

Problematizing lesson study and its impacts: studying a highly contextualised approach to professional learning

Mark Hadfield, Cardiff University

Michael Jopling, Northumbria University

Abstract

Researching the efficacy of lesson study has been complicated by the manner in which it has been reinterpreted in different settings and contexts. Drawing on research into a national school improvement programme in England, this paper looks at how various reconfigurations of lesson study were affected by the different collaborative arrangements among schools and practitioners in the programme. Utilising a mixed methods approach, the paper provides and problematizes new empirical evidence concerning the effectiveness of lesson study. It concludes with a consideration of how those researching lesson study initiatives would benefit from adopting a more critical contextual analysis.

1. Introduction

Lesson Study originated in Japan and became a global phenomenon after being brought to worldwide attention by the Third International Trends in Mathematics and Science Study (TIMSS, 1998). Its popularity has grown steadily since. Xu and Pedder (2015) state that the World Association of Lesson Studies, established in Hong Kong in 2005, has members from over 60 countries and a recent estimate was that it has been used in over 2500 primary and secondary schools in England (Dudley, 2014). Since the 1990s, lesson study¹ has been applied internationally to a range of curriculum areas and reinterpreted in numerous professional development programmes (Dudley, 2013; Lewis, Perry, & Hurd, 2004; Lieberman, 2009).

Researching the potential impact of lesson study has been made problematic by its versatility (Xu & Pedder, 2009) and the variability of its reinterpretations, making it difficult to identify its common characteristics (Perry & Lewis, 2009; Murata, 2011). In the UK the widespread adoption, and reinterpretation, of lesson study has occurred against a backdrop of broader school improvement phases that have focused on integrating professional learning and school development (Harris & Chrispeels, 2008; Hopkins, 2016). The different phases have formed a key part of the “practice architecture” (Kemmis & Grootenboer, 2008), or preconditions, that have affected how lesson study has been enacted in classrooms and schools.

In this paper we consider the effect of a range of contextual influences on the lesson study model used within a national school improvement initiative. The initiative was the Leading Teachers (LT) Programme which set out to improve pupil attainment in primary schools in England by training almost 1000 lead teachers in lesson study. Programme data indicated that 1175 primary schools in 92 local authorities (LAs, which are equivalent to school districts), representing 61 per cent of all LAs in England, were involved in the programme between 2008 and 2010. The research is based on case studies of three LAs where lesson study was credited with playing a key role in improving pupil attainment. The cases describe how existing approaches to school improvement and collaborative conditions within LAs prefigured how lesson study was interpreted and enacted and in doing so shaped the nature of the impacts on teachers and pupils.

2. Problematizing the impact of lesson study: the role of context(s)

From a post-positivist perspective, problematizing the outcomes from engagement in lesson study involves developing a more complex conceptualisation of the role played by the context in which it takes place. This raises a number of methodological and epistemological issues primarily because adopting such a perspective undermines simplistic notions of linear causality, based on regularity and predictability, and treats the links between a practice and its effects as belonging to a complex set of relationships with the contexts in which they are achieved (Ling, 2012). The methodological implications of adopting a post-positivist perspective have been extensively debated in evaluation research (Mowles, 2014; Jolley, 2014). The aim of this section is not to repeat these general arguments but to present a reflective critique of how the relationship between practice and context has often been under-theorised in research on lesson study. In doing so, we distinguish between three levels of context that affect how a particular model of lesson study was implemented and

¹ Throughout this article “lesson study” in lower case is used to refer to the wide range of professional learning processes that have labeled themselves as Lesson Studies, although these reinterpretations might differ significantly with regard to their focus, process and/or underpinning theories.

received: the national/local system (macro), the organisation (meso), and lesson study process itself (micro).

In this critique we have drawn on the findings of what is so far the only systematic review of lesson study (Cheung & Wong, 2014), which considered what constitutes lesson study, the nature of its intended effects, and the relationship between actual effects and the practices of lesson study by examining the evidence base relating to Lesson Study and its effectiveness, as reported in journal articles published between 2000 and 2010.

2.1 System context (macro)

The macro level of contextual analysis focuses on differences between and within educational systems. The issue of variations between different education systems' interpretation of the practice known as *jugyou kenkyuu* in Japan and renamed Lesson Study elsewhere (Yoshida, 1999; Watanabe, 2002) has been the focus of considerable debate (Perry & Lewis 2009; Yoshida & Jackson, 2011). The issue is exacerbated by a paucity of case studies that set the process within its original Japanese context (Ermeling & Graff-Ermeling, 2014). This led Cheng and Ling (2013) to argue that widespread uptake of lesson study in North America:

is based more on the perception of an ideal rather than on a fidelity approach. This is true of all countries adapting the lesson study approach. This also explains why the interpretations and practices of lesson study vary so widely across countries (p. 2).

The debate over what constitutes a lesson study has resulted in various attempts to define the key characteristics that determine whether a particular professional learning process is a “lesson study” (Takahashi, Watanabe, & Yoshida, 2006; Perry & Lewis, 2009). The resulting definitions tend to be lists of core practice elements with few references to contextual features. Cheung and Wong's (2014) systematic review only identified 20 articles that they considered to be examples of lesson study or the related process of Learning Study. The criteria for rejecting studies were based on the practices described in the original research, with little consideration of the role lesson study played in the local education systems.

“Between system” analyses provide a better understanding of contextual differences in the purposes and intended outcomes that led to the adoption of lesson study. “Within-system” analyses are important when examining how lesson study has been positioned and reinterpreted over time, reflecting changes in national and local policy agendas, and how this affects the intentions behind its adoption, and the impacts sought. The need for more within-system contextual analyses of lesson study can be illustrated with reference to the macro policy context of the English education system, specifically the relationship between professional development and school improvement. Lesson study in England was initially popularised during the first phase of school improvement research (Harris, 2002; Hopkins, 2003). One of the first publications to draw attention to lesson study after TIMSS (1998) was Stigler and Hiebert's (1999) *The Teaching Gap* which focused on articulating, validating and disseminating what were considered effective teaching practices. The policy discourse at this point was concerned with identifying gaps in teaching quality, rather than variations in pupil performance, and lesson study was seen as a promising approach to professional development. As school improvement policy subsequently moved towards assessing the impact of school-based interventions on pupil test scores (Cordingley, Bell, Rundell, & Evans, 2003), lesson study became increasingly judged on this criterion. In later phases of school improvement in England, the focus has moved towards system-level change and leadership, set within a discourse of the “self-improving system” (DfE, 2010, Hargreaves 2012). System-level change was premised upon developing capacity at different levels. Schools were encouraged to generate sufficient

capacity to meet their own improvement needs before supporting others, leading to a range of collaborative initiatives (Author(s) et al., 2009). Along with approaches such as walk-throughs (Bickford, 2010) and instructional rounds (City, Elmore, Fiarman, & Teitel, 2009), lesson study became increasingly positioned as a means of engendering the forms of joint practice development, reciprocal knowledge transfer and distributed leadership regarded as key elements of capacity building within a self-improving system (Mourshed, Chijioke, & Barber, 2010; Hargreaves, 2012), and assessments of its effectiveness began to shift to accommodate these elements.

A system-level contextual analysis allows for consideration of whether the adoption of lesson study in a locality, posited on the achievement of broad system-level impacts, has an effect on the way it is reconstructed at other levels. Cheung and Wong's (2014) review included only nine articles in its final analysis, mainly due to lack of evidence of impact. The notion of impact used was unfortunately restricted to meso and micro level analyses even though some of the articles reviewed contained studies that aspired to effect wider change. For example, Stewart and Brendefur's (2005) research involved 50 teams of teachers in 13 US school districts and was set explicitly within a discourse of system-level change based on the perceived failure of previous large-scale reform:

The systems we were working with were fragmented and overloaded, and the timelines were too long. Some teachers and administrators felt that there was too much going on. Others wanted to opt out because they could see little direct connection between the reform efforts and their day-to-day working lives (p. 682).

System level analyses are the starting point for understanding the extent to which contextual factors interact with a lesson study process. Factors from a system's new policy priorities to historical structural relationships can directly affect the lesson study model being promoted or exert indirect influence by means of the criteria used to judge its effectiveness. A contextual analysis at this level address effects on system level practices and structures, enabling researchers to differentiate between forms of lesson study on more holistic grounds than essentialist definitions based on key process characteristics. This level of analysis also allows for a consideration of how lesson study affects, and potentially re-configures, systems by changing professional development structures.

2.2 School context (meso)

A meso-level contextual analysis considers the interaction between organisational context and the lesson study process. From a positivist perspective, organisational contexts are treated as external to practice, a medium through which they pass, made up of moderating and mediating factors. From a post-positivist perspective, the relationship between context and practice is a dynamic of prefiguration and re-configuration, as practice is shaped by the school context and simultaneously reshapes it (Kemmis et al., 2014).

