxwell, B.; Jay, T.; Stevens, A.; Demack, S. Evaluation of the Mathematics Teacher Exchange: China-England. Available online: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/536003/Mathematics_Teacher_Exchange_Interim_Report_FINAL_040716.pdf (accessed on 15 November 2018).
22. Boylan, M.; Wolstenholme, C.; Maxwell, B. Evaluation of the Mathematics Teacher Exchange: China-England. Available online: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/666450/MTE_third_interim_report_121217.pdf (accessed on 15 November 2018).
23. Demack, S.; Jay, T.; Boylan, M.; Wolstenholme, C. Longitudinal Evaluation of the Mathematics Teacher Exchange: China-England. Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/666449/MTE_second_interim_report_121217.pdf (accessed on 15 November 2018).
24. Askew, A.; Bishop, S.; Christie, C.; Eaton, S.; Griffin, P.; Morgan, D. Teaching for Mastery: Questions, Tasks and Activities to Support Assessment. Available online: https://www.ncetm.org.uk/resources/46689 (accessed on 12 November 2018).
25. ATM/MA Primary Working Group. What does 'mastery' mean to me? *ATM* **2016**, *251*, 28–29.
26. Boylan, M.; Townsend, V. Understanding mastery in primary mathematics. In *Learning to Teaching in the Primary School*, 4th ed.; Cremin, T., Burnett, C., Eds.; Routledge: London, UK, 2017.
27. National Association of Mathematics Advisors. Five myths of mastery in Mathematics. *ATM 2016, 20–23.* Available online: http://www.nama.org.uk/Downloads/Five%20Myths%20about%20Mathematics%20Mastery.pdf (accessed on 15 November 2018).
28. Dweck, C.S. *Mindset: The New Psychology of Success*; Random House Incorporated: New York, NY, USA, 2006.
29. DfE. Primary Assessment and Accountability under the New National Curriculum. Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/298568/Primary_assessment_and_accountability_under_the_new_curriculum_consultation_document.pdf (accessed on 13 November 2018).
30. Askew, M.; Hodgen, J.; Hossain, S.; Bretscher, N. *Values and Variables: Mathematics Education in High-Performing Countries*; Nuffield: London, UK, 2010.

31. Waldow, F.; Takayama, K.; Sung, Y.-K. Rethinking the pattern of external policy referencing: Media discourses over the 'Asian Tigers" PISA success in Australia, Germany and South Korea. *Comp. Educ.* **2014**, *50*, 302–321. [CrossRef]
32. Sellar, S.; Lingard, B. Looking East: Shanghai, PISA 2009 and the reconstitution of reference societies in the global education policy field. *Comp. Educ.* **2013**, *49*, 464–485. [CrossRef]
33. NCSL. *Report on Research into Maths and Science Teaching in the Shanghai Region*; NCSL: Nottingham, UK, 2013.
34. NCSL. *Report on the International Maths Research Programme, China 2014*; NCSL: Nottingham, UK, 2014.
35. Lopez-Real, F.J.; Mok, I.A.C.; Leung, F.K.S.; Marton, F. Identifying a pattern of teaching: Analysis of a Shanghai teacher's lessons. In *How Chinese Learn Mathematics: Perspectives from Insiders*; World Scientific: Hackensack, NJ, USA, 2004; Volume 1, pp. 382–410.
36. Miao, Z.; Reynolds, D. The effectiveness of mathematics teaching: a cross-national investigation in primary schools in England and China (Partial findings). In Proceedings of the BERA Conference 2014, London, UK, 23–25 September 2014.
37. Miao, Z.; Reynolds, D. Uncovering classroom-level factors for the performance gap in mathematics between England and China: A synthesis of results and findings from an international teacher effectiveness study. In Proceedings of the BERA Conference 2015, Belfast, UK, 15–17 September 2015.
38. Clarke, D.; Xu, L.H.; Wan, M.E.V. Student speech as an instructional priority: Mathematics classrooms in seven culturally-differentiated cities. *Procedia Soc. Behav. Sci.* **2010**, *2*, 3811–3817. [CrossRef]
39. Gu, L.; Huang, R.; Marton, F. Teaching with variation: A Chinese way of promoting effective mathematics learning. In *How Chinese Learn Mathematics: Perspectives from Insiders*; World Scientific: Hackensack, NJ, USA, 2004; Volume 1, pp. 309–347.
