

An exploration of practitioner-researcher collaboration on randomised controlled trials of complex interventions

Mary Barbara Sawtell

Thesis presented for the award of a Doctorate in Philosophy

UCL-Institute of Education

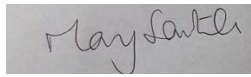
University College London

Submitted to Doctoral School 26/02/18

DECLARATION AND ACKNOWLEDGEMENTS PAGE

I, Mary Sawtell, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

A rectangular box containing a handwritten signature in cursive script that reads "Mary Sawtell".

Date: 27 February 2018

I would like to acknowledge all the support and thoughtful guidance of my two PhD supervisors, Professor Sandy Oliver and Dr Ginny Brunton. I also want to thank my colleague Meg Wiggins and of course my family for being there for me, in so many ways, during the writing of this thesis.

SUPPORTING STATEMENTS

Candidate Publication 1

Sawtell, M and Jones, C (2002). Time to listen: an account of the role of support health visitors. *Community Practitioner*, 75 (2), 461-463.

This paper is about the supportive Listening Programme, a public health intervention, for new mothers that aimed to reduce childhood injury and improve maternal psychological wellbeing intervention. It was evaluated in the Social Support and Family Health Study. During this study Mary was one of the Support Health Visitors (SHV) delivering the programme. The study ran from 1999-2002. Mary also contributed mixed method data via her participation in the process evaluation. Mary led on the writing of this paper; the second author (Carol Jones) was also an SHV on the study.

Candidate Publication 2

Sawtell, M, Jamieson L, Wiggins M, Smith F, Ingold A., Hargreaves K, Christie, D (2015). Implementing a structured education program for children with diabetes: lessons learnt from an integrated process evaluation. *BMJ open diabetes research & care*, 3 (1), e000065.

This paper describes the methods and findings from the process evaluation of the CASCADE study which was a cluster randomised controlled trial of a structured education programme for children with diabetes. Mary was the trial manager for the study which ran from 2008 – 2013. Meg Wiggins was a senior research officer working on the study. Mary managed both the trial and the integral process evaluation. This management role included: trial set up processes; designing mixed methods data collection tools; analysing process evaluation data; writing the process evaluation section of the final report for the funder. Mary led on the writing of the paper which included: writing the first draft; co-ordinating input from the other authors; producing a publishable version including responding to journal reviewer comments.

Candidate Publication 3

Jerrim J, Macmillan L, Micklewright J, **Sawtell M**, Wiggins M (2016). *Chess in Schools Evaluation report and executive summary*. London: University College London Institute of Education.

This publication is the peer reviewed final report produced for the evaluation of The Chess in Primary Schools programme. The study ran from 2013-2015. John Jerrim was the Principal

Investigator and provided overall management of the project. Mary's role was joint-lead on the process evaluation. This role included: contributing to process evaluation research question formation and protocol development; designing data collection tools and leading on data collection; contributing to process evaluation data analysis. She led on the writing of the process evaluation section of this publication.

Candidate Publication 4

Sawtell M, Sweeney L, Wiggins M, Salisbury C, Eldridge S, Greenberg L, Hunter R, Kaur I, Hatherall B, Morris J, Reading S, Renton A, Adekoya R, Green B, Harvey B, Latham S, Patel K, Vanlessen L, Harden A. Evaluation of community-level interventions to increase early initiation of antenatal care in pregnancy: Protocol for The Community REACH study, a cluster randomised controlled trial with integrated process and economic evaluations. Accepted for publication *BMC trials* 06.02.18.

This paper is the protocol paper of The UK Community REACH trial. This trial aims to assess the effectiveness of engaging communities in the co-production and delivery of interventions that address inequality in access to antenatal care. The study design is a matched cluster RCT with integrated process and economic evaluations. It started in 2015 and is due to finish in 2019. Professor Angela Harden is the Chief Investigator on the study. Mary's role on Community REACH is trial manager which involves planning, co-ordinating the project. Mary led on the writing of the protocol paper; she wrote the first draft and co-ordinated the contributions of all the other authors to produce a final draft. The paper has been accepted for publication.

As co-authors on these publications, we confirm that this information is correct and represents an original contribution to these research projects.

Carol Jones	12.2.18
Meg Wiggins	01.2.18
John Jerrim	22.1.18
Angela Harden	22.1.18

ABSTRACT

Background. The past two decades have seen great interest in the development and evaluation of complex social interventions. The randomised controlled trial (RCT) is the normative research design for these evaluations and academic-practitioner collaboration in the conduct of studies is increasingly common to maximise rigour and relevance. However, gaps remain in what is known about how collaboration can be most effective in the co-production of knowledge. Practical examples of academic-practitioner collaboration can address these knowledge gaps.

Aim: To demonstrate the development of academic-practitioner collaboration in the conduct of RCTs of complex health and education interventions through practical examples spanning two decades.

Methods: Insider research drew on: four of my publications; experience of working on RCTs; and wider experience, gained over 20 years, working as a health visitor.

Findings and conclusions: A general trend, in studies, across time is shown of: more relevant practitioners actively involved; in increasingly varied and influential study roles; with greater capacity to contribute to the research process. Improved management of the boundary between intervention and evaluation is also demonstrated. These trends have increased the potential for a more equal and effective blend of academic and practitioner knowledge and as such the co-production of more useful research. Key practitioner voices have been missing from decision-making processes in RCTs, however, which is likely to have had a negative impact on the utility of the findings. Creative approaches to collaboration, utilising skills in interpersonal relations, awareness of context and spanning of boundaries can bring these harder to reach voices into the research process. These are skills central to health visiting practice. Although health visitors are relatively new to RCTs they are well positioned to be part of the process of conducting the rigorous and relevant RCTs that are important in the development of services, including health visiting.

IMPACT STATEMENT

The randomised controlled trial (RCT) is the standard research design for evaluations of effectiveness of complex interventions, including those that are UK policy relevant. Academic-practitioner collaboration in the conduct of such studies is increasingly considered best practice in order to maximise both scientific value and applicability in practice. There is an acknowledged need for in-depth understanding and published practical examples of how to achieve such collaboration effectively. This thesis makes a significant and highly relevant contribution by providing accounts, reflections, observations and analyses of the practical reality of striving for meaningful academic – practitioner collaboration in four RCTs of complex interventions.

This thesis has a foundation in my learning during a career that started with my entry as a practitioner (health visitor) to an academic research setting and continued over several decades as I worked as a researcher, in an academic team specialising in RCTs of complex interventions, while still in practice as a health visitor. Being embedded in the worlds of academia and health care practice, at the same time, for such a sustained period, is unusual. This is even more the case for an allied health care practitioner working on large scale evaluations with an RCT design. This gives the findings in this thesis particular originality.

In order to maximise the impact of this thesis, I will continue to build on the learning demonstrated on how to co-produce knowledge between academic researchers and all practitioner groups, and particularly allied health practitioners. Specifically, this will include building on findings on how to support practitioners to: be site principal investigators, co-author articles, collaborate on small discrete tasks within large studies, recognise the relevance of skills used in clinical practice to the research environment. I will encourage, through my own research and that of others, meaningful collaboration at all stages of the research process, from identification of interventions to be evaluated through to dissemination activities, in order that research processes are truly blended academic/practitioner enterprises that include individuals with the most relevant knowledge. I will disseminate my work through a range of practitioner and academic focused channels including: journal articles, conference presentations, face to face networking and social media. For example, I will submit abstracts to conferences and articles to journals for which target audiences are nurses and health visitors and/or academic trialists and methodologists. Furthermore, I will use my role as fellow and

research champion for the Institute of Health Visiting to facilitate research involvement by health visitors.

As such I hope that my work will contribute to significant progress in co-production of research between practitioners and academics in order to support the generation of knowledge that is of maximum public benefit.

ABBREVIATIONS LIST

CGS – Community groups support

CI – Chief Investigator

Co-I – Co-Investigator

CLARHCs – Collaboration for Leadership in Applied Health Research and Care

EEF – Education Endowment Foundation

HTA – Health technology assessment

IHV – Institute of health visiting

MRC – Medical research council

NIHR – National Institute for Health Research

PPI – Patient and public Involvement

PSDN – Paediatric specialist diabetes nurse

RCT – Randomised controlled trial

SCPHN (HV) - Specialist Community Public Health Nurse (health visitor)

SHV – Support health visitor

Site PI – Site principal investigator

SSFH – Social Family Health

TMG – Trial management group

TSC – Trial steering committee

CONTENTS

ABSTRACT	5
IMPACT STATEMENT.....	6
ABBREVIATIONS.....	7
CONTENTS	8
LIST OF TABLES AND FIGURES.....	9
CHAPTER 1. Overview of this thesis.....	10
1.1 Aims and rationale.....	10
1.2 The health visiting context of this thesis.....	11
1.3 Methodology.....	13
1.4 Candidate publications	15
1.5 About this integrative summary	17
CHAPTER 2. Review of relevant literature	19
2.1 Introduction	19
2.2 Randomised controlled trials.....	19
2.3 Collaboration between practitioners and academics	21
2.4 The need for more research	26
CHAPTER 3. Trends in practitioner involvement	27
3.1 Introduction	27
3.2 Background	29
3.3 The four studies	29
3.4 Identifying the research question, developing the intervention and evaluation protocol.....	33
3.5 Conducting the evaluation.....	33
3.6 The learning from this view across the four studies.....	37
3.7 Conclusion.....	40
CHAPTER 4. Collaborating with practitioners to design core RCT components ...	42
4.1 Introduction	42
4.2 Collaborating on trial outcome decisions	44
4.3 How did we collaborate on activities 1 and 2?	45
4.4 Collaborating on the production of a logic model	50
4.5 Conclusion.....	54
CHAPTER 5. Discussion and conclusions	55
5.1 Introduction	55
5.2 Factors facilitating academic-practitioner collaboration.....	55
5.3 Implications of the findings from this thesis for health visiting	60
5.4 Conclusions	62

REFERENCES.....	67
APPENDIX 1. Candidate publication 1.....	76
APPENDIX 2. Candidate publication 2.....	80
APPENDIX 3. Candidate publication 3.....	97
APPENDIX 4. Candidate publication 4.....	99
APPENDIX 5. Candidate bibliography.....	120

LIST OF TABLES AND FIGURES

Tables

Table 1: Summary of key structural, organisational and research design features of the studies.

Table 2: Practitioner involvement in key evaluation phases and study roles

Table 3: Summary of practitioner numbers, types and roles

Table 4: Missing practitioner voices on the studies.

Table 5: Examples of impact of collaboration with consultant midwives and data analysts

Table 6: Example of impact of collaboration with chess tutors.

Figures

Figure 1: The relative separation and connection of intervention and evaluation teams

Figure 2: Sample text from Delphi process email one

Figure 3: Key characteristics of academic/practitioner influence at different levels of decision making over time.

CHAPTER 1: OVERVIEW OF THE THESIS

1.1 AIMS AND RATIONALE

The randomised controlled trial (RCT) is the reference standard for assessing the impact of interventions (Dechartes et al., 2017). In these studies, a population is randomly divided into two or more groups, with each group given a different intervention (including 'treatment as usual') (Connolly, 2017). The results within each group are compared at the end of the study. The process of randomisation is intended to minimise the influence of external factors on the results of the study. As such, it is argued, that any difference in outcome can be explained primarily by 'treatment' thereby demonstrating effectiveness (Roberts and Torgerson, 1998).

The importance of basing practice in public services on current best evidence was the focus of the evidence-based practice (EBP) movement which, in the UK, started in the early 1990s (Robinson, 2017). This movement had its roots in medicine and the RCT, which already had a long history in medicine, was at its centre in terms of preferred study design. Recently the RCT has been increasingly used for the evaluation of interventions designed to address social, problems (Oakley, 1998). These interventions are often described as 'complex interventions'. A complex intervention is defined as an intervention made up of various interconnecting parts that act both independently and inter-dependently (Datta and Pettigrew, 2013). Complex interventions are often used in public health, clinical health and education settings and a considerable literature has built up on the use of the RCT in their evaluation (see for example Datta and Petticrew, 2013 and Craig et al. 2008).

The privileged status afforded to the RCT as the key source of evidence, over and above other research designs, has not been without criticism. Critiques target both the assumption that RCT evidence is entirely without bias and the assumption that a treatment decision made by a practitioner, based on their experiential knowledge, is inferior to one based on evidence from an RCT (Hammersley, 2005). Despite these criticisms, the RCT continues to be the most valued research design for informing policy and practice in public services. There has however, been recognition of the need for a move from an emphasis on efficacy studies, to trials which aim to produce more generalisable, and hence potentially more useful, findings (Bauer et al 2015). Academics and practitioners conducting research together in order to combine the rigour of

science with the relevance of practice, as expressed by practitioners themselves, is considered a key way that research, including RCTs, can be both more useful and more used (Heaton et al., 2015).

This thesis is underpinned by theories of collaboration and co-production. Collaboration in the context of this thesis is best understood as academic researchers working with practitioners to produce something (Martin, 2010). Co-production was originally defined by Ostrom in the 1970s, *'as a process through which inputs used to produce a good or service is contributed to by individuals who are not "in" the same organisation'* (Ostrom 1996 p.1073).

It is widely recognised that practitioner- academic collaboration is challenging to achieve and there is much that is not known about its meaning and practice (Hewison et al., 2010). My aim in this thesis is to demonstrate how my experience of working on RCTs of complex interventions over the last 20 years is situated within, and contributes original insights to, wider debates about collaboration and co-production between academics and practitioners on these types of studies. I do this with reference to: my four candidate publications which span this 20-year period; the studies from which they arose; and my wider experience as both an academic and as a practitioner.

1.2 THE HEALTH VISITING CONTEXT OF THIS THESIS

I have extensive 'lived' experience both as a university researcher and as a practitioner. I have worked as the former for almost 20 years and the latter, as a NHS Specialist Community Public Health Nurse (HV) (from here on 'health visitor') over the last 30 years. I am therefore well-positioned to reflect on the past, current and future involvement of practitioners, and specifically health visitors, in the process of conducting RCTs of complex social interventions.

Health visitors provide a professional public health service for individuals, families, groups and communities; enhancing health and reducing health inequalities through a proactive, universal service for all children 0-5 years and for vulnerable populations, according to need (iHV, 2012). Specifically, their role includes: preventing and detecting development problems in early childhood, supporting vulnerable families, improving breastfeeding and safeguarding children (Cowley et al. 2013). A thoughtful questioning approach to one's own practice, known as reflexivity, has been identified as an important practice skill for practitioners working with

children and families, as a means of facilitating ethical working in the uncertain contexts that are often encountered (Cruz et al. 2007).

My experience as an academic researcher has primarily been on RCTs of complex interventions, with integrated process evaluations (see the four candidate publications and also Wiggins, Sawtell, Jerrim, (2017)). I have also worked on a number of other large studies with quasi-experimental designs (for example Wiggins, Bonell, Sawtell et al., 2009) as well as mixed methods studies with non-experimental designs (Sawtell et al., 2009; Sawtell et al., 2010; and Oakley, Wiggins, Strange, Sawtell, & Austerberry, 2011). The RCTs I have worked on have been in health care and education settings and have been intended for informing UK policy. My various roles on these studies have included: practitioner (health visitor) delivering a health visiting intervention being evaluated; practitioner as process evaluation participant; research officer conducting the research; trial manager overseeing the research.

There is a long-held consensus that services focussed on the early years of life, including health visiting, should adopt evidence-based practice in order to maximise impact and cost effectiveness of provision (iHV, 2018; Cowley and Bidmead, 2009). Despite this, the number of high-quality evaluative studies, conducted in the UK, specifically about health visiting or health visiting related activity, is low. So too is the level of evaluative research activity by health visitors, such that the health visitor voice is little heard in the production of evidence (Robinson 2017; Cowley et al., 2013). While health-visiting operates in areas of great relevance to UK policy, not least the reduction of health inequalities, the impact of health visiting at policy level is limited (Cowley and Bidmead, 2009). It is argued that the lack of evaluative research about health visiting and by health visitors, using an RCT design, as opposed to qualitative and observational work, is linked to this (ibid). This lack of evidence is illustrated by the fact that the RCT on which candidate publication 1 for this thesis is based (see Appendix 1) which was published over 15 years ago, remains one of the few RCTs that is directly focussed on health visiting in the UK. The need for specific support and action to enable health visiting to develop from this '*very low base*' has been strongly advocated (Cowley et al., 2013; p.23).

I argue in this thesis that health visitors can, and should, be part of the process of conducting the rigorous and relevant RCTs that are important in the development of health services, including health visiting. The writing of this thesis resulted in new personal understanding of how specific knowledge and skills, seemingly acquired and applied separately in my two distinct roles of health visitor and researcher, do in fact closely align. I have long appreciated

the application of the technical-rationale skills of the research environment to my health visiting practice; it is more the greater understanding of how skills central to health visiting have 'seeped into' my practice as a researcher, and more specifically, a trialist working from a co-production perspective. In addition to the reflexive practice skills mentioned above, such health visiting skills include: relationship building, working with context, the spanning of boundaries and brokering of knowledge in order to align different perspectives and facilitate transactions (Kislov et al., 2016). An example of the practice of these skills in health visiting is the trust building work routinely carried out with families where domestic violence has occurred (Litherland, 2012), and knowledge brokering and boundary spanning as discussed in relation to health visitors using an online Community of Practice (Ikioda and Kendall, 2016).

1.3 METHODOLOGY

My thesis includes the four candidate publications listed below, in chronological order, and this integrative summary. The methodological approach adopted in the production of this integrative summary is 'insider research'. This term describes projects where the researcher has a direct involvement or connection with the research setting (Greene, 2014). This 'situatedness' provides a unique position from which to study the issue of interest, in depth, and with special knowledge about that issue. For the purposes of this thesis, to reflect and provide learning on academic-practitioner collaboration, I have drawn on: my insider knowledge of the studies from which the candidate publications arose; my knowledge from my wider practitioner and academic research experience; and, as argued above, my reflexive practice skills which as stated above are central to the health visiting process (Robinson, 2017).

Reflexivity in the research context is described as concerning thoughtful, analytic self-awareness of researchers' experiences, reasoning, and overall impact throughout the research (Raheim et al., 2016). In this thesis, using a reflexive approach, I: make clear my specific involvement in the work described; consider tacit knowledge and iterative adaptations used by me at the time of this work based on self-awareness; explore the implications of involvement and particularly my insider (practitioner)/outsider (academic researcher) positioning in the work and render more visible the voices of practice-based stakeholders.

1. I therefore brought to this thesis both my methodological understanding acquired as an academic and my health visiting skills, including my inclination and skills for reflection and

reflexivity, to address the following research questions: What do we know about collaboration between academics and practitioners in the co-production of research, and in RCTs of complex interventions specifically?

2. How has the nature of practitioner active involvement in the conduct of evaluations of complex interventions using a RCT design changed over that last two decades? How does it differ between health and education research?
3. How can the voices of harder to reach practitioners be brought into the process of shaping decisions about key elements of a RCT?
4. What is the learning from these analyses, how does this advance knowledge in the field of study and where are health visitors situated in the on-going process of improving the knowledge production process through academic-practitioner collaboration on RCTs?

To support the reflexive process in this integrative summary I carried out the following:

- Purposive documentary analysis of a range of sources of information related to the four studies on which the candidate publications were based including: study protocols; other publications on the studies; minutes of meetings; emails; raw data and my own reflective field notes.
- Discussions with other academic researchers who worked on the studies where they took the role of informed critical friend.

These two approaches were iteratively intertwined. For example, a discussion would raise an issue that was then further explored through reference to various study documents and vice versa. This allowed me to gain, and validate, a wider picture and concomitant interpretations.

The candidate publications are all substantive papers (as opposed to theoretical pieces) arising from four different studies. The papers have been selected purposively to include examples of evaluations conducted at different time points in my career as a means of illustrating a journey, both personal and disciplinary. The publication selection reflects diversity in terms of: the nature of the practitioner research involvement (as passive data providers and active decision-makers); research settings (education, clinical and public health); types of practitioner/stakeholder; the practitioner-academic mix in terms of roles and responsibilities on studies. While these were predominantly academic researcher led studies, one for the studies (CASCADE) was practitioner led. This integrative summary uses 'tales from the field' to

provide rich description, alongside various relevant theoretical conceptualisations and in-depth analysis. As such I aim to further understanding of some key issues in this field of study.

1.4 CANDIDATE PUBLICATIONS

The four candidate publications in this thesis (see appendices 1-4) are as follows:

Publication 1

Sawtell M, and Jones C (2002). **Time to listen: an account of the role of support health visitors.** *Community Practitioner*, 75 (2), 461-463.

This paper describes the Social Support and Family Health Study and specifically the Supportive Listening Programme, a public health intervention, delivered by health visitors (of which I was one). This was delivered to new mothers and aimed to reduce childhood injury and improve maternal psychological wellbeing. The research design for the evaluation of this intervention was an individual RCT with integrated process and economic evaluations. The study ran from 1999-2002 and was led by Ann Oakley as part of her pioneering work using this research design in the UK. In this thesis, this paper provides an opportunity, in conjunction with the other more recent publications selected, to present sequential developments over nearly two decades in involvement of practitioners, as collaborators, on studies with this design. I led on the writing of this paper; the second author (Carol Jones) was also an SHV on the study. For this paper see appendix 1.

Publication 2

Sawtell M, Jamieson L, Wiggins M, Smith F, Ingold A, Hargreaves K, Christie, D (2015). **Implementing a structured education program for children with diabetes: lessons learnt from an integrated process evaluation.** *BMJ open diabetes research & care*, 3 (1), e000065

This paper describes the methods and findings from the process evaluation that ran alongside a cluster randomised controlled trial of a structured education programme for children with diabetes. The intervention, CASCADE, was delivered in a clinical setting by paediatric diabetic nurse specialists (PDNS) and aimed to improve the health and quality of life of participants. The intervention was developed by clinicians, one of whom was the study chief investigator (CI). The study provides the opportunity to explore the benefits and challenges of working as a multi-disciplinary, multi-institutional study team that included: university researchers; practitioners who were also the intervention developers; process evaluation and outcome

evaluation teams. I was trial manager of on the CASCADE study and I led on the writing of this paper. For this paper see appendix 2.

Publication 3

Jerrim J, Macmillan L, Micklewright J, **Sawtell M**, Wiggins M (2016). ***Chess in Schools Evaluation report and executive summary***. London: University College London Institute of Education.

This publication is the peer reviewed final report produced for the evaluation of The Chess in Primary Schools programme. This educational intervention was developed by the charity Chess in Schools and Communities and delivered in schools by chess tutors. The aim was to improve performance in maths and English SATS tests. The evaluation was an RCT with integrated mixed method process evaluation. In this thesis, as well providing the opportunity to compare and contrast evaluations in education and health settings, and the part practitioners play in them, this publication provides evidence on including practitioners in collaborative work on the production of the logic model on which the evaluation was based. My role on this publication was to lead on writing the process evaluation section. On the study, I was joint-lead on the conduct of the process evaluation. For this paper see appendix 3.

Publication 4

Sawtell M, Sweeney L, Wiggins M, Salisbury C, Eldridge S, Greenberg L, Hunter R, Kaur I, Hatherall B, Morris J, Reading S, Renton A, Adekoya R, Green B, Harvey B, Latham S, Patel K, Vanlessen L, Harden A (2018). **Evaluation of community-level interventions to increase early initiation of antenatal care in pregnancy: Protocol for The Community REACH study, a cluster randomised controlled trial with integrated process and economic evaluations**. *Trials*. 19(1):163.

This publication is the protocol paper of The UK Community REACH trial. This study aims to assess the effectiveness of engaging communities in the co-production and delivery of interventions that address inequality in access to antenatal care. The study design is a matched cluster RCT with integrated process and economic evaluations with trial outcomes assessed using routine maternity data. As part of a National Institute of Health Research applied research programme grant there is an expectation of a co-production approach to the research as well as the intervention. The paper therefore supports both description of, and reflection on, co-production processes in a study of this type. It also provides the opportunity to describe in detail specific pieces of collaborative work with midwives and NHS data analysts to

determine both the specific outcomes to be measured and the processes used to access the routine data. This study was on-going at the time of writing this thesis. My role on Community REACH is trial manager and I led on the writing of this paper. For this paper see appendix 4.

1.5 ABOUT THIS INTEGRATIVE SUMMARY

Structure

This integrative summary is presented as follows:

- Chapter one contains this introduction.
- Chapter two provides an overview of the current evidence.
- Chapter three provides a view across the four studies, from which the candidate publications arise, considering change in the nature and scale of practitioner involvement over time and sector. The learning from this historical view is presented, including the impact of missing practitioner voices.
- Chapter four builds on this emergent theme of missing voices. Using a case study approach, the social processes, at individual level, used to involve practitioners in co-producing three core components of an RCT are described and discussed.
- Chapter five considers key themes arising from the analyses in order to draw conclusions. It presents the unique contribution that this thesis makes to wider knowledge, with particular reference to the skills health visitors can bring to the research process in studies of this type.

Terminology in this thesis

Generally, the assumption in the literature is that the term ‘practitioner’ refers to professionals. Professional practitioners have been defined as those engaged in: paid employment; a profession with formalised continuous professional development and a professional body; and a direct client relationship (Bell et al, 2010:15). It is appreciated that for a few types of stakeholders discussed in this thesis (e.g. data analysts) ‘practitioner’ is not the most obvious description and the term ‘professional practice partner’ might be more appropriate. However, these types of stakeholders are closely associated with service provision, work as part of multi-disciplinary teams, and support the work of the clinicians delivering care (<https://www.healthcareers.nhs.uk/explore-roles/healthcare-science/roles-healthcare-science/clinical-bioinformatics/clinical-bioinformatics-health-informatics>). A

theme in this thesis is the risk of excluding the local knowledge of ‘seldom heard’ groups in the research process; therefore, I have actively taken an inclusive approach in my thesis. It is likely

that others searching for evidence in this area of study will start with the term 'practitioner'. Therefore, for simplicity, and in the interests of maximising the opportunities for dissemination of the learning from this thesis, I use the term 'practitioner' in this integrative summary.

Also, many of the individuals referred to in this integrative summary exist somewhere along a practitioner-academic spectrum in terms of their experience. As the chief focus here is on the influence of practitioner knowledge on academic led research, I use the term 'practitioner' if they are currently or recently in practice (however part-time). So, for example a midwife working part-time in both a research associate role and a clinical role is defined as a practitioner, as is an individual currently working only as an academic who was recently employed part-time as a general practitioner (GP). Conversely, a qualified doctor who has not practiced clinically for some time (e.g. five years) and is unlikely to return to practice is defined for the purposes of this thesis as an academic.

Finally, in the interests of brevity and 'flow' of the text the four studies are in general referred to as SSFH, CASCADE, CHESS and Community REACH rather than for example 'the SSFH study'.

CHAPTER 2

REVIEW OF RELEVANT LITERATURE

2.1 INTRODUCTION

This chapter provides background information, located through extensive searching of the literature, on collaboration between academics and practitioners in evaluations of complex interventions using a randomised controlled trial (RCT) design. It aims to establish the context for the ensuing chapters. The chapter starts with information on the RCT design, how it has developed methodologically and its use in health and education research. This is followed by a review of the concept of collaboration for co-production including: how practitioners can collaborate; the challenge of collaboration in research in general and specifically in RCTs; what is known and not known about overcoming challenges; and the need for more research.

2.2 RANDOMISED CONTROLLED TRIALS

Randomised controlled trials for the evaluation of complex social interventions

The RCT has long been considered the optimal method for obtaining estimates of effectiveness of interventions in the field of medicine due to the minimisation of selection bias achieved through randomisation. Oakley argued in the 1990s that the RCT research design could and should be used in the evaluation of complex interventions designed to address social, as opposed to, medical problems (Oakley, 1998). Complex interventions are extensively used in the health service, in public health practice, and in areas of social policy that have important health consequences, such as education, transport, and housing (Craig et al., 2008). The strong focus on the use of evidence to inform policy and practice in public services, has led to the evaluation of effectiveness of complex interventions being of increasing importance to practitioners, policy makers and researchers from the social sciences. The term 'applied research' is now widely used to denote this type of research that is concerned with finding solutions for practical problems and societal challenges (Campbell and Vanderhoven, 2016).

Randomised controlled trials of complex health interventions over time

The growth in use of the RCT has been particularly evident in research on health interventions, compared with interventions for use in other public service sectors. To support this growth,

an impressive national infrastructure has developed in the UK over the last few decades. Prior to 2006, patient-based research in the NHS was conducted and funded through multiple funding programmes and schemes managed by the Department of Health. In 2006, however, the National Institute for Health Research (NIHR) (<https://www.nihr.ac.uk/>) was established in the UK with the aim of transforming research in the NHS. Much of the work of the NIHR has been directed at providing structural support for increased capacity for high quality applied research, including RCTs, with the involvement of a range of different types of stakeholder groups as integral to this. Principles adopted by the NIHR for building research capacity at individual and organisational level include: developing skills and confidence; supporting linkages and partnerships; ensuring the research is 'close to practice'; investing in infrastructure, and supporting clinician led research (Cooke, 2005).

Randomised controlled trials of complex education interventions over time

Despite the dominance of RCTs in health research, the RCT also has a long history in education, which in fact predates use in medicine (Oakley, 2005). However, the use of this research design, with its roots in positivist traditions, was strongly contested within the education research community from the 1970s on the grounds it lacks predictive power. As a result, there was much less emphasis on the RCT as the preferred type of evidence, over the ensuing decades, than in the health sector (Connolly, 2017) with less progress in terms of quality (Torgersen et al., 2005). However, an increased interest in the use of RCTs, as a basis for UK public policy making beyond the health sector, emerged in 2012. This included in education research (Pearce and Ramen, 2014). The Department of Education invested in various education research-capacity building initiatives via the charity the Education Endowment Foundation (EEF) (<https://educationendowmentfoundation.org.uk>). The vision of the EEF, which was launched in 2011, is to break the link between family background and educational achievement. The RCT research design has been extensively used in evaluations funded by the EEF, to test effectiveness of innovative interventions in education settings with a view to levelling the achievement gap. Despite some arguing for even greater use of the RCT in education research, the structural support for this type of effectiveness research in education in the UK remains limited compared with health research (Goldacre, 2013; Connolly, 2017).

Increasing the utility of the RCT

Key methodological developments in the RCT design have been progressed since Oakley made her argument in the late 90s. These developments are intended to increase the utility of

evaluations as a form of applied research (Wolff, 2000). Examples of developments include the use of the integrated mixed methods process evaluation and pragmatic design. Process evaluations aim to further our understanding of the 'how' and 'why' of intervention effectiveness or ineffectiveness by addressing issues such as fidelity of implementation and the contextual factors that affect it (Moore et al., 2015). Pragmatic trials are effectiveness studies conducted in real-world settings to answer questions relevant to patients, clinicians, and healthcare decision makers. They use rigorous scientific methods but include adjustments in the study protocol that make them acceptable and feasible to conduct in real world service settings (Tunis et al., 2003). Another key development, and the focus of this thesis, is the drive to move practitioners beyond being recipients of research knowledge to having an active role in its generation (Lunt et al., 2010).

2.3 COLLABORATION BETWEEN PRACTITIONERS AND ACADEMICS

Rationale for wider involvement in the research process

The dominant metaphor underpinning research has, in the past, been one of knowledge passing from academic to policy and practice domains, thereby delivering evidence of what works, often using an RCT design, from academic creators of evidence to intended users in practice settings (Walker, 2010). It is argued that those on the 'inside' of public services (such as health and education) have a different epistemology from those on the outside, including academics, and this different way of knowing needs to be taken into account in the research process (Frankham, 2009) in order to increase both the usefulness and the use of research. Hence there has been a shift in focus to researching *with*, as well as, *on* research users (i.e. policy makers and practitioners) and services users (for example pupils and patients) (Hewison et al., 2010). Attempts to reorder the social relations of research production, i.e. who researches and who is researched, have been particularly evident in the drive for greater public involvement in research, as demonstrated by the INVOLVE initiative. INVOLVE is a government funded programme, established in 1996, that supports active public involvement as an essential part of the research process, including in RCTs, in NHS, public health and social care research (<https://www.invo.org.uk>).

Co-production of research between academic researchers and practitioners

In the quest to generate research that achieves both scientific rigour and public benefit there has been considerable interest in the co-production of research by academic researchers and

21

practitioners (Heaton et al., 2015). Co-production (or co-creation) of knowledge is defined, in this context, as collaborative knowledge generation, through research, by academics working in partnership with other key stakeholders (Eyre et al., 2015). Heaton and colleagues (Heaton et al., 2015) describe five core features of co-production in order to demonstrate that it is more than just involving and engaging, but directly contributing and thereby influencing the decision-making processes within the research. These authors are writing in the context of co-producing with service users, as opposed to research users (i.e. practitioners). Their analysis can, however, be usefully applied to co-production between practitioners and researchers. These five features are: users of services are regarded as active agents, not passive recipients; relationships between the various stakeholder groups become more equal and knowledge and experience have equal value; relationships become reciprocal and mutually beneficial; transformative service change results from service users' active involvement; and networks and organisations encourage and facilitate service user involvement.

Ways that practitioners have been shown to work together with researchers in the multifaceted and iterative processes of knowledge production include:

- shaping research agendas at national policy making or funding body level (e.g. determining research priorities; deciding who is awarded funding);
- designing and developing new interventions;
- designing and conducting evaluations at research programme or individual study level (e.g. as members of bodies overseeing studies, conducting the day to day research activities as co-researchers) (McCabe et al. 2016).