As a collaborative process of co-construction, lesson study is a social learning process that is close to practice and practitioners' needs. The adaptive nature of the lesson study cycle makes it sensitive to how practice issues are constructed and improvement is defined in a specific context (Maughan 2012; Koelner & Jacobs, 2015). As a complex set of practice arrangements, lesson study will therefore be prefigured to different degrees by schools' existing "practice architectures" (Kemmis & Grootenboer, 2008). These architectures are composed of the values and beliefs that underpin practice; the language used to describe it and justify its outcomes; the material conditions within which it is enacted; and the time, resources and practical arrangements provided, as well as the formal and informal roles and relationships of those involved.

The dynamic nature of the relationship between context and practice means lesson study has the

potential to change a school's practice architecture over time. The potential outcomes of lesson study therefore need not be restricted to either pupils or teachers but could also include effects on the contexts in which it takes place. Cheung and Wong (2014) only include two outcome categories in their systematic review: "effect on teacher or teaching" and "effect on students". Reference to broader organizational outcomes are only to be found in one study (Matoba, Shibata, Reza, & Arani, 2007), the only Japanese Lesson Study included, where reference is made to "higher % evaluation on school environment from teacher questionnaire" (Cheung & Wong, 2014, p.141). It is significant that at four years duration this was also one of the longest studies included in the review.

2.3 Lesson study process context (micro)

The final context that needs to be considered is that created by the lesson study process itself, its complex arrangement of practices and interactions are a "social site" (Chia and Holt, 2006) for learning. As with the other levels, there is a strong temporal dimension to the achievement of certain impacts on practice and professional understanding at this level. These sites for learning need to be sustained over time to bring about deep professional learning that results in the development of new professional theories and substantive professional growth (Clarke & Hollingsworth, 2002). In their review of 67 articles relating to lesson study, Xu and Pedder (2015, p. 38) noted that the short timespans and atheoretical nature of many the studies resulted in a narrow focus on:

questions of benefits and implementation challenges and so little research with questions of how teachers learn and develop practice through participation in [lesson study].

The issue of lesson studies being undertaken over insufficient number of cycles to prompt deeper learning was also reflected in Cheung and Wong's (2014) review in which only the longest study (Matoba et al., 2007) discussed any aspect of longer-term professional growth.

2.4 The reconstruction of Lesson Study in the context of a major government initiative

The LT programme was a later phase school improvement initiative that set out to improve practice in classrooms and schools, and build the leadership capacity required to reshape local education systems. To understand the approaches to lesson study that were enacted within it and the nature of their impacts required a degree of contextual analysis.

The LT programme was part of the National Strategies policy that ran in England from 1998 to 2011. It was the first national improvement policy that focused on changing how the core subject areas in the National Curriculum were taught, initially targeting literacy and numeracy and poorly performing LAs (Earl et al., 2003). Launched against a backdrop of concerns about the failure to improve pupil attainment and England's low standing in comparative studies of international performance (Keys, Harris, & Fernandes, 1996; Brooks, 1998), the National Strategies expanded to cover first all primary schools, then all of compulsory education. Initially, the Strategies were marked by a relatively hierarchical and centralised approach to reform based on a three-tier delivery model comprising the Department of Education's developing the training and materials, then LA advisers, who engaged the third tier: schools. The initial implementation model left relatively little space for innovation but over time more collaborative approaches were developed, including the LT programme.

Following a pilot in 2007, the LT programme ran from 2008 to 2010. Focusing on the final two years of primary school, Years 5 and 6, its objectives included:

training lead teachers to use the Lesson Study coaching cycle in order to help teachers increase the rate of progress of particular children.

Lead teachers were described in documentation from one of the case study authorities as classroom teachers with:

two or more years' worth of experience in a particular year group and their quality of teaching will have been judged either by Ofsted² or their own headteacher as consistently good or better over the preceding academic year. (Role description, Authority E)

The programme began with 14 LAs regarded as “coasting” in Spring 2008. By its conclusion 92 LAs had engaged in the programme. The guidance provided to LAs for the national roll-out contained materials on planning and evaluating lesson study in Mathematics and Literacy and a broad outline of the lesson study cycle (Figure 1). Lead teachers were funded to make an average of four visits to each teacher. While overall take-up and implementation of lesson study varied across LAs (Dudley, 2012), this research focuses on three LAs which adopted the guidance and implemented it strategically.

Insert Fig 1

The LT programme was constructed on the basis of collaborative professional learning, or school-led improvement, in that schools with acknowledged expertise in Mathematics and Literacy would be prepared to share this with other schools. The belief that primary schools would be prepared to engage in these types of reform was based on the experience of numerous collaborative initiatives that the English system had recently implemented, such as Excellence in Cities (Ofsted, 2003). The legacy of these programmes was not only a patina of collaborative arrangements in many areas but also a growing awareness among headteachers of the potential benefits of networking (Deloitte & Touche, 2000; Sammons, Mujtaba, Earl, & Gu, 2007). These collaborative programmes were not without their critics and participation in them was motivated by a complex mixture of moral purpose, voluntarism and instrumental pursuit of funding (Ainscow, Muijs, & West, 2006).

3. Purpose of the study

The overall purpose of the study was to assess the LT programme’s effectiveness as a “third wave” school improvement initiative (Fullan, 2009; Hargreaves & Shirley 2010), which meant that to be deemed effective it needed to raise teaching and learning standards, while simultaneously developing a local education system’s capacity to improve. Capacity development was based on creating cohorts of lead teachers to form supportive relationships between schools using lesson study as a collaborative professional learning process.

The research study had two sets of objectives. The first was to assess the short-term impact of the programme on pupil attainment and the medium-term impact on local systems’ capacity to support collaborative professional learning. The impact research questions were:

What evidence was there of short-term effects on pupil outcomes?

What evidence was there of any medium-term effects on the professional learning structures and habits in the case study authority areas?

² The school inspection agency in England.

The second set of objectives addressed the implementation of the programme, particularly the role context played in determining the way in which lesson study was enacted. The key contexts considered in this instance were the LA; collaborative relationships among the schools involved in the programme; and the interactions between the lead teacher and colleagues. The research questions that underpinned this were:

Local Authority: What were key factors that affected LAs' development and implementation of the programme?

Between and Within schools: What were the key factors that informed headteachers' decisions to engage with the Leading Teachers programme?

Within classrooms: What aspects of the lesson study cycle were seen as key in achieving changes in classroom practice?

4. Methodology

The research was designed as a multiple site case study exploring how the LT programme was implemented in three LAs. Each case was bounded by the development of a non-recursive linear path model (Desimone, 2009) that described the key steps between the programme's inception, as designed by the National Strategies, the enactment of the lesson study cycle in classrooms, and its potential impacts. This is outlined in Figure 2.

Insert Fig 2

The use of a logic model approach with a series of illuminative cases (Parlett & Hamilton, 1987) was an attempt to provide two types of insights: firstly, overarching insights into how the interaction of multiple factors, both internal and external to the programme, affected its intended outcomes, including improved pupil attainment; and secondly, context specific insights into the complexity of implementing the programme at a local level (Jolley, 2014).

4.1. Analytical framework

The analytical framework for the three case reports treated the development and implementation of the programme as a process of reconstruction at the three levels, or contexts, which framed the research questions (Bowe, Ball, & Gold, 1992). The cases focused on "enactment" (Ball, Maguire, & Braun, 2012): how a programme or policy is shaped by the interaction of individual and collective agency in overlapping contexts. The case reports were based on a two-stage analytical process adapted from framework analysis (Braun & Clarke, 2006). The initial stage consisted of a thematic analysis of the similarities and differences in the stated experiences of participants at the same level in each context in which the programme was reconstructed. The second stage related these themes to the existing research literature. The intention behind this two-stage analysis process was to capture variations between participants' initial reaction to, and reconstruction of, the programme across the three cases before exploring their significance with reference to the existing knowledge base.

Different aspects of research literature were used to explore each contextual level. At the LA level (macro), the effects of the local history of collaboration on the programme were related to a review of collaborative school improvement programmes (*Author(s), 2007*). Differences in how lesson study was enacted in schools (meso level) used a framework based on a systematic review of collaborative professional development (Cordingley et al., 2003) and research into school-based mentoring of teachers (Hobson, Ashby, Malderez, & Tomlinson, 2009). The in-classroom analysis (micro) was guided by Clarke and Hollingsworth's (2002) "Interconnected Model" of professional

learning. The Interconnected Model provided a theoretical formalisation of the lesson study cycle as a process of stimulating links, or interconnections, between different domains of teachers' professional worlds. The model views professional learning as resulting from a complex, non-linear combination of improved understandings, changes to practice and inferences about outcomes for pupils. Its key analytical power lay in the distinction it made between brief professional experimentation that resulted in short lived change sequences and sustained engagement that led to professional learning and growth.