40. An, S.; Kulm, G.; Wu, A. Pedagogical content knowledge of middle school mathematics teachers in China and the U.S. *J. Math. Teach. Educ.* **2004**, *7*, 145–172. [CrossRef]
41. Huang, R.-J.; Leung, K.S. Cracking the paradox of Chinese learners: Looking into the mathematics classrooms in Hong Kong and Shanghai. In *How Chinese Learn Mathematics: Perspectives from Insiders*; World Scientific: Hackensack, NJ, USA, 2004; Volume 1, pp. 348–381.
42. Al-Murani, T.; Kilhamn, C.; Morgan, D.; Watson, A. Opportunities for Learning: The Use of Variation to Analyse Examples of a Paradigm Shift in Teaching Primary Mathematics in England. Available online: https://doi.org/10.1080/14794802.2018.1511460 (accessed on 13 November 2018).
43. Hallam, S.; Parsons, S. Prevalence of streaming in UK primary schools: Evidence from the Millennium Cohort Study. *Br. Educ. Res. J.* **2013**, *39*, 514–544. [CrossRef]
44. Bruner, J.S. *Toward a Theory of Instruction*; Harvard University Press: Cambridge, MA, USA, 1966.
45. Hoong, L.Y.; Kin, H.W.; Pien, C.L. Concrete-Pictorial-Abstract: Surveying its origins and charting its future. *Math. Educ.* **2015**, *16*, 1–18.
46. Ball, S. New philanthropy, new networks and new governance in education. *Political Stud.* **2008**, *56*, 747–776. [CrossRef]
47. Ball, S.; Exley, S. Making policy with 'good ideas': Policy networks and the 'intellectuals' of new labour. *J. Educ. Policy* **2010**, *25*, 151–169. [CrossRef]
48. Noyes, A.; Adkins, M. The impact of research on policy: A case of qualifications reform. *Br. J. Educ. Stud.* **2016**, *64*, 449–465. [CrossRef]
49. Jerrim, J.; Vignoles, A. The link between East Asian 'mastery' teaching methods and English children's mathematics skills. *Econ. Educ. Rev.* **2016**, *50*, 29–44. [CrossRef]
50. Vignoles, A.; Jerrim, J.; Cowan, R. Mathematics Mastery Primary Evaluation Report. Available online: https://v1.educationendowmentfoundation.org.uk/uploads/pdf/Mathematics_Mastery_Primary_(Final)1.pdf (accessed on 13 November 2018).
51. Hall, J.; Lindorff, A.; Sammons, P. Evaluation of the Impact and Implementation of Inspire Maths in Year 1 Classrooms in England: Findings from a Mixed-Method Randomised Control Trial. Available online: https://ore.exeter.ac.uk/repository/handle/10871/24265 (accessed on 13 November 2018).
52. PM to Set out Ambitious New Approach to UK-China Education. Available online: https://www.gov.uk/government/news/pm-to-set-out-ambitious-new-approach-to-uk-china-education (accessed on 13 November 2018).

53. Creswell, J.W.; Plano Clark, V.L.; Gutmann, M.L.; Hanson, W.E. Advanced mixed methods research designs. In *Handbook of Mixed Methods in Social and Behavioral Research*; Hammersley, M., Ed.; SAGE Publishing: Newbury Park, CA, USA, 2003.
54. Ethical Guidelines for Educational Research. Available online: https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2011 (accessed on 13 November 2018).
55. Brandon, P.R.; Fukunaga, L.L. The state of the empirical research literature on stakeholder involvement in program evaluation. *Am. J. Eval.* **2014**, *35*, 26–44. [CrossRef]
56. Layder, D. *Sociological Practice: Linking Theory and Social Research*; Sage: London, UK, 1998.
57. Ryan, G.; Bernard, H. Techniques to identify themes. *Field Methods* **2003**, *15*, 85–109. [CrossRef]
58. Blase, K.; Fixsen, D. Core Intervention Components: Identifying and Operationalizing What Makes Programs Work. ASPE Research Brief. Available online: http://files.eric.ed.gov/fulltext/ED541353.pdf (accessed on 13 November 2018).