In health research, practitioner involvement in effectiveness studies intended for informing practice, originated with doctors. Considerable structural support continues to facilitate this involvement of medical clinicians, in applied research, as part of their career pathway (Tooke, 2008). Nurses and allied health professionals have, historically, been more likely to conduct small scale qualitative research (Richards et al, 2014) and this includes within the health visiting profession (Cowley et al, 2013). There are for example relatively long traditions of practitioner involvement in action research (Hegney and Francis, 2015) and case study approaches (Heale R and Twycross, 2018). While there are pragmatic reasons for this, in terms of scale and resource, these types of research also have appeal for practitioners because of the capacity they offer for dealing with contextualisation and interpretation for context (Bell et al., 2010). However, while action research and case studies suit knowledge production for particular contexts, they

are less well suited for producing the generalizable knowledge, that demonstrates impact, resulting from RCTs that underpins evidence-informed decision-making.

Challenges in academic-practitioner collaboration

Despite the logic of wider involvement in applied research, and the growth in interest and skills in knowledge use and production by practitioners, collaborative research is recognised as very difficult to conduct effectively (Krebbekx et al., 2012). Academics and practitioners, whether in healthcare or education, exist in different social and professional worlds with diversity in languages, needs and interests, resources and incentive structures. The ‘two communities’ concept has been widely used to illustrate the extent of the gap between these worlds (Newman et al., 2015).

Specific barriers to collaboration with academic research, from a practitioner perspective include: lack of time and/or confidence; perceived direct limited relevance of the research being carried out; work cultures where research is not given priority as a valuable way to spend employed time (Verhoef et al., 2009; Hewison et al., 2012); and limited scope for practice and career development within roles (Kunhunny and Salmon, 2017). In terms of barriers to collaboration on RCTs specifically, a key factor at systems level, as discussed above (see page 10 of this thesis), has been the historical association of the medical profession with this research design. This has influenced the availability of key facilitative factors, such as funding and capacity for evaluative research, for less powerful professional groups that have traditionally been associated with research methods considered to be less trustworthy (Glaziou et al. 2004). At individual study level, the dominance of quantitative evaluative research methods within large mixed-methods studies (such as RCTs with integral process evaluations) and of different disciplines, most notably doctors, within the multi-disciplinary teams that conduct these large research studies, has been demonstrated (O’Cathain et al, 2008). Conflict, lack of shared decision making and ultimately dysfunctional teams were identified in O’Cathain and colleague’s investigation into team working. These power differentials understandably can lead to those from professional groups holding less power and influence, particularly individuals new to research, feeling de-valued and disempowered and thus limited in the extent to which they function as decision-makers on such studies.

From a researcher perspective, practical difficulties with active practitioner involvement include the sustained work, described as ‘continuous investment’ (Reeve et al., 2016), required to manage the co-ordination of this diverse knowledge and experience in a meaningful way. A

particular specific challenge of co-production in trials, that is important for the credibility of evidence, is achieving a balance of the processes that gives it legitimacy as the best means of examining intervention causality, with factors that promote collaboration (Pearce and Raman, 2014). Describing the challenges in collaborating for research priority setting, Madden and Morley (2016) write *'it is an arena in which 'hard' evidence-informed ideals meet 'soft' participatory practices'*.

Thus, barriers to practitioner participation as collaborators in research, can be summarised as falling into a number of overlapping categories. These include: organisational resource, culture and environment; individual confidence and motivation; and logistical and epistemological challenges. As a result of these multiple challenges, co-production has been described as a *'risky method of social enquiry'* in that it is time-consuming, emotionally demanding and inherently unstable (Flinders et al., 2016).

Solutions to collaboration challenges

Co-production in research is now widely considered to be here to stay (Martin, 2010). As such, there is considerable interest in how to build and evaluate practitioner research capacity (Cooke, 2005; Matus et al., 2018). Research capacity building is defined as *'a process of individual and institutional development which leads to higher levels of skills and greater ability to perform useful research'* (Trostle, 1992; p 1321). Key factors for improving research capacity include enabling institutional infrastructures and the confidence of individuals (Cooke, 2008). While the level of research activity and ability to conduct research has been limited for nurses and allied health care practitioners increasingly attention is being paid to improving opportunities (Matus et al., 2018). Examples of mechanisms for this include: the NIHR supported Collaborations for Leadership in Applied Health Research and Care (CLAHRCs) (Heaton et al., 2015); and the Health Education England and NIHR Integrated Clinical and Academic Programme for non-medical healthcare professionals (<https://www.nihr.ac.uk/funding-and-support/documents/ICA/TCC-ICA-Eligible-Professions-and-Registration-Bodies.pdf>). Recent developments connected with these initiatives include: the NIHR/Council for Allied Health Professions Research (CAHPR) AHP Research Champion scheme and The NIHR CRN Allied Health Professionals Health Strategy 2018 – 2020 (<https://www.nihr.ac.uk/our-research-community/clinical-research-staff/allied-health-professionals.htm>); The NIHR Directory for Clinical Research Practitioners (<https://www.nihr.ac.uk/>). These opportunities recognise the desire for practitioners to

remain in clinical practice while conducting applied research, raising the profile of the clinical research practitioner and providing them with support.

In education, though less extensive and developed than health, there are examples of partnerships and networks which focus on raising the profile of research and are likely to increase capacity for collaborative research. One such example is the School Health Research Network, a policy–practice–research partnership established in Wales in 2013 (Hewitt et al., 2018).

Other solutions to collaborative challenges include alternative models of creating research evidence that maintain rigour while increasing relevance. Examples include the researcher in residence model where an academic researcher is situated in a practice setting (Marshall et al., 2014) and the incorporation of an intermediary, a knowledge broker, to span the boundaries between research and practice communities to facilitate partnership working (Bornbaum et al., 2015). Increasing recognition of this requirement for greater sensitivity in research to context, has informed the growing interest in multi-paradigm research (Day et al, 2016) and the realist approach. Realist research has as a central tenet the acknowledgement that programmes work differently in different contexts (Bonell et al, 2012). This methodology thus demands the involvement of multiple stakeholders, including practitioners, in order to answer the key questions of: what works for whom, under what circumstances, why and how? Furthermore, increasingly, study teams using an RCT design are starting to use, and disseminate details, on involvement processes. One such example in health care research is that of Day and colleagues (2016). These authors describe how they embedded action research techniques, which in contrast to the RCT are based on interpretivist and critical principles, in an RCT of telemonitoring on health-related quality of life. The action research element allowed the ‘different ways of knowing’ of practitioners to influence the execution of the RCT, enhancing rather than compromising the quality of the RCT.

Unsurprisingly, it has been suggested that academic researchers, used to drawing on their skills as analysts, may require new skills and behavioural approaches in order to embrace the role of collaborator/partner (Baumsbusch, 2008). It is recognised that while technical knowledge is important, social skills are required to manage ambiguous loyalties, reconcile different interests and motivations, as well as to negotiate competing goals (Orr and Bennett, 2009).

2.4 THE NEED FOR MORE RESEARCH

Although evidence on co-production of research has advanced the debate, further analyses and comment have been called for (Nutley, 2010; Felipe et al., 2017). In terms of RCTs it has been noted that there is a paradox. While trials are considered the cornerstone of evidence-based services, there is very little evidence available on methods and infrastructure for conducting these complex studies (Treweek et al., 2015; Datta and Pettigrew, 2013). Despite the increased priority given to using collaborative methods within trials, and a growing evidence base on what contributes to success, it is widely acknowledged that there is a need for more reporting of practical examples and concomitant discussion of the learning this generates. This is what this thesis attempts to do.

CHAPTER 3

TRENDS IN RESEARCHER - PRACTITIONER COLLABORATION

3.1 INTRODUCTION

This chapter presents an exploration of, and critical reflection on, changes over time in the scale and nature of researcher-practitioner collaboration in the conduct of evaluations of complex interventions using a randomised controlled trial (RCT) design. Specifically, I focus on the issue of practitioners as members of the various groups that oversee and run these studies. I do this with reference to all four of the evaluations featured in my candidate publications and capitalise on my experience as an insider on all the studies. The methods used include examination of study documents, for example protocols and minutes of meetings, and focussed reflexive discussions with other members of the study teams. My aim is, through an analytical approach, to contribute knowledge on issues in this field that have been explored generally in the literature, but where details are relatively ill-defined.

The chapter starts with background information to provide context. I start with information on study teams and issues related to the various roles and responsibilities of the members of these teams. I then provide detail on the four studies, with a particular focus on elements that illustrate the scale and nature of the practitioner collaboration. This is followed by mapping of the trends in practitioner involvement across the four studies. The chapter continues with an assessment of the meaning of these trends in terms of the potential for effective collaboration, on which co-production of knowledge is reliant. A particular emergent issue that is explored, and analysed, in the chapter is that of missing practitioner voices in the evaluations.

3.2 BACKGROUND

Study teams

Traditionally RCTs have been run by groups of academic experts with the required technical skills to achieve a research product that will have maximum scientific value. These experts have included statisticians, subject experts and methodologists, including more recently researchers with mixed methods skills (O’Caithin, 2008). Over the period that the four studies

have run, it has increasingly been expected that practitioners will be key members of these teams.

For those conducting health studies, the formation of several inter-disciplinary decision-making groups are recommended to conduct the work of the trial and ensure good governance (MRC, 2017). These groups include the trial management group (TMG) and the executive Trial Steering Committee (TSC). The TMG is responsible for the day-to-day delivery and conduct of the trial and the TSC is the executive decision-making group (Harman et al., 2015). While TSCs will generally meet a few times over the course of a study, the intensity of the intellectual and practical work required for successful conduct of these complex studies demand TMGs meet much more frequently (MRC 2017). Individuals with key oversight roles on studies with this design are also required. These include the trial Chief Investigator/s (CI) and, in health studies, site Principal Investigators (PI). The latter hold responsibility for the conduct of the research at a study site and the former the overall responsibility for the conduct of the whole study, including leading the TMG. Trials of education studies have not had the same explicit institutional expectations around trial management processes, but have generally adopted similar structures (Connolly 2017).

Membership of these decision-making groups is often a voluntary activity but increasingly opportunities for practitioners to work as paid co-researchers on RCTs of a complex intervention have become available. In this role, practitioners carry out the day to day work of the study alongside academic researchers. Benefits to the study of this type of involvement by practitioners include both the subject and practical knowledge they bring to the day to day conduct of research processes. Furthermore, these practitioner co-researchers build research capacity at the level of the individual and the profession (Stainton et al., 1998).

Relationship between intervention developers and evaluators

The research process in evaluations of complex interventions involves the identification of the research problem. Next steps include: the design and development of the potential solution (the intervention); the implementation of the intervention; the design and conduct of the evaluation to test the effectiveness, or otherwise, of intervention. The extent of the connectivity, or separation, of these 'next steps' and the teams responsible for them, varies between studies. Separation assures scientific credibility, through independence; connection or overlap maximises scope for detailed understanding of the intervention from those with the

most in-depth knowledge of it, but risks conflict of interest. When practitioners are involved with the intervention and the evaluation, this complexity is particularly apparent not least because of the likely greater threat to reputation and livelihoods if an intervention fails to effect change (Bonell et al., 2012). The theory of managing this issue of intervention/evaluation boundaries has been the subject of much debate, but practical guidance for those running these studies only emerged in 2015 (Moore et al, 2015). This guidance makes clear, with particular reference to process evaluation data, the importance of transparency of reporting of relationships with policy and practice stakeholders, and being mindful of how may these affect the evaluation. Understandably the importance of the leadership skills of those overseeing these complex group processes have been identified (Krebbekx et al., 2012). The next section describes my four studies, with a particular focus on the makeup of the study teams, and locates them in this history.

3.3 THE FOUR STUDIES

Overview

All four studies were examples of applied research using an RCT design. They were funded by UK national research programmes and all were the first UK effectiveness studies of the intervention in question. The funding bodies involved had multi-stakeholder groups (including practitioners) involved in the selection processes for funding. This was the case even for those operating in the late 90s (Stein and Milne, 1998). SSFH, CASCADE and CHESS were all single studies that are complete at the time of writing this thesis. Community REACH is different in that it is currently on-going and is part of a programme of work comprised of four work packages, one of which is Community REACH. These four packages function as a whole, as well as individual studies. While Community REACH was clearly different structurally, in terms of scale, it was not so different that comparisons are unreasonable. For example, CASCADE was a single study, not a programme, yet it was large with 28 study sites and a TMG of approximately 20 members. Community REACH was a single study (within the programme of four studies) with six sites and a TMG for the programme that had approximately 10 members (see table 1).

The studies

Table 1: Summary of key structural, organisational and research design features of the studies.

Study	Funder	Sector	Date	Research design	Grant holding institution (partner institutions)	Chief investigator
SSFH	NHS R and D Health Technology Assessment (HTA) Programme	Public health	1998 - 2002	Individual RCT; 3 trial arms.	IOE (now UCL IOE) (3 other institutions)	Academics x 2
CASCADE	NIHR HTA programme	Clinical health	2008 - 2013	Cluster RCT: 2 trial arms	UCLH (2 other institutions)	Practitioner (clinical psychologist)
CHESS	Education Endowment Fund (EEF)	Primary school education	2013 - 2015	Cluster RCT: 2 trial arms	IOE (now UCL IOE) (1 other institution)	Academic
Community REACH	NIHR – Applied research for patient benefit programme grant	Public health	2014 - 2019	Cluster RCT: 2 trial arms	UEL/Barts NHS Trust (4 other institutions)	Academic

The Social Support and Family Health Study (SSFH) 1998 - 2002. This study aimed to measure the effectiveness and cost effectiveness of two different strategies for providing support to mothers in disadvantaged inner-city areas, of London, compared with standard care delivered by NHS health visitors. The first strategy was a programme of monthly supportive visits delivered by five support health visitors (SHVs), trained in supportive listening, held primarily in the mother's homes over a one-year period. The structure of the visits was informal with a focus on listening to the woman, who chose the topics for discussion. The health visitors, of which I was one, all had at least two years' experience of working as an NHS health visitor, post qualification. The second strategy was Community Group Support delivered by eight local community support organisations (CGS). Seven hundred and thirty-one women were recruited to the trial.

The SSFH study had two joint Chief Investigators (CI) who were both academics, and who designed the intervention and the trial. In addition, on the team there were four academic researchers who were experts in mixed methods research. Five of these six researchers, who also oversaw the delivery of the intervention, originated from one academic institution. The

five support health visitors (SHVs) were also employed by this institution. Two health economists from another institution conducted the economic evaluation. The only practitioner contributing to the oversight of the conduct of the study was a senior doctor in the public health department of the Health Authority of the area where the intervention was delivered. She was a Co-investigator on the study and acted as an advisor on an ad hoc basis. (See candidate publication 1 and also: Wiggins M et al., 2004; Wiggins M et al., 2005).

The Child and Adolescent Structured Competencies Approach to Diabetes Education study (CASCADE) 2008 - 2013. This study aimed to test the effectiveness and cost effectiveness of a structured education programme for groups of children with diabetes. The intervention, CASCADE, incorporated psychological approaches to improve long-term glycaemic control, quality of life and psychosocial functioning in a diverse range of young people. The intervention consisted of four group education sessions delivered primarily by paediatric diabetes specialist nurses (PDSN). Twenty-eight paediatric diabetes services across London, South-East England and the Midlands took part in the evaluation. Forty-three practitioners were trained in the intervention, prior to intervention delivery.

The CI on the CASCADE study was a practitioner, a clinical psychologist, who had designed and developed the intervention with a senior specialist diabetes nurse. These two practitioners were employed by a clinical NHS Trust in London; a dietitian from a different Trust also provided advice on developing the project. Two other practitioners, who were both senior doctors from the same clinical team as the intervention developers, were co-applicants and members of the research team. Additionally, the research team was comprised of four user representatives and approximately 15 academics (exact numbers fluctuated during the course of the study). The academics were: a team of statisticians, data managers and health economist from a clinical trials unit; a number of researchers with mixed methods expertise from two different universities, one of whom was also a public involvement expert. Each of the 28 participating paediatric services (trial clusters) had a Principal Investigator for the trial; at the start of the trial all of these PIs were doctors (a nurse took over this role in one site when it became vacant mid-trial). (See candidate publication 2 and also: Christie, Thompson, Sawtell, et al., 2013; Christie, et al., 2016).

The Chess in Primary Schools evaluation (CHESS) 2013 - 15. The CHESS study aimed to test the effectiveness, in improving maths and English attainment, of a classroom-based programme of teaching children to play chess using an approach developed by the charity

Chess in Schools and Communities (CSC). The intervention involved children in Year five of primary school taking 30 hours of chess lessons delivered by a tutor, who was an experienced chess player and trained in the CSC programme, and supported by the class teacher. One hundred schools in 11 Local Authorities in England took part, 50 were allocated to the intervention and 50 to the control arms. Twenty-three chess tutors delivered the intervention across the 50 intervention schools.

The charity CSC, which had developed the intervention, oversaw intervention delivery during the evaluation. The evaluation team was completely independent of the intervention team and consisted of five academics, all employed by the same institution. Of the five, three were quantitative researchers (one of whom was the study CI) and two mixed methods experts (of which I was one) who ran the process evaluation. There were no practitioners on the research team. (See candidate publication 3 and also: Jerrim, Macmillan, Micklewright, Sawtell, Wiggins, 2017).

The Community REACH study (Community REACH) 2014- 2019. This study aims to assess the effectiveness of engaging communities in the co-production and delivery of an intervention that addresses inequality in access to antenatal care and aims to increase early initiation of antenatal care. Volunteers from the local communities who have been trained, as part of the research, as 'Antenatal Care Champions' deliver the intervention messages through engagement with women and wider family members, as well as local community groups and organisations. Six NHS Trusts in north and east London and Essex have been recruited to the study. The intervention is being delivered in 10 electoral wards; 10 comparator wards have normal practice. No participants will be individually recruited to the trial as outcomes are measured using the anonymised routine maternity data of two cohorts of women.

Community REACH is one component of a wider programme of research, the REACH (Research for Equitable Antenatal Care and Health) Pregnancy Programme. The CI on the study is an academic with a social science research background. Three practitioners are co-applicants on the programme (two doctors and a senior maternity service manager) as are two maternity service users. Three other midwives are employed on the programme, in addition to the service manager. Their roles include: research associate (employed by one of the universities named in candidate paper 4); research midwives (seconded full or part-time from the lead NHS trust). The site PIs for Community REACH are consultant midwives.

3.4 IDENTIFYING THE RESEARCH QUESTION, DEVELOPING THE INTERVENTION AND EVALUATION PROTOCOL

Across the four studies the role of practitioners changed in terms of identifying the research question and an intervention to address this (see Table 2). In the first study, SSFH, identification of the research question and intervention development were primarily an academic exercise. Ten – fifteen years later, in both CASCADE and CHESS, these processes were led by practitioners who had been delivering the intervention, unevaluated, in the preceding period prior to the evaluation. In Community REACH, the intervention was co-designed by academics with input from co-investigators who were practitioners and also a diverse mix of other stakeholder groups, as described in publication four. Thus, across the four studies the general trend of decision-making is from academe to practice to a shared approach.

With a research problem and potential solution established, the next key issue is to determine who develops the study protocol and the methodology. In CHESS this was done by an independent university-based evaluation team. In contrast, in the three health studies the boundaries between those responsible for the intervention and the evaluation were blurred. The general trend was from the same group of academics doing both (SSFH), to practitioners and academics working together on the protocol (CASCADE and Community REACH).

Table 2: Practitioner involvement in key evaluation phases and study roles

	Identifying the research question/ developing the intervention	Developing the study protocol/ methodology
SSFH	Academics (+ limited input from practitioner co-investigator)	
CASCADE	Practitioners	Practitioners and academics
CHESS	Practitioners	Academics
REACH	Academics/practitioners	Academics and practitioners

3.5 CONDUCTING THE EVALUATION

Across the studies, five dimensions of practitioner engagement in the conduct of the RCTs can be mapped. These can be summarised as: numbers involved; voluntary or paid position on the study; types of practitioner; study roles; nature of intervention/evaluation responsibilities. These are discussed below.

Numbers – voluntary or paid

As shown in table 3, SSFH had no practitioners in paid roles on the research team and only one practitioner in a voluntary role, as a co-investigator (Co-I). Ten years later, CASCADE had a practitioner as study CI and three other practitioners as key members of the TMG. These four practitioners had some research experience, but for most this was relatively limited and they were actively practising in a clinical role at the time of the study, with some funded research time. For REACH, the CI role was held by an academic but there were multiple other practitioners involved, in different types of roles. Some worked in paid research roles carrying out day to day research tasks alongside academic researchers. In CHES, the only study of the four in an education setting, practitioner involvement was very limited, with no practitioners formally involved in the conduct of the research. Thus, despite being conducted in the period 2013-15, CHES was much more similar in this respect to SSFH which was carried out over fifteen years earlier, than to CASCADE which ran at a similar time.

Table 3: Summary of practitioner numbers, types and roles

Study	TMG practitioner member (n = approx.)	Practitioner CI (n)	Practitioner Co-I (n)	Paid; type of practitioner (role on the study)	Voluntary: type of practitioner (role on the study)
SSFH	1	0	1	0	Doctor (Co-I)
CASCADE	4	1	3	Clinical psychologist (CI; intervention developer); specialist diabetic nurse (Co-I; intervention developer)	2 doctors (Co-I); 1 dietitian (Co-I); 28 doctors (site PIs)
CHES	0	0	0	0	0
REACH (whole programme – includes 2 trials)	3	0	3	Midwife (research fellow); 2 midwives (seconded research midwives)	2 managers/ midwives; 1 doctor (all Co-I); 1 consultant midwife (steering group member); GP (chair of steering group); 6 consultant midwives (site PIs)

Types of practitioners and their roles on the study

While comparisons of numbers of individuals provides some information, it is not an entirely clear cut 'measure'. Individual practitioners worked in different configurations on the research trials which impacted on their capacity for involvement such as: full-time/part-time; for the whole study/for part of the study; paid as co-researchers, unpaid in an advisory role. Furthermore, many were researcher/practitioner 'hybrids' with weaker or stronger practitioner identity depending on their closeness to practice at the time of the study. Therefore, trends in which professional groups were involved and in what types of roles is a more reliable way of charting change.

On the health studies, none of the interventions being evaluated were delivered by doctors. SSFH was delivered by health visitors and community organisations, CASCADE by specialist nurses and Community REACH by lay volunteers. Yet, in the earlier health studies, doctors were strongly represented in those overseeing the evaluation. On Community REACH, in contrast, the involved practitioners were predominantly midwives. Some were contributing to the study in a voluntary capacity (for example as site PIs and steering group members) while working full time in a service delivery capacity and others were in paid research roles.

As demonstrated in CASCADE and Community REACH, by this time in the historical trajectory (2008 and 2015 respectively) practitioners were holding key study roles. The difference between the two studies is in diversity of roles and capacity of individuals for input. In CASCADE practitioners were on various study management and advisory groups but not doing the day to day work. On Community REACH practitioners were involved across the various levels of the study including working as paid co-researchers on the day to day conduct of evaluation tasks. As such, on Community REACH midwives were involved with key study tasks, such as designing process evaluation research instruments, ensuring that they addressed pertinent service issues.

Comparison of the incumbents of the site PI role in CASCADE and Community REACH illustrates the trend in allied health professionals increasingly holding key study roles. In CASCADE, at the start of the study, this role was conducted by doctors (paediatricians) in all the 28 sites. This was the normative arrangement at the time. However, during the course of the study the departure of a site PI required a replacement to be found. The search for another doctor to take on the role proved fruitless; and though concern was expressed by various decision

makers (for example the site Research and Development office) about the relevance of a non-medical/non-research active practitioner taking on the role, eventually a pragmatic decision was made for a senior specialist nurse to fill the position. This nurse was a senior practitioner and had been heavily involved in delivering the intervention and supporting data collection, from the start of the trial two years earlier. Six years later, in Community REACH, senior midwives (rather than obstetricians) took the PI role in all six sites, from the start of the study.

Intervention/evaluation responsibilities

As stated above, managing the boundary arrangements, of those with different roles and interests, in order to balance close observation of implementation processes, and independence of evaluators is a challenge in evaluation studies. The four studies illustrate the historical development of how this was approached in practice. In the SSFH study, as shown, the intervention was developed primarily by academics who were also the evaluators. Furthermore, the intervention implementers, the practitioner SHVs, were employed, trained, managed and supervised by this academic team. In this early pioneering study there were no explicit boundaries between these various groups/activities. Reflective discussions for this thesis, with one of the SSFH research team members, and subsequent interrogation of study documents (e.g. meeting minutes) suggest that some research team meetings did take place without the SHVs. Thus, while information flowed freely from the SHVs to the research team it is likely that some information in the other direction was withheld. Examination of the final report demonstrates, however, that while other potential limitations of the study were acknowledged, this was not one of them (Wiggins et al., 2005).

In CASCADE the intervention originated primarily from practice. The senior clinician who developed the intervention and delivered a component of it in the trial (the training for the nurse intervention implementers), was also the trial CI. As such, this individual was both implementing a key aspect of the intervention and evaluating it. Recognition of the risk to objectivity of this crossing of domains was not considered prior to the study, but rather emerged once it started and the conflicts of interest came into sharp view. This led the TMG to impose boundaries to communication between different groups within the TMG. In practical terms this required process evaluators not to share emerging findings with TMG colleagues during the course of the study.

In Community REACH, seven years after the CASCADE study, intervention development, oversight of delivery and evaluation were also primarily the responsibility of the research team. However, as described in candidate publication 4, and as per good practice guidelines on boundary management (Moore, 2015), adaptive loosening and strengthening of boundaries occurred through the study. Critically this was an acknowledged process, regularly discussed by the research team, with respect to the perceived relative importance of objectivity versus closeness in terms of the balance of rigour and relevance.

CHES had the most separation between intervention processes and evaluation (see figure 1). The intervention and evaluation were managed by completely separate teams and there were no formal structures in place for regular contact between them. The little contact that took place across the practice/research divide included two meetings between the evaluators and the developers and occasional ad hoc contact, from researchers to the intervention team, to address specific issues that arose, for example, commenting on a draft questionnaire.

Figure 1: The relative separation and connection of intervention and evaluation teams



3.6 THE LEARNING FROM THIS 'VIEW' ACROSS THE FOUR STUDIES

Trends in the scale and nature of practitioner involvement

The general trend over time was for more practitioners, in more and varied roles, with more protected time to dedicate to the research. Importantly, involved practitioners were increasingly more representative of the types of practitioner groups that would be expected to use the knowledge generated by the studies. Thus, by the time of Community REACH, the study was immersed in a 'pool of representative practitioner knowledge' across the programme. This 'pool' was made up of the simultaneous contributions of many practitioners, at different stages and levels of the research process. Furthermore, greater understanding of how to robustly manage boundaries, intervention/evaluation boundaries, to ensure a balance of scientific rigour and practice relevance was observed over time.

The exception to these trends was in the education-based CHES where the conduct of the trial was almost entirely academically driven, prioritising objectivity through a relatively disembodied approach.

A key question is what the impact of this was i.e. did these trends improve collaboration and as a result co-production of valuable research knowledge. The analysis in this chapter has considered quantitative issues and not attempted to consider quality, or impact, of collaborative activities within the study teams. Detailed assessment of the quality, or effectiveness, of the interactions that occurred within the study teams described above, is beyond the scope of this thesis. Bowen et al., (2005) state that one of the key criteria for selecting effective team members is the ability to work collaboratively with peers. The need for further research on this issue has been identified but acknowledged to be complex requiring dedicated ethnographic work to adequately explore the intricate webs of issues at play (Frankham, 2009). However, using the classifications of collaboration of McCabe et al., 2016 (low, high and deep), the Community REACH appears to demonstrate the most advanced form, i.e. 'deep collaboration'.

A variety of mechanisms appear to have informed the changes over time in practitioner involvement with these RCTs. These were a combination of: increasing evidence on involvement theory and practice; structural support mechanisms at an institutional level; and continuity of some researchers (for example myself and other methodologists) bringing benefits of accumulated experience through what has been described as enduring connectivity (Armstrong and Alsop, 2010). Mechanisms supporting practitioner-academic collaboration are discussed in more detail in chapter 5.

Missing practitioner voices

As shown above, allied health practitioners came relatively late 'to the party', compared with doctors. The reflection in this thesis has revealed, however, despite a belief by myself at the time that the process was an inclusive one, there was in fact an on-going absence of some crucial practitioner voices from the decision making about the research process, in all three completed studies (SSFH, CASCADE and CHES). This finding emerges from the rich data of the integral process evaluations, which shows these voices were missing for a range of reasons related to the research capacity of practitioner groups and/or individuals. As shown in candidate publications 1-3, and in table 4, the interventions failed, or were significantly

inhibited, in all of the three completed studies. Whilst the analysis undertaken for this thesis suggests this failure cannot categorically be attributed to this missing information, it also demonstrates that the gaps in what was known by myself and the other evaluators about key aspects of the interventions, were significant.

Table 4: Missing practitioner voices on the studies

	Outcomes from the process evaluation (examples for each study)	Implications	Crucial missing independent practitioner voices as collaborators	Difference these voices could have made
SSFH	Poor uptake of Community Group Service by participants (19% compared with 94% for SHV).	Intervention implementation failed for 1 of the 2 trial arms.	Community group experts; health visitors.	Anticipated the need for more language support for CGS delivery & provided for this in the model.
CASCADE	Poor uptake of the intervention by families (30%). Key barrier was poor administrative support; insufficient numbers meant intervention not delivered as intended.	Intervention implementation failed.	Paediatric Diabetes Specialist Nurses (PDSNs), clinic administrative staff.	Anticipated the problem of the excessive administrative burden posed by the intervention & added such support to the model.
CHESS	Sub-optimal uptake (31%) by class teachers of training to support intervention implementation (see p 42 candidate publication 3).	Intervention implementation inhibited.	Classroom teachers, head teachers; chess experts.	Anticipated the burden of the training requirements for teachers & minimised this by arranging training locally rather than at a distance.

For example, the mixed methods process data in CASCADE, showed that administrative tasks and additional costs associated with running the intervention, created an unacceptable burden to the PDSNs. This contributed to low uptake of the intervention by the participating families and hence to reduced intervention fidelity. This was because the groups relied on a minimum number of attendees and the numbers often fell short of this. Had there been relevant practitioners actively involved in the study, who brought objectivity through independence, early identification of administrative barriers as a risk to implementation could have expected and timely and appropriate adaptations made, potentially resulting in implementation success.

While implementation success does not equate to effectiveness, it is a more conclusive answer about the intervention than the intervention failure that was the outcome of the trial. Similarly, in CHESS, process data revealed that teacher engagement in the intervention training was poor, and this detracted from the quality of intervention delivery. If teachers (independent of the trial) had been actively involved in the research process, they would be expected to have identified the problems with the training component in time for this to be addressed, thus salvaging a key factor compromising the chances of the intervention proving successful.

It is important to point out that that these omissions would have probably emerged through thorough intervention development work, conducted prior to the main trial (see for example practitioner data anticipating such barriers on: candidate publication 2, p. 4; candidate publication 3, p. 44). However, SSFH and CHESS had no funded formative work, despite this being strongly recommended for evaluation studies (Craig et al., 2008). CASCADE was preceded by a pilot study but this was conducted in a different context and hence failed to identify some key practical barriers to intervention delivery (Christie et al., 2008). This said, even the highest quality intervention development process will miss key information, not least because practice settings are inherently unstable. The nature of the context of applied and pragmatic studies requires that studies have the flexibility and skills within the research team to identify setting changes and respond dynamically using collaborative approaches. Whilst Community REACH also lacked a formal development phase prior to the trial commencing, the need for this was identified early in the study. As a result, the following fully integrated formative stage was instigated. This is described in candidate paper 4 as follows:

We will use a staggered approach, with implementation taking place first in three 'pathfinder' sites before starting in the remaining seven sites, in order that intervention development and delivery can benefit from some initial learning (p.113, appendix 4).

3.7 CONCLUSION

In this chapter I have explored the scale and nature of practitioner involvement in the conduct of the four RCTs of complex interventions from which the candidate publications arose. In terms of change over time, i.e. the last 20 years, the trend observed in the health research is one of: more practitioners involved, in more varied and central roles, with more dedicated time to spend on the research. Importantly practitioners coming into the research process have

become more representative of the practitioner groups most likely to be future users of the knowledge produced, as opposed to being predominantly doctors. Furthermore, the management of intervention/evaluation boundaries has become more robust, ensuring the credibility of the knowledge produced both in terms of scientific rigour and practice relevance. As such, potential for practitioner impact on researcher processes has increased considerably. In terms of difference in sectors, an RCT of a health intervention conducted at the same point in time as an RCT of an education intervention was much more progressive in terms of the nature and scale of practitioner involvement. A key theme to emerge from this retrospective exploration across the studies, is the absence of certain crucial but hard to reach practitioner voices, from the research process and the implications of this for the utility of the research.

CHAPTER 4

COLLABORATING WITH PRACTITIONERS TO DESIGN CORE COMPONENTS OF A TRIAL

4.1 INTRODUCTION

This chapter considers three examples of specific, short-term collaborative work between researchers and particular practitioner groups, within the previously described studies. Work at micro-level is characterised by the interactions that take place in the day to day conduct of an RCT i.e. at the 'on the ground' or direct-action level (Pfadenhauer et al., 2017). These examples were selected because they demonstrate attempts to co-produce three different core technical components of the RCT process. These episodes of collaborative work were conducted between practitioners, situated in practice-based settings, and academics situated in research settings. This chapter picks up on the theme of who is missing from research processes. The aim of the work described was to mobilise the experiential knowledge of particular practice-based stakeholders, whose voices might otherwise have been missing from decision making about research processes. In this chapter, using the rich description of a case study approach, I articulate the mechanisms and social processes used to engage practitioners in the collaborative activities. From the research 'side' the collaborative work was operationalised by myself supported by a small number of research team colleagues.