4.2. Sampling

Purposive sampling (Mason, 2002) was used at two levels: LA and school. Authorities were selected on the basis of being involved in the programme for a minimum of two years, improvements in pupil test scores, and their approach to implementation. Secondary criteria were the size of the authorities and the number of schools where pupil performance at Key Stage 2³ were of concern. Three LAs were selected using these criteria:

- Authority E was selected for the funded programme in December 2007 because it was a large authority of 464 primary schools in which a high percentage of pupils were judged as having made poor progress across Key Stages 1⁴ and 2, assessed using contextual value added data (CVA, a statistic developed by government to assess the progress of children over time, incorporating external factors such as age and first language). 70 schools participated in the programme in this LA.
- Authority N was selected for the funded programme in December 2007 because it was a large authority of 295 primary schools, also with poor progress across Key Stages 1 and 2. 51 schools participated in the programme.
- Authority B joined the programme in September 2008. It was a medium-sized LA with 156 schools with over a third of schools (55) identified as making poor progress. 13 schools took part in the programme.

Five supported schools in each authority were selected for the research on the basis of having been involved in the programme for more than one year and improved pupil attainment. The views of LA staff about which schools had typified the approach to the programme were also taken into account. A secondary criterion was that schools which provided lead teachers were also willing to take part in the research. This resulted in a final sample of 22 participant schools.

4.3. Data collection and analysis

In each of the three LAs data collection took a similar form, which involved:

- Semi-structured interviews with LA officers with strategic and operational responsibility for the implementation of the LT programme in their authority. Seven officers were interviewed in the three LAs.
- Semi-structured interviews with headteachers and teachers from "leading" and "supported" schools in each LA. As far as possible these were partnered schools. A total of 21 headteachers, 12 lead teachers and 10 supported teachers were interviewed across the three LAs.
- Documentation, including programme materials, national and local school performance data and existing case studies, were collected from each LA and analysed.

³ For children aged 8-11.

⁴ For children aged 5-7.

A researcher was allocated to each LA, undertook all the interviews and partially transcribed them using an analytical note approach (Miles & Huberman, 1994). As the research took place after the final iteration of the programme, it was not possible to observe classroom practice directly. Evidence of changes to practice was based on documentation provided by LAs and schools and interviews with the teachers involved. Data analysis was undertaken by the three researchers in collaboration. The overall analytical flow followed the general structure of the logic model.

5. Results

The results section is divided into two parts addressing the two overarching objectives of the research results. The first part (5.1) analyses the programme as a “third wave” school improvement initiative with a focus on short-term effects on pupil outcomes and medium-term effects on capacity improvement, in this instance the dispositions of participants to continue leading collaborative professional learning and cross-school working. The second part (5.2) is a contextual analysis of how the LT programme was reconstructed at different levels in each authority and the ways in which this affected how lesson study was enacted across the three LAs and the outcomes that resulted.

5.1 Impact of the programme

5.1.1 What evidence was there of short-term effects on pupil outcomes?

Short-term effects on pupil outcomes were examined by analysing national attainment data supplied by the National Strategies. This indicated that 1175 primary schools were involved in the programme nationally in 2009 and 2010. Analysis was restricted by the fact that the Key Stage 2 data were aggregated at LA level, giving mean percentage scores in each LA for schools in and outside the LT programme. School and pupil level data were not available. This meant that analysis was restricted to calculating weighted mean scores comparing schools in the programme to those not in the programme in the same LAs and comparing the outcomes at LA level. Significance testing was not undertaken because the LA data was population data and the unavailability of school level data meant that effect sizes could not be calculated. LAs were included in the analysis where they had schools both in the programme (LT schools) and outside it (non-LT schools). This resulted in data being analysed for a total of 1147 LT schools and 5728 non-LT schools in 91 LAs. Cross-tabulation with national assessment data for 2010 indicated that the analysis involved included 265,412 pupils eligible for Key Stage 2, representing 63 per cent of all such pupils in England in 2010.

The analysis focused on Key Stage 2 results in English, Mathematics, and English and Mathematics combined at Level 4 or above as this was the expected attainment level for all children at the end of Key Stage 2. Analysis indicated that both programme and non-programme schools improved their results in Mathematics and English and Mathematics combined between 2009 and 2010, albeit marginally in some cases (see Appendix A Table 1). Results decreased slightly for non-LT schools in English (by 0.20 percentage points) but increased for LT schools (0.63). The biggest increases in weighted mean scores were for LT schools in Mathematics (2.86) and English and Mathematics combined (2.62). Comparing the weighted mean percentage of children achieving Level 4 in LT and non-LT schools in 2009 and 2010 indicated that while non-LT schools continued to outperform LT schools, LT schools narrowed the gap in all three areas. Again, slightly greater improvements were seen in Mathematics, narrowing the gap by 1.81 percentage points compared to 1.68 in English and Mathematics combined and 0.83 in English, reflecting improvements in Mathematics at Level 4 overall.

The analysis also looked at weighted means for the percentage of pupils achieving Level 5, that is children performing above expected levels. Here, although non-LT schools continued to outperform LT schools as was expected, results increased in all three areas for LT schools, while decreasing for non-LT schools in Mathematics (by 0.83). The biggest increases at Level 5 were in English for both non-LT schools (2.94) and LT schools (2.62) and the increase for LT schools in Mathematics was marginal (0.17). Consequently, while LT schools narrowed the gap in Mathematics at Level 5 (by 1.0), the gap increased slightly between 2009 and 2010 in both English (0.33) and English and Mathematics combined (0.06).

In terms of valued added scores, measured by the percentage of pupils making two or more levels of progress from Key Stage 1, results improved for both LT and non-LT schools. The gap in the mean percentage of children making such progress between LT and non-LT schools narrowed in both English (0.52) and Mathematics (1.74) in 2010 (see Appendix A Table 2). Thus, at a national level LT schools' Key Stage 2 results improved in all measures used and narrowed the achievement gap at Level 4 in the three subject areas and in Mathematics at Level 5. This all occurred against a national background in which Ofsted found that teaching and learning was good or outstanding in 69 per cent of primary schools in 2008-09 (Ofsted, 2009) and that teaching was good or outstanding in "over two-thirds" of primary schools in 2010 (Ofsted, 2010). However, the quality of teaching was also found to be highly variable: teaching was found to be only satisfactory in 39 per cent of the 4620 primary schools inspected between September 2009 and August 2010 (Ofsted, 2010).

Scatterplots were created to show the spread of outcomes, comparing the differences between results in 2010 and 2009 for English, Mathematics, and the subjects combined. Figures 3 and 4 indicate the most striking outcomes, both of which are at Level 5. Differences between outcomes for LT schools in each LA are indicated on the x axis, differences for non-LT schools are indicated on the y axis. Figure 3 illustrates the outcomes for English, which had the highest proportion of LAs with mean overall improvements for both LT and non-LT schools (66% - upper right quadrant) and the lowest proportion where results declined for both groups (1 school – lower right quadrant).

Insert Fig 3

Figure 4 indicates the outcomes for Mathematics at Level 5, which had the lowest proportion of LAs with improvements for both LT and non-LT schools (14% - upper right quadrant) and the highest proportion of negative outcomes for both groups (36% - lower right quadrant). Outliers are mostly LAs with low numbers of schools in the programme. The outcomes were more mixed at Level 4 (see Appendix A Figs 5-7) with Mathematics attracting a much higher proportion of overall positive scores for all LT and non-LT schools (60%) than English (34%). This underlines the large amount of variation in the system nationally and the challenges of trying to track and effect change across only two years of measurement.

Insert Fig 4

As already indicated, the fact that the Key Stage 2 data supplied by the National Strategies were not disaggregated below the LA level limited the analysis that could be undertaken for the three case study LAs to exploratory analysis. Of the case study LAs, N had the highest proportion of schools in the programme (27%), compared with 20 per cent in E and 18 per cent in B. All narrowed the performance gap between LT and non-LT schools in English and Mathematics combined at Level 4 or above between 2009 and 2010, by just over 7 per cent in the case of Authority B, and outperformed its geographical region as a whole. Only Authority E narrowed the gap in English and Mathematics combined at Level 5 (see Appendix A Table 4).

The picture was also mixed when the case study LAs' results were examined in more detail. Improvements in Authority B were greatest with LT schools outperforming non-LT schools in all areas (except English and Mathematics combined at Level 5). A similar pattern was evident in Authority N, whereas in E programme schools performed well in Mathematics and English and Mathematics combined at Level 4 but less so in English at both levels. This followed data produced in 2009, when Authority B had been in the programme for only a short time, when Level 4 results for programme schools were significantly below schools not involved (and below programme schools nationally). The selection of schools for inclusion was made more rigorous in Authority B in the second year of the programme to try to improve results. Following this, results improved for programme schools in Authority B in 2010 but the rate of progress in closing the gap slowed in E and N. However, LT schools showed improvements between 2009 and 2010 on all measures in the three LAs. Thus, all of the LAs appear in the upper right quadrant of Figure 3, indicating improvements for both LT and non-LT schools in English at Level 5, but in the lower right quadrant of Figure 4 because while the mean score for LT schools improved in Maths at Level 5, it declined for non-LT schools in all three LAs.