59. Cooksy, L.J.; Gill, P.; Kelly, P.A. The program logic model as an integrative framework for a multimethod evaluation. *Eval. Program Plan.* **2001**, *24*, 119–128. [CrossRef]
60. Cordingley, P.; Higgins, S.; Greany, T.; Buckler, N.; Coles-Jordan, D.; Crisp, B.; Saunders, L.; Coe, R. *Developing Great Teaching: Lessons from the International Reviews into Effective Professional Development*; Teacher Development Trust: London, UK, 2015.
61. Hattie, J. *Visible Learning for Teachers: Maximizing Impact on Learning*; Routledge: Abingdon, UK, 2012.
62. Cohen, J. *Statistical Power Analysis for the Behavioural Sciences*; Erlbaum: Hillsdale, NJ, USA, 1988.
63. Higgins, S.; Katsipataki, M.; Kokotsaki, D.; Coleman, R.; Major, L.E.; Coe, R. *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit*; Sutton Trust: London, UK, 2014.
64. Sammons, P.; Taggart, B.; Siraj-Blatchford, I.; Sylva, K.; Melhuish, E.; Barreau, S. Variations in Teacher and Pupil Behaviours in Y5 Classes. Available online: https://ro.uow.edu.au/cgi/viewcontent.cgi?referer=https://scholar.google.co.uk/&httpsredir=1&article=2693&context=sspapers (accessed on 13 November 2018).
65. Hattie, J. *Visible learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*; Routledge: Abingdon, UK, 2009.
66. Stockard, J.; Wood, T.W.; Coughlin, C.; Khoury, C. The effectiveness of direct instruction curricula: A meta-analysis of a half century of research. *Rev. Educ. Res.* **2018**, *88*. [CrossRef]
67. Good, T.L.; Grouws, D.A.; Ebmeier, H. *Active Mathematics Teaching*; Longman: New York, NY, USA, 1983.
68. Ruthven, K. Using international study series and meta-analytic research syntheses to scope pedagogical development aimed at improving student attitude and achievement in school mathematics and science. *Int. J. Sci. Math. Educ.* **2011**, *9*, 419–458. [CrossRef]
69. Dunlosky, J.; Rawson, K.A.; Marsh, E.J.; Nathan, M.J.; Willingham, D.T. Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychol. Sci. Public Interest* **2013**, *14*, 4–58. [CrossRef] [PubMed]
70. Alexander, R.J. *Towards Dialogic Teaching: Rethinking Classroom Talk*, 4th ed.; Dialogos: York, UK, 2015.
71. Jay, T.; Willis, B.; Thomas, P.; Taylor, R.; Moore, N.; Burnett, C.; Merchant, G.; Stevens, A. *Dialogic Teaching: Evaluation Report and Executive Summary*; Education Endowment Foundation: London, UK, 2017.
72. Mercer, N.; Sams, C. Teaching children how to use language to solve maths problems. *Lang. Educ.* **2006**, *20*, 507–528. [CrossRef]
73. EEF. *Improving Mathematics in Key Stages Two and Three*; EEF: London, UK, 2017.
74. Hodgen, J.; Foster, C.; Marks, R.; Brown, M. *Improving Mathematics in Key Stages 2 and 3: Evidence Review*; EEF: London, UK, 2018.
75. Woodward, J.; Beckmann, S.; Driscoll, M.; Franke, M.; Herzig, P.; Jitendra, A.; Koedinger, K.R.; Ogbuehi, P. *Improving Mathematical Problem Solving in Grades 4 Through 8: A Practice Guide (NCEE 2012-4055)*; National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education: Washington, DC, USA, 2012.
76. Barmby, P.; Bolden, D.; Raine, S.; Thompson, L. Developing the use of diagrammatic representations in primary mathematics through professional development. *Educ. Res.* **2013**, *55*, 263–290. [CrossRef]
77. Sowell, E.J. Effects of manipulative materials in mathematics instruction. *J. Res. Math. Educ.* **1989**, *20*, 498–505. [CrossRef]

78. Gersten, R.; Chard, D.J.; Jayanthi, M.; Baker, S.K.; Morphy, P.; Flojo, J. Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Rev. Educ. Res.* **2009**, *79*, 1202–1242. [CrossRef]
79. Gersten, R.; Beckmann, S.; Clarke, B.; Foegen, A.; Marsh, L.; Star, J.R.; Witzel, B. Assisting Students Struggling with Mathematics: Response to Intervention (RtI) for Elementary and Middle Schools. Available online: https://ies.ed.gov/ncee/wwc/PracticeGuide/2 (accessed on 13 November 2018).