Two of the examples are from Community REACH and one from CHES. They are referred to in the two candidate publications (papers 3 and 4; see appendices 3 and 4) from these studies. As is common in research outputs, there is a reporting gap. Actions used to operationalise these important trial processes, including the social processes orientated to maximising collaboration, are reduced to a few descriptive sentences with little analysis. This chapter (and the further analysis in chapter 5) aims to open the 'black-box' and make explicit the processes used. In order to ensure evaluation as well as description, using the methodology described (see pages 13 -14 of this thesis) I conduct a reflexive analysis of both the 'ends (or findings/knowledge)' and the 'means' of the collaborative work conducted within the messy context of the real (practice) world (McCabe et al., 2016).

The three collaborative activities are:

1. A piece of work between researchers, NHS Trust data analysts and consultant midwives on Community REACH to determine trial outcome measures, where outcomes are to be assessed using routine NHS maternity data;
2. A related piece of work, on Community REACH, between the same three groups of stakeholders to design and conduct a pilot study to test and refine the process of transferring the routine maternity data required for outcomes assessment to the research team (for activities 1 and 2 see candidate publication 4, appendix 4);
3. An exercise between researchers and chess experts to construct a logic model for CHESS (see candidate publication 3, appendix 3).

This chapter starts with a brief overview of the rationale for collaborating on shaping these key RCT components. I then examine the three discrete collaborative activities including: the background to the role of each task in RCTs and specifically in the studies discussed here; how the collaborative work was operationalised; what was achieved and learnt in terms of the trial component and the collaborative process.

Rationale for collaborating at micro level on key components of the RCT design

RCTs are complex processes involving multiple different discrete, but interdependent, technical components. The factors that are the focus of this chapter i.e. outcome measures, data collection (in this case routine data) and a logic model, are all examples of these technical components. Traditionally the work to develop these has been carried out by academic researchers. They are all areas, however, where it can be conceived that practitioners have knowledge and skills that if harnessed, could transform the trial process.

The multiple potential barriers to practitioners participating actively in research are described in chapter 2. and include organisation and individual factors. The work described in this chapter was underpinned by organisational (i.e. the schools and hospitals) commitment to the studies. The focus therefore was on encouraging individuals to collaborate with the research.

4.2 COLLABORATING ON TRIAL OUTCOME DECISIONS

The two trial outcomes tasks in Community Reach study are covered (together) first, followed by the logic model construction task in CHESS.

Activity 1 - Determining trial outcomes in the Community Reach Study

Trial outcomes are the meaningful end points of a study and much of the work carried out in a trial is driven by the nature of the outcomes. Pragmatic trials aim to evaluate outcomes that genuinely reflect real-world settings and concerns. Yet defining relevant outcome measures for trials on social interventions is complex, and many trials measure and report outcomes that fall short of the requirement to reflect the concerns of diverse stakeholder groups (Heneghan et al., 2017). Careful selection of outcomes, and management of the processes involved in measuring them, is therefore a critical aspect of the trial (Williamson et al., 2012).

In Community REACH the primary outcome is the proportion of pregnant women attending their antenatal booking appointment by the 12th completed week of pregnancy. This was established at the proposal stage and written into the protocol. Some secondary outcomes were also proposed at this early stage, including for example antenatal admissions, emergency caesarean rates, gestation and weight at delivery, maternal and infant death. However, it was anticipated that the final full secondary outcome set would be confirmed in collaboration with local stakeholders. The primary and some secondary outcomes in Community REACH are measured using routine maternity data collected by the NHS Trusts where the participants receive their care.

Activity 2 – Transferring routine NHS data for outcomes measurement in the Community Reach Study

Routinely collected health data are defined as data collected without specific a priori research questions developed prior to utilization for research (Spasoff, 1999). Vast amounts of data are now collected by health care (and education) providers. Increasingly this is captured, stored electronically and released into the public domain by the Government to support public understanding of policy and practice. The Maternity Services Data Set was introduced in 2014 and is a patient-level data set that captures key information at each stage of the maternity service care pathway in NHS-funded maternity services

(<http://content.digital.nhs.uk/maternityandchildren/maternity>). The potential for large digital routine data sets to support improvements in research has been seen as a clear benefit of the digitalisation process (NHS England et al., 2014). Using routine data for research purposes is perceived to have substantial cost saving and resource-use advantages. This is based on the assumption that demands of data collection on NHS and research staff are reduced and long-term data collection on clinical outcomes is relatively unobtrusive (Staa et al., 2012). Thus, it is a growth area in trials (Benchimol et al., 2015). However, evidence suggests it is not the panacea that some had expected. Complications arise from the fact that data is collected primarily for practice and policy purposes and not for research. This influences what and how data are recorded and while improvements to hospital electronic data are noted it is still often considered 'messy and imperfect' (Read et al., 2013). Aside from data quality, there are other challenges in the use of routine patient data; including ensuring access arrangements, data processing and secure transfer that do not put data protection at risk (Staa et al., 2012).

The use of routine data for outcome measurement in Community REACH means that pregnant women may experience the intervention but are not recruited to the study; instead any impact of this experience is assessed using anonymised routine data. The anonymity of the data is a key requirement of the ethics permission for the trial, as described in candidate publication 4. The NHS Trusts, when they agreed to be Community REACH trial sites, were required to provide anonymised routine data for randomisation and outcomes analysis. While it could be expected that most, if not all, the individual data items required for the study are routinely collected, the specific items required for the study is a bespoke set and not, as with many trials of maternity services, a standard set of national data. As such the close involvement of local experts in managing this data, the NHS data analysts in the research sites, was imperative. Not only are they the gatekeepers of this resource, they have the potential to contribute to bringing meaning to it and to ensuring the best possible data quality and transfer processes.

4.3 HOW DID WE COLLABORATE ON ACTIVITIES 1 AND 2?

Reaching out to establish relationships through face to face encounters

At the beginning of the trial one or more face to face meetings took place with the consultant midwives as part of the process of recruiting NHS Trusts to the study. These senior midwives agreed to be the site PIs and gave us the contact details of the practitioner to liaise with about the routine data. A face to face meeting was then arranged, at each Trust site, with these data

experts to start building relationships and to share relevant information. It is important to note that this brief description of how relationships were established represents a lot of work; frequently a slow process of finding the right people, gaining their interest, negotiating practicalities such as meetings. All this early engagement work was orientated by the researchers, and particularly myself who carried out most of the day to day work, to making it as easy as possible for practitioners to be involved. For example, all meetings were held in practice settings and I asked practical questions such as *'what days, months etc. do you anticipate being busiest i.e. when should we try and avoid making demands on your time?'*

Prior to conducting the two outcomes activities, it was a trial requirement that the electoral wards that are the unit of randomisation were selected and randomised. This was also to be done using routine maternity data transferred by the data analysts. It is not detailed in this thesis, as it was a relatively straight forward process that did not require a significant investment of time/collaborative effort by the analysts. The site selection process is reported in publication 4 (see appendix 4). Apart from achieving its essential trial function, this site selection task provided useful insights for the research team, into local contexts of practitioners. These insights were critical foundations for the next steps of the work.

Next step –group meeting of the researchers, consultant midwives and data analysts

A key next step was to bring all the stakeholders together, from the different study sites. I arranged a meeting that included: site PIs, the data analysts and member of the research team, including the CI and the health economist. This was at a University site. Specific aims of the meeting were to finalise the outcome set and address the logistics of the data transfer. We also hoped that bringing the practitioners into this academic space would be motivating in that it: underlined the value the research team attached to their knowledge; allowed us to be hospitable and friendly (an important enabler reported by Bowen and Martens, 2005); conveyed the high academic status of the research.

Five (of the six) site PIs and two (of the six) site data analysts attended. I structured the meeting to include a presentation of the study and progress to date by the research team followed by a group discussion around the trial outcomes issues. Several key additions to the proposed outcomes list were strongly advised by the midwives to increase policy relevance (see example in Table 5). In general, there were strong shared levels of agreement between the research team and the midwives. However, the midwives enthused by the potential opportunity to answer some enduring questions in their work, suggested data items that did not contribute

to the scientific value of the trial. Researchers had to explain that these ideas could not be included. This brings into view uncomfortable issues of trust and control. Practitioner involvement and ownership of the research were encouraged until the point at which they put the scientific value of the trial at risk, whereon researcher priorities ‘trumped’ those of practitioners.

Table 5: Examples of impact of collaboration with consultant midwives and data analysts

	Example of contribution	Why is this contribution important?	Impact on research
Consultant Midwives	Added ‘booking at 10 weeks gestation’ as secondary outcome measure	10-week gestation expected to replace current ‘12 weeks gestation’ as Government target	Increased policy relevance
Data analysts	Added data item ‘category of type of delivery of baby’ to previous researcher determined data item of ‘method of delivery of baby’	‘Category’ is broader than ‘method’. Use of former increased accuracy of reporting of ‘type of birth’; a key secondary outcome measure	Increased internal validity

The discussions at the meeting exposed considerable uncertainty about the routine data including: completeness; consistency across research sites; accessibility; safe transfer processes that met accountability standards. It was also clear that there was a large gap in understanding by the research team, in the practices and language of health data analysis which put at risk the success of the data capture, and ultimately the whole trial. As such a key outcome of the meeting, which was a change to the planned process, was the decision to run a small collaborative pilot to address the uncertainties.

Collaborating with the data analysts to develop and pilot data transfer processes

The specific aims of this piece of work were to: assess the feasibility of, and processes for, collecting routine maternity data that is sufficiently complete and consistent across the trial sites for measuring the required study outcomes; establish the processes for secure transfer of the anonymous individual level data adhering to protocols that ensure the scientific integrity of the study through ‘blinding’ processes (Akobeng, 2005). As stated it was also an aim for the research team that the data analysts were active partners in the process so that we were

piloting 'with' rather than 'on' them. However, the likely challenges to this aspiration were illustrated by the fact that only two of the six NHS Trusts were represented at the meeting by an analyst, although all had been invited. I therefore proceeded by asking one analyst who appeared particularly engaged with the research to collaborate closely with me in developing and testing processes. This development work took place over two weeks and involved intensive communication via email and over the phone between the analyst and myself. Together we produced a pack of written support documents to structure and formalize the different roles and responsibilities of the various participating stakeholders and provide meaningful guidance on the technical aspects of the work and, critically, on a range of data security requirements. I then 'rolled out' the refined process to the analysts in the other five Trusts. Further communication from thereon was by email and phone. When analysts failed to respond and key deadlines were missed I learnt that making appeals to the site PIs to communicate with the analysts on our behalf, was generally effective in encouraging engagement by the analysts.

What did we achieve and learn from the two collaborative activities?

Two key outputs were achieved; a finalised list of outcome measures, and a process that resulted in the successful receipt of a data set from all six sites that were of good quality with respect to completeness and consistency. Furthermore, these outputs were co-produced; they reflected the concerns and interests of the practitioners while maintaining, and even enhancing, the scientific integrity of the trial (see examples in Table 5).

A key finding was that the data transfer task took much longer than expected - nine months from the initial formal request for the data at the start of the pilot to the receipt of the data from all the six sites. The expectation when the study was designed was that all the outcomes data, for the baseline, first follow up and second follow up cohorts, would be collected retrospectively at the end of the trial. As discussed above, the ability to collect data retrospectively is considered a key advantage of using routine data for trial outcome measurement. Our findings, with this bespoke data set, suggested that in the real world it is not the simple, linear, efficient and cost-effective route to assessing trial outcomes that is often implied. All the uncertainty and delay experienced during the pilot suggested a considerable degree of risk which at worst would result in the study failing due to a lack of outcomes data. Based on this analysis, a decision was made, therefore, to collect the data as it became available for each whole cohort at three successive time points in the trial. This could be

expected to maintain some momentum with the process (including building on relationships with our practitioner collaborators). Critically, it would also allow outcomes data to be sequentially acquired during the study when potential delays could be absorbed within the timeline, as opposed to relying on obtaining it all at the end when time pressures were likely to be more acute.

In terms of the midwives, the methodology for collaboration was on the surface simple; a series of face to face meetings. However, the evidence of others shows achieving 'buy-in' even on this level is extremely difficult (Hewison et al., 2012). Therefore, when planning and iteratively adapting the engagement activities based on progressive learning, I took into consideration factors such as location, timing, content and 'feel' of the meetings. As a result, the midwives fully engaged with the process of shaping the trial outcomes. Furthermore, they collaborated as co-authors on the protocol paper (candidate publication 4), thereby accruing evidence of personal research activity.

Attempts to collaborate with the data analysts were less successful. Despite researcher efforts to motivate individual analysts, it was hard work achieving timely provision of data transferred using specified safe processes and active participation in developing the processes. Key learning in terms of success collaborating with the data analysts was to nurture a relationship with one enthusiast, rather than expecting everyone to be interested in this role. The input of the enthusiastic data analyst who worked closely with us in the first instance was critical. She translated the trial requirements into meaningful language for data analyst colleagues which can be expected to have had a positive impact on the quality of the data. Also, we found that the consultant midwives could broker the researcher – analyst relationship, to positive effect.

It is also important to consider that while my insider knowledge, as a practitioner, may have brought some advantages; it is questionable how significant these were, in terms of credibility, working with the data analysts where I did not share a knowledge base. With these data analysts I used my practitioner skills to nurture mutually respectful relationships, but my lack of understanding of the specifics of their job did not confer the advantages I had with the midwives in terms of access and acceptability. This learning suggests that having an NHS data analyst as a member the research team, where a study relies on routine data, would be a useful approach.

4.4 COLLABORATING ON THE PRODUCTION OF A LOGIC MODEL

Introduction

Logic models in the form of diagrams depict underlying assumptions or logic of the potential mechanisms driving interventions. Making explicit in this way the theory of change mechanism of an intervention is important for informing the design and conduct of evaluations. This includes the collection of data on the possible causal pathways and the selection of appropriate outcomes (Bonell et al., 2014). Logic models include four key components: inputs, activities, outputs and outcomes. A logic model should be a dynamic tool that can be changed as needed; it is not a rigid framework that imposes restrictions on what can be done.

The production of a logic model, by evaluators, is increasingly expected at proposal and reporting stages. Best practice, in constructing a logic model, is considered to involve both intervention developers and evaluators in order that respective operational definitions are taken in to account thereby generating a common perspective (ibid). The funder for this evaluation did not specify a logic model in the protocol for the study (interestingly they now do; a reflection of evaluations of education interventions taking on the learning from those in health). Furthermore, the separation of intervention and evaluation teams (as described in chapter 3) reduced the opportunities for easy collaboration on this. As experienced evaluators, we were aware of the benefits of clarifying assumptions around the theory of change of a programme being investigated at its start. A decision was made therefore, within the research team, to create a logic model in collaboration with practitioner experts. For budgetary and time reasons, this ad hoc exercise, had to be a rapid, resource-light process. The research team decided to use a three-stage Delphi exercise, which is described below. My role was to plan and carry out the Delphi exercise.

Using a Delphi approach to construct the logic model

The Delphi approach is described as a technique for anonymous structured collaboration (Becker et al., 2009). It solicits opinions from groups in an iterative process of answering questions without individuals conferring or seeing the responses of others in the group. The consultation was carried out in this case by email. This facilitated anonymity and thus reduced the risk of stronger members dominating the group discussion. CHES was a national study with research sites spread over a wide geographical area. The Delphi approach allowed for geographical spread of participants, while minimising cost. The technique has been used for a

range of purposes, for example, providing information for service development (Bryar et al., 2013) and in a few studies for logic model construction (for example Pessoa and Nora, 2015).

The names of eight potential participants were provided by the intervention team at Chess in Schools and Communities. These individuals were all good chess players and had expertise in one or more of the following: delivering the intervention in primary schools; developing the intervention; delivering intervention training. I had not met most of the participants and was aware therefore that my approach had to be sufficiently engaging to overcome this potential barrier. With this in mind I focussed on creating high quality supporting documents for the process that were intended to be clear, engaging and credible as a means of maximising capacity and capability for undertaking the work. To illustrate this, the text of my first email, which asked for views on what the different components of the logic model should be, is shown in the Figure 2 below.

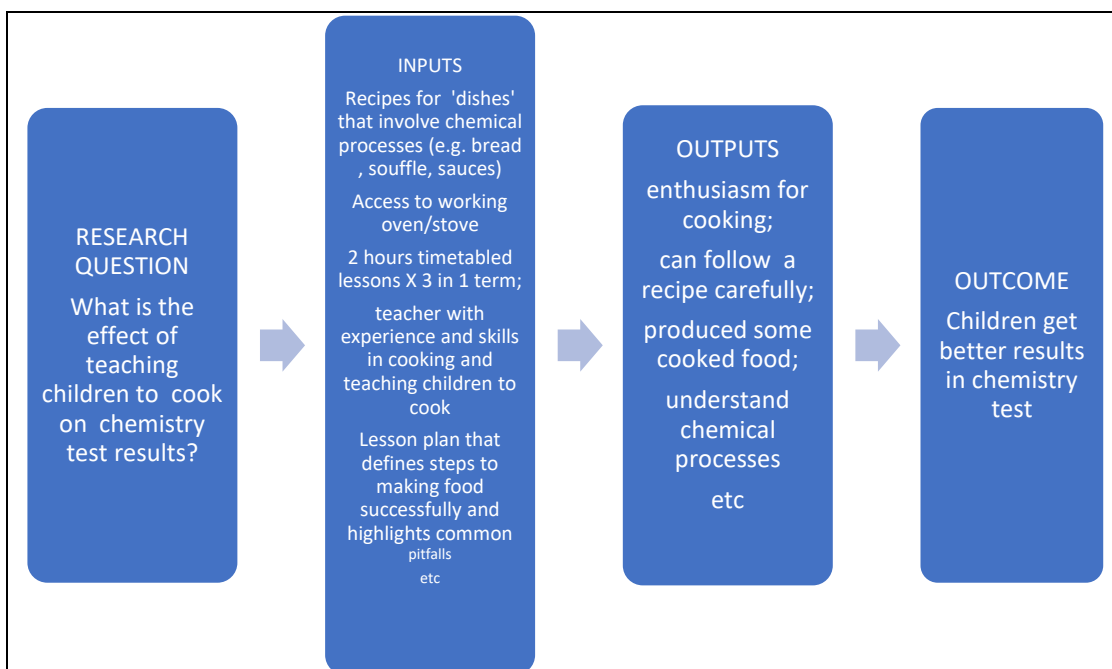
Figure 2: Sample text from Delphi process email one

The goal of this consultation exercise is to develop a “logic model” for the CSC programme in schools. A logic model graphically and simply depicts how you, who are experts in the programme, believe that the activities and resources of your programme will lead to changes that will subsequently lead to the intended outcomes. In the EEF project, these outcomes have been defined as improved Maths and English attainment (KS2 SATS).

```
graph LR; A[Taking part in CSC classroom programme] --> B[?]; B --> C[Better KS2 maths & english results]
```

What we would like you to do now

We would like you to tell us what you think the key inputs and outputs of the CSC intervention are for achieving the intended outcome of impact on Maths and English attainment. The example below for the effect of cooking on chemistry attainment is intended to clarify what we mean by inputs and outputs. We have deliberately not completed all the inputs and outputs in order not to influence your thinking too much!



For the purposes of the EEF funded study, the CSC ‘intervention’ is defined as: delivery of the CSC curriculum for one hour a week over 30 weeks to children in year 5 during normal school hours.

1. Inputs required for the CSC intervention

In no particular order, please suggest inputs (resources and activities) that you think are key components of the CSC intervention in EEF schools. If you think there are particularly important sub-components within these inputs, please detail these as well. For instance, the CSC tutor is likely to be listed as an ‘input’. If there are characteristics of the tutor that you think are key (e.g. years playing chess), then we’d be keen to know these views.

2. Outputs required for the CSC intervention

Please do the same for outputs. By outputs we mean the changes that you think occur (to children, teachers, the school) as a result of the programme.

3. The theory of change

Please tell us briefly here what your assumptions are about the change mechanism that makes the CSC programme lead to better attainment in Maths/English. How do you think the inputs you have listed lead to the outputs you have listed and how do the outputs lead to the expected outcomes?

For the second stage I consolidated the components submitted by the participants and then redistributed a list of the components to the participants by email again. The eight participants were asked to rank the listed components in order of importance. Myself and a research team colleague analysed the ranked lists that were returned and constructed a draft logic model that reflected the combined views of the experts. This draft was emailed to the participants with a request for any amendments to be provided. Thus, through a process of convergence, from the identification of common trends and analysis of outliers, a consensus was reached.

What did we achieve and learn on activity 3?

In terms of the technical requirements of the trial, a logic model was built that was acceptable to all the experts that participated in the Delphi process and that functioned as intended to support the evaluation and analysis.

Table 6: Example of impact of collaboration with chess tutors

	Example of contribution	Why is this contribution important?	Impact on research
Chess tutors	Identified potential significance of class teacher ‘modelling’ chess playing as a change mechanism	The potential significance of this would not have been identified by research team alone	Supported explanation of findings around class teacher involvement

In terms of the effectiveness of the collaborative process, seven of the eight invited experts participated fully within the requested timelines; many added notes on suggestions and insights beyond our basic requests. The Delphi mechanism created an accessible structure through which to communicate at a distance. While this approach exploited the separation of the participants it is reasonable to question how close the collaborative relationship could be when there was no face to face contact. Furthermore, arguably it denied the practitioners equal power as a group over myself and my academic colleagues who ultimately took control over how participant’s views were taken into account and responded to (Frankham, 2009). However, within the constraints of the study context, the Delphi process did provide a ‘communicative space’ that allowed divergent and contradictory views, from those with experience of the intervention, to emerge.

This said, the two-way process across the research/practitioner divide flowed in a timely and (easy) manner with the shared intent, to complete an important piece of work, appearing to facilitate the participative process. The results as shown above were transformative and feedback from some participants showed that they found this unique way of examining ‘their’ intervention interesting and maybe even empowering. It is undeniable, however, that the commitment of the Chess Tutors to this piece of work was closely linked, by them, to implied rewards this inclusion would bring to their chess tutoring work as a result of positive trial findings. The sense of a betrayal of trust I felt when delivering the ultimate null result of the trial was considerable. As a result, I have learnt that clear management of expectations about what may, or may not be found in the trial, should be part of the ‘conversation’ when inviting collaboration.

4.5 CONCLUSION

This chapter has demonstrated that it is feasible to carry out short-term collaborative activities, between researchers and practitioners, to co-produce key components central to RCTs of complex interventions. While these were relatively small pieces of work, in the context of the study as a whole, and they had limitations in terms of the level of co-production achieved, they had a transformative effect on the studies. As such, these discrete collaborations have the scope to contribute to a co-produced 'whole' in a manner that minimises risks to, and maximises benefits for, the study as well as the individuals involved. This said, the time and effort required to work collaboratively even on this small scale is not to be under-estimated. Furthermore, the analysis undertaken highlights the likely advantage of my practitioner status, knowledge and reflexive skills, for example in negotiating access and gaining acceptance and credibility. My practitioner status was, however, less influential with groups with whom I did not share a knowledge base. The analysis also reveals the tensions and discomforts that arise from the 'in-between' position I inhabited, providing opportunities for learning for improved future practice.

While on one level they may not appear particularly remarkable, these exercises all required continuity of action and interaction at the practice/research boundary and as such dealt with multiple potential barriers. As others have commented (for example, Hewison et al., 2012), this focus on the specifics of how communication was carried out by researchers with practitioners may seem prosaic, but the importance of getting this right for sustaining processes and ultimately achieving co-production goals cannot be under-estimated.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

5.1 INTRODUCTION

This final chapter focusses on the learning about key factors that influenced the nature and scale of researcher practitioner–collaboration described in this thesis. Firstly, I consider factors at the level of society and institutions and then those at individual level. The writing of this thesis has led me to realise that attributes and skills employed in my practice as a health visitor align closely to those required to support effective collaboration across the research-practice divide. To demonstrate this, this chapter continues with a brief overview of the theory and practice of health visiting, demonstrating the specific continuities between health visiting and collaborative research. I then conclude by summarising the distinct contribution this thesis makes to the knowledge in this field of study, including implications for the health visiting profession.

5.2 FACTORS FACILITATING ACADEMIC-PRACTITIONER COLLABORATION

Higher level factors

The trends in the scale and nature of practitioner involvement across the four studies, described in chapter 3, reflect the mechanisms at the societal and institutional level designed to facilitate improved conduct of applied research at the time of each study, including widening the involvement of different types of experts. As an aside, it is worth noting that the studies were all considered high quality, cutting edge examples of trials at their time. This is particularly with respect to the integral process evaluations. As such, they are likely to have shaped, as well as reflected mechanisms at this level.

At the societal level, for example, the launch of the NHS R&D programme, in 1991, provided a needs-led programme of commissioned research with a new emphasis on consumer involvement. This represented a counterbalancing of the historical reliance primarily on researchers suggesting potential research projects to funders (Peckam, 1991). Also, at this higher level, particular UK professional drivers in the fields of nursing and allied health professions can be expected to have been influential. For example, the greater orientation to

involvement by practitioners from professions relevant to the intervention, in Community REACH, compared to the earlier studies, where most roles were automatically filled by doctors. For example, various policy documents called for improved support for sustainable research and innovation with a focus on demystifying research and its importance to service innovation and improvement (see for example Willis, 2014). At the institutional level, there are also likely multiple influences that will have had an impact on the trends in involvement observed; these include factors in universities, health care organisations and the research councils funding the studies. Considering the latter briefly, in terms of the three health studies (SSFH, CASCADE and Community REACH), the NHS R and D Health Technology Assessment (HTA) Programme had little formal expectation of significant practitioner active involvement at study level. The establishment of the NIHR in 2006 brought successive requirements and initiatives (such as the research design service and the CLARHCs) to facilitate the co-production of research by staff in the health service and public health departments with academic researchers (<https://www.nihr.ac.uk>). Reflecting these trends, SSFH (funded by the HTA in 1999) had very little practitioner involvement and CASCADE (funded by the NIHR in 2008) had a large multidisciplinary team which included a significant number of practitioners and was practitioner led. By the time of the REACH Programme (and within it Community REACH), further facilitating mechanisms were in place. These included the requirement for applications to be from both an NHS and a university applicant with the funding award to be held by the NHS partner. This set the context for the expectation of a co-produced, academic-practitioner research process, and facilitated its operation through mechanisms such as secondments of NHS clinicians to research roles. Other institutional mechanisms, in place by the time of Community REACH, were specific requirements of NIHR funded study leaders as follows:

A lead applicant that is suitable to lead a programme of applied health research, as indicated through an excellent track record in this area of research. Eminence solely in a clinical area, or in more basic research, is not considered, on its own, to be sufficient (NIHR, 2012).

The context was different for the CHESSE education study, as demonstrated in chapter 3. There was little expectation by the funding body of, or resource for, practitioner involvement in the conduct of the study. Thus, while practitioners were involved in identifying the research problem, developing and implementing the intervention and endorsing the final report, the approach to research execution was not a shared one. The logic model process was the exception, but as explained in chapter 4, this was an ad hoc activity devised by the research team to introduce at least some level of collaboration. The CHESSE study was a relatively early study to be funded by the EEF. Personal experience of subsequent evaluations, funded by the

56

EEF, suggests increasing recognition of the merits of practitioner perspectives in the conduct of studies.

The combined impact of these higher level structural mechanisms, on what happened at study level, is well illustrated in the example of the site PIs. As described in chapter 3 in the CASCADE study, this key research role was occupied entirely by doctors (despite the intervention deliverers being nurses) and in Community REACH, six years later, by midwives. While the CASCADE PIs were encouraged to get actively involved (for example attending the CASCADE training) very few did. Conversely, while the REACH PIs were also not always easy to engage (due to pressure of work) for key study activities, as demonstrated, they committed time and focus. As a result, the relevant experiential knowledge influencing Community REACH, from practitioners in this key role, was much greater than that on CASCADE.

Team and individual level factors

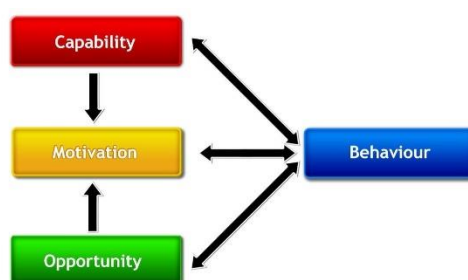
Research in general and an orientation to co-produced research knowledge, in particular, can be viewed as predominantly social (or human) processes. As such, an emphasis on interpersonal relationships and the interactions, sense-making and dialogue (Eyre et al., 2015) that occur within these relationships is required. The importance of prioritising working on relationship development and team-building has particularly been identified in recent evidence on achieving effective public involvement in research studies (Howe et al., 2017). Whereas inter-personal communication generally involves the exchange of information between individuals, the goal in this context is to encourage the kind of communication that fosters collaborative working relationships, for example encouraging joint problem solving (Zwarenstein, 2007). Furthermore, the requirement is constructive and focused dialogue that flows in both directions across the research-practice boundary (Kitson et al., 2013). Empirical evidence on factors facilitating this include: a careful and planned approach to communication (Bennett and Gadlin, 2012); establishing an appropriate communicative space (Day et al., 2016); showing humility and patience (Martin, 2010); demonstrating the value of the study and the practitioner's unique contribution to it (ibid); regular communication and continuity of participation (Stokols, 2006); being a credible messenger (Lavis et al., 2003). The underlying objective of these approaches is building mutual respect and trust which can be expected to encourage collaboration.

It can be seen that many of these facilitating factors ran through the work described in the three case studies of collaborative activities in chapter 4. For example, the carefully planned

face to face contacts with the midwives and data analysts in the Community REACH study, and the content of the email messages to the CHESS experts that, aimed to convey both the value of the task and the value of these practitioners' contribution to it. The notion of 'rigour' is generally used in research with reference to the technical aspects of the design. However, it is argued that it is in fact these social processes, and the rigour with which they are carried out, that will ultimately influence the success of the research (Greenhalgh, 2017). The approaches used to facilitate collaboration with the three different groups of practitioners, described in chapter 4, can be viewed as mechanisms for enhancing the rigour of the inter-personal work.

Our decisions to organise the work in these various ways were underpinned by assessment of potential barriers to collaboration, within the local contexts of the practitioners (Bartunek et al., 2003) and the best ways to overcome these. The COM-B model (Capability, Opportunity, Motivation ----- Behaviour; see Figure 3) within the behaviour change wheel (BCW) (Michie et al., 2011), provides a useful theoretical lens for understanding context and considering changes needed to overcome contextual barriers. It is a primarily intended to support more efficient design of effective interventions. However, it also has value in this context, as a tool to explain how the methods used did or did not engender the required behaviour, i.e. collaboration. For example, the chess tutors and midwives, were already motivated to participate in a process that held the promise of producing high quality evidence to underpin products and service changes in which they were invested. The carefully planned and executed sequential meetings with the midwives, and the Delphi process emails sent to the chess experts, capitalised on this motivation by supporting capability and offering opportunity to collaborate. As such, collaboration and co-production were achieved with these two groups.

Figure 3: The Com-B model



The situation of the data analysts was different. The data analysts could be expected to have a limited sense of ownership and obligation to participate in a process (this particular research study) to which they had little commitment, other than to do their job. This job was to meet the data access obligations of the research site. Furthermore, the notion of being co-

learners/creators in the processes of thinking around the meaning and practice of the research, appeared to be something they were not familiar with. While they were entirely capable, in the context of a heavy workload, the approaches used were unsuccessful in motivating (or 'reaching') this group effectively. However, while collective action was not achieved, the interest of one individual was captured, and even this relatively limited input had an important impact on key decisions around outcome measurement. This relative lack of engagement, suggests that even greater creativity is required by researchers in order to engage such groups. Ways to decrease the burden of the request (i.e. increase capability and opportunity) and generate a greater sense of shared ownership of research, are required. The understanding of reciprocity in collaborative work is little explored, but has relevance here (Hewitt et al., 2018).

A final reflection on the situation with the data analysts, relates to the discussion in earlier chapters, of the challenge in RCTs of combining flexible and 'soft' engagement processes whilst maintaining scientific integrity arising from independence and clear boundaries. The use of 'blinding' is central to the rigour of the RCT design (Akobeng, 2005). The data analysts, were 'blinded' as part of the testing processes in the pilot as described in chapter 4. It is possible, that the 'hard' boundary to communication of specific study information, that blinding establishes, introduced a tension that adversely affected the collaborative process.

A recent realist evaluation which aimed to address the gap in evidence on the extent to which patient and public involvement (PPI) has been embedded within health-care research, identified a number of actions that were required for effective PPI (Wilson et al., 2015). These actions were characterised by: a shared understanding of moral and methodological purposes of PPI; a key individual co-ordinating PPI; ensuring diversity (i.e. finding the right people); a research team positive about PPI input and fully engaged with it; relationships that were established and maintained over time; and PPI being evaluated in a proactive and systematic approach. While clearly not directly transferable to the practitioner context, these findings provide valuable insights for consideration by researchers and practitioners considering conducting research together.

5.3 IMPLICATIONS OF THE FINDINGS FROM THIS THESIS FOR HEALTH VISITING

Health visitors - experts in inter-personal relations, context and spanning boundaries

As stated above, writing this thesis has led to personal reflections on the transferability of health visiting and collaborative research practices, and as an extension of this, contributions that my work can make to wider debates about research capacity in health visiting. This final section of my thesis explores this further.

Health visitors in the UK, are trained nurses with additional training in community public health nursing. They provide a universal health promotion service to families with young children in order to improve health outcomes (iHV, 2012). In a review of the health visiting literature the key values, attitudes and approaches in the health visiting profession are identified (Cowley et al., 2013). An underlying 'orientation to practice' is described which includes being person-centred (human valuing) and context-sensitive (human ecology-based). It is argued that health visitors express these concepts in their routine work through three intertwined approaches to practice: relationship-development, home visiting and assessment of individual/family needs.