There was a range of different forms of evidence of more general impacts on pupils, mainly generated by LAs' own evaluations and feedback from lead teachers and headteachers. These highlighted enhanced pupil confidence, increased engagement and greater enjoyment. The following free responses from a survey of headteachers in the final year of the programme were fairly typical:

- Improved attitude of pupils with an increase in self-esteem and confidence and a reduction in anxiety around taking the SATs⁵.
- Greater engagement on homework tasks to support preparation for SATs.
- Positive enjoyment of working with a teacher other than their class teacher.
(Headteacher Questionnaire Feedback, 2010 Authority N)

5.1.2. What evidence was there of any medium-term effects on the professional learning structures and habits in the case study authority areas.

The analysis of any medium-term impacts on professional learning was based on two types of evidence; the extent to which the programme had affected teachers' beliefs about the potential of collaborative professional learning, and the extent to which it had helped to develop sustained cross-school working and networking.

The extent to which the LT programme affected professional learning structures and networking between schools was difficult to ascertain because of the speed and depth of LA reforms that led to extensive restructuring of professional learning provision locally. However, changes in beliefs about, and dispositions towards, collaborative professional learning associated with the programme were more easily identifiable and there was some evidence of growing recognition of both benefits and costs.

Across all three cases, lead teachers were the most enthusiastic supporters of continuing to learn collaboratively, but their emphasis was on developing their professional networks rather than engaging in substantive school-to-school networking. As LAs withdrew or substantially reduced subject-specialist support, the LT programme had played an increasingly important part in maintaining their professional networks. The approach adopted in authority N, where groups of lead

⁵ Tests taken by 10 and 11-year olds before they leave primary school at the end of Key Stage 2.

and supported teachers attended cluster meetings, appeared particularly effective in developing networks that continued offering subject-specific support long after the programme had finished:

The project has ended but I would never think twice about contacting them and know whatever they had they would share. (Supported teacher, N)

Across all three cases, headteachers' commitment to lesson study appeared dependent on continued access to some form of collaborative structure that supported cross-school working, mainly to reduce the cost of working with other schools. Headteachers with more limited experiences of collaborative working were generally less convinced of its benefits and more insistent on continued LA involvement in maintaining cross-school collaborative structures:

You can collaborate but someone has to say this school can offer that sort of support. If you are not careful you could get within a family of schools and you could all be in the foothills and never reach the top of the mountain. (Headteacher, N)

5.2. A contextual analysis of the LT programme

A detailed contextual analysis was undertaken at three levels: LA (macro), school (meso), and within the lesson study cycle itself (micro). The aim was to understand how the ways in which the LT programme had been enacted in the three local education systems had affected the form lesson study had taken, and hence its impacts.

5.2.1. What were the key factors that affected LAs' development and implementation of the programme?

At the macro level, the first key factor in the reconstruction of the LT programme was how LA officers positioned it in their overall school improvement strategy. At the time of the research, central government policy was starting to change and affect the role of LAs. Although still responsible for "failing" schools, LAs were being encouraged to take a less direct role in school improvement and instead to broker enhanced school-to-school support; encourage leaders of high achieving schools to work across the authority and to "federate" with "failing" schools in a shared leadership and governance structure (Chapman & Muijs, 2013). The second key factor was the patina, consisting of school leaders' dispositions towards working collaboratively and the local structures that supported and coordinated improvement efforts, created by the history of school-to-school working in a local system. The third factor was the nature of the school improvement challenges faced within each LA, specifically the number of schools regarded as "failing" (missing national pupil attainment targets) or "coasting" (improving more slowly than other local schools).

Authority N had only recently started to adopt more collaborative forms of school-to-school improvement. Its main strategy was to develop cohorts of lead teachers, of which it already had over 150 before the LT programme was announced, to provide subject specialist support to schools selected by the LA. The LT programme provided an opportunity to expand this offer, mainly to "coasting" schools in the authority. The intention was to involve as many schools as possible in order to promote the benefits of collaborative learning and school-to-school support. Authority E had well-established collaborative approaches to school improvement based around networks of schools which had been involved in cross-school inquiry and improvement initiatives, including lesson study. Authority E's history enabled it to integrate the LT programme within these collaborative structures, through which schools were invited to join the programme. Authority B had a much higher level of "failing" schools than the other two authorities and constructed the programme as means of supplementing its existing improvement work with schools that were failing

or at risk of failing, rather than extending existing collaborative structures. During the period of the programme, lead teachers were used to replace LA officers and consultants as part of an external teams that worked intensively with such schools until their performance improved.

The ways in which the LT programme was reconstructed at this level affected the type of schools that took part, the basis on which they had been recruited and partnered with other schools, and even the structure of the lesson study process. “Lead” and “supported” schools were positioned differently in each LA. In Authority E, existing school networks committed their own resources to extend the period schools worked together on lesson study to nine months, with lead and supported teachers visiting each other’s schools. In authority N, the programme was restricted to a period of around three months, prior to pupil testing, with lead teachers working only in the supported school. In Authority B, the cycle normally ran over a mutually agreed number of sessions, often six, as part of more general support framework for schools “at risk”.

5.2.2. What were the key factors that informed headteachers’ decisions to engage with the Leading Teacher programme?

Each LA presented the programme differently to school leaders. Authorities N and E explicitly used it to develop more collaborative, school-led approaches to improvement, building on headteachers’ prior history of collaboration. In both these LAs, headteachers and lead teachers were concerned that it should not be a “quick fix”, booster programme targeted cynically at teachers in the year before national testing, but a genuine opportunity for professional learning. This concern was addressed in Authority E by using the existing collaborative structures of school networks to not only extend the lesson study cycles to nine months but also to apply them to the whole of Key Stage 2. Experienced headteachers in these networks tended to construct the programme in capacity-building terms, seeing clearly identifiable benefits which ranged from improved teacher motivation and reflection to being better able to offer feedback and professional learning in their own school:

A lot of my staff are more valuable out of the classroom than they are in because they’re working with other schools. (Headteacher, E)

Authority N’s lack of school-to-school collaborative structures or headteachers with extensive experience of collaborative working meant it did not have the capacity to reshape the programme substantively and so presented it as a pilot, hopeful that the benefits that accrued to individuals and schools would eventually lead to more sustained, school-funded collaboration. In Authority B, building capacity for collaboration was never a key objective and the programme was subsumed into their existing approach to “failing” schools. Headteacher engagement tended to be mandated due to their school’s “at risk” status and the programme became part of a short-term external support package focused on immediate impacts on the supported teacher:

There is nothing better than to have someone with you in the classroom in a kind of hands-on approach. (Headteacher, B)

5.2.3. What aspects of the lesson study cycle were seen as key in achieving changes in classroom practice?

The final (micro) level of analysis, the relationships between lead and supported teachers in classrooms during the lesson study process, was nested within the previous levels. Practitioners across all three authorities entered into a lesson study process that had been positioned differently in the local approach to school improvement. Although schools had become engaged for different reasons, the case studies revealed that practitioners could still exert considerable agency as to how they reconstructed lesson study in the classroom. The following analysis focuses on those aspects of

the relationships between “lead” and “supported” teachers, and the practices they supported, that were seen as key in facilitating practice experimentation and professional learning.

The LT programme was targeted across LAs to varying degrees which meant that differences in expertise, and effectiveness as measured by pupil attainment, between lead and supported teachers varied. At one end of the spectrum of the relationships between lead and supported teachers they were involved in a mutual process of joint practice development, at the other end the supported teacher was expected to expose far more of their practice and professional concerns than the lead. A consistent theme was that the potential for tension between participating teachers, because of perceived differences in status within their relationships, were ameliorated if **professional autonomy** was valued, particularly when deciding on the focus of the lesson study:

There were no preconceived ideas or anything. You sit down with your partner school and you discuss the areas you'd like to develop [...] It's not just a one-way thing, it's a sharing of ideas. (Lead teacher, Authority E)

It was this sense of professional autonomy and collaborative endeavour that appeared key to the acceptance of the programme in the classrooms of supported teachers and disposed them towards risk-taking and innovation. Autonomy was more likely to be expressed in relationships and practices based on creating opportunities for mutual learning, rather than deficit or one-sided expert coaching models:

The success was mainly due to the quality of the leading teacher. The leading teacher confirmed their good practice and the positive and successful relationship with the leading teacher was the best resource of all. (Leading Teacher Evaluation, 2010, Authority N)

The second cross-case theme was the importance of facilitating **professional dialogue**, linked to the reflecting on practice and its outcomes:

When do you ever sit down and talk to someone about your lesson you did that day? (Lead teacher, Authority E).

A lesson study cycle can potentially support professional dialogue around a number of issues, such as specific children's learning needs, meeting these needs, creating teaching materials, and evaluating interventions. The short periods of co-teaching, and the use of reference groups, which were present more frequently within the extended lesson study cycles in LA E were regarded as particularly effective in structuring dialogues that prompted reflection and the articulation of professional theories around the links between classroom practices and pupil learning:

You pick up so much from each other - the skill of focusing on target children and making sure they're progressing. (Lead teacher, Authority E).