80. Hodgen, J.; Coe, R.; Brown, M.; Küchemann, D. Improving students' understanding of algebra and multiplicative reasoning: did the ICCAMS intervention work? In Proceedings of the 8th British Congress of Mathematics Education, University of Nottingham, Nottingham, UK, 14–17 April 2014.
81. Searle, J.; Barmby, P. Evaluation Report on the Realistic Mathematics Education Pilot Project at Manchester Metropolitan University. Available online: http://mei.org.uk/files/pdf/rme_evaluation_final_report.pdf (accessed on 13 November 2018).
82. Carbonneau, K.J.; Marley, S.C.; Selig, J.P. A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *J. Educ. Psychol.* **2013**, *105*, 380. [CrossRef]
83. Doias, E.D. The Effect of Manipulatives on Achievement Scores in the Middle School Mathematics Class. Ph.D. Thesis, Lindenwood University, Saint Charles, MI, USA, 2013.
84. Kontas, H. The Effect of Manipulatives on Mathematics Achievement and Attitudes of Secondary School Students. *J. Educ. Learn.* **2016**, *5*, 310–320. [CrossRef]
85. Ojose, B.; Sexton, L. The effect of manipulative materials on mathematics achievement of first grade students. *Math. Educ.* **2009**, *12*, 3–14.
86. Fan, L.; Zhu, Y.; Miao, Z. Textbook research in mathematics education: development status and directions. *ZDM* **2013**, *45*, 633–646. [CrossRef]
87. Slavin, R.E.; Lake, C. Effective programs in elementary mathematics: A best-evidence synthesis. *Rev. Educ. Res.* **2008**, *78*, 427–515. [CrossRef]
88. Shield, M.; Dole, S. Assessing the potential of mathematics textbooks to promote deep learning. *Educ. Stud. Math.* **2013**, *82*, 183–199. [CrossRef]
89. Williams, P. *Independent Review of Mathematics Teaching in Early Years Settings and Primary Schools: Final Report*; DCSF: London, UK, 2008.
90. Hill, H.C.; Rowan, B.; Ball, D.L. Effects of teachers' mathematical knowledge for teaching on student achievement. *Am. Educ. Res. J.* **2005**, *42*, 371–406. [CrossRef]
91. Black, L. Interactive whole class teaching and pupil learning: Theoretical and practical implications. *Lang. Educ.* **2007**, *21*, 271–283. [CrossRef]
92. Boaler, J. Setting, social class and survival of the quickest. *Br. Educ. Res. J.* **1997**, *23*, 575–595. [CrossRef]
93. Watson, A.; De Geest, E. Principled teaching for deep progress: Improving mathematical learning beyond methods and materials. *Educ. Stud. Math.* **2005**, *58*, 209–234. [CrossRef]
94. Hallam, S.; Ireson, J. Secondary school teachers' attitudes towards and beliefs about ability grouping. *Br. J. Educ. Psychol.* **2003**, *73*, 343–356. [CrossRef] [PubMed]
95. Ramaprasad, A. On the definition of feedback. *Behav. Sci.* **1983**, *28*, 4–13. [CrossRef]
96. Wiliam, D. *Embedded Formative Assessment*; Solution Tree Press: Bloomington, IN, USA, 2011.
97. Hattie, J.; Timperley, H. The Power of Feedback. *Rev. Educ. Res.* **2007**, *77*, 81–112. [CrossRef]
98. *Children's Errors in Mathematics*, 4th ed.; Hansen, A., Ed. Sage: London, UK, 2017.
99. Ryan, J.; Williams, J. *Children's Mathematics 4-15: Learning from Errors and Misconceptions*; McGraw-Hill Education: London, UK, 2007.