The skills required by health visitors for relationship-development and home visiting can particularly be seen to map closely to those required for effective collaboration between academics and practitioners. Considering relationships first, features of professional relationships with parents emerging from research on health visiting include: the enabling and mediating function these have (de la Cuesta, 1994); the importance of a professional model that recognises partnership as a central notion (Bidmead and Cowley, 2005); the sophisticated communication skills that under-pin them (Donetto, et al., 2013); and a strengths based approach that acknowledges the capacities, skills and knowledge of the client (Whittaker, 2014). The ability of health visitors to use these 'tools of the trade' to reach out to engage with those who might otherwise not take up the service, is stressed by Cowley and colleagues (2013). This reaching out component is conceptualised most obviously in the home visit. This requires the ability to function effectively 'outside the institution' (Alaszewski, 2006: 4) which requires a sensitivity to context (or place) (Poland et al., 2005). Related to this sensitivity to context is the ability to work across boundaries, most obviously physical/geographical boundaries, but also others such as power differentials. The impact of good health visiting is demonstrated in evidence from parents. This reveals that it makes them feel 'known',

respected and listened to, increasing their willingness to engage with the service, with an associated potential impact on maternal and child health outcomes (Donetto et al., 2013).

It is interesting to note that Cowley and colleagues (2013) conclude, that while individual elements of health visiting are generic, in that other practitioner groups provide them, what is unique about health visitors is the overall approach. Practitioner knowledge is described by Titchen (2000) as 'professional craft knowledge' or 'practical know-how'. Clearly it cannot be assumed that the craft knowledge of health visitors, that appears to work favourably in the context of routine health visiting practice, can simply be transferred to another, the research environment, with the same effect. However, the potential continuities exist and more exploration of this apparent overlap of health visiting and collaborative research epistemologies would be useful.

Building research capacity in health visiting

Despite the importance attached to a sound scientific base to inform policy and practice of all public services, with effectiveness studies using an RCT design as the reference standard for this, the level of research activity, research experience and research quality, in many of the allied health professions, is relatively limited (Cooke, 2005). This certainly applies to health visiting. For example, Robinson (2017 p. 30) explains: '*health visiting is rarely seen as a valid subject either for scientific research or for practice narratives*' and Cowley et al., (2013; p.20) conclude from their extensive review of the literature: *research [about health visiting] is characterised by small-scale, single studies, (i.e., one-off, not part of a programme of research), often under-theorised or forming part of masters or doctoral work that is not then followed up to create a convincing body of work.* A rare systematic review of health visiting was conducted almost two decades ago (Elkan, 2000), and candidate publication 1 for this thesis, is one of a relatively few RCTs that is directly focussed on health visiting in the UK.

Reasons for this 'research immaturity' are likely to include the well documented problem that the role of the health visitor is poorly understood, or agreed on, outside of the profession (Baldwin, 2013). Peckover (2009) relates this partially to the invisibility that arises from the home visiting dimension as well as the professions' struggle to establish itself as a discrete discipline, separate from nursing, medicine and social work. Others link it to the known challenges of measuring effectiveness of public health interventions (Luker and McHugh, 2017). A cause, and effect, of this low level of research capacity, is the limited academic

infrastructure for health visiting (Cowley et al., 2013) with the historical challenges faced by nurse-academics, in general, described by Bryar et al., (2013).

The need for specific support and action to enable health visiting to develop from this ‘*very low base*’ is advocated (Cowley et al., 2013; p.23). More RCTs directly related to health visiting will be an important part of this development. Evidence on effectiveness is particularly vital in the current political climate where, as Luker and McHugh state (2017; p. 314), value for money in terms of proven health gain is the currency of the marketplace in which the service exists. The Institute of Health visiting (iHV), launched in 2012, has a core purpose of raising professional standards in health visiting practice by ‘*promoting and supporting a strong evidence base for health visiting*’ (iHV, 2012). A recent acceleration in the process of building research capacity in health visiting is evidenced by a joint iHV/NIHR conference held in January 2018, which aimed to be ‘*very focused on the research process and how to embrace it*’ (iHV email communication, 15.12.17). This thesis has demonstrated that although health visitors are relatively new to RCTs, skills central to routine health visiting practice, such as interpersonal communication, sensitivity to context and spanning boundaries are particularly well suited to conducting the collaborative work required to maximise the rigour and relevance of the RCT process.

5.4 CONCLUSIONS

The distinct contribution of this thesis.

This final section presents key findings, learning for future design and conduct of RCTs and potential future research directions arising from this work for each of the four research questions (stated in Chapter 1) that this thesis aimed to address.

Finding 1: (research question 1)

What do we know about collaboration between academics and practitioners in the co-production of research, and in RCTs of complex interventions specifically?

Key finding: A general trend across time is shown, in the four publications, of: more practitioners; in increasingly varied and influential study roles; with more capacity for involvement; and being more representative of future users of the research. Furthermore, over time, management of the intervention/ evaluation boundary is adapted iteratively through the study, with explicit acknowledgement of this made. These two trends have


increased the potential for a more equal and effective blend of academic and practitioner knowledge and as such the production of more useful research (see summary in figure 3).

Learning for future trial design. It should be ensured, from the start of a study, that the right mix of individuals are in place, throughout all the study structures, in order that the research process is a truly blended academic/practitioner enterprise. Furthermore, all members of the team should be fully engaged with prioritising co-production.

Research directions: What additional forms of support are required for practitioners to be actively involved in RCTs of complex interventions? What are the distinctive requirements of leaders of these studies? What are the social processes within study teams that facilitate effective collaboration?

Figure 3: Key characteristics of academic/practitioner influence at different levels of decision making over time in relation to RCTs.

Structures and institutions (macro level)		
<i>Academic independence</i>		<i>Needs led</i>
Research team and individual study level – framing research question, make up of team		
<i>Academics frame RQ</i>		<i>Academic/practitioner framing of RQ</i>
<i>Academics/doctors on teams</i>		<i>Blended teams, include allied health practitioners</i>
Research team and individual study level – formative work, boundary management		
<i>Limited formative work</i>	<i>Basic formative work</i>	<i>Fully integrated formative work</i>
<i>Boundary management not acknowledged</i>	<i>Boundary management acknowledged, not adapted</i>	<i>Boundary management fully acknowledged and adapted</i>
Individual (micro level)		
<i>Technical research skills</i>		<i>Technical research skills + collaborative research skills</i>
Overall trend over time		
<i>Focus on scientific rigour</i>		<i>Balance of scientific/practice relevance</i>
<i>1998</i>		<i>2018</i>



Finding 2: (research question 2)

How has the nature of practitioner active involvement in the conduct of evaluations of complex interventions using a RCT design changed over that last two decades? How does it differ between health and education research?

Key finding: Comparison of an evaluation of an education intervention with a study in the health sector, at an equivalent time point, showed that practitioner involvement in the former was relatively limited compared with the latter. However, there is a suggestion of more recent trends towards greater researcher-practitioner collaboration in education research.

Learning for future trial design and research directions: The comparison was encouraging in that it highlighted the progress made in this field in health research. With the apparent progress now in education research, there is likely to be value in more extensive cross domain research comparing collaborative methods in the conduct of RCTs education and health RCTs.

Finding 3: (research question 3)

How can the voices of harder to reach practitioners be brought into the process of shaping decisions about key elements of a RCT?

Key finding: Examination of potential causal links between research processes and study outcomes, in completed studies, reveal that specific key practitioner voices are often missing from the research process in RCTs. These practitioners are from groups that are seldom heard in the conduct of research and 'harder to reach'. The absence of this specific practitioner knowledge, from key decision-making processes in the trial, is likely to have had a negative impact on the utility of the findings of the studies. This thesis has demonstrated that it is possible to conduct small, short-term but intensive pieces of collaborative work with practitioner groups as part of the conduct of RCTs. Furthermore, this work can be on the core components of the trial process and as such has the potential to be transformative to the trial. While traditionally the definition of a collaborative relationship has assumed a sustained process, the key criterion is active involvement in decision making processes, even if this is short-term. These small, creative pieces of work are not a substitute for, but complementary to, the involvement of practitioners at the centre of the whole study process.

Learning for future trial design. This finding highlights the importance of study teams searching for, listening to and valuing this knowledge; taking an adaptive and iterative approach to allowing it to impact on the research process, during the study.

Research directions: How can the practitioners whose knowledge is important to the success of an RCT be identified early in the research process? What incentives are most effective in encouraging those groups least able or willing to get involved as collaborators? How can the theory and practice of boundary maintenance work in RCTs be developed?

Finding 4: (research question 4)

What is the learning from these analyses, how does this advance knowledge in the field of study, where are health visitors situated in the on-going process of improving the knowledge production process through academic-practitioner collaboration on RCTs?

Key finding: The skills central to effective collaboration, in research, are those used by health visitors in their routine service delivery role. At a time when health visiting leaders are prioritising developing research capacity this thesis has demonstrated the scope for health visitors not only to lead as subject experts but as experts in collaborative methods.

Learning for future trial design. This expertise should be recognised by those responsible for research capacity building in relation to the health visiting profession and by those identifying the ‘right’ practitioners to be members of research collaborations.

Research directions; What additional incentives will facilitate research involvement, by health visitors, at the individual and organisational levels? What additional skills do health visitors need in order to participate in research and in particular in large research collaborations conducting RCTs? What more can be learnt about health visitors as researchers by exploring further the arguments about continuities and discontinuities in this thesis?

Taken together the above findings show that practitioners can be involved as collaborators for different reasons, at different stages and to different degrees, in RCTs. Approaches to enhance practitioner collaboration therefore need to be tailored to suit different types of: research and research questions, decisions within studies; groups of practitioners; individual practitioners with varying capacity. Collaboration should be viewed as a multi-faceted approach with a range of different creative mechanisms employed to bring different voices into the knowledge production process.

Strengths and limitations of this thesis

This thesis has an unusual methodology – a longitudinal, retrospective, reflective approach, conducted by an ‘insider’ (myself) who inhabits both academic researcher and clinical practitioner worlds. These various dimensions of closeness, over a twenty-year period working at the cutting edge of this field of research, have allowed a unique view of the development of academic-practitioner collaboration.

There are of course limitations, in terms of lack of objectivity, from this ‘closeness’. Furthermore, the historical element of this thesis requires reflections and recollections of processes that occurred up to 20 years ago. As such, I have been reliant on memory which is dependent on social context, and subject to change over time (Oakley, 2017). While my personal recall has been checked and challenged through the methodology employed for this thesis the findings should be viewed in this context.

I am also aware that in most of the work described, even where I claim improvements in the direction of travel in terms of collaboration, academics retained responsibility for the process of enquiry and had the ultimate control over the final decision making. Linked to this is my recognition of the fact that while coming from both academic and practitioner worlds myself, my positioning in the work described is in the former not the latter. While I have striven to use my 'hybrid' status to achieve balance, I appreciate that I risk perpetuating the traditions of academic superiority; assuming the transfer of knowledge, from the research community, to that of practitioners for the latter to learn from.

Conclusion

This PhD offers observations, reflections and analyses from a personal journey starting with my entry as a practitioner to an academic research setting, through my increasing experience as a researcher in an academic setting while still in practice as a health visitor. On this journey I learnt that greater researcher-practitioner collaboration in RCTs creates better knowledge and I developed approaches to support the process of blending different types of knowledge required. In particular I have focussed on ways to address inequities of whose voice is heard in the conduct of RCTs, in the knowledge that this impacts on the utility of the findings. It is hoped that this work will encourage openness to, and creativity in, collaboration between academic researchers and practitioners in evaluations of complex social interventions using an RCT design. It has also been an intention to provide greater understanding of practices involved which others might find helpful in preparing for, conducting and assessing collaborative work. It is my particular hope that my findings will support the capacity for more effectiveness research about health visiting and by health visitors.

REFERENCES

- Alaszewski, A. (2006). Managing risk in community practice: nursing, risk and decision-making. in: Godin, P. ed. *Risk and Nursing Practice*. United Kingdom: Palgrave Macmillan, pp. 24-41. Available here: <https://kar.kent.ac.uk/3578/1/godinch2Andyjune.pdf> [Last accessed 3.1.18].
- Akobeng A. (2005). Understanding randomised controlled trials. *Archives of Disease in Childhood*; **90**:840-844.
- Armstrong, A. and Alsop, A. (2010). Debate: Co-production can contribute to research impact in the social sciences. *Public Money & Management; Volume 30, 2010 - Issue 4 pp 208-10*.
- Baldwin, S. (2013). Advancing professional practice: Why is it necessary for health visiting?. *Journal of Health Visiting*, 1(8), pp.468-473.
- Bartunek, J., Trullen, J., Bonet, E. and Sauquet, A. (2003). Sharing and expanding academic and practitioner knowledge in health care. *Journal of Health Services Research & Policy*, 8 (2_suppl), pp.62-68.
- Bauer, M., Damschroder, L., Hagedorn, H., Smith, J., and Kilbourne, A. An introduction to Implementation Science for the non-specialist. *BMC Psychology* 3 (32) pp 1-12.
- Baumbusch, J., Kirkham, S., Khan, K., McDonald, H., Semeniuk, P., Tan, E. and Anderson, J. (2008). Pursuing common agendas: A collaborative model for knowledge translation between research and practice in clinical settings. *Research in Nursing & Health*, 31(2), pp.130-140.
- Becker, G. and Roberts, T. (2009). Do We Agree? Using a Delphi Technique to Develop Consensus on Skills of Hand Expression. *Journal of Human Lactation*, 25(2), pp.220-225.
- Bell, M., Cordingley, P., Isham., C. & Davis., R. (2010). *Report of Professional Practitioner Use of Research Review: Practitioner engagement in and/or with research*. Coventry: CUREE, GTCE, LSIS & NTRP. Available here: <http://www.curee.co.uk/node/2303>. [Last accessed 03.02.18]
- Benchimol, E., Smeeth, L., Guttman, A., Harron, K., Moher, D., Petersen, I., Sørensen, H., von Elm, E. and Langan, S. (2015). The Reporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Medicine*, 12(10), p.e1001885.
- Bennett, L. and Gadlin, H. (2012). Collaboration and Team Science. *Journal of Investigative Medicine*, 60(5), pp.768-775.
- Bidmead, C. and Cowley, S. (2005). A concept analysis of partnership with clients. *Community Practitioner* 78: 6, 203–8.
- Bonell, C., Fletcher, A., Morton, M., Lorenc, T. and Moore, L. (2012). Realist randomised controlled trials: A new approach to evaluating complex public health interventions. *Social Science & Medicine*, 75(12), pp.2299-2306.
- Bonell, C., Jamal, F., Melendez-Torres, G. and Cummins, S. (2014). 'Dark logic': theorising the harmful consequences of public health interventions. *Journal of Epidemiology and Community Health*, 69(1), pp.95-98.

Bornbaum, C., Kornas, K., Peirson, L. and Rosella, L. (2015). Exploring the function and effectiveness of knowledge brokers as facilitators of knowledge translation in health-related settings: a systematic review and thematic analysis. *Implementation Science*, 10 (1).

Bowen, S. and Martens, P. (2005). Demystifying knowledge translation: learning from the community. *Journal of Health Services Research & Policy*, 10(4), pp.203-211.

Bryar, R., Anto-Awuakye, S., Christie, J., Davis, C. & Plumb, K. (2013). Using the delphi approach to identify priority areas for health visiting practice in an area of deprivation. *Nursing Research and Practice*, 2013, 780315.

Bryar, R., Closs, S., Kendall, S. and Griffiths, J. (2013). Comment on: Thompson D.R. & Darbyshire P. (2013) Is academic nursing being sabotaged by its own killer elite? *Journal of Advanced Nursing* 69 (1), 1-3. *Journal of Advanced Nursing*, 69(5), pp.1210-1211.

Campbell, H. and Vanderhoven, D. (2016) *Knowledge That Matters: Realising the Potential of Co-Production*. N8/ESRC Research Programme. Available from: <https://www.n8research.org.uk/media/Final-Report-Co-Production-2016-01-20> [Last accessed 05.09.17]

Christie, D., Romano, G. M., Thompson, R., Viner, R. M., & Hindmarsh, P. C. (2008). Attitudes to psychological groups in a paediatric and adolescent diabetes service - implications for service delivery. *Pediatric Diabetes*, 9(4), pp 388-392.

Christie D, Thompson R, Sawtell M, Allen E, Cairns J, Smith F, et al. *Structured, intensive education maximising engagement, motivation and long-term change for children and young people with diabetes: a cluster randomised controlled trial with integral process and economic evaluation – the CASCADE study*. *Health Technology Assessment* 2014; 18(20).

Cooke, J. (2005). A framework to evaluate research capacity building in health care. *BMC Family Practice*, 6(1).

Cooke, J., Nancarrow, S., Dyas, J. and Williams, M. (2008). An evaluation of the 'Designated Research Team' approach to building research capacity in primary care. *BMC Family Practice*, 9(1).

Connolly, P. (2017). *Using randomised controlled trials in education*; BERA/SAGE Research Methods in Education. London: SAGE Publications.

COWLEY, S. & BIDMEAD, C. (2009) Controversial questions (part three): is there randomised controlled trial evidence for health visiting? *Community Practitioner*, 82, 24-28.

Cowley, S., Whittaker, K., Grigulis, A., Malone, M., Donetto, S., Wood, H., Morrow, E. & Maben, J. (2013). *Why health visiting? A review of the literature about key health visitor interventions, processes and outcomes for children and families*. (Department of Health Policy Research Programme, ref 016 0058). National Nursing Research Unit, Florence Nightingale School of Nursing and Midwifery, King's College London.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I. and Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*, p.a1655.

Cruz, H., Gillingham, P. and Melendez S. (2007). [Reflexivity: A Concept and its Meanings for Practitioners Working with Children and Families](#). *British Journal of Social Work* 37, pp. 73–90

Datta, J. and Petticrew, M. (2013). Challenges to evaluating complex interventions: a content analysis of published papers. *BMC Public Health*, 13(1).

Day, K., Kenealy, T. and Sheridan, N. (2016). Should we embed randomized controlled trials within action research: arguing from a case study of telemonitoring. *BMC Medical Research Methodology*, 16(1).

De La Cuesta, C. (1994). Relationships in health visiting: enabling and mediating. *International Journal of Nursing Studies*, 31(5), pp.451-459.

Dechartres, A., Trinquart L, Atal I., Moher D, Dickersin K., Boutron I., Perrodeau E., Altman DG, Ravaud P. (2017). Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study. *BMJ*, p.j 3806.

Donetto, S., Malone, M., Hughes, J., Morrow, E., Cowley, S. and Maben, J. (2013). *Health Visiting: the voice of the service users. Learning from service users' experiences to inform the development of UK health visiting practice and services*. National Nursing Research Unit, King's College London [online] [Last accessed: 1 November 2017].

Elkan, R., Kendrick, D., Hewitt, M., Robinson, J. et al. (2000). The effectiveness of domiciliary health visiting: a systematic review of international studies and a selective review of the British literature. *Health Technology Assessment*, 4 (13). i-v, 1-339.

Eyre, L., George, B. and Marshall, M. (2015). Protocol for a process-oriented qualitative evaluation of the Waltham Forest and East London Collaborative (WELC) integrated care pioneer programme using the Researcher-in-Residence model. *BMJ Open*, 5(11), p.e 009567.

Filipe, A., Renedo, A. and Marston, C. (2017). The co-production of what? Knowledge, values, and social relations in health care. *PLOS Biology*, 15(5), p.e2001403.

Flinders, M., Wood, M. and Cunningham, M. (2016). The politics of co-production: risks, limits and pollution. *Evidence & Policy: A Journal of Research, Debate and Practice*, 12(2), pp.261-279.

Frankham, J. (2009). *Partnership research: a review of approaches and challenges in conducting research in partnership with service users*. Discussion Paper. NCRM. (Unpublished). Available at: <http://eprints.ncrm.ac.uk/778/> [Last accessed 15.11.17].

Greene, M. J. (2014). *On the Inside Looking In: Methodological Insights and Challenges in Conducting Qualitative Insider Research*. *The Qualitative Report*, 19(29), 1-13. Available here: <http://nsuworks.nova.edu/tqr/vol19/iss29/3> [Last accessed 03.01.18].

Greenhalgh (2017) *Report on a research conference on June 13th and 14th 2017, Green Templeton College, Oxford*. Available here: <http://www.gtc.ox.ac.uk/whats-on-calendar/lectures-seminars/complexity-workshop.html> [Last accessed 03.01.18].

Goldacre, B. (2013). Building Evidence into Education. Available here: <https://www.gov.uk/government/news/building-evidence-into-education> [Last accessed 15.11.2017].

Hammersley, M. (2005). Is the evidence-based practice movement doing more harm than good? Reflections on Ian Chalmer's case for research-based policy making and practice. *Evidence and Policy*, 1 (1), pp 85-100.

Harman, N., Conroy, E., Lewis, S., Murray, G., Norrie, J., Sydes, M., Lane, J., Altman, D., Baigent, C., Bliss, J., Campbell, M., Elbourne, D., Evans, S., Sandercock, P. and Gamble, C. (2015). Exploring the role and function of trial steering committees: results of an expert panel meeting. *Trials*, 16(1).

Heale R and Twycross A, 2018. What is a case study? *Evidence-Based Nursing* 2018; **21**: pp 7-8.

Health Education England/NIHR(<https://www.nihr.ac.uk/funding-and-support/documents/ICA/TCC-ICA-Eligible-Professions-and-Registration-Bodies.pdf>). [Last accessed 18 Feb 2019]

Heaton, J., Day, J. and Britten, N. (2015). Collaborative research and the co-production of knowledge for practice: an illustrative case study. *Implementation Science*, 11(1).

Hegney D, and Francis K. (2015). Action research: changing nursing practice. *Nursing Standard*. Jun 3;29(40):36-41.

Hewison, A., Gale, N. & Shapiro, J. (2012). Co-production in research: some reflections on the experience of engaging practitioners in health research. *Public Money & Management*, vol 32, no. 4, pp. 297-302. DOI: 10.1080/09540962.2012.691311.

Heneghan, C., Mahtani, K., Goldacre, B., Godlee, F., Macdonald, H. and Jarvies, D. (2017). Evidence based medicine manifesto for better healthcare. *BMJ*, p.j2973.

Hewitt, G., Roberts, J., Fletcher, A., Moore, G. and Murphy, S. (2018). Improving young people's health and well-being through a school health research network: Reflections on school-researcher engagement at the national level. *Research for All*, 2(1), pp.16-33.

Howe, A., Mathie, E., Munday, D., Cowe, M., Goodman, C., Keenan, J., Kendall, S., Poland, F., Staniszewska, S. and Wilson, P. (2017) Learning to work together - lessons from a reflective analysis of a research project on public involvement. *Research Involvement and Engagement*, 3 (1). pp. 1-12.

Ikioda, F. and Kendall, S. (2016). Transformation of health visiting services in England using an online community of practice. *Health Policy and Technology* 5, pp 298-306.

iHV (2012). *What is a health visitor?* Institute of Health Visiting. <https://ihv.org.uk/> [Last accessed 12.02.18].

iHV (2017). 'Do you aspire to do research?' iHV email communication, 15.12.17.

iHV (2018). *Health visiting and the NHS in the next 10 years*. Institute of Health Visiting. <https://ihv.org.uk/news-and-views/news/ihv-publishes-its-10-year-plan-for-health-visiting/> [Last accessed 16.02.19]

INVOLVE. <https://www.invo.org.uk>. [Last accessed 10 February 2019].

Jerrim, J., Macmillan, L., Micklewright, J., Sawtell, M. and Wiggins, M. (2017). Does teaching children how to play cognitively demanding games improve their educational attainment? Evidence from a Randomised Controlled Trial of chess instruction in England. *Journal of Human Resources*, pp.0516-7952R.

Kislov, R., Hodgson D, Boaden R. (2016). Professionals as knowledge brokers: The limits of authority in healthcare collaboration. *Public Administration* 94 (2), 472-489.

Krebbekx, W., Harting, J. and Stronks, K. (2012). Does collaborative research enhance the integration of research, policy and practice? The case of the Dutch Health Broker Partnership. *Journal of Health Services Research & Policy*, 17(4), pp 219-229.

Kunhunny, S. and Salmon, D. (2017). The evolving professional identity of the clinical research nurse: A qualitative exploration. *Journal of Clinical Nursing*, 26(23-24), pp.5121-5132.

Lavis, J. N., Robertson, D., Woodside, J. M., McLeod, C. B., & Abelson, J. (2003). How can research organizations more effectively transfer research knowledge to decision makers? *Milbank Quarterly*, 81(2), 221-222.

Litherland, R (2012). The health visitors role in the identification of domestic abuse. *Community Practitioner*, 85 (8), 20-3

Luker, K. and McHugh, G. (2017). Evaluating practice. *In Health Visiting: Preparation for Practice* (fourth edition; pp 285-385); Luker, K., McHugh, G. and Bryar, R (eds). John Wiley & Sons, Inc., Hoboken, NJ, USA.

Lunt, N., Shaw, I. and Fouché, C. (2010). Practitioner research: collaboration and knowledge production. *Public Money & Management*, 30(4), pp.235-242.

Madden, M. and Morley, R. (2016). Exploring the challenge of health research priority setting in partnership: reflections on the methodology used by the James Lind Alliance Pressure Ulcer Priority Setting Partnership. *Research Involvement and Engagement*, 2(1).

Marshall, M., Pagel, C., French, C., Utlely, M., Allwood, D., Fulop, N., Pope, C., Banks, V. and Goldmann, A. (2014). Moving improvement research closer to practice: The Researcher-in-Residence model: Table 1. *BMJ Quality & Safety*, 23(10), pp.801-805.

Martin, S. (2010). Co-production of social research: strategies for engaged scholarship. *Public Money & Management*, 30(4), pp.211-218.

Matus, J., Walker, A., Mickan, S. (2018). Research capacity building frameworks for allied health professionals – a systematic review. *BMC Health Services Research* 18(716) pp1 -11.

McCabe, Angela C., Parker, Rachel L., & Cox, Stephen (2016). The ceiling to coproduction in university-industry research collaboration. *Higher Education Research & Development*, 35(3), pp. 560-574.

Michie, S., van Stralen, M. and West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6(1).

Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D. and Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, 350(mar 19 6), pp.h1258-h1258.

MRC (2017); MRC guidelines for management of global health trials Involving Clinical or Public Health Interventions; MRC Clinical Trials Series. <https://www.nihr.ac.uk>. [last accessed 10 Feb 2019].

National Institute Health Research (NIHR) (2012). *Programme Grants for Applied Research: six competitions, 100 awards*. Available here: <https://www.nihr.ac.uk/funding-and-support/documents/funding-for-research-studies/research-programmes/PGfAR/CCF-PGfAR-ProgrammeReview.pdf>.

NHS England, Public Health England, Health Education England, Monitor, Care Quality Commission, NHS Trust Development Authority (2014). Five Year Forward View. Available here: <http://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf>. [Last accessed 30 Oct 2017]

NIHR/ Allied Health Professionals. Integrated Clinical Academic Programme. Available here: <https://www.nihr.ac.uk/our-research-community/clinical-research-staff/allied-health-professionals.htm>. [Last accessed 18 Feb 2019]

Newman, J., Cherney, A. and Head, B. (2015). Do Policy Makers Use Academic Research? Reexamining the "Two Communities" Theory of Research Utilization. *Public Administration Review*, 76(1), pp.24-32.

Nutley, S 2010, 'Debate: Are we all co-producers of research now?' *Public Money & Management*, vol 30, no. 5, pp. 263-265.

OAKLEY, A. (1998). Gender, methodology and people's ways of knowing: some problems with feminism and the paradigm debate in social science. *Sociology*, 32(4), pp.707-731.

Oakley, A. (2005). *Experiments in knowing*. Cambridge: Polity Press.

Oakley, A. (2017). *A small sociology of maternal memory: an interview with Ann Oakley. Monday 13th November, 2017*. Available here: <https://www.thesociologicalreview.com/blog/a-small-sociology-of-maternal-memory-an-interview-with-ann-oakley.html>

Oakley, A., Wiggins, M., Strange, V., Sawtell, M., & Austerberry, H. (2011). Becoming a Mother: continuities and discontinuities over three decades. In E. F. J. M, R. MPM, H. J (Eds.), *Birth Rites and Rights* (pp. 9-27). Oxford: Hart Publishing.

O'Cathain, A., Murphy, E. and Nicholl, J. (2008). Multidisciplinary, Interdisciplinary, or Dysfunctional? Team Working in Mixed-Methods Research. *Qualitative Health Research*, 18(11), pp.1574-1585.

Orr, K. and Bennett, M. (2009). Reflexivity in the co-production of academic-practitioner research", *Qualitative Research in Organizations and Management: An International Journal*, Vol. 4 Issue: 1, pp.85-102, <https://doi.org/10.1108/17465640910951462>

Ostrom, E. (1996). Crossing the great divide: Coproduction, synergy, and development. *World Development*;24: 1073–1087. doi: [10.1016/0305-750X\(96\)00023-X](https://doi.org/10.1016/0305-750X(96)00023-X)

Pearce, W. and Raman, S. (2014). The new randomised controlled trials (RCT) movement in public policy: challenges of epistemic governance. *Policy Sciences*, 47(4), pp.387-402.

Peckham M. (1991). Research and development for the National Health Service. *Lancet*; 338:367–71.

Peckover, S. (2011). From 'public health' to 'safeguarding children': British health visiting in policy, practice and research. *Children & Society*, 27(2), pp.116-126.

Pfadenhauer, L., Gerhardus, A., Mozygemba, K., Lysdahl, K., Booth, A., Hofmann, B., Wahlster, P., Polus, S., Burns, J., Brereton, L. and Rehfues, E. (2017). Making sense of complexity in context and implementation: The Context and Implementation of Complex Interventions (CICI) framework. *Implementation Science*, 12(1).

Poland, B., Lehoux, P., Holmes, D. & Andrews, G. (2005). How place matters: unpacking technology and power in health and social care. *Health and Social Care in the Community*, 13(2), pp 170–80.

Råheim M., Magnussen L., Sekse R., Lunde A., Jacobsen T and Blystad, A. (2016). Researcher–researched relationship in qualitative research: Shifts in positions and researcher vulnerability. *International Journal Qualitative Studies in Health Well-being*. 11: 30996 - <http://dx.doi.org/10.3402/qhw.v11.30996>

Read, S., Wild, S. and Lewis, S. (2013). Applying missing data methods to routine data: a prospective, population-based register of people with diabetes. *Trials*, 14(Suppl 1), p.P113.

Reeve, J., Cooper, L., Harrington, S., Rosbottom, P. and Watkins, J. (2016). Developing, delivering and evaluating primary mental health care: the co-production of a new complex intervention. *BMC Health Services Research*, 16(1).

Richards, D., Coulthard, V. and Borglin, G. (2014). The State of European Nursing Research: Dead, Alive, or Chronically Diseased? A Systematic Literature Review. *Worldviews on Evidence-Based Nursing*, 11(3), pp.147-155.

Robinson, K. (2017). Managing knowledge in health visiting. In *Health Visiting: Preparation for Practice* (fourth edition; PP 33-72). Luker, K., McHugh, G. and Bryar, R (eds); John Wiley & Sons, Inc., Hoboken, NJ, USA.

Sawtell, M., Austerberry, H., Ingold, A., Strange, V., Stevens, M., & Wiggins, M. (2009). Evaluation of the Teenage Health Demonstration Sites Programme: Final Report. London: Social Science Research Unit, Institute of Education.

Sawtell, M., Ingold, A., Wiggins, M., & Austerberry, H. (2010). *Innovative services to improve the health and well-being of vulnerable young people*. London: Social Science Research Unit, Institute of Education.