In contrast, the truncated forms of lesson study that took place in the more pressurised context of “failing” schools in LA B resulted in lead teachers feeling they had to share as much of their expertise as possible, rather than engaging in dialogue, and mutual observation tended to lose out to extended modelling of practices. Lead teachers in authority B felt that such exchanges were less likely to result in sustained changes to classroom practice or new professional insights.

The final theme of lead teachers' **professional disposition** arose from analysis of their perceived role and how this compared with supported colleagues' views of the types of interactions they had

found supportive. Initially, the analysis revealed relatively little beyond broad a recognition of what one lead teacher called:

a conglomeration of roles [...] Partly you are going in as a troubleshooter, partly you are going in as a friend and colleague. (Lead teacher, Authority N).

Further analysis highlighted a mixture of attitudes, values and demeanours that participants believed constituted the dispositions required to establish productive learning relationships. Three dispositions were highlighted across the case studies:

- adopting a non-confrontational stance to professional learning characterised by questions such as “*have you ever tried this way of doing it?*” (Lead teacher, Authority N).
- being open and sharing failures: “*You need to be a risk taker and not afraid to have a go*” (Lead teacher, Authority B).
- motivating teachers through demonstrating enthusiasm for their subject and learning in general: “*He was so passionate about Mathematics [...] that you wanted to take that idea and run back to the classroom and start using it there and then*” (Supported teacher, N).

These dispositions were consistent with research on effective mentoring (Hobson et al., 2009), although a key missing element was challenging low expectations or poor practice. Lead teachers, especially those involved in truncated cycles, reported feeling uncomfortable about challenging certain practices and norms if there was insufficient time to address them in practice.

6. Discussion

The research set out to answer two sets of questions that addressed different aspects of the impact of involvement in lesson study on pupils, teachers and the contexts in which they worked. Did it make a difference to pupil attainment and did it enhance the capacity for profession learning in local systems?

In each LA and nationally, analysis of pupil test results indicated that LT schools demonstrated gains, significant in some subject areas, in pupil attainment and narrowed the performance gap with non-programme schools in most areas measured. These outcomes need to be seen in the context of the factors LAs used to identify schools for the LT programme. Alongside issues such as insufficient capacity to lead improvement, schools were targeted for inclusion following poor progress in attainment at Key Stage 2 or CVA scores. In such LAs, it would not be expected that programme schools would be able to close the gap completely with non-programme schools in such a short space of time. In Authority B, which had large numbers of schools with pupils making poor progress across Key Stage 2, the programme was used to supplement and expand its “traditional” work with low performing schools. Despite operating in some of the most pressurised contexts with truncated lesson study cycles, some of the largest initial gains in pupil achievement were achieved in this authority. As programme schools were working from a lower baseline these initial gains might be easier to achieve than in non-programme schools, and the relative levelling-off in programme schools’ progress was indicative that these initial gains were unsustainable. Even so, the continuing improvement in both results and programme schools relative performance to non-programme schools in authorities E and N suggest that involvement in lesson study was associated with gains in pupil attainment over time.

The extent to which the LT programme had any medium-term effects on a local system’s capacity to improve was difficult to ascertain due to national reforms to LAs, which disrupted existing professional learning structures and were likely to have overwhelmed any impacts of engagement in

lesson study. The research indicated that attempts to embed cross-school lesson study cycles benefitted from pre-existing collaborative structures and supportive system contexts. There was some evidence that involvement in lesson study could at least temporarily reconfigure practice architectures in schools, making them more supportive of this form of professional learning. Positive changes in leaders and practitioners' beliefs about, and dispositions towards, lesson study as a collaborative professional learning process were identified and revealed a growing recognition of both its benefits and costs.

The overarching findings raised further questions about the extent to which a lesson study process that had resulted in increased pupil attainment necessarily prompted deep professional learning. This question arose in part because of how in LA B accountability pressures had restricted the nature of the relationships between "lead" and "supported" teachers within a truncated lesson study cycle, but this had still contributed to significant improvements in pupil outcomes. The use of Clarke and Hollingsworth's (2002) model of professional learning to analyse the case studies had undermined any simplistic notions linking improvements in pupil outcomes to professional learning.

The multi-level contextual cross-case analysis provided insights into how the reconstruction of the LT programme in each authority had affected the way in which the lesson study was enacted, and hence the nature of professional learning it supported. Firstly, the LT programme was positioned differently within each LA's overall school improvement strategy. The strategies ranged from short-term interventions by LA based teams that targeted schools at risk of failure, to those based on mutual support brokered by the LA within networks of local schools. Positioning the programme within these strategies directly influenced the length of lesson cycles and patterns of classroom visits between participants, and indirectly shaped their interactions by surrounding them in contrasting professional discourses, from being a "quick fix" to deal with low pupil attainment to an opportunity to extend collaborative professional learning.

Setting lesson study in the context of a targeted school-based intervention in a high accountability system increased the pressures and constraints on both lead and supported teachers. When the outcomes being sought were heavily influenced by the requirement to hit pupil floor targets, and expressed solely in improved levels of attainment, there was an unwillingness to engage in innovation, risk-taking and the forms of dialogue required to prompt professional learning (Harrison, 2004). The cases illustrated how in some contexts the relationship between participants became one of expert coaching (Lord, Atkinson, & Mitchell, 2008), with a focus on raising attainment that limited the nature of professional learning. Lead teachers involved in a lesson study cycle with a "failing" school felt under pressure to share their existing expertise and struggled to create the types of professional relationships where they could prompt supported teachers to reflect on misconceptions and poor practice.

The macro and meso level analysis illustrated the extent to which the lesson study process was prefigured by contextual factors in each LA. The micro level analysis revealed that there was still considerable leeway for practitioners to express their professional agency as to how lesson study was enacted in classrooms. As a site of professional learning, lesson study appeared relatively resilient to the influence of certain external factors if the relationships between participants were marked by professional autonomy, opportunities for focussed dialogue, and a shared disposition towards openness and non-confrontation.

7. Conclusions

International research has provided some evidence that lesson study is a powerful professional learning process (Xu & Pedder, 2015) that can be adapted to meet the needs of a range of teachers

and fulfil a number of professional learning outcomes (Saito, 2012). The versatility and adaptability of lesson study means that the question of what are its “core” elements still remains of theoretical and practical interest.

Lesson study as a global phenomenon has been taken up across numerous education systems, resulting in White and Lim (2008, p. 916) believing that it has become “*an umbrella term for a variety of adaptations*”. As lesson study programmes attract government attention, and funding, the issues around what constitutes a lesson study have become re-cast. The adaptation of lesson study would now appear to be driven by wider systemic and contextual factors than simply its re-interpretation by enthusiastic researchers and early adopting practitioners. As this research and others have indicated (Akiba & Wilkinson, 2016) when lesson study is positioned within governmental improvement strategies it can result in different forms of modification. Direct modifications to the nature of the process can arise due to structural factors within systems and organisations, such as the availability of funding or existing power relationships between groups. Indirect modifications, as to what being involved in a lesson study means to a practitioner and the status it is given, occur due to socio-cultural factors such as the discourses surrounding a governmental improvement strategy or traditional practices around professional learning.

A growing trend towards the deliberate adaptation, or modification, of lesson study, due to its increased integration within government policy agendas gives a degree of urgency to our critique of current research. The benefits of adopting a more critical contextual analysis of a lesson study initiative would include:

- avoiding a narrow debate around whether it is a lesson study by focussing on core practices by giving consideration to the wider characteristic effects being sought at different levels of a system by those who are promoting it;
- greater recognition of how the impact of lesson study arises from the emergence and interaction of various effects or outcomes, both within and across the levels of a system;
- moving research towards a consideration of the contextual factors, at different levels within a system, which support the development and sustainability of lesson study.

Calls for further research into how lesson study helps teachers learn (Xu & Pedder, 2015) will need to be set within a critical understanding of how as a social learning process it is inextricably linked to the contexts in which it is enacted. The modifications of lesson study being promoted should be examined as to whether they will provide teachers with the opportunity to develop the necessary professional agency, individual and collective, to support each others learning and reconfigure the practices that surround them in order improve teaching and learning.

8974 words (excluding references tables and figures)

References

- Ainscow, M. Muijs, D., & West, M. (2006). Collaboration as a strategy for improving schools in challenging circumstances. *Improving Schools*, 9(3), 192-202.
- Akiba, M., & Wilkinson, B. (2016). Adopting an international innovation for teacher professional development: State and district approaches to Lesson Study in Florida. *Journal of Teacher Education*, 67(1), 74 –93.
- Author(s) et al. (2009). Details removed for peer review.