100. Dowker, A. What Works for Children with Mathematical Difficulties? DfES Publications: London, UK, 2004; Volume 554.
101. Blatchford, P.; Russell, A.; Bassett, P.; Brown, P.; Martin, C. The role and effects of teaching assistants in English primary schools (Years 4 to 6) 2000–2003. Results from the Class Size and Pupil–Adult Ratios (CSPAR) KS2 Project. *Br. Educ. Res. J.* **2007**, *33*, 5–26. [CrossRef]
102. Torgerson, C. Randomised controlled trials in education research: A case study of an individually randomised pragmatic trial, Education 3-13. *Int. J. Prim. Elem. Early Year. Educ.* **2009**, *37*, 313–321. [CrossRef]
103. Torgerson, C.; Wiggins, A.; Torgerson, D.; Ainsworth, H.; Hewitt, C. Every child counts: Testing policy effectiveness using a randomised controlled trial, designed, conducted and reported to CONSORT standards. *Res. Math. Educ.* **2013**, *15*, 141–153. [CrossRef]

104. Smith, T.M.; Cobb, P.; Farran, D.C.; Cordray, D.S.; Munter, C. Evaluating math recovery: Assessing the causal impact of a diagnostic tutoring program on student achievement. *Am. Educ. Res. J.* **2013**, *50*, 397–428. [CrossRef]
105. EEF (n.d.). Same Day Intervention. Available online: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/same-day-intervention/ (accessed on 13 November 2018).
106. Cardelle-Elawar, M. Effects of metacognitive instruction on low achievers in mathematics problems. *Teach. Teach. Educ.* **1995**, *11*, 81–95. [CrossRef]
107. VanDerHeyden, A.; McLaughlin, T.; Algina, J.; Snyder, P. Randomized evaluation of a supplemental grade-wide mathematics intervention. *Am. Educ. Res. J.* **2012**, *49*, 1251–1284. [CrossRef]
108. Huang, R.; Peng, S.; Wang, L.; Li, Y. Secondary mathematics teacher professional development in China. In *Reforms and Issues in School Mathematics in East Asia*; Leung, F.K.S., Li, Y., Eds.; Sense Publishers: Rotterdam, The Netherlands, 2010; pp. 129–152.
109. Yoon, K.S.; Duncan, T.; Lee, S.W.Y.; Scarloss, B.; Shapley, K.L. Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement. Available online: https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf (accessed on 13 November 2018).
110. Stoll, L.; Harris, A.; Handscomb, G. *Great Professional Development Which Leads to Great Pedagogy: Nine Claims from Research*; National College for School Leadership: Nottingham, UK, 2012.
111. Timperley, H.; Wilson, A.; Barrar, H.; Fung, I. *Teacher Professional Learning and Development*; International Academy of Education: Brussels, Belgium, 2008.
112. Garet, M.S.; Porter, A.C.; Desimone, L.; Birman, B.F.; Yoon, K.S. What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* **2001**, *38*, 915–945. [CrossRef]
113. Cordingley, P.; Bell, M.; Rundell, B.; Evans, D. The Impact of Collaborative CPD on Classroom Teaching and Learning. In *Research Evidence in Education Library*; EPPI Centre, Social Science Research Unit, Institute of Education: London, UK, 2003.
114. Lomos, C.; Hofman, R.H.; Bosker, R.J. Professional communities and student achievement—A meta-analysis. *Sch. Eff. Sch. Improv.* **2011**, *22*, 121–148. [CrossRef]
115. Yang, Y. How a Chinese teacher improved classroom teaching in teaching research group: A case study on Pythagoras theorem teaching in Shanghai. *ZDM Math. Educ.* **2009**, *41*, 279–296. [CrossRef]
116. Orton, J.D.; Weick, K.E. Loosely coupled systems: A reconceptualization. *Acad. Manag. Rev.* **1990**, *15*, 203–223. [CrossRef]
117. Lewis, C.; Perry, R.; Murata, A. How should research contribute to instructional improvement? The case of lesson study. *Educ. Res.* **2006**, *35*, 3–14. [CrossRef]
118. OECD. PISA 2012 Results in Focus: What 15-year-Olds Know and What They Can Do with What They Know. Available online: https://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf (accessed on 13 November 2018).
119. Jerrim, J.; Wyness, G. Benchmarking London in the Pisa Rankings. *Lond. Rev. Educ.* **2016**, *14*, 38–65. [CrossRef]
120. EEF Report Template. Available online: https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluating-projects/evaluator-resources/writing-a-research-report/ (accessed on 13 November 2018).