- Spasoff, R. (1999). *Epidemiologic Methods for Health Policy*. New York: Oxford University Press, Inc.
- Staa, T., Goldacre, B., Gulliford, M., Cassell, J., Pirmohamed, M., Taweel, A., Delaney, B. and Smeeth, L. (2012). Pragmatic randomised trials using routine electronic health records: putting them to the test. *BMJ*, 344(feb07 1), pp.e55-e55.
- Stainton, M., Harvey, S., McNeil, D., Emmanuel, E., Johnson, C. (1998). The transforming power of research on practice: involving practitioners in analysis. *International Journal of Nursing Practice* 4 (4), pp. 220–224.
- Stokols, D. (2006). Toward a Science of Transdisciplinary Action Research. *American Journal of Community Psychology*, 38(1-2), pp.79-93.
- Titchen, A. (2000). *Professional Craft Knowledge in Patient-Centred Nursing and the Facilitation of Its Development*. D.Phil. Linacre College Oxford. Oxfordshire: Ashdale Press Tackley.
- Tooke J. (2008). *Aspiring to excellence: final report of the independent inquiry into Modernising Medical Careers*. London: MMC Inquiry. Available here: www.mmcinquiry.org.uk. [Last accessed 01/12/17].
- Torgerson, C., Torgerson, D., Birks, Y. and Hall, J. (2005). A comparison of randomised controlled trials in health and education' *British Educational Research Journal*, vol 31, no. 6, pp. 761-785.
- Treweek, S., Altman, D., Bower, P., Campbell, M., Chalmers, I., Cotton, S., Craig, P., Crosby, D., Davidson, P., Devane, D., Duley, L., Dunn, J., Elbourne, D., Farrell, B., Gamble, C., Gillies, K., Hood, K., Lang, T., Littleford, R., Loudon, K., McDonald, A., McPherson, G., Nelson, A., Norrie, J., Ramsay, C., Sandercock, P., Shanahan, D., Summerskill, W., Sydes, M., Williamson, P. and Clarke, M. (2015). Making randomised trials more efficient: report of the first meeting to discuss the Trial Forge platform. *Trials*, 16(1).
- Trostle, J. (1992). Research capacity building in international health: Definitions, evaluations and strategies for success. *Social Science & Medicine*, 35(11), pp.1321-1324.
- Tunis, S., Stryer, D. and Clancy, C. (2003). Practical Clinical Trials. *JAMA*, 290(12).
- Verhoef, M., Mulkins, A., Kania, A., Findlay-Reece, B. and Mior, S. (2010). Identifying the barriers to conducting outcomes research in integrative health care clinic settings - a qualitative study. *BMC Health Services Research*, 10(1).
- Walker, D. (2010). Debate: Do academics Know better or merely different? *Public Money & Management*. 30:4, pp. 204-206.
- Whittaker, K. (2014). Supporting parents and parenting practices: The health visiting context. *Journal of Health Visiting*, 2(5), pp.250-259.
- Willis (2014). Shape of caring (2014); Theme 7. Available here: <https://www.hee.nhs.uk/sites/default/files/documents/2348-Shape-of-caring-review-FINAL.pdf>

- Wiggins, M. (2005). Postnatal support for mothers living in disadvantaged inner city areas: a randomised controlled trial. *Journal of Epidemiology & Community Health*, 59(4), pp.288-295.
- Wiggins, M., Bonell, C., Sawtell, M., Austerberry, H., Burchett, H., Allen, E., Strange, V. (2009). Health outcomes of youth development programme in England: Prospective matched comparison study *British Medical Journal* **339**: Jul.
- Wiggins, M., Oakley, A., Roberts, I., Turner, H., Rajan, L., Austerberry, H., Mujica, R. and Mugford, M. (2004). The Social Support and Family Health Study: a randomised controlled trial and economic evaluation of two alternative forms of postnatal support for mothers living in disadvantaged inner-city areas. *Health Technology Assessment*, 8(32).
- Wiggins, M., Sawtell, M., Jerrim, J. (2017). *Learner Response System Evaluation report and executive summary*. Available here:
https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Learner_Response_System.pdf
- Williamson, P., Altman, D., Blazeby, J., Clarke, M., Devane, D., Gargon, E. and Tugwell, P. (2012). Developing core outcome sets for clinical trials: issues to consider. *Trials*, 13(1).
- Wilson, P., Mathie E., Keenan J., McNeilly E., Goodman C., Howe A., [Poland F](#), [Staniszewska S.](#), [Kendall S.](#), [Munday D.](#), [Cowe M.](#) & [Peckham S.](#). ReseArch with Patient and Public involvement: a Realist evaluation the RAPPORt study. *Health Service Delivery Research* 2015;3(38)
- Wolff, N. (2000). Using randomized controlled trials to evaluate socially complex services: problems, challenges and recommendations. *The Journal of Mental Health Policy and Economics*, 3(2), pp.97-109.
- Zwarenstein, M., Reeves, S., Russell, A., Kenaszchuk, C., Conn, L., Miller, K., Lingard, L. and Thorpe, K. (2007). Structuring communication relationships for interprofessional teamwork (SCRIPT): a cluster randomized controlled trial. *Trials*, 8(1).

APPENDIX 1

CANDIDATE PUBLICATION 1

Candidate publication 1 - Sawtell, M and Jones, C (2002). Time to listen: an account of the role of support health visitors. *Community Practitioner*, 75 (2), 461-463. [Not available through open access.]

RANDOMISED CONTROLLED TRIAL

Time to listen: an account of the role of 'support' health visitors

Abstract

The Social Support and Family Health Study is a randomised controlled trial (RCT) of three different forms of support for families with infants living in a deprived inner city area. The overall aim of the study is to quantify effectiveness and cost-effectiveness of additional health visitor input, in the form of supportive listening, compared with either input from non-professional family support services in the community or statutory services. Primary outcome measures such as child injury and maternal psychological ill-health will be used in the study. The trial results will be available in 2003. The aim of this paper is to describe the study and our role as support health visitors using a supportive listening model. Two case-studies are presented followed by a discussion of the key points about this type of health visitor intervention.

Key words: Randomised controlled trial, social support, supportive listening, client centred, nondirective

Community practitioner 2002; 75, 12: 461-463

There is strong evidence to show that mothers dislike it when health visitors take a more directive, authoritarian approach and give unsolicited advice

Mary Sawtell

Carol Jones

Support health visitors
Social Science Research Unit, Institute of
Education, University of London

It is a well accepted fact that exposure to disadvantage in childhood results in an increased risk of future ill-health. There is now an established research literature that shows that early interventions are effective in improving the health of high risk populations.¹

The importance of support for parents and children, and in particular present and future mothers, as a means of tackling inequality in child health, has been emphasised.² As a result of such evidence, a huge financial commitment has been made by the government in recent years to enhancing social support to deprived families, and reducing social exclusion, principally through the Sure Start programme.³ Great importance is attached in the Sure Start literature to the involvement of a wide range of voluntary and statutory agencies in this task and it is stated that efforts will be targeted at individual,

family and community level.⁴ Bidmead reports that this is actually happening in the Sure Start trailblazer sites, albeit slowly.⁵ Despite the weight of evidence in favour of this overall approach to improving public health, many important unanswered questions remain as to the exact nature of the support offered.¹ The search for rigorous evidence on this issue is therefore a matter of great concern.

The Social Support and Family Health Study is a randomised controlled trial (RCT) of three different forms of support for families with infants living in a deprived inner city area. The eventual aim of the study is to quantify effectiveness and cost-effectiveness of additional health visitor input, in the form of supportive listening, compared with either input from non-professional family support services in the community or statutory services. Primary outcome measures such as

Case-study one

Stacey was recruited to the study when her baby George was three months old. At the first visit I met Carl, partner of Stacey and father of George. It was clear at this first encounter that this was a very young couple, living in deprived and chaotic circumstances but delighted with their new baby. As a health visitor accustomed to searching for health needs from the first encounter, there were many that were easily seen. These included home safety, hygiene, maternal smoking and feeding. These issues were not raised by Stacey in the first visit however.

What she clearly wanted to talk about were her feelings about a traumatic birth and postnatal period in the hospital, and George's congenital leg deformity. It was the insensitive attitude of health professionals to these two issues that was preoccupying her. She attributed this to the fact that she and Carl were young and therefore perceived to be both relatively unlikely to challenge poor standards of care, and not to be trusted in their ability to care for their baby well.

I visited Stacey about once a month for between one and two hours each time. It emerged that between 11 am and 1 pm was the optimum time of day for visits. This was after she had got up and had 'the first cigarette of the day' and before Carl's mother visited in the early afternoon.

Stacey was diagnosed with postnatal depression when George was about five months old. Within a couple of months she was markedly better, but when George was 11 months old she was dismayed to find that she was pregnant again.

During my visits Stacey talked mainly about her own life: her time in care as a child; her persistent truanting from school; a previous termination; her relationship with Carl and the members of her family. The common theme that ran through all these discussions was her determination that she was going to do the best she possibly could for George and that his childhood was going to be very different from hers. She was confident that she could do this but felt that the professionals around her were expecting her to fail.

Towards the end of my year of visiting, Stacey was beginning to seek and act on advice from me on issues such as appropriate discipline, breastfeeding the new baby, coping with two young children, managing her relationship with Carl and sources of support for herself in the community. By the time my contact finished she was apprehensive but excited about the new baby's imminent arrival and she was involved with a community support organisation that was seeking her help in establishing a group for young parents.

Case-study two

When I first met her, Stephanie was 23 years old. She had two daughters aged three months and 17 months. At the first meeting she was able to tell me that she had been depressed, but she was glad this was now in the past. She attributed the depression to living on the 10th floor of a tower block and to the traumatic delivery of her baby. I visited her monthly. After four visits Stephanie said that she was depressed again. She and her husband had argued, and he had moved out to stay with a relative temporarily.

Using the supportive listening model, I had learnt a little about Stephanie's childhood. During one visit I was able to reflect that this must be a difficult time for her, as she was at the same stage in her life as her parents were when their marriage broke down. At the following visit Stephanie commented on how surprised she was that I had made a connection between her current and past family situations. She felt this was something she would not have appreciated the enormous significance of herself. Furthermore she had used my comments about her own situation and similarities with her parent's life to talk to her husband. She described how this 'changed everything' allowing them to really start to talk and understand each other more fully.

I continued to visit and listen to Stephanie's concerns, but to offer no unsolicited health promotion advice. However it emerged on one visit that Stephanie was worried about Mary, the baby. She wasn't doing what her older daughter had been doing at the same age. Stephanie asked what she could do about this. We discussed various strategies she could try but I was reassuring about Mary's general development.

On my next visit there was an enormous change in Stephanie and Mary's relationship with much more direct eye contact and other communication between them. Stephanie attributed this to following my suggestions. She said in general my visits had been very beneficial and wondered how other women coped who didn't have this support.

All names have been changed in case-studies one and two.

child injury and maternal psychological ill-health will be used in the study. The trial results will be available in 2003.

This paper aims to describe the study and our role as support health visitors (SHVs) using a supportive listening model. Two case-studies are presented followed by a discussion of the key points about this type of health visitor intervention.

Study format

The study aims to compare different forms of social support offered to families with a newborn child over a one-year period. Seven-hundred-and-thirty-one women living in deprived wards of Camden & Islington were recruited to the study, when their baby was approximately nine weeks old. Thirty-nine per cent did not have English as their first language and interpreters were provided for the 100 women who needed them. The women were randomly allocated to one of three groups. All three groups had the standard health visitor service provided by Camden & Islington Community Health Service NHS Trust and the usual access to other support services in the community. The groups were as follows:

- a) control – 364 women. These women received standard services.
- b) community group support intervention – 184 women. In addition to standard services, members of this group were offered one of three types of service provided by community groups operating in the locality.
- c) support health visitor intervention – 183 women. In addition to standard services, these women were offered a supportive listening visit from a SHV approximately once a month for one year.

The study was designed as an RCT. This methodology is generally acknowledged to be the gold standard for demonstrating cause and effect. Using random allocation of research participants to two or more groups, one of which is a control group, interventions can be reliably tested. The RCT is commonly used in medicine. Oakley argues it can and should be used to evaluate social interventions.⁶ The RCT methodology has been used in studies of intensive home visiting to high risk families by appropriately trained nurses in North America.⁷ The results of these studies show a positive impact on a range of health and social outcomes in the immediate period after the intervention and up to 15 years later.⁸ RCTs of aspects of health visiting practice in the UK, such as effectiveness in reducing unintentional injury in under fives, have been conducted.⁹ In general though, as Robinson points out, there are few RCTs of health visiting in this country, this one is therefore timely.¹⁰

Supportive listening

The Acheson report states:

'Health visitors already aim to provide social and emotional support to parents and are acceptable in disadvantaged settings. We consider that an enhancement of their role and capacity to provide social and emotional support to disadvantaged parents is needed.' (p 76)

The empirical evidence on whether or not health visitors are seen as supportive by their clients is varied. Studies such as those by Machen show women like the support that health visitors can give.¹¹ She concludes that it is the non-directive approach of health visitors that is particularly valued. This finding concurs with the evidence from a study

into additional social support for pregnant women from midwives by Oakley. She found that a non-judgmental listening approach is perceived as supportive by mothers; it significantly increases their self-esteem and confidence, empowering them to make the best use of the resources available, to provide good care for their children.¹² Similarly Holman discovered that a course of non-directive counselling by health visitors was effective in managing postnatal depression.¹³

There is strong evidence to show that mothers dislike it when health visitors take a more directive, authoritarian approach and give unsolicited advice. For example, Heritage and Sefi¹⁴ taped advice-giving sequences that occurred in health visitors' first visits to clients. They found that advice giving was predominantly initiated and delivered by the health visitor in a unilateral manner with little effort to accommodate the circumstances of the individual mother. When this occurred, the mother generally met it with either passive or active resistance.

Similar results emerge from Murphy's¹⁵ work on feeding intentions in pregnant women. Murphy sympathises with the professional's dilemma when faced with an aspect of parenting which contradicts the knowledge base on which their professional practice is based. However, like Heritage and Sefi, she argues that a surveillance approach, where the professional can be seen as interfering and insensitive, is counter-productive in that the client is unlikely to turn to the professional as a source of support and advice.

Supportive listening in the study

Evidence such as this led to a supportive listening model being adopted for this study. This involved providing non-judgmental maternal support. The primary aim of the visits by the SHVs was not to give information or advice but to offer a listening ear for whatever the mother deemed significant issues. The visits were entirely client led. Thus, when a four-month-old baby was seen sleeping on her front, or a baby was seen being fed solids before the recommended age, no comment was made unless the client gave an appropriate cue. If advice was clearly sought by the client, it was to be given. We gave an assurance of complete confidentiality stressing that we would not be communicating with other professionals in the field, outside the study team, other than to let them know of our involvement with the family. The client was informed that the only exception to this rule was if child protection issues arose.

Wherever possible the women decided how long and how frequent the visits were. They also chose the location. Evening and weekend visits were made if that was what suited the client best. Within reason ►

◀ were able to continue visiting a family if they moved. So, while at the start of the study each SHV had a geographical 'patch', given the high rate of mobility in the study population, by the end of the intervention period each SHV was visiting all over the boroughs and sometimes beyond, for example, Essex and Hertfordshire.

There were five SHVs on the study, two full time and three working three days a week. The full time case-loads were made up of approximately 50 families and the part time, 30. Our principle task was our supportive role; we did however collect data from each woman on our case-load in the form of three taped interviews during the course of the year.

Discussion

In our work as SHVs on this study we had the heady combination of time and management support to be truly client centred in our work, commodities that are often rare and luxurious in health visiting today. In terms of time, our case-loads were at the most a fifth of the size of the average health visiting case-load. None of the core health visiting tasks such as new birth visits, clinics and case conferences impinged on our time and administrative tasks were minimal. Thus the bulk of our working week was spent home visiting. We could allow on average one to two hours for each visit if necessary and generally we could offer a lot of choice in terms of when, where and how often visits took place. The women's self-expressed needs were listened to and accommodated, which meant visits could also involve accompanying them to, for example, the solicitor, the housing office or to view a school.

One woman was spending five hours, five days a week travelling by train to see her partner in prison. Her SHV accompanied her on this journey on several occasions. Being able to accommodate the practical preferences of individual families helped to convey to them the egalitarian ethos of the work.

We were also free from any pressure to make an assessment of health needs at the first encounter, as ample time was available to make a relatively well-informed assessment in the weeks and months ahead. Thus, the 'suspected but hidden' category of health need identified by Chalmers¹⁶ as one of the four types of need that health visitors see in their work, could be held in mind in the hope that the right cue to pursue the issue was given to the SHV at some future visit. Chalmers stresses the importance of getting the timing right in this way, a luxury not always available to the health visitor with a case-load of several hundred families.

It appeared to us that as our clients got used to the fact that we had time, they gained in confidence in using us as a resource. It was common for us to hear the comment from a

client that her local health visitor was very nice but she was clearly very busy and thus not apparently available. For those living in a disadvantaged community this may be a particular problem. As Reading's study shows, the scale of need in the community is often not reflected in the size of health visitor case-loads.¹⁷

The attitude of those running the study was that the supportive listening model, with its focus on building a relationship, was a powerful therapeutic tool. This is not new to health visitors. The search for health needs, involving listening and partnership and acknowledging the uniqueness of each client, is a fundamental principle of the health visiting service. What is unusual as a health visitor is being given permission to not have to combine support and surveillance. While practitioners may be acutely aware that a directive approach can alienate the client, rendering efforts to be supportive pointless, there is pressure to continue with this approach. This pressure can arise from the service requirements of the employing authority, often in the form of a core programme of tasks, and the anxiety arising from a sense of responsibility for the welfare of the woman and the child. Also as Heritage and Sefi suggest, the health visitors' readiness to deliver advice during a visit gives them a sense of purpose, a 'ticket of entry'.¹⁴

Our ticket of entry had been secured for us on this study by the recruitment process. Both parties were prepared for the fact that listening was the principle purpose of the visit. This undoubtedly made our task easier. It was clear however that some of the women had consented to take part in the study out of altruism rather than a perceived personal need. For some this remained the case as the visits progressed and they chose to have visits relatively infrequently. The majority in this position seemed to be surprised, however, at how therapeutic they found being listened to.

There were some drawbacks to working in this way. The most obvious one was the lack of inter-agency contact. This professional isolation, even with clinical supervision, was most challenging for us when concerns about a family arose and an assessment of risk had to be made without consulting other service providers with knowledge of the family. It was also clear that some of the local health visitors were confused and anxious about our involvement with 'their clients'.

Remaining faithful to the supportive listening model was an additional challenge. As experienced health visitors, accustomed to a surveillance component in our work, it was frequently tempting to offer unsolicited advice. As the study progressed however we were encouraged by the fact that the majority of the women were clearly enjoying the visits and were defining and frequently meeting their own health needs through being

listened to. This gave us the confidence to continue and enabled us to fully enjoy the exciting and professionally liberating experience of health visiting in this way.

Acknowledgments: We would like to thank our support health visitor colleagues Hermine Barry-Clarke, Carolyn Dawson and Margaret Killen and the other members of the study team: Ann Oakley, Ian Roberts, Meg Wiggins, Helen Austerberry, Lyn Rajan, Helen Turner, Sandra Stone, Miranda Mugford, Ruben Mujica Mota.

This study was funded by the Health Technologies Assessment Programme of the NHS R&D programme and by Camden & Islington Health Authority. The views and opinions in this paper are those of the authors and do not necessarily reflect those of the Department of Health.

References

- 1 Cowley S. Early interventions: evidence for implementing Sure Start. *Community Practitioner* 1999; 72: 6, 162-165.
- 2 Acheson D. Independent inquiry into inequalities in health. London: The Stationery Office, 1998.
- 3 Department of Health. Supporting families. London: The Stationery Office, 1998.
- 4 Department of Health. Making a difference. London: The Stationery Office, 1999.
- 5 Bidmead C. Slowly but surely. *Community Practitioner* 2000; 73: 12, 859-860.
- 6 Oakley A. Experimentation and social interventions: a forgotten but important history. *British Medical Journal* 1998; 317: 1239-1242.
- 7 Olds DL *et al.* Preventing child abuse and neglect: a randomised controlled trial of home visitation. *Pediatrics* 1986; 78: 65-78.
- 8 Olds D *et al.* Long term effects of home visitation on maternal life course and child abuse and neglect. Fifteen year follow up of a randomised controlled trial. *Journal of the American Medical Association* 1997; 278: 637-643.
- 9 Watson M *et al.* Injury prevention: working together on an RCT. *Community Practitioner* 2002; 75: 5, 172-175.
- 10 Robinson J. Domiciliary health visiting: a systematic review. *Community Practitioner* 1999; 72: 2, 15-18.
- 11 Machen I. Clients perceptions of the health visiting service. In: Appleton J, Cowley S (eds). The search for health needs: research for health visitor practice. London: Macmillan, 2000.
- 12 Oakley A. Social support and pregnancy outcome. *British Journal of Obstetrics and Gynaecology* 1990; 97: 155-162.
- 13 Holman J *et al.* Counselling in a general practice setting: controlled study of health visitor intervention in treatment of postnatal depression. *British Medical Journal* 1989; 298: 223-226.
- 14 Heritage J, Sefi S. Dilemma of advice: aspects of the delivery and reception of advice in interactions between health visitors and first-time mothers. In: Drew P, Heritage J (eds). Talk at work: interaction in institutional settings. Cambridge: Cambridge University Press, 1992.
- 15 Murphy E. Breast is Best. *Sociology of Health and Illness* 1999; 21: 187-208.
- 16 Chalmers K. Searching for health needs: the work of health visiting. *Journal of Advanced Nursing* 1993; 18: 900-911.
- 17 Reading R, Allen C. The impact of social inequalities in child health on health visitors' work. *Journal of Public Health Medicine* 1997; 19: 424-430.

APPENDIX 2

CANDIDATE PUBLICATION 2

Implementing a structured education program for children with diabetes: lessons learnt from an integrated process evaluation

Mary Sawtell,¹ Liz Jamieson,² Meg Wiggins,³ Felicity Smith,² Anne Ingold,³ Katrina Hargreaves,³ Meena Khatwa,³ Lucy Brooks,⁴ Rebecca Thompson,⁵ Deborah Christie⁵
To cite: Sawtell M, Jamieson L, Wiggins M, et al.

BMJ Open Diabetes Research and Care 2015;3: e000065. doi:10.1136/bmjdr-2014-000065

ABSTRACT

Background: There is recognition of an urgent need for clinic-based interventions for young people with type 1 diabetes mellitus that improve glycemic control and quality of life. The Child and Adolescent Structured Competencies Approach to Diabetes Education (CASCADE) is a structured educational group program, using psychological techniques, delivered primarily by diabetes nurses. Composed of four modules, it is designed for children with poor diabetic control and their parents. A mixed method process evaluation, embedded within a cluster randomized control trial, aimed to assess the feasibility, acceptability, fidelity, and perceived impact of CASCADE.

Methods: 28 pediatric diabetes clinics across England participated and 362 children aged 8–16 years, with type 1 diabetes and a mean glycosylated haemoglobin (HbA1c) of 8.5 or above, took part. The process evaluation used a wide range of research methods.

Results: Of the 180 families in the intervention group, only 55 (30%) received the full program with 53% attending at least one module. Only 68% of possible groups were run. Staff found organizing the groups burdensome in terms of arranging suitable dates/times and satisfactory group composition. Some staff also reported difficulties in mastering the psychological techniques. Uptake, by families, was influenced by the number of groups run and by school, work and other commitments. Attendees described improved: family relationships; knowledge and understanding; confidence; motivation to manage the disease. The results of the trial showed that the intervention did not significantly improve HbA1c at 12 or 24 months.

Conclusions: Clinic-based structured group education delivered by staff using psychological techniques had perceived benefits for parents and young people. Staff and families considered it a valuable intervention, yet uptake was poor and the burden on staff was high. Recommendations are made to inform issues related to organization, design, and delivery in order to potentially enhance the impact of CASCADE and future programs.

Key messages

- The Child and Adolescent Structured Competencies Approach to Diabetes Education (CASCADE) structured education program is perceived by young people and parents who attend as having benefits but practical challenges associated with attendance result in low uptake.
- Staff are positive about the potential of the program but organizational aspects are unacceptably burdensome.
- CASCADE is potentially deliverable to families as part of routine care and could be a useful intervention. However, improvements in clinical and administrative support, staff training, program content, and service structures are required to ensure fidelity to the program and feasibility and acceptability to key stakeholders.

INTRODUCTION

Type 1 diabetes mellitus (T1DM) in children and young people is increasing worldwide. Fewer than one in six children and young people achieve glycosylated fraction of hemoglobin (HbA1c) values in the range identified as providing best future outcomes.¹ It has been recognized that there is an urgent need for clinic-based pragmatic, feasible, and effective interventions that improve both glycemic control and quality of life, with a particular emphasis on structured education programs.² In recent years, a number of large multicentre studies have trialled a standard education intervention.³⁻⁵ Findings published, to date, report no significant positive impact on glycemic control as measured by HbA1c and only limited impact on a wide range of secondary measures.⁴⁻⁵ Nevertheless, the recent Best Practice Tariff for Paediatric Diabetes for diabetes services in the UK⁶ requires the provision of structured educational programs for young people and their families and, as a consequence, there is an urgent need for high-quality evidence to inform the implementation of this recommendation.

The CASCADE (Child and Adolescent Structured Competencies Approach to Diabetes Education) pragmatic cluster randomized controlled trial (RCT) with integral process and economic evaluation is the most recent study. It was undertaken by a team that included clinicians from a London-based paediatric diabetes clinic, a representative from a diabetes patient organization and researcher teams from three universities in London. The CASCADE intervention is a structured education program designed for children and young people with T1DM aged between 8 and 16 years and their parents or carers.⁷ The intervention underwent phase 1 pilot work and a non-randomized trial, in which the delivery was carried out by a psychologist.⁸ The CASCADE intervention was then modified to be delivered by two members of a diabetes multidisciplinary team (MDT) who receive 2 days of training to enable them to become 'site educators'. CASCADE is a manual-based program. It is delivered in four modules over 4 months, each lasting approximately 2 hours, to groups of three to four families with children and young people grouped according to age (8–11 or 12–16 years). Two psychological approaches, motivational interviewing and solution-focused brief therapy, shown to have potential with children with diabetes are

central to the CASCADE intervention.⁹⁻¹⁰ These aim to engage participants to identify and develop their own positive approaches and consequent behavior change relevant to the management of their condition. The intervention thus offers both structured education, to ensure young people (and their parents) know what they need to know, and a delivery model designed to motivate self-management through empowerment techniques (see table 1).

Table 1

Outline of the CASCADE program (as set out in the manual)

The teaching plan	Session activities, objectives, time guides, and resources including key information essential for the educator, learning objective for the family, and brief descriptions of each activity
Each module starts with a review of, and since, the previous session, creating an opportunity for families to highlight any changes that have taken place and to congratulate young people on successes	
Module 1	Focuses on the relationship between food, insulin, and BG (<i>eg, considering the pros and cons of matching insulin to food to attain better glycemic control</i>)
Module 2	Reviews BG testing and factors influencing BG fluctuation (<i>eg, identifying factors that cause BG to rise and fall and explore hypoglycemia definitions, reviewing symptoms according to severity</i>)
Module 3	Looks at the pros and cons of adjusting insulin (<i>eg, a brainstorming session considers when, how, and who to contact for help managing hyperglycemia</i>)
Module 4	Addresses aspects of living with diabetes, including managing BG levels and exercise (<i>eg, young people and families complete a 'blueprint for success'. This marks the end of the sessions and acknowledges the steps into the future the young person has already made</i>)
Homework tasks are given to families to consolidate learning after each module	

The intention is that delivering CASCADE to groups will provide staff with an alternative mode of working with young people in the clinic setting to improve outcomes, rather than requiring additional work.

CASCADE TRIAL SUMMARY

The trial involved young people with T1DM and family members in 28 English pediatric diabetes clinics (randomly assigned at clinic level to intervention or control) in London, South East England, and the Midlands. Clinics eligible to participate were staffed by at least one paediatrician and pediatric nurse with an interest in diabetes. Other inclusion criteria included not running a group education program at time of recruitment and not participating in a similar pediatric diabetes trial within the past

12 months. It was approved by the University College London (UCL)/UCLH Research Ethics Committee (REC) reference number 07/HO714/112. Site-specific approval was granted at each site. Three hundred and sixty-two young people were recruited to the study. Inclusion criteria included: diagnosis with a duration ≥ 12 months; mean 12-month HbA1c of 8.5 or above; aged 8–16 years. Clinical staff identified eligible young people from their patient list. Researchers sent letters and information sheets to these young people and their parents or carers inviting them to participate in the research and to speak to a researcher at their next clinical appointment. Recruitment was primarily carried out by members of the process evaluation team who attended clinics at which eligible young people had an appointment. Signed consent forms were collected from parents and children wishing to participate.

The primary outcome measure was venous HbA1c at 12 and 24 months. Secondary outcomes included: knowledge, skills and responsibilities associated with diabetes management; emotional and behavioral adjustment; quality of life. Two staff members from each intervention site clinical team participated in the 2 days CASCADE training program. These site educators then took responsibility for organizing the modules at their clinics and delivering the intervention. The extensive and integral process evaluation was designed to enable an understanding of the implementation of CASCADE and examination of the interaction of causal mechanisms and contextual factors that may be determinants of the intervention's success or failure, as assessed by the trial.¹¹ Given that the trial found no evidence of benefits on venous HbA1c at 12 and 24 months and little evidence of benefits on secondary outcomes, the focus of this paper is to use the findings of the process evaluation to suggest how future structured education may be more effectively implemented.¹²

PROCESS EVALUATION METHODS

The process evaluation aimed to assess the feasibility, acceptability, fidelity and perceived impact of the CASCADE intervention. It ran for the 4-year life of the trial and included the multiple methods shown in table 2. Researchers from the process evaluation teams at the Institute of Education (IOE) and the School of Pharmacy (SOP) conducted the fieldwork.

Table 2

Process evaluation methods and response rates

Phase of the study	Methods	Purpose of methods	Response rates
Two-day training of site educators	Unstructured observation of training of site educators by a member of the research team	Fidelity of training	6 training days observed

Phase of the study	Methods	Purpose of methods	Response rates
	Participant questionnaires (completed 2 weeks after training)	Description of participants Participant experience/acceptability of training	27 participant questionnaires from 18 nurses, 8 dietitians, and 1 doctor (63% of participants)
	Semistructured interviews with the two trainers	Background to intervention development; views on training days	Both trainers
Delivery of CASCADE modules with patients/carers	Observation of modules carried out by a member of research team including rating of fidelity to psychological techniques and content of manual	Fidelity of delivery Experience/acceptability of delivery of program to site educators Experience/acceptability of participation in the program by young people/parents	47 CASCADE modules observed across 13 intervention sites (12 each of modules 1, 2, and 4; 11 of module 3)
	Self-complete feedback proformas for site educators	Who delivered each module; who attended each module Self-assessment of delivery fidelity and general feedback on each module	Site educators returned 125 feedback proformas (94% of 131 completed modules)
Following delivery of all CASCADE groups	Young person and parent 12 and 24 month questionnaires in intervention arm	Perceptions of impact Acceptability of the intervention	Process questions were completed on questionnaires by 135 young people (82%) and 121 parents (66%) at 12 months; 121 young people (66%) and 114 (63%) parents at 24 months
	Semistructured interviews (audio-recorded) with site staff (nurses and dietitians), young people, and parents/carers in both trial arms	Description of standard care—including any structured education currently delivered Intervention arm only—experiences of the intervention (training and delivery)	30 site staff (16 intervention sites; 14 control) 53 young people (32 intervention/21 control) and 52 parents were interviewed. Of the young people, 31 were female; 17 were 10–11 years old; and 36 were 12–18 years old

- **CASCADE, Child and Adolescent Structured Competencies Approach to Diabetes Education.**

PROCESS EVALUATION DATA ANALYSIS

Qualitative data analysis was carried out by the process evaluation teams at IOE and SOP (all the authors except LB, RT, and DC). Qualitative analysis of the interview data, supported by the use of NVivo software, identified key topics and issues that emerged through familiarization with transcripts.¹³ Pertinent excerpts were coded and memos written to summarize and synthesize emerging themes. Researchers refined their analysis ensuring that themes were crosschecked with other data, first within and then between transcripts. Analysis of each training workshop observation was carried out by a researcher, who was not the observer, reading through the notes made by the observer and identifying key themes and fidelity issues emerging from the data. Quantitative data were analyzed by MW using Excel and the SPSS V.19 software for statistical tests. In terms of the CASCADE modules delivered in the sites, composite fidelity delivery scores were created for content and for technique from individual researcher observer and site educator self-rated scores. A further composite variable was then calculated which summed the content and technique scores for each site across all four modules, allowing comparison across sites and modules.

PROCESS EVALUATION RESULTS

The results are structured under the following themes: recruitment and training of site educators; organizing the groups; delivery of the modules; uptake and acceptability of the modules; and perceptions of impact. Response rates are reported in table 2.

Recruitment and training of site educators

The National Institute for Health and Care Excellence (NICE) requirement,² that structured education programs are delivered as part of routine care was widely recognized by clinic staff and, as a consequence, it proved relatively straightforward to recruit two members of the MDT from each of the 14 intervention sites to become site educators. The majority of site educators were experienced pediatric diabetes specialist nurses (PDSNs); in approximately half of the sites one of the educators was a dietitian. The diabetes specialist nurse and psychologist who developed the intervention delivered the 2 day CASCADE training for site educators in four workshop sessions. In general, it was feasible for sites to send the required minimum of two staff to the core workshops. A few sites sent additional interested members of the MDT though only four consultants attended some or all of the training. The training was delivered in a central London location, except for one site where following a request, training was delivered locally. Site staff reported this change in location to be helpful. The majority of staff who completed the questionnaire following the workshops indicated they had been 'extremely' or 'very' keen to participate. Most staff thought the training was very good, motivating, and comprehensive. The most common concern raised in staff interviews about becoming site educators and

running the CASCADE program, both before and after the training, was additional workload. Other concerns included practical constraints such as finding available rooms in which to run the groups and ability to rapidly change their practice to employ the psychological approaches underpinning CASCADE.

One site educator commented.

It [the training] was a lot in the few days. Teaching people theories and expecting them to suddenly change their behaviour I think is very difficult.

The two trainers, and some attendees, expressed concern about levels of diabetes knowledge among the site educators.

Some of it [the training] was ending up teaching them the content as opposed to teaching them the style of delivery. (UCLH trainer)

At the time [of the training] I'd got very little diabetes knowledge so, for me, I was actually learning from it and I know that's not really what it was about but a lot of it was that...I found it quite intimidating because of my lack of knowledge. (Site educator)

Organizing the groups

A total of 30 complete CASCADE groups, comprising all four modules, were run across 12 of the 14 intervention sites. A post hoc calculation, based on the number of study recruits in a site and the optimum group size of 3–4 young people, suggested 44 groups should have been run across the 14 intervention sites. Thus, 68% of possible groups ran, with only three clinics completing the maximum number of groups possible for their site. A key reason for this limited delivery was difficulties with organizing the groups. The organization was undertaken by the site educators in all the sites. This involved: deciding which participants should be grouped together using similar ages as a key criterion; setting dates and times; inviting families to attend; and booking a room. Interviews revealed that site educators found these processes frustrating and very time-consuming. One site educator commented:

I didn't notice that it saved me any time because I was constantly chasing them [families] up to be there.