- Author(s) (2007). Details removed for peer review.
- Ball, S.J., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary schools*. London: Routledge, Taylor and Francis Group.
- Bickford, H.L. (2010). One school's cautious use of classroom walk-throughs. *The School Administrator*, 10(67), 20-21.
- Bowe, R., Ball, S. J., & Gold, A. (1992). *Reforming education and changing schools: Case studies in policy sociology*. London: Routledge.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Brooks, G. (1998). Trends in standards of literacy in the United Kingdom 1948- 1996. *Topic*, 19(1), 1-10.
- Chapman, C., & Muijs, D. (2013). Collaborative school turnaround: A study of the impact of school federations on student outcomes. *Leadership and Policy in Schools*, 12(3), 200-226.
- Cheng, E.C.K., & Ling, L.M. (2013). *The approach of learning study: Its origin and implications. OECD education working papers No. 94*. Paris: OECD Publishing.
- Cheung, W.M., & Wong, W.Y., (2014). Does lesson study work? A systematic review on the effects of lesson study and learning study on teachers and students. *International Journal for Lesson and Learning Studies*, 3(2), 137-149.
- Chia, R., & Holt, R. (2006). Strategy as practical coping: A Heideggerian perspective. *Organizational Studies*, 27(5), 635-655.
- Chokshi, S. & Fernandez, C. (2004). Challenges to importing Japanese lesson study: Concerns, misconceptions, and nuances. *Phi Delta Kappan*, 85(7), 520-5.
- City, E., Elmore, R., Fiarman, S., & Teitel, L. (2009). *Instructional rounds in education: A network approach to improving teaching and learning*. Boston, MA: Harvard Education Press.
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18(8), 947-967.
- Cordingley, P., Bell, M., Rundell, B., & Evans, D. (2003). *The impact of collaborative CPD on classroom teaching and learning*. In Research Evidence in Education Library. London: EPPI Centre, Social Science Research Unit, Institute of Education, University of London.
- Deloitte and Touche (2000). *Evaluation of European school partnerships under Comenius action 1 & Lingua action E*. London: Deloitte and Touche.
- Department for Education (2010). *The importance of teaching: The schools white paper 2010*. London: DfE.
- Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199.
- Dudley, P. (2012). Lesson study in England: From school networks to national policy. *International Journal for Lesson and Learning Studies*, 1(1), 85-100.
- Dudley, P. (2013). Teacher learning in lesson study: What interaction-level discourse analysis revealed about how teachers utilised imagination, tacit knowledge of teaching and fresh evidence of pupils learning, to develop practice knowledge and so enhance their pupils' learning. *Teaching and Teacher Education*, 34,107-121.
- Dudley, P. (2014). *Lesson study: Professional learning for our time, a resume*. Retrieved from <http://lessonstudy.co.uk>.
- Earl, L., Watson, N., Levin, B., Leithwood, K., Fullan, M., & Torrance, N., (2003). *Watching and learning 3: Final report of the external evaluation of England's national literacy and numeracy strategies*. London: DfES.
- Ermeling, B.A., & Graff-Ermeling, G. (2014). Learning to learn from teaching: A first-hand account of lesson study in Japan. *International Journal for Lesson and Learning Studies*, 3(2), 170-191.
- Fullan, M. (2009). Large scale reform comes of age. *Journal of Educational Change*, 10, 101-113.

- Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Cochran-Smith, M., McDonald, M., & Zeichner, K. (2005). How teachers learn and develop. In L. Darling-Hammond (Ed.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 358-389). San Francisco: John Wiley & Sons.
- Hargreaves, A., & Shirley D. (2010). *The fourth way*. Thousand Oaks, CA: Corwin Press.
- Hargreaves, D.H. (2010) *Creating a self-improving system*. Nottingham: National College for School Leadership.
- Harris, A. (2002). *School improvement. What's in it for schools?* London: Routledge.
- Harris, A., & Chrispeels, J. (Eds.), (2008) *International perspectives on school improvement*. London: Routledge Falmer.
- Harrison, J.K. (2004). Encouraging professional autonomy: Reflective practice and the beginning teacher. *Education 3-13*, 32(3), 10–18.
- Hopkins, D. (2003). *School improvement for real*. London: Routledge.
- Hopkins, D. (2016). School improvement and system reform. In C. Chapman, D. Muijs, D. Reynolds, P. Sammons, & C. Teddlie (Eds.), *The Routledge international handbook of educational effectiveness and improvement* (pp. 124-148). London: Routledge.
- Hobson, A.J., Ashby, P., Malderez, A., & Tomlinson, P.D. (2009). Mentoring beginning teachers: What we know and what we don't. *Teaching and Teacher Education*, 25(1), 207-216.
- Jolley, G. (2014). Evaluating complex community-based health promotion: Addressing the Challenges. *Evaluation and Program Planning*, 45, 71-81.
- Kemmis S., & Grootenboer, P. (2008). Situating praxis in practice: Practice architectures and the cultural social and material conditions for practice. In S. Kemmis & T.J. Smith (Eds.), *Enabling praxis: Challenges for education* (pp. 37-62). Rotterdam: Sense.
- Kemmis, S., Wilkinson J. Edwards-Groves, C., Hardy, I., Grootenboer, P., & Bristol, L. (2014). *Changing practices, changing education*. Singapore: Springer.
- Keys, W., Harris, S., & Fernandes, C. (1996). *Third international mathematics and science study: First national report part 1. Achievement in mathematics at age 13 in England*. Slough: NFER.
- Koelner, K., & Jacobs, J. (2015). Distinguishing models of professional development: The case of an adaptive model's impact on teachers' knowledge, instruction, and student achievement. *Journal of Teacher Education*, 66(1), 51-67.
- Lewis, C., Perry, R., & Hurd, J. (2004). A deeper look at lesson study. *Educational Leadership*, 61(5), 18–23.
- Lieberman, J. (2009). Reinventing teacher professional norms and identities: The role of lesson study and learning communities. *Professional Development in Education*, 35(1), 83-99.
- Ling, T. (2012). Evaluating complex and unfolding interventions in real time. *Evaluation*, 18(1), 79-91.
- Lord, P., Atkinson, M., & Mitchell, H. (2008). *Mentoring and coaching for professionals: A study of the research evidence*. London: The Training and Development Agency for Schools.
- Mason, J. (2002). *Qualitative researching* (2nd ed.). London: Sage.
- Matoba, M., Shibata, Y., Reza, M., & Arani, S. (2007). School-university partnerships: A new recipe for creating professional knowledge in school. *Educational Research, Policy and Practice*, 6 (1), 55-65.
- Maughan, S. Teeman, D., & Wilson, R. (2012). *What leads to positive change in teaching practice?* Slough: National Foundation for Educational Research.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Mourshed, M. Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better*. London: McKinsey & Company.
- Mowles, C. (2014). Complex, but not quite complex enough: The turn to the complexity sciences in

- evaluation scholarship. *Evaluation*, 20(2), 160-175.
- Murata, A. (2011). Introduction: Conceptual overview of lesson study. In L.C. Hart, A. Alson, & A. Murata, (Eds.), *Lesson study research and practice in mathematics education* (pp. 1-12). Heidelberg, London, New York: Springer.
- Ofsted (2003). *Excellence in cities and education action zones, Management and impact*. London: Ofsted, HMI 1399.
- Ofsted (2009). *The annual report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2008/09*. London: The Stationery Office.
- Ofsted (2010). *The annual report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2009/10*. London: The Stationery Office.
- Opfer, V., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, 81(3), 376-407.
- Parlett, M. & Hamilton, D. (1987). Evaluation as illumination: A new approach to the study of innovatory programs. In R. Murphy, & H. Torrance (Eds.), *Evaluating education: Issues and methods* (pp. 57-73). London: Paul Chapman.
- Perry, R.R., & Lewis, C.C. (2009). What is successful adaptation of lesson study in the US? *Journal of Educational Change*, 10(4), 365-91.
- Saito, E. (2012). Strategies to promote lesson study in developing countries. *International Journal of Educational Management*, 26(6), 565-576.
- Sammons, P., Mujtaba, T., Earl, L., & Gu, Q. (2007). Participation in network learning community programmes and standards of pupil achievement: Does it make a difference? *School Leadership and Management*, 27(3), 213-238.
- Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Stewart, R.A., & Brendefur, J.L. (2005). Fusing lesson study and authentic achievement: A model for teacher collaboration. *Phi Delta Kappan*, 86(9), 681-687.
- Stigler, J., & Hiebert, J. (1999). *The teaching gap*. New York: Free Press.
- Takahashi, A., Watanabe, T., & Yoshida, M. (2006) Developing good mathematics teaching practice through lesson study: a U.S. perspective. In Center for Research on International Cooperation in Educational Development (CRICED), (Ed.), *Progress report of the APEC project: A collaborative study on innovations for teaching and learning mathematics in different cultures among the APEC member economies* (pp. 129-136). Tsukuba, Japan: CRICED.
- TIMSS (1998). *Third international trends in mathematics and science study*. Wellington, NZ: Comparative Education Research Unit.
- Watanabe, T. (2002). Learning from Japanese lesson study. *Education Leadership*, 59(6), 36-39.
- Watterman, S. (2011). *A study of lesson study's impact on student achievement*. Morgan Hill, CA: Silicon Valley Mathematics Initiative.
- White, A.L., & Lim, C.S (2008). Lesson study in Asia Pacific classrooms: Local responses to a global movement. *ZDM International Journal of Mathematics Education*, 40, 915-925.
- Xu, H., & Pedder, D. (2015). Lesson study: an international review of the research. In P. Dudley (Ed.), *Lesson study: Professional learning for our time* (pp. 24-47). Abingdon: Routledge.
- Yoshida, M. (1999). Lesson study (Jugyokenkyu) in elementary school mathematics in Japan: a case study. Paper presented at *On the threshold of the 21st Century: Challenges and opportunities, the annual meeting of the American Educational Research Association*, Montreal, 19-23 April 1999.
- Yoshida, M., & Jackson, W.C. (2011). Response to part V: Ideas for developing mathematical pedagogical content knowledge through lesson study. In L.C. Hart, A. Alson, & A. Murata (Eds.), *Lesson study research and practice in mathematics education* (pp. 279-288). Heidelberg, London, New York: Springer.