One site delivered no modules because the lead site educator left her PDSN post soon after the training. Another site delivered only the first module because of a number of challenges which included: the small number of potential eligible patients on the clinic list; poor uptake of the first module by young people/ parents; practical organizational constraints. All the sites ran the groups in addition to routine clinics where standard care continued to be received by patients on an individual basis. Staff interview data revealed that the pressure on hospital clinic facilities was too great to make running the groups feasible during clinics. Establishing a date and time for the group sessions that was acceptable to the families was extremely challenging. To maximize attendance, some site educators tried a range of timings including during

school hours, after school, weekends, and school holidays. Communication with families, about groups, was via a combination of letter, telephone, and (occasionally) text messages. No sites used email or online meeting booking sites. Despite all the negotiation and careful planning by site educators, late cancellation or non-attendance by participants was reported as common.

Some didn't even bother to get back to us and some did and said they were still gonna come but still didn't come. It is frustrating and I think that's what was time consuming, which I hadn't really accounted for...(Site educator).

As a result of these difficulties, compromises were made to the intended group size and composition. Groups often had small numbers (sometimes one family only) and/or a wide age range among the young people attending. Although the intention was to run four modules with the same participants, the composition of many groups changed.

Delivery of the modules

The site educators believed they were appropriate individuals to deliver the intervention because they knew patients well, although familiarity with patients was not a requirement. Participating families appeared to support this view. All sites had continuity of at least one trained site educator, but complications in sustaining the availability of a second educator in a few sites resulted in some lack of continuity of trainer pairs. Site educators reported that the time required to organize sessions meant that they often had little or no time for planning and practising delivery of the modules. Observation data and some staff interviews suggested that this lack of practice time was particularly challenging when staff had limited experience in group work. Researcher observation of the modules and site educator feedback forms indicated that site educators generally delivered activities as described in the manual. However, less time than was recommended was spent on some of the key exercises due to staff finding them difficult to deliver and/or not well received by groups. One such example was the 'review since the previous session' exercise at the beginning of each module. Also, while researcher observation and staff feedback showed fidelity of CASCADE psychological techniques was good across sessions in half the sites, it was not optimal in the remainder. Difficulties in delivering the intervention particularly occurred when sessions had groups of participants with a wide age range or group numbers were very small.

The first group that we ran had two girls and a boy and the boy was at the younger end of the teenage years and the girls were at the older, it was unfortunate because we didn't have that many patients as part of the study so it was very difficult then to get the groups sorted out so we kind of had to put them together. [...] He was just a bit of a silly boy in that...I don't mean horribly, he was lovely, but just kind of played the fool a little bit whereas the girls were older and a similar age and a lot more grown up about it all. (Site educator)

Staff reported that the organization and delivery of the intervention was affected by the research context in a number of ways. First, having to restrict the education groups to a subset of recruited patients, instead of offering them to the entire clinic list, was perceived as making the organization of the groups more challenging. This meant that natural groupings of patients (by age or geographical area) often proved too difficult to achieve. Second, delays encountered in the recruitment of families to the trial in many sites (see 12 for detail on this), meant site educators often had to wait several months after their training before they could start to organize groups and deliver the intervention. Third, some site educators reported that additional trial-related tasks, such as organizing research blood samples added to their workload and took time away from organization of, and preparation for, groups.

Uptake and acceptability of the modules

Of the 180 young people recruited to the intervention arm, only 55 (30%) received the full education program of four modules with just over half of the original recruits (53%) attending at least one module. Eighty-four young people (47%) failed to attend any modules. Those who attended had significantly lower mean baseline HbA1c scores than those who were offered the sessions but did not attend (9.52 vs 10.33, $p < 0.01$). Significantly more children (8–12 years) attended at least one module compared with teenagers (13–16 years; 64% vs 44%, $p < 0.01$). Clinics were permitted to offer sessions at a time of their choice. If out of school hours sessions were not offered, the main reason given for young people not attending modules was that they did not want to miss school. For parents, taking time off work during the day was a barrier to attendance. Other reasons for non-attendance cited by children and parents included holidays and other extracurricular activities.

On most occasions a parent/carer attended with the young person. Parents and young people reported that joint attendance was a very positive aspect of the experience (see [table 3](#)). Staff also, in most instances, found it helpful to include parents.

Table 3

Acceptability of CASCADE to parents and young people attending at least one CASCADE module (12-month questionnaire)

	Young people	Parents/carers
Themes	‘Quite a lot’ or ‘A great deal’ n/N (%)	‘Quite a lot’ or ‘A great deal’ n/N (%)
Group dynamic		
Liked parents/young people being together in modules	81/90 (90)	81/84 (96)

Themes	Young people	Parents/carers
	'Quite a lot' or 'A great deal' n/N (%)	'Quite a lot' or 'A great deal' n/N (%)
Felt learnt something from other people in the group	64/93 (69)	60/85 (71)
Teaching style and length		
Liked the way the trainers taught	74/93 (81)	81/86 (94)
Felt the sessions were too long	22/94 (23)	7/85 (8)
Content		
Felt that some of the things covered were too complicated	7/92 (8)	8/85 (9)
Felt that some of the things covered they knew before	48/91 (53)	42/86 (49)

Perceptions of impact

The majority of parents and young people who attended CASCADE groups described some positive impacts, including improved family relationships, wider knowledge and understanding of diabetes, greater confidence, and increased motivation to manage the disease (see [table 4](#) and young person's comment below).

I've been more happier...yeah, like around the house I've been more happier. Not so many strops... 'cause my readings are better and we've been given a lot more information about the ketones and how to treat it....I found it really good. [Young person]

Table 4

Parents' and young people's perceptions of influence of CASCADE (12-month questionnaire)

Questionnaire items	Answered 'Quite a lot'/'A great deal'	
	Parent N=90	Young person N=97
Question: After attending some or all of the CASCADE diabetes education sessions, how much did your child/you...?		
Knowledge		

Questionnaire items	Answered 'Quite a lot'/'A great deal'	
	Parent N=90	Young person N=97
Question: After attending some or all of the CASCADE diabetes education sessions, how much did your child/you...?		
Understand better how insulin works	66 (74%)	70 (73%)
Understand better which foods contain CHO	73 (81%)	68 (70%)
See why counting the CHO in the food your child/you eat(s) can be helpful	80 (90%)	73 (75%)
Intention to change		
Want to stop your child's/your glucose levels from going too low or high	85 (94%)	85 (87.5%)
Want to test your child's/your BG levels more often	56 (43%)	40 (42%)
Control		
Feel more in charge of your child's/your diabetes	61 (69%)	65 (68%)
Feel able to change your child's/your insulin dose when they are exercising	68 (77%)	66 (69%)
Feel you are able to control your child's/your BG levels better	70 (78%)	64 (67%)
Access to care		
Feel more able to ring/contact your diabetes nurse/GP/hospital if your child/ you need(s) help	72 (82%)	52 (54%)
Family dynamic		
Feel you had a better understanding of how diabetes affects your family	67 (75%)	69 (72%)

- **BG, blood glucose; CASCADE, Child and Adolescent Structured Competencies Approach to Diabetes Education; CHO, carbohydrate; GP, general practitioner.**

A number of young people and parents mentioned that timing of the CASCADE sessions would be more appropriate and useful sooner after diagnosis; site educators

90

also commented that this may lead to better uptake of the sessions and have greater impact.

I felt they were of little use to me as I already knew everything however this kind of session would be useful to someone who had just been diagnosed. (Young person)

They're a bit sort of more 'do as they're told' for the first 12 months, they're more likely to attend and perhaps take it on board, it gets them in the right frame of mind early. (Site educator)

Twenty-four months after the intervention, when asked in the questionnaire what effect the program had had, nearly half of the young people selected the response "The sessions made me want to try harder and I have carried on trying". However, these impacts were not reflected in the primary or secondary outcome measures, even for the subgroup of those who attended.

DISCUSSION

The CASCADE intervention aimed to train PDSNs and other members of diabetes teams to deliver a manualised, structured education program, based on behavior change methods, to groups of families. Training of these site educators took place over 2 days. Few members of the MDT, other than PDSNs, attended the training. Trainee educators expressed enthusiasm for the program but highlighted concerns including that: CASCADE would increase their workload; there would be practical constraints to setting up and running groups; and that incorporating the CASCADE psychological model into their practice would be challenging.

Following delivery of CASCADE in the sites, PDSNs and other clinical staff were positive about the program. Having PDSNs and dietitians, who knew the patients, as site educators worked well for both the educators and families. There were, however, feasibility issues with regard to running the program in its current form in the 'real world' of the National Health Service. These were evidenced by low uptake by families and staff feeling unacceptably burdened by organizational aspects of the intervention. Organizing groups was, as anticipated by staff, challenging and time-consuming and many groups did not comprise the recommended number or age range of young people. This affected group dynamics and made it difficult to run the sessions as set out in the manual. It was also difficult to keep a group together for the planned four modules. Delivery of the modules was further compromised by: the gap in time between training and delivering sessions; time spent on organizing group sessions at the expense of practising delivery of the modules; and finding some exercises consistently hard to deliver.

Despite the fact that families and staff reported that they liked the program and felt that it offered benefits, the trial found no evidence of impact on venous HbA1c at 12 and 24 months and little evidence of benefits on secondary outcomes, even with the subgroup who attended the training. We think the reasons behind this are twofold. First the organizational difficulties that made the intended group composition problematic and second the difficulties with delivery, especially the lack of fidelity to

the psychological techniques. To address these issues, and to support the development of other structured education programs, we make a range of recommendations.

RECOMMENDATIONS

To reduce the burden on the site educators more members of the MDT, including consultants, could attend the program training to foster greater buy-in and a team approach to facilitate sharing of the workload. To make this feasible, including containing cost, training of teams could be conducted at local sites rather than centrally in London. Furthermore, dedicated administrative support to organize venues, appointments, groups, and effective reminder systems would increase the likelihood of improved overall uptake, and would help with grouping the young people by age, as intended. Additional support for site educators in practising and sustaining quality of delivery would have been beneficial. Possible approaches could include: those associated with the successful DAFNE program,¹⁴ such as longer training, a greater focus in the training on improving group work skills, and an observation of CASCADE experts delivering the program; site level mentoring from CASCADE experts including feedback on site educators delivering trial runs; face-to-face mentoring from local colleagues, such as psychologists. In addition, before undertaking structured education programs, there may be a need to improve the knowledge base of some of the current pediatric diabetes service workforce, as levels of knowledge were very variable. Raising knowledge levels may be addressed by the development of a curriculum for professionals specifically in diabetes, ranging from a core curriculum (basic knowledge that all team members would be expected to know) to an extended curriculum (covering high level application of knowledge specific to individual team members). This finding may have relevance to other medical specialisms where structured education programs are being considered.

The uptake of the education sessions was low. For families the key issue was the challenge of fitting attendance into busy day-to-day routines. The education modules were offered in sessions independent of routine clinic appointments. Our data suggest that to improve accessibility it could have been advantageous to make the modules an integral part of routine clinic appointments, thereby overcoming the need for families to make additional hospital visits, with the implications this has for time away from school and work. This would require those in organizational administrative roles to assist with sustainable organizational adjustments required for extending clinic services. This finding and the suggestion that there should be greater 'buy-in' from the wider clinic team echo those in the broader literature on group-based programs.¹⁵ Furthermore in the study, participants had to have been diagnosed with diabetes for more than a year to meet the inclusion criteria for participation. Our data suggest that if the program was offered to families sooner after the initial diabetes diagnosis, this might lead to improved motivation to attend the groups. Additionally, offering this structured group education more universally might be more successful, including making the organization of groups by age more feasible, than targeting those with the poorest control of their blood glucose levels. It may be more

realistic to assume that those with the very poorest control might also require the greater flexibility and intensity that individualized interventions with a psychologist would offer. A summary of the key recommendations is presented in [box 1](#).

Box 1

Summary of key recommendations to improve training in, and delivery of, structured education sessions

- More involvement of the wider clinical team facilitated by local training;
- Greater mentoring of site educators by trainers;
- Practice sessions with feedback from trainers for site educators before going 'live' and time between training and delivery of first session kept to a minimum;
- More diabetes-specific training for the pediatric diabetes service workforce to guarantee a basic level of diabetes knowledge prior to training in the program;
- Dedicated administrative support to assist with organizing the sessions;
- Education sessions to be held within clinic time;
- Offer the sessions to all young people on clinic lists and soon after diagnosis.

STRENGTHS AND LIMITATIONS OF THE STUDY

It is a strength of the study that the process evaluation was unusually extensive and fully integrated into the main trial. Data were collected from all key stakeholders through a range of different methods throughout the different phases of the implementation of the intervention. Triangulation of findings enabled an evaluation of the implementation, barriers, and facilitators in relation to all aspects of implementation, operation, and perceived impact to be examined. It was also a strength that as a pragmatic RCT this intervention was evaluated in 'real-life' and representative settings. One limitation of the study was the impact of the research context on implementation, but steps were taken in the information and reassurance provided, methods, and timing of data collection to minimize effects as much as possible. Additionally, a major hindrance to the intervention was the lower than expected number of CASCADE groups run and the poor uptake of these groups by families. This might suggest a weakness in the intervention's pilot, which was not carried out within the same clinical contexts as the main trial. As such, opportunities to address challenges in organization and delivery were missed prior to, or through carefully managed processes within, the full trial.¹⁶ Experience from pragmatic studies of complex interventions such as CASCADE has yielded valuable new learning on the importance of particular investment in the developmental and piloting stages of complex interventions.¹⁷

CONCLUSION

The extensive multimethod process evaluation showed that the CASCADE structured education program was deliverable; however, improvements in clinical and administrative support, staff training, program content, and service structures to improve accessibility for families were required. The suggested improvements identified in this study all have resource implications, and thus any future research requires cost-benefit considerations. These findings give valuable information on what is required not only in CASCADE but also other similar programs to achieve their aims.

ACKNOWLEDGMENTS

The authors would like to thank all the young people, parents, and staff who took part in the study. They acknowledge the wider CASCADE evaluation team who measured the impact and economic aspects of the intervention and provided a user perspective.

REFERENCES

↵Royal College of Paediatrics and Child Health. National Paediatric Diabetes Audit 2010–2011. London: Royal College of Paediatrics and Child Health, 2012. Google Scholar

↵National Institute for Clinical Excellence (NICE). Diagnosis and management of type 1 diabetes in children, young people and adults. London: National Institute for Clinical Excellence, 2004. Google Scholar

↵Price KJ, Wales J, Eiser C, et al. Does an intensive self-management structured education course improve outcomes for children and young people with type 1 diabetes? The Kids In Control OF Food (KICK-OFF) cluster-randomised controlled trial protocol. *BMJ Open* 2013;3:e002429. doi:10.1136/bmjopen-2012-002429 Abstract/FREE Full Text Google Scholar

↵Murphy HR, Wadham C, Rayman G, et al. Approaches to integrating paediatric diabetes care and structured education: experiences from the Families, Adolescents, and Children's Teamwork Study (FACTS). *Diabet Med* 2007;24:1261–8. doi:10.1111/j.1464-5491.2007.02229.x CrossRef PubMed Web of Science Google Scholar

↵Coates V, Chaney D, Bunting B, et al. Evaluation of the effectiveness of a structured diabetes education programme (CHOICE) on clinical outcomes for adolescents with type 1 diabetes: a randomised controlled trial. *J Diabetes Metab* 2013;4:280. Google Scholar

↵National Health Service. Improving Quality. Best Practice Tariff for Paediatric Diabetes (cited 18 July 2013).

http://www.diabetes.nhs.uk/networks/paediatric_network/best_practice_tariff_for_paediatric_diabetes/Google Scholar

↵Christie D, Strange V, Allen E, et al. Maximising engagement, motivation and long term change in a Structured Intensive Education Programme in Diabetes for children, young people and their families: Child and Adolescent Structured Competencies Approach to Diabetes Education (CASCADE). *BMC Pediatr* 2009;9:57. doi:10.1186/1471-2431-9-57CrossRefPubMedGoogle Scholar

↵Viner R, Christie D, Taylor V, et al. Motivational/solution-focused intervention improves HbA1c in adolescents with type 1 diabetes: a pilot study. *Diabet Med* 2003;20:739–42. doi:10.1046/j.1464-5491.2003.00995.xCrossRefPubMedGoogle Scholar

↵Channon S, Huws-Thomas MV, Rollnick S, et al. The potential of motivational interviewing. *Diabet Med* 2005;22:353. doi:10.1111/j.1464-5491.2005.01321.xPubMedGoogle Scholar

↵George E, Iveson C, Ratner H. Problem to solution: brief therapy with individuals and families. London: BT Press, 1990.Google Scholar

↵Lewin S, Glenton C, Oxman AD. Use of qualitative methods alongside randomised controlled trials of complex healthcare interventions: methodological study. *BMJ* 2009;339:b3496.Google Scholar

↵Christie D, Thompson R, Sawtell M, et al. Structured, intensive education maximising engagement, motivation and long-term change for children and young people with diabetes: a cluster randomised controlled trial with integral process and economic evaluation the CASCADE study. *Health Technol Assess* 2014;18:1–202. doi:10.3310/hta18200CrossRefPubMedGoogle Scholar

↵Lofland J, Lofland H. Analyzing social settings: a guide to qualitative observation and analysis. Belmont, CA: Wadsworth, 1995.Google Scholar

↵DAFNE Study Group. Training in flexible, intensive insulin management to enable dietary freedom in people with type 1 diabetes: dose adjustment for normal eating (DAFNE) randomised controlled trial. *BMJ* 2002;325:746. doi:10.1136/bmj.325.7367.746Abstract/FREE Full TextGoogle Scholar

↵Novik G, Sadler S, Knafel K, et al. In a hard spot: providing group prenatal care in two urban clinics. *Midwifery* 2013;29:690–7. doi:10.1016/j.midw.2012.06.013CrossRefPubMedGoogle Scholar

↵Cooper C, O'Cathain A, Hind D, et al. Conducting qualitative research within Clinical Trials Units: avoiding potential pitfalls. *Contemp Clin Trials* 2014;38:338–43. doi:10.1016/j.cct.2014.06.002CrossRefPubMedGoogle Scholar

↵Medical Research Council. *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council, 2006.Google Scholar

FOOTNOTES

Contributors MS wrote the first draft of the manuscript and LJ, MW, FS, and KH made substantial contributions to the text of subsequent drafts. AI, LB, MK, RT, and DC reviewed and edited the manuscript. MS, LJ, KH, AI, and MK collected the data. These authors and MW, FS, and LB carried out data analysis. DC and RT developed the intervention and assisted with interpretation of data. MS is the guarantor and, as such, had full access to all of the data reported in this paper and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding This work was supported by NIHR Health Technology Assessment programme—HTA—06/44/05.

Competing interests None declared.

Patient consent Obtained.

Ethics approval The study was approved by the University College London (UCL)/UCLH Research Ethics Committee. Site-specific approval was granted at each site.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

APPENDIX 3

CANDIDATE PUBLICATION 3

Chess in Schools

Evaluation report and executive summary

July 2016



Education
Endowment
Foundation

Independent evaluators:

John Jerrim, Lindsey Macmillan, John Micklewright, Mary Sawtell
and Meg Wiggins





The Education Endowment Foundation (EEF) is an independent grant-making charity dedicated to breaking the link between family income and educational achievement, ensuring that children from all backgrounds can fulfil their potential and make the most of their talents. The EEF aims to raise the attainment of children facing disadvantage by:

- identifying promising educational innovations that address the needs of disadvantaged children in primary and secondary schools in England;
- evaluating these innovations to extend and secure the evidence on what works and can be made to work at scale; and
- encouraging schools, government, charities, and others to apply evidence and adopt innovations found to be effective.

The EEF was established in 2011 by the Sutton Trust as lead charity in partnership with Impetus Trust (now part of Impetus - Private Equity Foundation) and received a founding £125m grant from the Department for Education.

Together, the EEF and Sutton Trust are the government-designated What Works Centre for improving education outcomes for school-aged children.



About the evaluator

The project was independently evaluated by a team led by University College London, including John Jerrim, Lindsey Macmillan, John Micklewright, Mary Sawtell and Meg Wiggins.

The lead evaluator was John Jerrim.

e. j.jerrim@ucl.ac.uk

Contents

About the evaluator	1
Contents	2
Executive Summary	3
Introduction	5
Methods	10
Impact evaluation	20
Process evaluation	36
Conclusion	54
References	57
Appendix A. School consent form to access the National Pupil Database (NPD)	58
Appendix B. Questions CSC were asked to estimate costs	59
Appendix C: Padlock rating.....	60
Appendix D: Cost rating	60

Executive Summary

The project

Chess in Primary Schools is a whole-school approach to teaching primary school children how to play chess. Children take 30 hours of chess lessons delivered by a tutor who is an experienced chess player, and the school is given the option to set up a chess club as a lunchtime or after-school activity. Chess classes are delivered during the school day and are expected to replace subjects such as music or PE.

The intervention was evaluated using a two-armed randomised controlled trial. The trial took place over the 2013/2014 academic year and assessed the impact of one year of Chess in Primary Schools on the mathematics attainment of pupils in Year 5. It was an effectiveness trial, with the intervention tested under realistic conditions in a large number of schools. This study looks at whether the intervention had an impact on attainment one year after the intervention had ended in June 2015. One hundred schools across 11 local education authorities (LEAs) in England participated in the trial, a total of 4,009 pupils. A process evaluation was also carried out to answer questions about implementation and to help explain the findings of the trial. The programme was delivered by the education charity Chess in Schools and Communities (CSC).

Key Conclusions

1. There is no evidence that the intervention had a positive impact on mathematics attainment for the children in the trial, as measured by Key Stage 2 scores one year after the intervention ended. The same is true for science and reading.
2. There is no evidence that the intervention had a positive impact on Key Stage 2 scores for children eligible for free school meals (FSM).
3. Although a current school teacher is allocated to every chess class, it is desirable for the tutors themselves to have strong class management and teaching skills. Without these, it was difficult to ensure that all children were suitably engaged in the chess lessons.
4. For successful implementation, class teachers need to work closely with the tutor and actively contribute to the intervention. It was felt that classes were less effective if the teacher did not actively take part, or was present only at the beginning and end of the class.
5. Half of the pupils who participated in the trial said that they liked the chess lessons a lot, and only 8% reported that they didn't like them. School teachers were very positive about the intervention and its impact on pupils' skills and behaviour.

How secure are the findings?

Security rating awarded as part of the EEF peer review process

Findings from this study have high security. The study was a large and well-designed clustered randomised controlled trial (RCT). It was an effectiveness trial, which means it aimed to test the intervention under realistic conditions in a large number of schools. Relatively few pupils were lost to the analysis and the pupils who were allocated to receive the intervention were similar to the pupils in the comparison group. There were no substantial threats to the validity of the results.

What are the findings?



Pupils, headteachers and class teachers were generally very positive about the intervention. In particular, pupils liked playing games of chess with their friends, and class teachers welcomed the enthusiasm of the tutors for sharing their expertise. School staff perceived that the chess lessons had a positive impact on maths ability, as well as on a range of important skills for learning such as concentration. What pupils liked least was tutors 'talking too much' and some teachers had concerns about the level of tutors' teaching skills. There were some departures from the intended delivery of the programme—primarily the level of class teacher engagement, which was lower than expected. Moreover, some schools reduced the number of maths lessons in the timetable in order to accommodate the chess lessons. Two key areas for intervention improvement emerged from the study. These were: (a) improving the teaching skills of tutors to help them keep all children engaged—specifically, improving their ability to manage difficult behaviour and manage classes where pupils had varying levels of ability; and (b) increasing the amount of tutor/class teacher liaison.

Despite the generally positive feedback received from schools from the process evaluation, the impact evaluation results found no evidence that the Chess in Primary Schools programme raised children's attainment in their Key Stage 2 exams. Indeed, the difference between the treatment and control arms was essentially zero. A similar impact was found for pupils eligible for free school meals, and for boys and girls. This is in contrast to the only other large-scale RCT of the impact of chess on educational attainment by Boruch and Romano (2011), who detected a substantial effect for primary school children in Italy and to another recent study by Gumede and Rosholm (2015), which found a positive effect of chess on primary school children's achievement in Denmark (effect size 0.15). The reasons for the differences could include the fact that this study measured the impact after one year, that this study used high stakes national tests, and the English setting.

How much does it cost?

The cost of delivering the intervention to two classes of Year 5 pupils is approximately £1,900, or £32 per pupil. The majority of this is to contribute towards CSC's costs of delivering the chess lessons (£1,200) and setting up the after-school chess club (£600).

Summary Table

Group	Effect size (95% confidence interval)	Estimated months' progress	Security rating	Cost rating
Pupils in CSC schools	0.01 (-0.15, 0.16)	0 months		£
Pupils in CSC schools eligible for FSM	0.01 (-0.18, 0.19)	0 months		£

Introduction

Intervention: The Chess in Primary Schools programme

The programme involved the CSC charity introducing chess lessons as part of a standard school day to Year 5 children within primary schools in England. The rationale behind the evaluation was that chess may help increase children's concentration, their ability to think strategically, and their selfconfidence (see the 'logic model' in Figure 4 for further details). This would, in turn, lead to a long-term improvement in their academic achievement.

The intervention was delivered by fully trained tutors following a standardised 30-hour curriculum, details of which can be found at http://www.chessinschools.co.uk/sample_curriculum.htm. All chess tutors were chess specialists and did not necessarily have a teaching background. CSC regularly run one-day training courses aimed at anyone involved in school chess, including CSC tutors and class teachers. During the study intervention year CSC also ran a number of weekend seminars for tutors, at which the intervention was examined and a discussion process initiated on methods of enhancing the classroom delivery. Tutors were also able to exchange teaching methods and received useful tips on classroom management from some schoolteachers who attended.

Each participating school was asked to designate a teacher (or teaching assistant) who would help the CSC tutor to run the intervention in the class. This teacher or teaching assistant was required to attend a training seminar run by the CSC charity, and had full access to the CSC curriculum. The tutor was encouraged to discuss each lesson in advance with the class teacher / teaching assistant, in person or by email. Each school was also sent chess sets for classroom use, workbooks and curriculum books and also, later in the year, each child received a chess set and chess book.

The chess lessons were delivered as part of a regular school day. This meant that schools were to replace one regularly scheduled lesson to make room for the Chess in Primary Schools intervention. Schools were asked by CSC not to replace a maths or English lesson. Common lessons to be replaced were 'topic' or humanities, music and PE.

Whole-class teaching was used to deliver the Chess in Primary Schools programme. During lessons, material was presented using either a chess demonstration board or via the whiteboard. In order to use the whiteboard, each tutor was given specialist chess software, with the curriculum converted into a proprietary file format. Tutors had learning plans and objectives for each lesson, as well as worksheets for pupils. In each lesson, children shared a chess set on the desk to practise moves or, later, to play complete games in pairs. Tutors were encouraged to talk for no more than 15 minutes before allowing children to practise what they had been taught. In each school a chess club could optionally be set up at lunchtime or after school during the intervention period. Time at this chess club was additional to the 30 hours' taught curriculum time. Schools were encouraged to do this themselves. However, because of a lack of expertise, input from the CSC tutor was often required.

The game was taught piece by piece and visualisation of moves was required from lesson 2. By lesson 10, more abstract concepts such as 'check' and 'checkmate' were introduced. By the end of the first term, children were expected to be able to begin to play chess. Then, by the end of the second term, most children were expected to be able to play a reasonable game of chess. At the end of the school year, CSC organised competitions locally for groups of schools or within individual schools.

A 'business as usual' approach was used in control schools. This meant that no formal chess lessons were to be delivered (though if an after-school chess club already existed, this would continue to run). These schools were not allowed to access the intervention in either the 2013/14 or 2014/15 academic year. There was a small amount of crossover between treatment and control groups; six treatment schools did not deliver the intervention, while one control school gained partial access to the chess treatment. This was accounted for with a contamination-adjusted intention to treat (CA-ITT) analysis to supplement the main intention to treat (ITT) analysis.

Box 1 provides a summary of the intervention, including details of the materials and procedures used, how the intervention was delivered, and the amount of chess instruction it was intended children would receive.

Box 1: TIDieR checklist

1. **Brief name.** Chess in Primary Schools
2. **Why: Rationale, theory and/or goal of essential elements of the intervention.** Chess would help to increase children's level of concentration, self-confidence and ability to think strategically. This would, in turn, lead to an improvement in their academic achievement.
3. **Who: Recipients of the intervention.** Year 5 (age 9/10) pupils within selected local education authorities in England.

- 4. What: Physical or informational materials used in the intervention.** Chessboards, chess workbooks, chess software for whiteboard, CSC developed curriculum hanging demonstration board, classroom tables/chairs.
- 5. What: Procedures, activities and/or processes used in the intervention.** Pupils are taught chess, as part of the school curriculum, by a trained chess tutor using the CSC curriculum. The CSC curriculum contains detailed 1-hour lesson plans that include mini-games and worksheets.
- 6. Who: Intervention providers/implementers.** The intervention was provided by the charity Chess in Schools and Communities.
- 7. How: Mode of delivery.** Face-to-face whole class delivery to children.
- 8. Where: Location of the intervention.** Within primary school classrooms in England.
- 9. When and how much: Duration and dosage of the intervention.** During the 2013/14 academic year. Children were to receive 30 chess lessons of 1 hour spread over the academic year.
- 10. Tailoring: Adaptation of the intervention.** The tutors were provided with the CSC curriculum as the foundation for lessons, but were allowed to adapt lesson plans to suit individual classes.

Background evidence

The majority of studies that link chess to academic attainment have been conducted outside of England and include self-selecting intervention groups (e.g. Achiego et al., 2012). They have also tended to use 'low-stakes' tests (for which children in the control group are likely to be less motivated than those in the treatment group, as neither they nor their schools have anything riding upon the results). To our knowledge, only one randomised controlled trial of chess has been conducted (Boruch and Romano, 2011). This tested how 30 hours of chess tuition, provided by qualified tutors, influenced 8–9-year-olds' educational achievement in Italy. The study included 123 classes, randomly assigned to receive chess in either the 3rd or 4th grade. The intervention was found to increase mathematics achievement by an effect size of 0.34, though again 'low-stakes' tests were used. However, another recent quasi-experimental study by Gumede and Rosholm (2015) also found a positive effect of chess on primary school children's achievement in Denmark (effect size 0.15).

The rationale of this evaluation was to test the Chess in Primary Schools programme within the English setting. It was an effectiveness trial, with the intervention delivered at scale. As the Chess in Primary Schools programme is already well developed and widely used in schools, it was decided that a large-scale effectiveness trial was appropriate. The trial has a number of advantages over existing studies, including the use of 'high-stakes' tests, and a focus upon *medium term* effects of the intervention.

Evaluation objectives

The main question that the impact evaluation attempted to address was '*What is the impact of chess in schools on children's achievement in mathematics?*'. This was supplemented by a series of additional questions, including:

- What is the impact of teaching Year 5 children chess upon their Key Stage 2 reading and science test scores?
- What is the impact of teaching Year 5 children chess upon different mathematics sub-domains (e.g. mental arithmetic)?
- What impact does teaching FSM children how to play chess have upon their Key Stage 2 attainment?
- Does the impact of teaching chess differ between boys and girls?
- Is there any evidence that the Chess in Primary Schools programme has differential effects across the achievement distribution?

The process evaluation sought to answer the following questions:

- How feasible and acceptable is it for chess tutors to implement a 30-week classroom chess intervention in Year 5 of primary school? Could teachers who attended training and helped with the intervention continue to teach chess afterwards?
- How feasible and acceptable do teachers and headteachers feel it is for primary school children to play chess in class as part of the curriculum?
- What are the views, on the intervention, of the children who were offered it? How do these views vary by subgroup (e.g. boys vs. girls)?
- What are staff perceptions of the current and possibly sustained impact of the intervention on children's educational attainment? How do they think it affects different subgroups? How do they think it impacts on other matters such as class cohesion and school ethos? What are their perceptions of facilitators and barriers to impact? How scalable do they think the intervention is? What are their suggestions for change if the intervention was to be more widely implemented?

Project team

John Jerrim: Principal Investigator. Led the trial design, data analysis and writing of the final report. Overall management of the project.

Lindsey Macmillan: Assisted with trial design, data analysis and production of final report.

John Micklewright: Assisted with trial design.

Mary Sawtell: Joint-lead on the process evaluation design and analysis.

Meg Wiggins: Joint-lead on the process evaluation design and analysis.

Ethical review

The evaluation of the Chess in Primary Schools project was submitted to the Institute of Education ethics committee. Ethical approval was granted on 17 May 2013 (code FPS 504). School level consent has been obtained to conduct the trial and to access pupils' data from the National Pupil Database (NPD).

Trial registration

The protocol for this study is published online at:

<http://educationendowmentfoundation.org.uk/library/chess-in-schools-protocol/>

The trial has been registered with the independent ISRCTN website at: <http://controlled-trials.com/ISRCTN33648117>

The trial registration number is ISRCTN33648117 and the DOI is 10.1186/ISRCTN33648117

Methods

Trial design

The study was designed as a clustered randomised controlled trial (RCT). At the start of the project, the evaluation team considered three options for randomisation of the intervention: (a) at the *pupil* level, (b) at the *class* level, and (c) at the *school* level. Option (a) was immediately ruled out due to the Chess in Primary Schools programme being designed as a group activity. We therefore focus on options (b) and (c).

Randomisation at the class level was deemed likely to be a powerful statistical design. (This was the approach taken in the Italian study of the impact of chess on attainment by Boruch and Romano, 2011.) However, the evaluation team decided this was outweighed by the following limitations.

First, concerns remained over possible ‘contamination’ between treatment and control classes. As each school would contain children in the same year in the two groups, it was deemed possible that children learning to play chess could encourage friends or siblings in the control group to also play chess outside of school. If chess does indeed have a positive effect on the outcome, such contamination would downwardly bias the estimated impact of the intervention. Second, parents may object to children receiving different ‘types’ of education within the same school as a result of random assignment of classes. For example, a parent who believes that chess will have a positive impact upon attainment may have been upset that their child had been assigned to the control group, when the child next door was getting the treatment in another class. Third, the need to alter the curriculum for one class but not another within the same year could present schools with an organisational problem. These second and third issues might have reduced the willingness of schools to take part in the trial, threatening both attrition and external validity.