Fig. 1 The Lesson Study Cycle

Suggested timeline for the lesson study cycle – Year 6 Leading Teacher programme 2009 -10

Suggested breakdown of time for each leading teacher:-

- Initial joint meeting with headteachers and partner teachers ½ day
- Support to be provided for 4 schools – 4 ½ day visits in each school - 8 days
- Flexible time to be used by LT to work in own school/write up www case Study/prepare for school visits/training etc. 1½ days

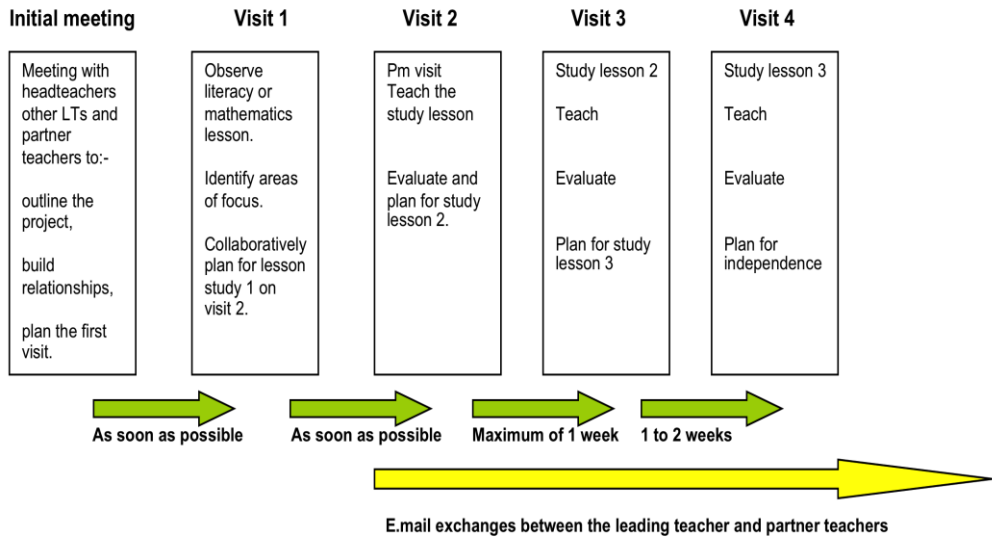


Fig. 2 Outline logic model for the evaluation of the Primary Leading Teachers programme

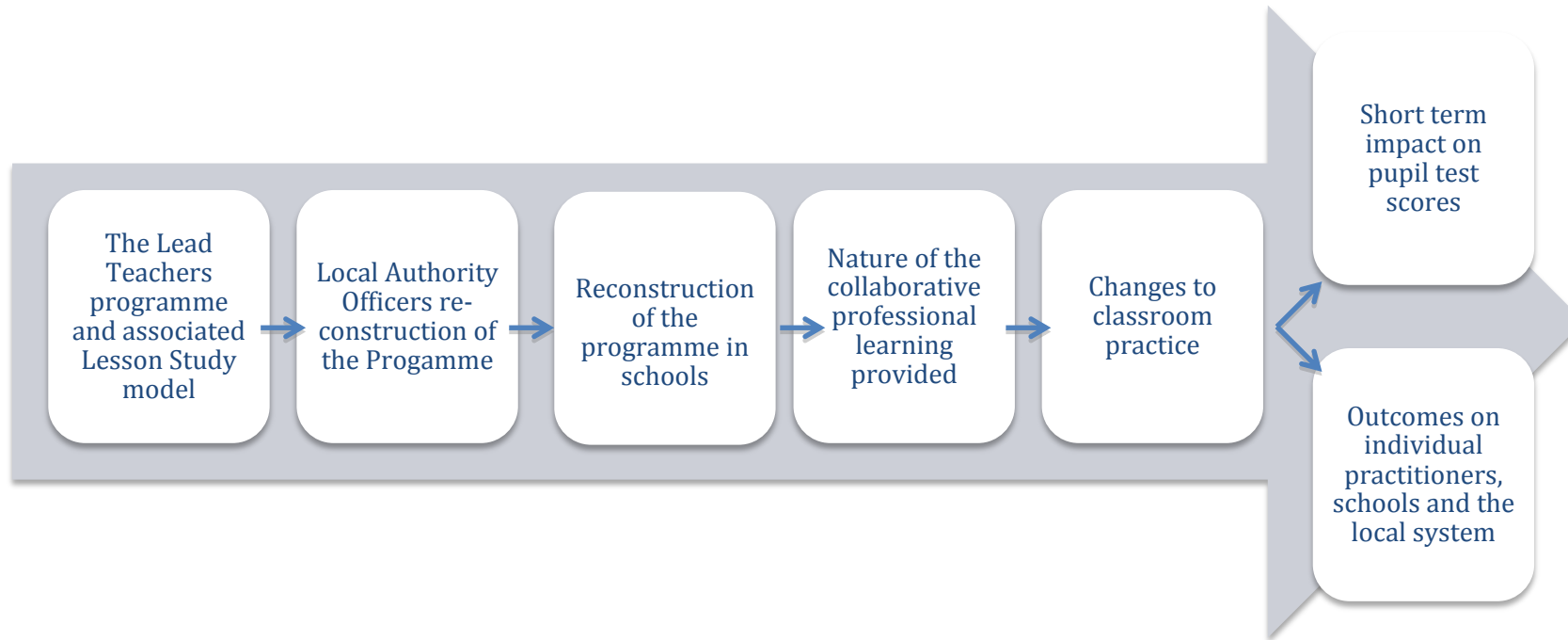


Fig 3. Differences between results in English at Level 5 between 2009 and 2010 comparing LT schools (x) and non-LT schools (y)

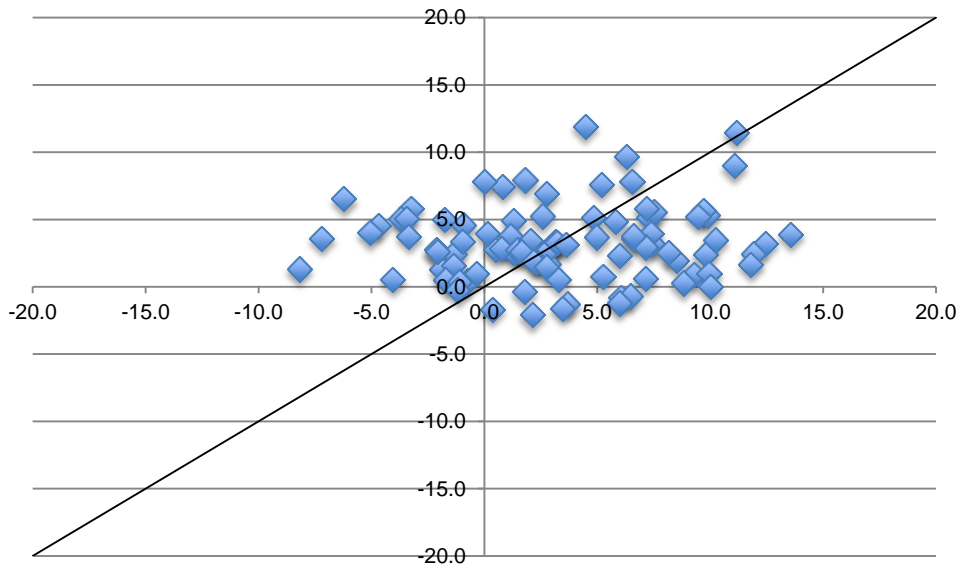
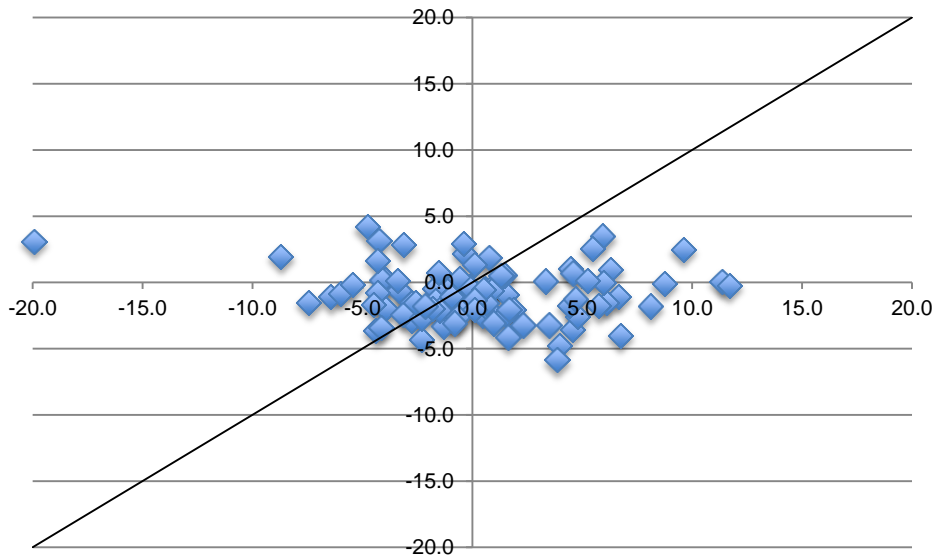