Thus option (c) was chosen: a clustered randomised controlled trial, with randomisation at the school level. All forms within the selected year in a treatment school would receive the intervention; none would in the control schools. Moreover, control schools would continue to use ‘business as usual’ teaching, with the Chess in Primary Schools programme becoming available to them two years after the intervention began. This design is likely to provide less statistical power—but all three potential problems with class randomisation were likely to be greatly diminished.

Outcome measures

The primary outcome is pupils’ Key Stage 2 maths test scores. KS2 scores are derived from a national examination that children sit in England at the end of primary school (when pupils are typically age 10 or 11). It is a reliable, externally valid measure that is a strong predictor of children’s later educational outcomes. It is also a ‘high stakes’

test for schools, who are ranked in publicly available league tables by their pupils' performance. This test is not specific to the Chess in Primary Schools intervention and is marked blind to treatment. This outcome was pre-specified as part of the evaluation protocol. Maths was chosen because this is the academic area where Boruch and Romano (2011) reported a substantial effect.

Secondary outcomes include (i) performance in Key Stage 2 English tests, (ii) performance in Key Stage 2 Science tests (where available), and (iii) performance in sub-domains of the Key Stage 2 Maths test. The latter are known as 'paper A', 'paper B', and 'mental arithmetic', with the following links providing the three test papers that children took in June 2015:

Paper A:

http://www.satspapers.org/SATs%20papers/SATs%20Papers%20pdf%20format/Maths%20SATs%20papers/2015%20Maths/2015_KS2_L3-5_mathematics_paper1_PDFa.pdf

Paper B:

http://www.satspapers.org/SATs%20papers/SATs%20Papers%20pdf%20format/Maths%20SATs%20papers/2015%20Maths/2015_KS2_L3-5_mathematics_paper2_PDFa.pdf

Mental arithmetic:

http://www.satspapers.org/SATs%20papers/SATs%20Papers%20pdf%20format/Maths%20SATs%20papers/2015%20Maths/2015_KS2_L3-5_mathematics_mentalmathematics_transcript_PDFa.pdf

Note that the Chess in Primary Schools intervention was delivered while children were in Year 5 (age 9/10). Key Stage 2 tests (the outcome) were conducted at the end of Year 6. Hence outcomes have been measured one year after the intervention finished. The trial has therefore been designed to detect a *medium term* effect of the intervention. **Baseline test**

Children's Key Stage 1 (KS1) maths, reading, writing and science test scores were used to measure children's academic achievement prior to the Chess in Primary Schools intervention. These are based upon teacher assessments of pupils when they were age 7—and thus before schools were randomly assigned to treatment and control groups. Indeed, at the point these baseline tests were conducted, teachers would have been unaware that the Chess in Primary Schools trial would take place. These baseline scores are used to (i) investigate balance between treatment and control groups in terms of prior attainment, and (ii) increase power and reduce any imbalance between treatment and control groups in the statistical analysis.

Participant selection

The Institute of Education (IoE) and Chess in Schools and Communities (CSC) teams first identified specific local education authorities (LEAs) in England where CSC had capacity to deliver the intervention. The LEAs selected were:

- City of Bristol
- Hackney
- Hammersmith and Fulham
- Leeds
- Liverpool
- Middlesbrough
- Newham
- Sefton
- Sheffield
- Southwark
- Thameside

The Institute of Education then produced a list of all primary schools within these LEAs. Private schools and schools where CSC already operated were excluded. For logistical reasons, it was also agreed that any primary school with four-form entry would be excluded from the evaluation. Schools with more than 90 pupils aged 11 were thus removed from the sampling frame. This was working on the assumption that there were approximately 30 pupils per class within primary schools, and that year group size within schools would not significantly change within a short space of time.

The sampling frame was further restricted to schools with a high intake of disadvantaged pupils, based upon the percentage of children receiving free school meals (FSM). Schools were only selected if at least 37% of their children had either been eligible for FSM within the last six years or had been looked after by the local authority continuously for six months¹. Thus the population of interest was defined as Year 5 state school pupils within the selected LEAs, who attended a one, two or three form entry primary school, which had a high proportion of disadvantaged pupils, and was not currently enrolled in the Chess in Primary Schools programme.

This final list contained 442 schools. CSC were then asked to recruit 100 of these schools by the third week of July 2013. CSC sent all schools an information pack. Those that agreed to take part in the trial completed a consent form to participate in

¹ The figure of 37% was decided upon so that the population list given to CSC would contain a population of approximately 450 schools from which they could recruit into the trial. ² A value of 0.15 for the ICC was chosen after the team conducted an analysis of within and between school variation in Key Stage 2 test scores within the National Pupil Database. ³ The figure of 60 pupils was based on the assumption of most recruited schools being twoform entry, with each form containing 30 pupils.

the study and to allow access to data from the National Pupil Database (NPD). Ninety-two schools were recruited into the trial by this date. A further eight schools were recruited by September 2013, and were also included in the trial (bringing the total to 100). School-level consent to participate in the trial, and to allow the evaluation team access to the NPD data, was obtained from schools prior to randomisation. All children in the Year 5 treatment schools were required to participate in the programme to avoid selection problems. **Sample size**

The evaluation team regarded 100 schools as the minimum necessary to detect an effect of approximately 0.18 of a standard deviation in Key Stage 2 (KS2) mathematics test scores. This calculation assumed:

- i. an intra-cluster correlation (ICC) of 0.15 at the school level²;
- ii. equal cluster sizes of 60 Year 5 pupils per school³;
- iii. 40% of the variation in KS2 maths test scores would be explained by baseline covariates²; and
- iv. 80% power for a 95% confidence interval.

Table 1 provides estimates of the ICC for the actual sample of schools/pupils that took part in the study. Estimates are presented for baseline (KS1 average points score) and follow-up (KS2 maths) tests, when using either a fixed or random school-level effect. The ICC for KS1 average point scores (APS) was 0.08 when using a fixed effects model. The analogous ICC for KS2 maths was 0.13. In the results section, we illustrate that 45% of the variance in KS2 maths test scores can be explained by the baseline covariates. Using these figures in place of (i) and (iii) above, we calculate the minimum detectable effect in this trial was approximately 0.16 (see Table 3 below for further details).

Table 1: Estimated inter-cluster correlation

	Fixed effect	Random effect
Key Stage 1 APS	0.08	0.05
Key Stage 2 Maths	0.13	0.11

Note: Figures refer to the proportion of the variation in pupils' test scores occurring between schools.

Randomisation

The trial was designed as a stratified, clustered randomised controlled trial—with random allocation occurring at the school level. Schools were first separated (stratified) into groups based upon (i) the percentage of pupils achieving level 4 or above at the end of KS2 in both English and mathematics, and (ii) the percentage of

² A value of 0.4 was chosen after the team conducted an analysis of the association between Key Stage 1 and Key Stage 2 test scores within the National Pupil Database.

current KS2 pupils who had been eligible for FSM in the last six years. The schools were categorised into three strata for each variable and then the variables were cross-tabulated to create the following nine strata:

1. Low achieving–low FSM
2. Low achieving–middle FSM
3. Low achieving–high FSM
4. Middle achieving–low FSM
5. Middle achieving–middle FSM
6. Middle achieving–high FSM
7. High achieving–low FSM
8. High achieving–middle FSM
9. High achieving–high FSM

A tenth stratum was then included to incorporate the eight schools that were recruited into the trial between the end of July and September 2013:

10. ‘Late’ recruited schools

The number of schools within each stratum can be found in Table 2.

Table 2: Number of schools within each stratum

Strata ID	Average achievement	% Free school meals	Schools per strata
1	Low	Low	7
2	Low	Medium	12
3	Low	High	12
4	Medium	Low	13
5	Medium	Medium	11
6	Medium	High	7
7	High	Low	11
8	High	Medium	10
9	High	High	9
10	Late recruitment	Late recruitment	8

Within each stratum a random number was then drawn from a uniform distribution for each school. The schools in the bottom half of the random draw distribution, *within each stratum*, were assigned to the control group. Schools with a number in the top half of the random draw distribution were assigned to treatment. If the stratum contained an odd number of schools, then the school with the median random draw was randomly assigned to treatment or control. STATA version 12 has been used to generate all random numbers. Note that all schools in strata 1 to 9 were

randomly assigned on the same day in the third week of July 2013. Schools in stratum 10 were randomised on a separate day in August 2013.

The creation of the random number sequence and allocation of participants was done by Dr John Jerrim.

Analysis

The analysis strategy used intention to treat. Analysis of whether the intervention was effective or not was based upon the following OLS regression model:

$$Y_{ij}^{Post} = \alpha + \beta.Treat_j + \gamma.Y_{ij}^{Pre} + \varepsilon_{ij} \quad (1)$$

where:

Y^{pre} = children's KS1 test scores in maths, reading, writing and science

Y^{post} = children's KS2 maths test score

Treat = a binary variable indicating whether the child was enrolled in a treatment or control school (0 = control; 1 = treatment).

ε = error term (with children clustered within

school) i = child j = school j

To account for the clustering of pupils within schools, the STATA survey (svy) command is used to make Huber-White adjustments to the estimated standard errors. The coefficient of interest from equation 1 is β – is there a positive effect of the Chess in Primary Schools treatment?

After our main analysis, we re-estimate model 1 (i) separately for boys and girls, and (ii) separately for FSM pupils. The same analysis process has been followed for the secondary outcomes (sub-components of the Key Stage 2 maths tests, Key Stage 2 English scores, Key Stage 2 Science scores)³.

The evaluation team has also conducted an 'on-treatment' analysis, where we investigate whether the effectiveness of the intervention varies by how it was implemented within schools. This part of the analysis was not pre-specified in the study protocol, but was undertaken in order to investigate whether there was any difference in the effect of the treatment by the fidelity of the treatment. In order to conduct this analysis, children within treatment schools were asked how much they liked the chess lessons that were delivered. Each chess tutor was then assigned to one of three categories (high, medium, low) depending upon the proportion of

³ If an observation is missing Key Stage 2 test score data, it has been excluded from the analysis.

children that they taught who responded that they ‘liked the lessons a lot’⁴. We label this variable ‘chess tutor quality’. In our analysis, mean post-test scores for children taught by teachers within these three ‘tutor quality’ groups are compared to mean post-test scores for the control group. The intuition is that liking chess is one of the key ‘change mechanisms’ through which we anticipate an effect to occur; thus greater levels of pupil enjoyment is likely to indicate a more fertile treatment. In an additional ‘on-treatment’ analysis, we also illustrate whether the intervention was more effective when the class teacher attended the CSC one-day training workshop (as anticipated in the study protocol).

In addition to the intention-to-treat (ITT) analysis described above, we also undertake a contamination adjusted intention-to-treat analysis. This was done in order to test the sensitivity of our results to the small amount of crossover between treatment and control groups that occurred during the trial. Further details are provided below.

Costs

Information of costs was gathered directly from the programme developer (i.e. the charity Chess in Schools and Communities). Specifically, the evaluation team asked CSC a series of questions regarding various aspects of the resources needed to run the intervention (see Appendix B for a list of questions asked).

The response of CSC to these questions was then used to calculate the cost of a school participating in the programme next academic year. This figure varies by the size of the school, as CSC requests a larger contribution to their overheads from bigger schools. As regular class teachers were expected to complete a one-day training course, a figure of £200 was added to this value for each teacher from a school who would attend (under the assumption this would cover the costs of employing a supply teacher)⁷. These figures were then added together to give a total cost. The total cost was then divided by the number of pupils, under the assumption of 30 pupils per school form.

Implementation and process evaluation

The process evaluation was integral to the trial. It was designed with three key purposes: to assess the fidelity of delivery of the intervention, to answer questions related to the feasibility of the intervention, and to support understanding of the results of the impact evaluation.

⁴ If less than a third responded positively, the tutor was assigned to the low group. If between one-third and two-thirds responded positively, the tutor was assigned to the middle group.

Constructing a logic model

A logic model was developed to clarify assumptions on CSC's views of the theory of change of the programme and to provide a framework to support the evaluation including the assessment of fidelity and explanation of findings.

The construction of a logic model was undertaken by the research team, using a three-stage Delphi consultation exercise, designed to achieve consensus within a group of eight experts in the CSC programme. The consultation was carried out by email without individuals conferring or seeing the responses of others in the group. The first stage asked for views on what the different components of the logic model were at each stage of the causal pathway. The components submitted by the participants were consolidated by the research team. The eight participants were then asked to rank the listed components in order of importance. The research team analysed the ranked lists and constructed a draft logic model that reflected the combined views of the experts. This draft was emailed to the participants with a request for

Teachers in the high group had at least two-thirds of children reporting that they liked the chess lessons a lot.

⁷ Following EEF guidance, we spread this cost over three years.

any amendments to be provided. The final version of the logic model can be found in the process evaluation results section (see Figure 4).

Pre-intervention data collection

Baseline headteacher survey

A headteacher survey was conducted with all the schools enrolled in the study in July 2013, immediately prior to randomisation. This short, paper-based survey asked about any current or recent chess playing in the school, headteachers' own chess playing experience and their keenness for taking part in the trial. The rationale for this survey was to be able to assess the general level of chess interest and exposure in each school prior to intervention delivery. Non-respondents were sent two email reminders which included an invitation to complete the survey with a researcher over the telephone.

Observation of CSC training for tutors and teachers

During the set-up phase of the trial, two members of the research team observed, together, a CSC one-day training course, aimed at teachers, prospective tutors and anyone involved in school chess. The two researchers then carried out one observation each of two further oneday training courses to which teachers from treatment schools had been invited. Free-form observation notes were taken by the researchers.

*Data collection during and immediately after intervention delivery**Observations of chess lessons*

Observations were carried out in four schools approximately half-way through the intervention delivery period (March/April 2014). The four schools were purposefully selected to ensure a range of the following:

- location in the country;
- number of classes in a year group;
- previous levels of chess exposure and interest (as assessed by the baseline headteacher survey); and
- tutor factors (general teaching experience; current level of chess playing; number of years employed by CSC; whether they worked individually or as a pair; and gender).

The aims of the non-participatory observations were to provide information on: how the intervention was delivered, with a particular focus on fidelity; the acceptability by all stakeholders; and barriers and facilitators to delivery. Two evaluation team researchers carried out observations of one-hour chess lessons in eight Year 5 classes in the four selected schools. The researchers completed a semi-structured proforma during the observation, which included prompts for the various inputs listed in the logic model. Immediately after the lesson the observer had a brief discussion with the class teachers and the CSC tutor to clarify any issues arising from the observation.

Teacher survey

An online survey of all Year 5 teachers in intervention and control schools was carried out in June and July 2014 when intervention delivery was nearly complete. Teachers in schools in both trial arms were asked questions about themselves (e.g. gender and years of teaching experience); their class, including the amount of support they had in the classroom; other interventions during the year aimed at raising maths and literacy attainment; and numbers of pupils with particular needs such as special educational needs (SEN) and English as an additional language (EAL). Intervention teachers were also asked questions on the acceptability, feasibility and sustainability of the intervention and on their perceptions of impact. Two email reminders were sent to all non-responders which offered the option of completion over the telephone. An additional paper version of the questionnaire was sent by post with an accompanying prepaid reply envelope to all remaining control teachers who had not responded.

Stakeholder interviews

In-depth telephone interviews were conducted during July 2014 with two tutors, two headteachers, and five class teachers (from four different treatment schools). These were audio taped with the permission of the interviewee. Participants were

purposively selected, based on survey responses, to provide insights into key themes emerging from the survey and observation data. Notes were made during and immediately after the telephone interview to capture the key points. The audio recording was used as a check where there was uncertainty or to extract a particular quote. A further five teachers, two headteachers and five tutors were briefly interviewed, face to face, during observation site visits. Notes of key points were made by the researchers immediately after these discussions.

Observation of CSC seminars for tutors

A further evaluation activity was a one-day observation at each of two weekend seminars organised by CSC for tutors. The aim of these seminars was to support tutors in developing their teaching and classroom management skills. The programmes for these events included presentations by external experts, and tutors and others involved with CSC sharing personal experiences and tips. Free-form notes were taken by the researcher observing the seminars.

Data collection in the year following intervention delivery

Pupil survey

A short pupil survey was carried out in treatment schools in February 2015, approximately seven months after the chess lessons ended. Participation was optional for the students, and parents were given the opportunity to opt them out of this exercise. The survey was paper based and self-completion, with administration by class teachers. Packs of questionnaires, with accompanying guidance for teachers, were sent to the current teacher (Year 6) of each class that had received the chess lessons the previous year. The survey included closed questions on acceptability of the lessons; chess playing prior to the lessons; any chess playing since the chess lessons ended. Free text boxes were provided for pupils to write about their views on the best and worst aspects of the chess lessons.

Interview with intervention provider

A face-to-face semi-structured audio-taped interview with a senior member of the CSC head office team was conducted in May 2015. The main aim of this final data collection exercise was to explore themes that had emerged from other data sources from the perspective of the providers of the intervention. Questions asked covered: views on the process of overseeing the delivery of the intervention (including training and support for tutors); what was learned from the process; and whether any associated subsequent changes had been made or planned. As with the other stakeholder interviews, notes were made immediately after the interview. The tape recording was used for reference as required.

Analysis

Framework analysis was used for the analysis of the qualitative data from interviews and observations. This involved constructing frameworks based on key themes that answered the main research questions. This method allowed exploration of the data by both theme and respondent-type, enabling identification of patterns and associations across themes and types of respondents.

Descriptive statistical analyses of the teacher, headteacher, tutor and pupil surveys was carried out using SPSS V22. Chi-square tests were used to measure statistical significance.

Using data from across process evaluation sources, measures of intervention dose and quality were constructed for each school.

Timeline

Date	Activity
June 2011	Sample children's Key Stage 1 tests conducted
March – July 2013	Schools recruited (8 late schools recruited in August 2013)
July 2013	Schools assigned to treatment or control group (8 additional schools in August 2013).
October 2013 – July 2014	Chess in Primary Schools programme delivered in treatment schools
June 2015	Key Stage 2 (post-tests) conducted
October 2015	Analysis conducted

Impact evaluation

Participants

Sample allocation

Figure 1 provides details of sample allocation and attrition. One hundred schools were recruited to participate in the trial. Schools were randomly allocated to treatment (n = 50) and control (n = 50) groups.

All Year 5 children enrolled in the 100 participating schools in the trial on 3 October 2013 were considered to be part of the Chess in Primary Schools trial. (This was the date of the autumn school census in 2013.) Information on school enrolment on this date was drawn directly from the National Pupil Database (NPD). A total of 1,954 children were enrolled in the 50 control schools and 2,055 in the 50 treatment schools.

Missing data at baseline

Pupils' KS1 maths, reading, writing and science test scores were taken directly from the NPD. Information was missing for a small number of pupils who were not enrolled in a school in England at age 7 or where there were problems linking NPD data over time⁵. KS1 data was available for a total of 3,775 (94%) of the 4,009 children within the 100 participating schools. A 'missing' dummy variable is included in the OLS regression model to ensure these observations are not dropped for our analysis.

Attrition between intervention and post-test

The schools and children recruited into the trial were tracked using the NPD. Pupils who moved to a different school could be tracked via their unique pupil number (UPN) and were included in the final analysis. KS2 test score data could be linked for 3,865 of the 4,009 pupils initially recruited into the trial (see Figure 1). This group of pupils forms our final analysis sample.

Contamination

Six out of the 50 schools assigned to the treatment group dropped out of the Chess in Primary Schools programme before the intervention had begun.

One control school was unwilling to accept their assigned group and delivered chess lessons to their Year 3 pupils. Although chess lessons were not provided to the Year 5 pupils who were the intended controls, there is nevertheless an element of non-compliance.

To summarise, six of the schools who were meant to receive the Chess in Primary Schools treatment did not, while one control school managed to (partially) gain access to the intervention. As per our study protocol, our main analysis will follow an intention-to-treat (ITT) approach. This is where treatment and control groups are defined based upon their initial random allocation. However, we also present alternative estimates applying a contamination adjusted intention-to-treat (CA-ITT) methodology. This is an instrumental variable (IV) approach, where initial treatment/control allocation is used as an IV for actual receipt of the intervention. The key assumption is that initial random allocation (the IV) is strongly associated with the probability of actually receiving the intervention, but is not independently associated with the outcome (KS2 scores). This assumption is likely to hold as the extent of non-compliance is relatively small, meaning that initial allocation will strongly predict who actually received the treatment, and there is no reason to believe the IV and the outcome are associated (as the IV is random assignment to treatment/control status). The CA-ITT methodology also assumes that that if non-

⁵ Note that children who move between schools can be tracked through the NPD—so children are not lost from the trial for this reason.

compliers had received the treatment, the treatment would have had the same effect as it did on the compliers.

It is important to recognise that ITT and CA-ITT address two different (though related) questions. Whereas ITT asks: *How much do study participants benefit from being assigned to a treatment group?*, CA-ITT considers: *What is the size of treatment benefit for someone who receives the treatment?* In other words, CA-ITT attempts to abstract from the problem of contamination. Thus a benefit of CA-ITT is that it leads to improved accuracy in estimating the size of treatment benefit for individuals who receive the treatment (Sussman and Hayward, 2010).

Figure 1: Participants in the Chess in Primary Schools trial

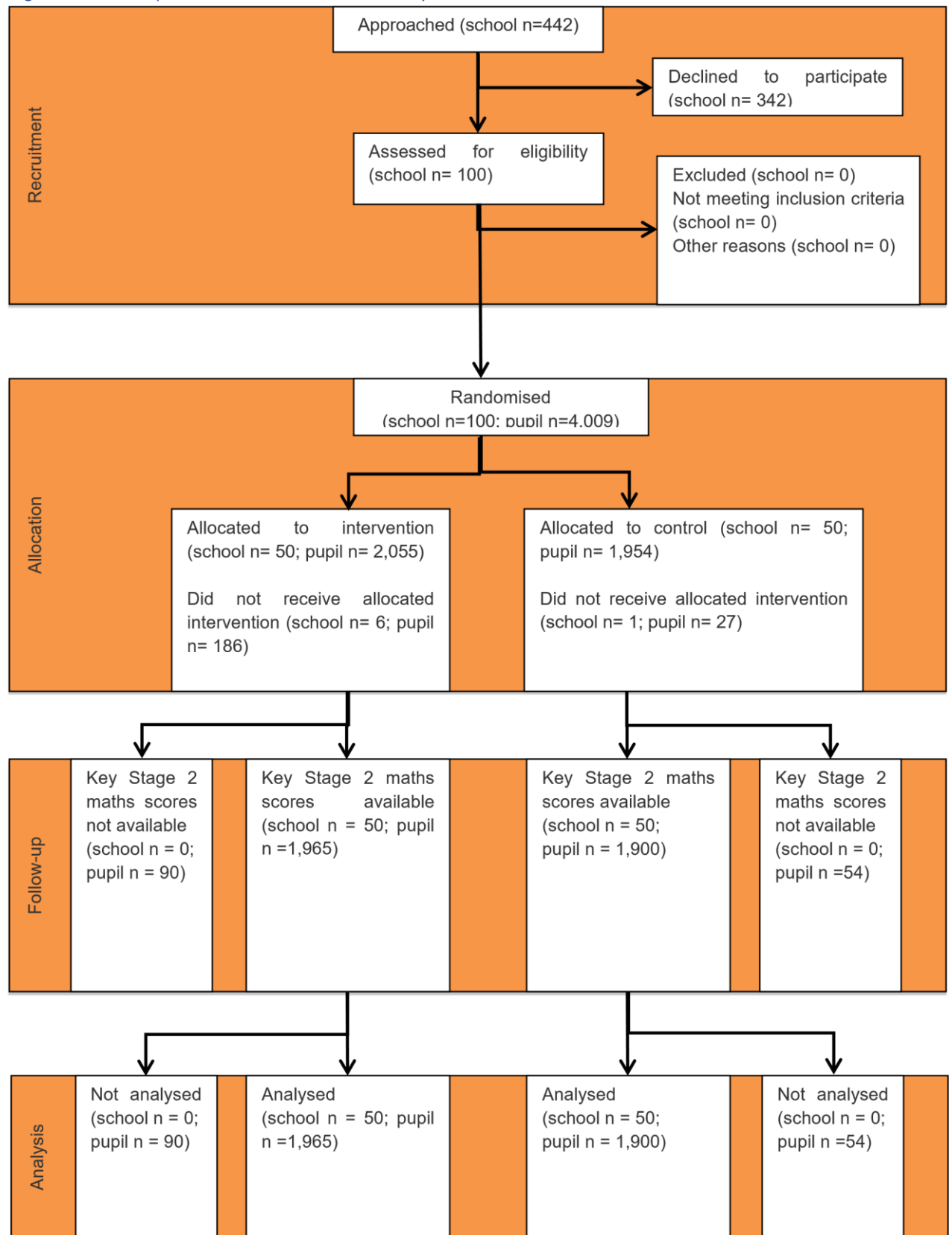


Table 3: Minimum detectable effect size at different stages

Stage	N [schools/pupils] (n=intervention; n=control)	Correlation between pretest (+other covariates) & post- test	ICC	Blocking/stratification on or pair matching	Power	Alpha	Minimum detectable effect size (MDES)
Protocol	100 schools (50 treatment, 50 control). 6,000 pupils (3,000 treatment,3,000 control)	0.63 (40% of variance explained)	0.15	10 Stratum based upon FSM and prior achievement	80%	0.05	0.18
Randomisation	100 schools (50 treatment, 50 control). 4,009 pupils (2,055 treatment,1,954 control)	0.67 (45% of variance explained)	0.11	10 Stratum based upon FSM and prior achievement	80%	0.05	0.16
Analysis (i.e. available pre- and post-test)	100 schools (50 treatment, 50 control). 3,865 pupils (1,965 treatment,1,900 control)	0.67 (45% of variance explained)	0.11	10 Stratum based upon FSM and prior achievement	80%	0.05	0.16

Note: Correlation between pre-test and post-test based upon OLS regression model including controls for gender, FSM, KS1 maths score, KS1 reading score, KS1 writing score, and KS1 science score.

Pupil characteristics

Table 4 compares KS1 scores for children in the treatment and control groups across four subject areas (numeracy, reading, writing, and science). All children for whom KS1 information could be linked are included in this comparison. The distribution of KS1 maths scores is very similar across the two groups, with differences at any given level typically just one or two percentage points. Similar findings hold for KS1 reading and writing. Indeed, the only instance where there is a difference of meaningful magnitude is KS1 science, where more children reach level 3 in the treatment group (14%) than in the control group (7%). We have additionally looked at mean KS1 average point scores (APS) for treatment and control groups. The difference is again small, standing at 0.05 standard deviations. Overall, Table 4 suggests that the sample is well balanced in terms of prior academic achievement.

Table 5 considers balance between treatment and control groups in terms of other observable characteristics. (These characteristics are presented for all children initially randomised.) There is broadly the same proportion of boys and girls in the two arms of the trial, though with slightly more children eligible for FSM in the control group (36%) than the treatment group (33%). Nevertheless, most of the differences observed between treatment and control groups in Table 5 are relatively small. Overall, Table 5 suggests that the treatment and control groups are also reasonably well balanced on a range of baseline characteristics.

Table 4: Comparison of baseline (Key Stage 1) test scores between treatment and control groups

Variable	Intervention group		Control group	
	n	Percentage (or standardised mean)	n	Percentage (or standardised mean)
Key Stage 1 maths				
Level 1	242	12%	236	12%
Level 2A	441	21%	450	23%
Level 2B	590	29%	567	29%
Level 2C	366	18%	356	18%
Level 3	246	12%	191	10%
Missing	170	8%	154	8%
Key Stage 1 reading				
Level 1	330	16%	309	16%
Level 2A	428	21%	457	23%
Level 2B	523	25%	491	25%
Level 2C	278	14%	280	14%
Level 3	304	15%	243	12%
Missing	192	9%	174	9%
Key Stage 1 writing				
Level 1	363	18%	373	19%
Level 2A	340	17%	319	16%
Level 2B	586	29%	509	26%
Level 2C	433	21%	453	23%
Level 3	116	6%	112	6%
Missing	217	11%	188	10%
Key Stage 1 science				
Level 1	306	16%	297	16%
Level 2	1,317	68%	1,369	74%

Level 3	266	14%	131	7%
Missing	166	2%	157	3%
Key Stage 1 average point score				
Standardised mean	1,932	0.024	1,843	-0.025
Missing	123		111	
School n				
Pupil n				

Notes: All figures refer to percentages, except KS1 average points score (which has been standardised to have a mean of 0 and standard deviation of 1 across the participating sample of 4,009 pupils). All analysis performed at the pupil level (i.e. figures refer to percentage of pupils – not percentage of schools).

Table 5: Comparison of demographic characteristics between treatment and control groups

Variable	Intervention group		Control group	
	n	Percentage	n	Percentage
Eligible for FSM				
No	1,376	67%	1,250	64%
Yes	679	33%	704	36%
Gender				
Female	1,007	49%	997	51%
Male	1,048	51%	957	49%

Notes: All figures refer to percentages. All analysis performed at the pupil level (i.e. figures refer to percentage of pupils—not percentage of schools). There is no missing data for these variables.

External validity

Schools were not randomly selected into the trial. Rather, the evaluators composed a list of 442 schools within the 11 local authorities who were eligible to participate in the trial (see ‘Method’ section above). This list of schools was then given to the CSC project team, who were asked to recruit 100 schools to participate in the trial. Put another way, the CSC team had to ensure that at least 22% of the 442 eligible schools were recruited.

Table 6 considers whether pupils within the 100 participating schools have similar baseline (KS1) test scores to pupils in the population of 442 schools who were eligible to take part in the trial. (Figures for all state school pupils in England are also provided for context, though the trial has not been designed to generalise to the country as a whole. This data has been drawn from the National Pupil Database.)

The percentage of children in each Key Stage 1 performance level is very similar across the ‘trial participants’ and ‘eligible’ samples. Standardised APS scores differ by less than 0.01 standard deviations between these two groups. A similar finding holds

for the distribution of KS1 levels across each of the four subject areas; differences between trial participants and the eligible population is never more than one or two percentage points. Thus, despite the absence of random sampling, children who took part in the trial were very similar to the population of pupils they were meant to represent in terms of prior academic achievement.

Table 7 presents a similar comparison for other demographic characteristics. There are slightly fewer children with English as an additional language (EAL) among trial participants (34%) than in the eligible population. Likewise, London is somewhat over-represented compared to the rest of the country. However, differences observed between eligible and participating pupils are nevertheless relatively small in terms of magnitude. Overall, this reinforces the main message of Table 6—the sample of trial participants is broadly representative of the population who were eligible to take part (at least in terms of observable characteristics).

Table 6: Comparison of Key Stage 1 test scores of trial participants to (i) the population of eligible pupils, and (ii) all state school pupils in England

Variable	Trial participants	All eligible pupils	England
Key Stage 1 maths			
Level 1	12%	12%	8%
Level 2A	22%	24%	27%
Level 2B	29%	30%	27%
Level 2C	18%	20%	15%
Level 3	11%	11%	20%
Missing	8%	3%	2%
Key Stage 1 reading			
Level 1	16%	16%	12%
Level 2A	22%	23%	25%
Level 2B	25%	27%	23%
Level 2C	14%	14%	12%
Level 3	14%	15%	26%
Missing	9%	4%	3%
Key Stage 1 writing			
Level 1	18%	20%	15%
Level 2A	16%	16%	20%
Level 2B	27%	29%	29%
Level 2C	22%	23%	20%
Level 3	6%	7%	13%
Missing	10%	5%	4%

Key Stage 1 science			
Level 1	15%	16%	10%
Level 2	67%	72%	68%
Level 3	10%	10%	20%
Missing	8%	2%	2%
KS1 average points score (standardised across the population in England)	-0.28	-0.29	0.00
School n	100	442	0
Pupil n	4,009	16,397	570,344

Notes: 'All eligible pupils' refer to all pupils in the schools that were eligible to be recruited into the trial. Trial participants includes both treatment and control group. England provides figures for all state school pupils. In this table, KS1 average points score has been standardised across the 570,344 pupils in the English state school population. Hence, for this variable, figures will not match between Table 4 and Table 6.