Fig 4. Differences between results in Mathematics at Level 5 between 2009 and 2010 comparing LT schools (x) and non-LT schools (y)



Appendix A. Data analysis tables and figures

Table 1

Weighted Means of the percentage of pupils achieving Level 4 or above and Level 5 at Key Stage 2 in 2009 and 2010 in LT and non-LT schools in LAs involved in the programme (n=91)

	<i>Level 4</i>	<i>Level 5</i>
LT English 2010	78.66	28.71
Non-LT English 2010	81.90	33.42
LT English 2009	78.03	26.10
Non-LT English 2009	82.10	30.47
LT Mathematics 2010	79.18	31.32
Non-LT Mathematics 2010	81.59	34.96
LT Mathematics 2009	76.32	31.15
Non-LT Mathematics 2009	80.54	35.80
LT English & Mathematics 2010	71.68	19.42
Non-LT English & Mathematics 2010	75.18	23.10
LT English & Mathematics 2009	69.06	17.59
Non-LT English & Mathematics 2009	74.24	21.21

Table 2

Weighted means of the percentage of pupils achieving two or more levels of progress from Key Stage 1 in 2009 and 2010 in LT and non-LT schools in LAs involved in the programme (n=91)

	<i>English</i>	<i>Mathematics</i>
LT 2010	80.44	78.11
Non-LT 2010	82.77	81.35
LT 2009	82.82	81.80
Non-LT 2009	84.63	83.29

Table 3

Differences in percentage improvement in pupil performance between LT schools and non-LT schools at Key Stage 2 Level 4 or above 2009-10 for the three case study LAs

	<i>Difference between LT and non-LT schools</i>		
	<i>LA N</i>	<i>LA E</i>	<i>LA B</i>
English & Mathematics	4.0	0.5	7.1
English	2.6	0.1	3.5
Mathematics	4.4	1.1	9.1

Fig 5. Differences between results in English at Level 4 between 2009 and 2010 comparing LT schools (x axis) and non-LT schools (y axis)

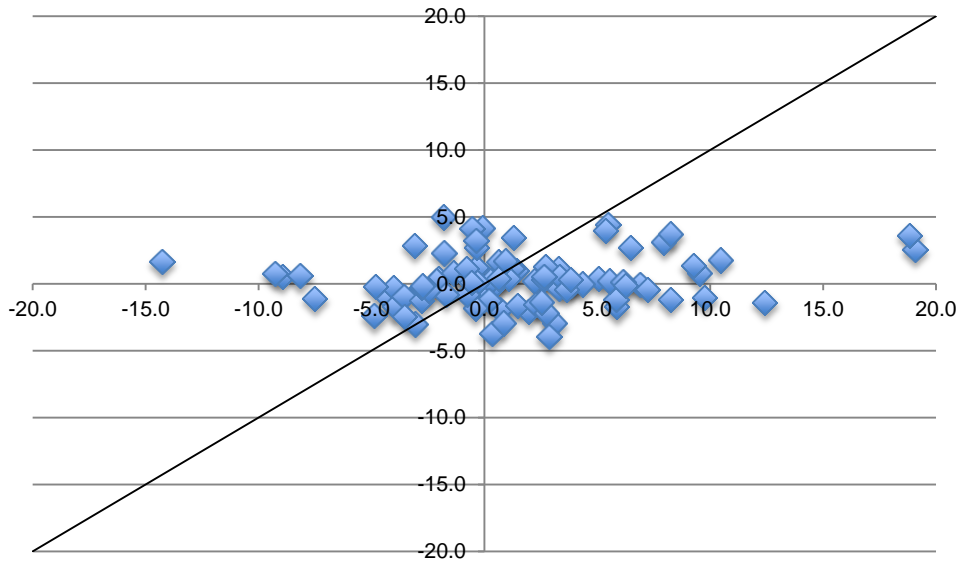


Fig 6. Differences between results in Mathematics at Level 4 between 2009 and 2010 comparing LT schools (x) and non-LT schools (y)

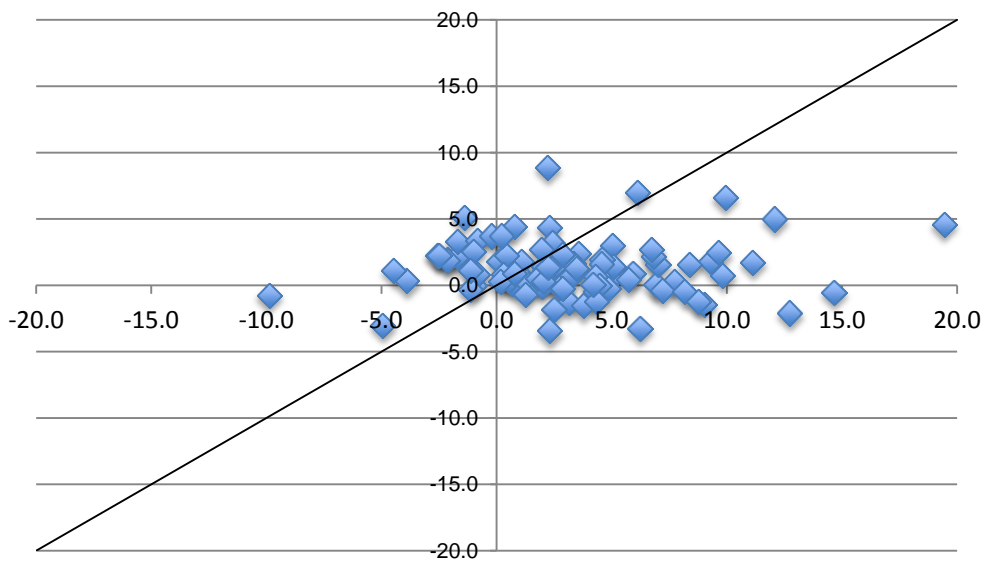


Fig 7. Differences between results in English and Mathematics combined at Level 4 between 2009 and 2010 comparing LT schools (x) and non-LT schools (y)

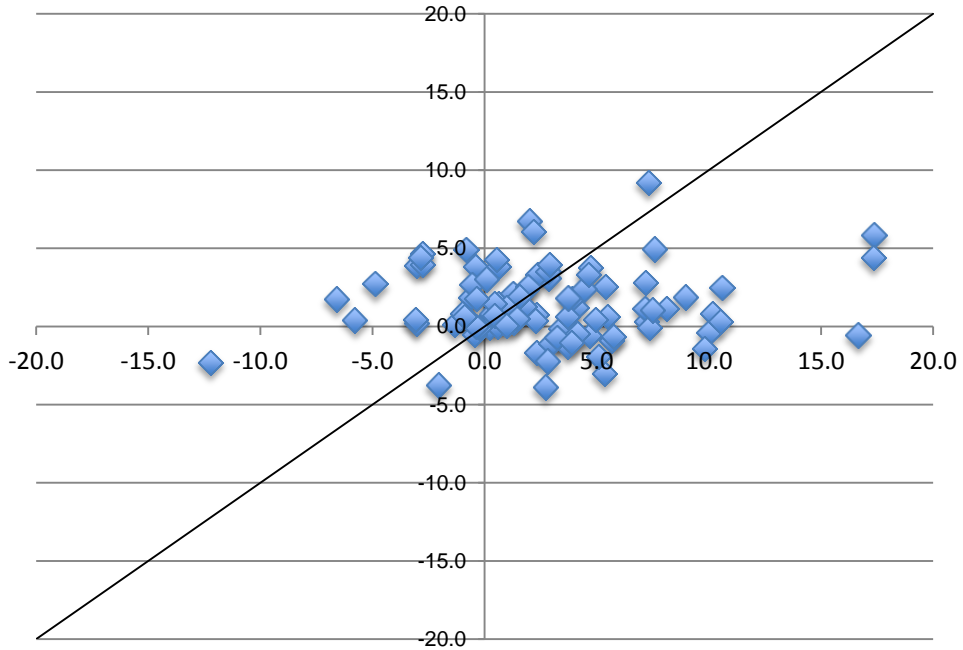


Fig 8. Differences between results in English and Mathematics combined at Level 5 between 2009 and 2010 comparing LT schools (x) and non-LT schools (y)