Table 7: Comparison of demographic characteristics of trial participants to (i) the population of eligible pupils, and (ii) all state school pupils in England

Variable	Trial participants	All eligible pupils	England
Eligible for FSM			
No	66%	65%	82%
Yes	35%	35%	18%
Gender			
Female	50%	50%	49%
Male	50%	51%	51%
Language Group			
English	65%	63%	82%
Other	34%	37%	18%
Local Authority			
Hackney	15%	10%	-
Hammersmith and Fulham	6%	4%	-
Southwark	17%	11%	-
Newham	13%	14%	-
Liverpool			
Sefton	5%	4%	-
Tameside	7%	5%	-
Sheffield	4%	9%	-

Leeds	16%	18%	-
Bristol	6%	6%	-
Middlesbrough	4%	5%	-
Ethnic Group			
White	52%	54%	77%
Black	22%	19%	5%
Asian	12%	14%	10%
Mixed	8%	7%	5%
Other	4%	4%	2%
Unclassified	1%	1%	1%
Chinese	0%	1%	0%
School n	100	442	0
Pupil n	4,009	16,397	571,733

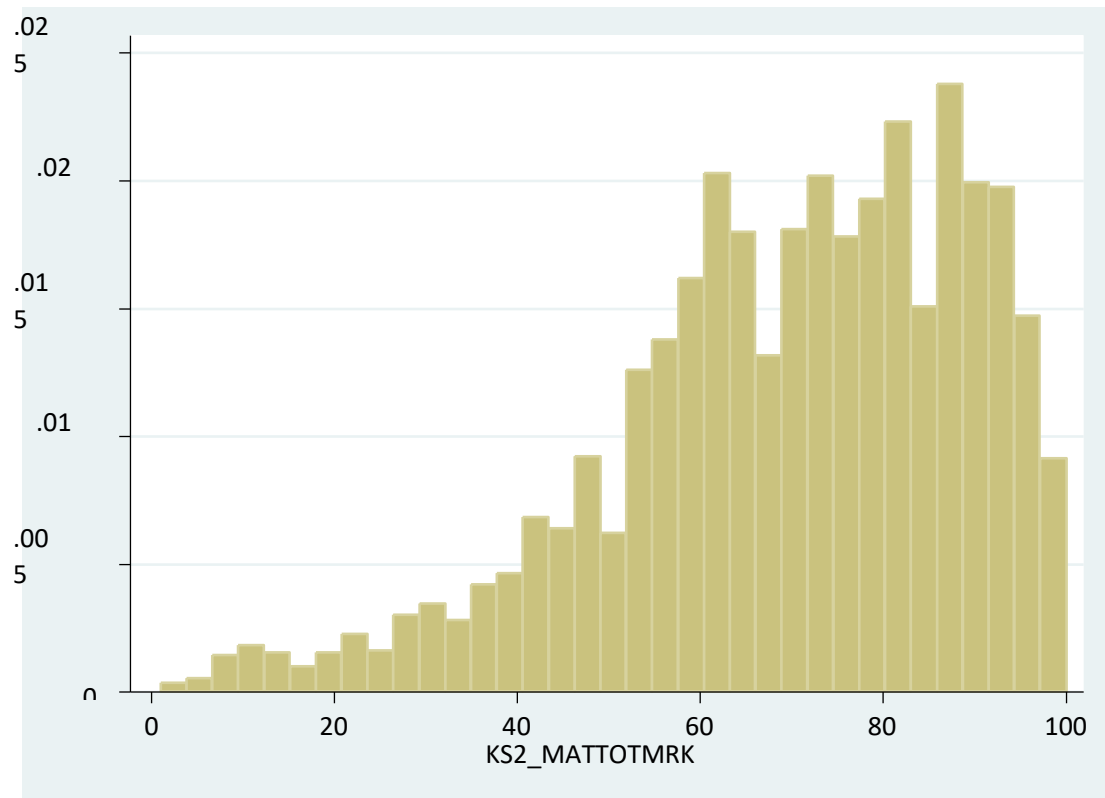
Notes: 'All eligible pupils' refer to all pupils in the schools that were eligible to be recruited into the trial. Trial participants includes both treatment and control group. England provides figures for all state school pupils.

Outcomes and analysis

Descriptive statistics

Figure 2 plots the distribution of Key Stage 2 test scores for the children in the analysis sample. There is little evidence of either floor or ceiling effects, though the distribution does have notable negative skew. The overall mean is 70 points, and the standard deviation is 20. We have also estimated the strength of the association between children's Key Stage 1 average points score and their marks in the Key Stage 2 maths exam. The correlation is 0.65, with around 40% of the variance in Key Stage 2 maths scores explained.

Figure 2: The distribution of children's Key Stage 2 raw scores



Notes: The y-axis refers to the probability density. The x-axis refers to the total score on the Key Stage 2 mathematics test.

Primary outcome: Overall Key Stage 2 maths scores

Results are presented in Table 8. The first row presents the intention-to-treat (ITT) estimates, while the second provides the contamination adjusted intention-to-treat (CA-ITT) estimates. Children who received the Chess in Primary Schools intervention achieved Key Stage 2 maths scores no higher than the control group, with an effect size of 0.01 and 95% confidence interval ranging from -0.15 to +0.16. Similar substantive conclusions hold for both the ITT and CA-ITT analyses. In additional analysis (results not presented), we have also reestimated the effect of the intervention having excluded the seven schools that removed a maths lesson in order to make room for the CSC curriculum. The effect size actually fell slightly, to -0.02 (95% confidence interval from -0.18 to +0.13), suggesting that this is unlikely to explain why no evidence of impact was found.

Table 8: Estimated effect of the Chess in Primary Schools intervention upon children's average maths test scores

Outcome	Raw means				Effect size		pvalue
	Intervention group		Control group		n in model (intervention; control)	Effect size (95% CI)	
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
ITT	1,965 (0)	70.0 (67.9 to 72.1)	1,900 (0)	69.2 (66.8 to 71.7)	3,865	0.01 (- 0.15 to +0.16)	0.93
CA-ITT	N/A	N/A	N/A	N/A	3,865	0.01 (- 0.16 to +0.17)	0.93

Notes: Authors' calculations. Effect size estimates based upon an OLS regression model, controlling for Key Stage 1 maths, reading, writing and science test scores. ITT refer to Intention-To-Treat estimates. CA-ITT refer to the instrumental variable (Contamination Adjusted Intention-To-Treat) results.

Differences in treatment effects by sub-group

Table 9 presents results for three sub-groups: boys, girls, and children who were eligible for FSM. The estimated effect of the intervention on the latter was 0.01 (95% confidence interval running from -0.18 to +0.19). For boys, the impact was -0.02 standard deviations (95% confidence interval running from -0.17 to +0.13) compared to +0.03 for girls (95% confidence interval -0.14 to +0.20). However, a formal test of the gender-by-treatment interaction failed to reject the null hypothesis of no difference between boys and girls at conventional thresholds. Overall, there is little evidence that the intervention had any impact upon the pre-specified sub-groups after one year.

Table 9: Estimated effect of the Chess in Primary Schools intervention upon subgroups

Outcome	Raw means				Effect size		pvalue
	Intervention group		Control group		n in model	Effect size (95% CI)	
	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)			
Boys	994 (0)	71.0 (68.8 to 73.2)	931 (0)	71.3 (68.6 to 74.0)	1,925 (0)	-0.02 (-0.17 to +0.13)	0.77
Girls	971 (0)	69.0 (66.7 to 71.4)	969 (0)	67.2 (64.6 to 69.9)	1,940 (0)	+0.03 (-0.14 to +0.20)	0.73
FSM	641 (0)	65.6 (62.6 to 68.5)	680 (0)	64.8 (61.8 to 67.7)	1,321 (0)	0.01 (-0.18 to +0.19)	0.95

Notes: Authors' calculations. Effect size estimates based upon an OLS regression model, controlling for Key Stage 1 maths, reading, writing and science test scores.

*Secondary outcomes**Mathematics sub-domains*

Table 10 provides the estimated impact of the treatment on each of the maths sub-domains (paper A, paper B, and mental arithmetic). The effect size is very close to 0 on each occasion. This further supports the finding that the intervention had no impact upon maths achievement after one year.

Table 10: Estimated effect of the Chess in Primary Schools intervention upon different components of the Key Stage 2 maths test

Outcome	Raw means				Effect size		
	Intervention group		Control group		n in model	Effect size (95% CI)	pvalue
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)				
Paper A	1,965 (0)	27.5 (26.7 to 28.3)	1,900 (0)	27.2 (26.1 to 28.2)	3,865 (0)	0.01 (-0.15 to +0.16)	0.91
Paper B	1,965 (0)	28.9 (28.0 to 29.9)	1,900 (0)	28.7 (27.6 to 29.7)	3,865 (0)	0.00 (-0.16 to +0.17)	0.96
Mental arithmetic	1,965 (0)	13.6 (13.1 to 14.0)	1,900 (0)	13.4 (12.9 to 13.9)	3,865 (0)	0.00 (-0.12 to +0.13)	0.94

Notes: Authors' calculations. Estimates based upon an OLS regression model, and are based upon ITT.

Reading and science

Table 11 turns to examine spillover effects into two other academic subjects: reading and science. The point estimate was -0.06 standard deviations for the impact on reading (95% confidence interval from -0.21 to +0.09) and -0.01 for science (95% confidence interval from 0.12 to +0.09). There is hence no evidence the intervention had any spillover impact upon these other subject areas.

Table 11: Estimated effect of the Chess in Primary Schools intervention upon children's Key Stage 2 reading and science test scores

Outcome	Raw means				Effect size		
	Intervention group		Control group		n in model	Effect size (95% CI)	pvalue
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)				

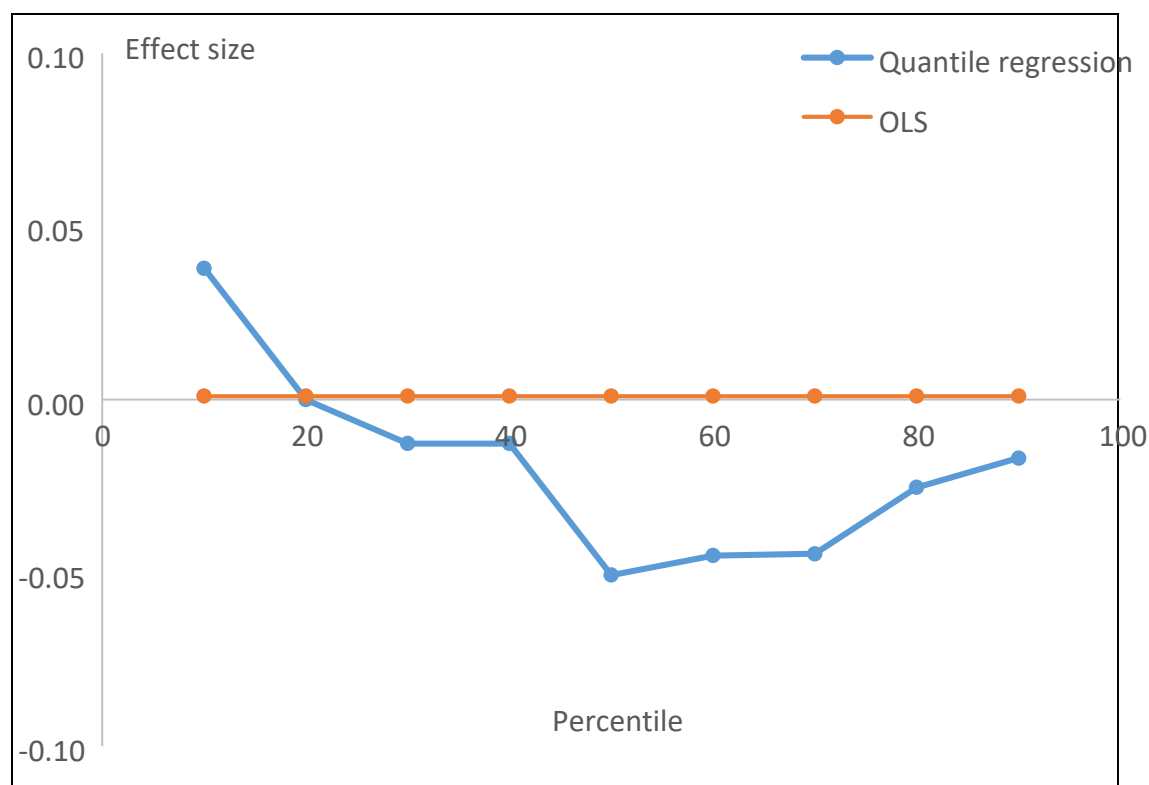
Reading	1,954 (11)	29.8 (29.0 to 30.6)	1,891 (9)	30.1 (29.0 to 31.2)	3,845	-0.06 (- 0.21 to +0.09)	0.44
Science	1,965 (0)	4.20 (4.16 to 4.24)	1,900 (0)	4.20 (4.14 to 4.25)	3,865	-0.01 (- 0.12 to +0.09)	0.82

Notes: Authors' calculations. Estimates based upon an OLS regression model, and are based upon ITT.

Quantile regression estimates (heterogeneous effects)

It could be that our finding of zero impact upon mean mathematics scores is driven by a large positive impact upon one group (e.g. low maths achievers) and a large negative impact upon another (e.g. high maths achievers). Consequently, Figure 3 presents quantile regression estimates of treatment effect at each decile of the post-test (Key Stage 2) distribution, in order to examine whether the Chess in Primary Schools programme had a different impact upon high and low academic achievers. Running along the x-axis is the percentile of the post-test score distribution where the quantile regression is estimated. The y-axis provides the estimated treatment effect. There is very little evidence that the intervention had any positive effect on either high or low maths achievers. Indeed, many of the point estimates are actually negative, though none are significantly different from zero at even the 10% level. Again, this further strengthens the evidence that the intervention had little medium-term impact upon pupils' maths achievement.

Figure 3: Quantile regression estimates of the impact of the Chess in Primary Schools intervention



Source: Authors' calculations. Solid line refers to the quantile regression estimates. Dashed line provides OLS estimates for comparison. Key Stage 2 total mathematics scores is the dependent variable.

'On-Treatment' analysis

Table 12 presents results from our on-treatment analysis, focusing upon whether the effectiveness of the intervention varied by chess tutor 'quality'. (Recall that 'tutor quality' has been defined using the proportion of children who reported that they liked the chess lessons run by the tutor 'a lot'.) All figures refer to differences in Key Stage 2 maths test scores (presented in terms of an effect size) relative to the control group. There is no clear evidence that children taught chess by tutors of higher quality achieved significantly higher KS2 test scores. Children taught by 'low quality' tutors achieved KS2 test scores slightly below the control group (-0.05 standard deviations) while children with 'medium quality' tutors scored a little higher than the control group (+0.11 standard deviations). However, there is no clear pattern of a 'dose-response' relationship, as the effect of having a high quality tutor was essentially zero. Moreover, none of the estimates presented in Table 12 reach statistical significance at conventional levels. Overall, there is no evidence that the effect of the Chess in Primary Schools intervention varied significantly by whether children liked a particular tutor's chess lessons.

Table 12: Estimated effect of the Chess in Primary Schools intervention upon children's Key Stage 2 mathematics scores, by chess tutor quality

Effect size

Outcome	n in model	Effect size (95% CI)	p-value
'low quality'	89 schools 3,498 pupils	-0.05 (-0.26 to +0.15)	0.63
'medium quality'	89 schools 3,498 pupils	+0.11 (-0.07 to +0.29)	0.25
'high quality'	89 schools 3,498 pupils	0.00 (-0.27 to +0.26)	0.99

Table 13 presents analogous results for whether the regular class teacher attended the CSC training workshop, as per the study protocol. All figures refer to differences compared to the control group, expressed in terms of an effect size. There is no evidence that the effect of the intervention varied by whether the regular class teacher attended the CSC workshop. For instance, children in treatment schools whose teacher did attend the workshop scored just 0.01 standard deviations higher on their KS2 maths test than children in the control group. This difference is very small and statistically insignificant at conventional levels. Likewise, the test scores of children in treatment schools where the class teacher 'did attend' is little different from the test scores of children where the class teacher 'did not attend'.

Table 13: Estimated effect of the Chess in Primary Schools intervention upon children's Key Stage 2 mathematics scores, by whether the class teacher attended the CSC training workshop

Effect size			
Outcome	n in model	Effect size (95% CI)	p-value
Did not attend	95 schools 3,714 pupils	+0.04 (-0.14 to +0.21)	0.69
Did attend	95 schools 3,714 pupils	+0.01 (-0.16 to +0.18)	0.92

Cost

If a primary school were to participate in the Chess in Primary Schools programme next year, they would incur two main costs. The first is that they would have to pay an annual contribution to the CSC charity. This varies by the size of the school, and whether an afterschool or lunchtime chess club is also set up⁶. For instance, a typical

⁶ Whether a school sets up a lunchtime or an after-school chess club is optional. It is included in the cost estimate presented here, as the after-school club formed part of the intervention evaluated.

two-form primary school with an after-school chess club would be asked to pay £1,800 per year.

The second main cost to schools is that the regular class teacher is expected to complete a one-day training course organised by Chess in Schools and Communities. For instance, for a two-form entry school which needs to pay for supply cover, we estimate this to require a oneoff cost of around £400 (assuming a figure of £200 per day for each supply teacher).

Following EEF guidance, we spread this cost over three years, to give an annual figure of £133.

In Table 14, we add these two costs together, and illustrate how the total cost varies by size of school. For instance, we estimate the average annual cost of a primary school to be £1,933 for a two-form entry school. This estimate of the total cost is then divided by the number of pupils (assuming 30 pupils per school form) to provide a cost per pupil. This varies from £52 per pupil in single-form entry schools to £22 per pupil for schools with four forms or more.

It should be noted that schools that participated in this evaluation were not expected to make a contribution to the Chess in Schools and Communities charity during the intervention year; rather, this was covered directly by the EEF grant.

Table 14: Cost to schools to participate in the Chess in Primary Schools programme (costs per year based on delivery over 3 years)

Number of forms	Class lessons	After school club	Teacher training	Total cost	Number of pupils	Cost per pupil
1	£900	£600	£67	£1,567	30	£52
2	£1,200	£600	£133	£1,933	60	£32
3	£1,500	£600	£200	£2,300	90	£26
4	£1,800	£600	£267	£2,667	120	£22
5	£2,400	£600	£333	£3,333	150	£22
6	£3,000	£600	£400	£4,000	180	£22

Process evaluation

Introduction

This section of the report covers the key findings of the process evaluation of the Chess in Primary Schools programme. The process evaluation aimed to explore aspects of the study that provide insight into effectiveness as well as issues such as perceptions of impact and potential improvements and sustainability of the programme.

This section covers:

- Overview of data sources and response rates
- Logic model—developer's' view of necessary conditions
- Implementation
 - Dosage
 - Fidelity to the model
 - Response to the intervention
 - Factors influencing implementation
- Perceptions of programme impact
- Sustainability of the programme □ Lessons for future implementation □ Control group activity.

Overview of data sources and response rates

Table 15 summarises the data sources and response rates. While some individual types of data were more complete than others, across the various sources a good picture of the key themes across the treatment schools has been achieved.

Table 15: Process evaluation data collection—methods and response rates

Method	Sample size	Response – number (rate)
Head teacher baseline survey	100 schools	78 (78%)
Class teacher post-intervention survey – treatment schools	44 treatment schools	28 (64%) schools – 36 teachers submitted data
Class teacher post-intervention survey – control schools	49 control schools	10 schools (20%) – 15 teachers submitted data
CSC tutor survey	24 tutors	23 (96%)
Head teacher treatment schools – postintervention survey	44 headteachers	18 (41%)
Pupil post-intervention survey – paper	75 classes across 44 treatment schools	776 pupils from 36 classes (48%) across 26 treatment schools (60%)
Treatment arm: observations and interviews	Numbers conducted	
Researcher observation of CSC training for study school teachers	2	
Researcher observation of CSC training events for tutors	3	
Researcher observation of intervention delivery	8 classes in 4 schools	
Teacher interviews – telephone or face-to-face following observation	10 (5 telephone; 5 face-to-face)	

Headteacher interviews – telephone or face-to-face in observation schools	3 (2 telephone; 1 face-to-face)
CSC head office team – interview	1 face-to-face

Logic model—developers' view of necessary conditions

The logic model (Figure 4) reflects the views of CSC experts on the necessary conditions (inputs and processes in school) for their intervention to be successful. The following sections will reflect on whether these conditions have been met.

Figure 4: Chess in Primary Schools logic model*

Inputs	Processes in school	Change mechanisms (*)	Pupil intermediate impacts	Outcomes
<ul style="list-style-type: none"> <input type="checkbox"/> Sufficient chess boards/pieces <input type="checkbox"/> Chess tutor—enthusiastic, reasonable player, good teacher <input type="checkbox"/> Team teaching with class teacher/assistant <input type="checkbox"/> Good classroom environment, e.g. sufficient space for pairs to play games of chess <input type="checkbox"/> CSC primary school curriculum plus CSC work book <input type="checkbox"/> One-day basic training course for tutors <input type="checkbox"/> One-day basic training course for teachers <input type="checkbox"/> Whiteboard or manual display board <input type="checkbox"/> Additional training for tutors 	<p>Learning chess 1 hour a week for 30 weeks as part of the school curriculum.</p> <p>A CSC tutor using the CSC curriculum for primary schools teaches the chess.</p> <p>The class teacher/assistant plays an active part in the chess lesson.</p>	<p>Education</p> <p>Children learn chess—chess playing requires skills that support maths and English ability.</p> <p>CSC tutor teaches using a graded chess curriculum—lessons are fun and interactive.</p> <p>Tutor (with teacher/assistant support) differentiates teaching to meet different learning needs of individuals within a class.</p> <p>Pupils play chess games together and learn new things about each other.</p> <p>Persuasion (or encouragement); modelling</p> <p>Pupils who do not excel academically can show an aptitude for chess and gain confidence/recognition.</p> <p>Teacher/assistant learns/plays chess with the children.</p> <p>Tutor promotes and models positive attitudes and behaviour in the context of teaching chess rules and etiquette, e.g. silence, concentration.</p>	<ul style="list-style-type: none"> • Improved concentration and perseverance • Improved logical thinking and problem solving • Improved confidence and self-esteem • Improved behaviour in school • Improved communication of complex ideas • Excelling of those with particular needs, e.g. gifted and talented, special educational needs, more solitary, etc. • Ability to play a reasonable game of chess • Improved creativity 	<p>Improved KS2 maths SATs results</p> <p>Improved KS2 English SATs results</p> <ul style="list-style-type: none"> <input type="checkbox"/> <input type="checkbox"/>

*Developed through a Delphi process with CSC experts. This reflects their opinions of what is required to achieve the desired outcome

Implementation

This section describes what was actually delivered in the treatment schools. It considers the dosage received, the fidelity to the model, the response to the intervention and key factors that affected implementation.

Dosage

Six treatment schools chose not to participate in the intervention. The headteachers for three of these six schools returned baseline survey data, prior to randomisation. At that stage two of these three school leaders had been very keen and one fairly keen on taking part in the study. Information provided by CSC suggested that the main reasons for the subsequent withdrawals were practical changes that impacted on the feasibility for the school of being involved, such as turnover of key staff.

The chess lessons were delivered in all 75 Year 5 classes in the 44 treatment schools which delivered the intervention. The one-hour chess lessons ran, in all participating schools, from midway through the autumn term (2013) to near the end of the summer term (2014). According to the tutor and teacher data, only one third of schools received the full intended dose of 30 hours—see Table 16.

Table 16: Amount of chess teaching delivered in a school

	30 hours	25–29 hours	Less than 25 hours
Number of schools	14 (33%)	26 (59%)	4 (9%)

The main reason for this shortfall in intended hours of delivery was that the classes commenced midway through the first term and not at the start. The cause of this delay was the completion of trial processes by CSC with schools. If there were few or no rescheduling requests by the school, then achieving the 30 hours was possible. But where chess lessons were occasionally cancelled by schools, which was not uncommon, completing the full 30 hours became difficult to achieve.

Reach

Teacher and tutor survey data shows that the programme reached its intended recipients in the 44 participating schools. While a few teachers reported via the survey that a few children did not routinely participate in the chess lessons, this was unusual. These children were removed from the lessons either due to other demands on their time (such as instrument lessons) or being perceived by the teacher not to be managing the chess. **Fidelity to the model**

In general, intervention delivery adhered fairly closely to the intended programme, as spelled out in the logic model, though there were some key areas where expected inputs and processes of delivery were not completely as intended and/or where there was variability across the sites. Table 17 summarises the adherence and deviation.

Equipment and curriculum

There was clear fidelity to the model in terms of equipment, facilities and the curriculum. Classrooms and equipment were deemed fit for purpose by tutors and the chess sets provided by CSC at the start of the intervention remained available throughout the year. All tutors used the CSC curriculum. Children were given worksheets to complete, as part of the curriculum and work books were introduced as an additional resource by most tutors. As intended by the developers, the tutors did make adaptations to the curriculum. Examples of adaptations included running a class competition over several lessons and using video clips of grand masters playing in a competition. While some tutors were using the curriculum in digitalized form via whiteboards or TV screens, many were not.

Those that weren't were using the manual display board provided by CSC.

Tutor background and skills

The Chess in Primary Schools programme was to be delivered by a chess tutor who was to be 'enthusiastic, a reasonable player and a good teacher'. Twenty-three CSC tutors taught the intervention across the participating schools; there was considerable continuity with very little turnover of tutors in schools. The tutors' survey provided the following information about their background and skills:

- All the tutors were proficient chess players, mainly describing themselves as average or strong club players.
- Just under half the tutors had some kind of teaching qualification, including one who was a qualified primary school teacher, five who had secondary/adult education teaching qualifications and six a specialist chess teaching qualification.
- Five tutors started working for CSC in the year of the study; ten had worked for CSC for 1–2 years and eight for 3 or more years.
- Before the study, 18 tutors had had a moderate amount or a lot of experience of teaching chess to primary school children (in clubs, schools, etc.) while 4 had no or a little experience (1 non response). Of the 23 tutors, 18 had been on the one-day CSC training for tutors, teachers and prospective tutors, which covered the CSC curriculum. During the intervention year, 17 of the tutors attended additional weekend training seminars, a new initiative organised by CSC to support and develop their teaching and classroom management skills.

However, while teachers and headteachers were generally very positive in interviews and surveys about many aspects of the tutors' input (including their enthusiasm), in many schools concern was raised, to varying degrees, about the tutors' teaching and classroom management skills. This theme also emerges from the pupil survey. This will be discussed further in the section 'Factors affecting implementation'.

Tutor delivery in the schools

Individual tutors usually worked in one or two treatment schools (range 1–6 schools). Where a school had two or more Year 5 classes the same tutor usually worked with all the classes,

generally in consecutive lessons. Variation to the original planned intervention came with the introduction of paired tutors delivering lessons in some schools. Tutor survey data indicated that pairs of tutors worked in 6 schools for the whole intervention period and in 15 other schools for part of the intervention period. CSC staff confirmed that these pairings generally occurred with relatively new CSC tutors or when the intervention was perceived to need a boost. Pairing of tutors is unusual for CSC but was possible for the trial due to the additional resource available.

Teacher involvement

The class teachers were to attend the CSC one-day training course and then be actively involved in the chess lessons, ideally team teaching with the tutor. In fact, only 31% of teachers attended the training. Furthermore, while most teachers did engage with the lessons their engagement was less extensive than hoped for by many tutors. These deviations are discussed in more detail in the section ‘Factors affecting implementation’.

Lesson replacement

Class teacher survey data received from teachers in 30 schools (68% of the treatment group) showed that the chess lesson most commonly replaced a ‘topic’/humanities lesson; others replaced included music or PE. However, seven from this group replaced a maths lesson—six wholly, and one partially—and one school said they replaced an English lesson for the whole of the intervention year. The replacement of a maths or English lesson with the chess lesson was a clear departure from the intended programme. The reasons for this replacement are not completely clear, though many teachers and headteachers reported difficulties fitting all aspects of the curriculum into the Year 5 timetable. Additionally, there appeared to have been some misunderstanding by, and within, schools of what was expected of them for the Chess in Primary Schools programme. This occurred despite contact from CSC with all treatment schools, when it emerged during the first few months of the programme that some schools were replacing maths lessons.

Table 17: Intervention delivery—achievements and variations from plan

Achieved as planned	Variation from plan
<p><i>Inputs</i></p> <ul style="list-style-type: none"> • Sufficient chess boards and pieces • Chess tutor—enthusiastic, reasonable player • Good classroom environment • CSC primary school curriculum plus CSC work book • One-day basic training course for tutors • Whiteboard or manual display board • Additional training for tutors <i>Processes in school</i> • A CSC tutor using the CSC curriculum for primary schools teaches the chess 	<p><i>Inputs</i></p> <ul style="list-style-type: none"> • Team teaching—<i>tutor/teacher planned collaboration often limited</i> • Chess tutor good teacher—<i>some lacked class differentiation skills</i> • One-day basic training course for teachers—<i>very low uptake Processes in school</i> • Learning chess 1 hour a week for 30 weeks as part of the school curriculum—<i>most schools achieved 25–29 hours</i> • Replacing a lesson other than maths—<i>at least 7 schools replaced maths lessons</i> • The class teacher/assistant plays an active part in the chess lesson—<i>happened in some but not all classes</i>

Response to the intervention

This section considers the views of the key stakeholder groups—pupils, teachers, headteachers— on the acceptability of the Chess in Primary Schools programme. A limitation of this part of the report is that it is based upon only 774 responses from the initial intervention sample of 1,900 (41%).

Pupils—response to the intervention

Pupils were asked in the survey: *How much did you like the chess lessons you had in your Year 5 class that were taught by a chess tutor?* Table 18 shows that very few disliked the lessons, with the majority liking them ‘a lot’.

Table 18: How much did children like the Chess in Primary Schools lessons?

Liked the chess lessons	n=774
A lot	53%
A little	39%
Didn't like them	8%

Prior to the chess lessons, over half (57%) of the pupils said they had never played chess before; whereas less than a fifth (17%) said they had played at least weekly before the chess lessons. The proportion of children who liked the chess lessons ‘a lot’ was significantly higher for those who had regularly played chess (72%) than those who had not played before the Chess in Primary Schools intervention began (44%, $p < .01$).

Children from schools in the study with the historically lowest attainment scores were significantly more likely to dislike the chess lessons ($p < .01$). More unexpectedly this also appeared to be the case in schools with the relatively lower FSM proportions ($p < .01$). One reason for this appears to be that more pupils in relatively less deprived communities had had more previous access to chess and as a result some of these children reported finding the lessons too slow and insufficiently engaging.

Tutor effects on how much pupils liked the chess lessons

Pupil data was received from 19 of the 23 tutors. A tutor effect emerged when considering the degree of liking lessons. Proportions of pupils who liked the lessons ‘a lot’ ranged across tutors from 15% to 76%. Five tutors had more than two-thirds of the pupils they taught like the chess lessons ‘a lot’, whereas five tutors had less than a third. No key background features (e.g. teaching qualifications, years working for CSC) varied significantly between those tutors where a greater proportion of children liked the sessions a lot and those where they did not.

When extent to which pupils liked the chess lessons was analysed by school, it became apparent that there was one school where there was a significant number of children who disliked the chess lessons (n=14 or 50% of the pupils who responded from this school). Process evaluation interviews indicated that the teacher of this class had been dissatisfied with the ability of the tutor to keep children of different abilities engaged with the lessons.

What pupils liked best about the chess lessons

Pupils were asked to write free text answers to the question: *What did you like best about the chess lessons?* Common responses were:

- **Playing chess games and/or mini games.** This was the most frequent answer to this question. Many who said this referred in particular to the fun of playing their friends/classmates. This theme was strongly reinforced by teachers who gave many examples, in interviews and surveys, of pupils choosing to play chess with their friends, rather than other activities available to them during free time in school.
- **Being taught the theory of the game**—such as the rules for the different chess pieces or strategy. This category of response implied a satisfaction with how these theoretical concepts were explained by the tutor.
- Particular **tutor attributes** such as being funny or having specific skills.
- **Learning a new skill** that was both fun and challenging.
- **Choosing friends to play against.** Many said that they disliked having opponents (who were not their friends) chosen for them by the tutor. However, there were children who prioritised having an opponent with whom they would have a good and challenging game over playing particular friends. These pupils welcomed tutors helping organise this.

'I liked that every lesson we had was clear and understandable.' (Year 5 pupil)

'I liked chess in Year 5 because it helps your brain to think and it is good for knowledge. I also like it because you play with your friends.' (Year 5 pupil)

Interestingly pupils rarely mentioned that they liked winning. This suggests that the tutors had succeeded in placing the emphasis on gaining satisfaction from the process of learning and playing rather than securing a victory at the end of a game.

What pupils liked least about the chess lessons

Pupils were also asked what they liked least about the chess lessons. Common themes were:

- **Too much talking by the tutor.** This was the most frequent response to this question. Common sub-themes to this response were too much: repetition of what they had learnt in a previous chess lesson; time on the carpet at the beginning of the lesson; talking by the tutor at the expense of time for playing games; interrupting of their games by tutors.
- **Losing.** For some it was evident that they lost frequently and unsurprisingly they disliked this.
- Finding the lessons either **too easy or too hard.**
- **Disliking the rules** set by the tutor, including the need to play quietly and not being allowed to choose who they played with, or conversely wanting stricter discipline to control the disruptive behaviour of their peers, during the lesson.

- **Not been chosen** as anyone's partner or generally finding the partnering process stressful.

'I felt that the teacher spends most of the time explaining things that he had already told us in the previous lesson which means we hardly get to play/finish our game.'
(Year 5 pupil)

The pupil acceptability data, both quantitative and qualitative, suggested that most children liked the chess lessons, but it also illustrated the extent of the challenge for tutors because of the diversity of pupil need within classes and the need to balance fun with applying clear rules.

Headteachers and teachers—response to the intervention

Teachers and/or headteachers from 39 treatment schools provided data for one or more of the different data collection activities aimed at them.

Teacher views before the intervention

Prior to randomisation, the baseline survey of headteachers showed that these school leaders were generally very enthusiastic about the Chess in Primary Schools intervention, regardless of personal chess-playing experience or level of chess activity in the school in the past five years. Any reservations were generally about the difficulties fitting the lessons into the timetable.

While the majority of class teachers said they were initially keen to try the intervention, there were some who admitted—retrospectively in the post-intervention teachers survey—to some initial ambivalence (29%) or reluctance (14%). Initial concerns were that the chess lessons would be difficult to fit in to a tight Year 5 timetable, and that low-achieving children would lose out most, through loss of core subject time and difficulties engaging with learning chess.

Teacher views during and after the intervention

The class teachers who engaged in interviews and observations and/or submitted post-intervention survey data were, in general, very positive about many aspects of the programme. Most of them rated the performance of the CSC tutors in their school as 'good'. Many teachers commented positively on particular qualities of the tutors—particularly their enthusiasm for chess and the impact this had on enthusing and engaging the children as well as their positive interactions with the pupils.

'[The tutor] has been brilliant for us, the manner in which he delivers the lesson, his engagement with the children and his ability to calmly solve their game problems has been great. The children look forward to seeing him each week.' (Class teacher)

Examples of high quality teaching practice by the tutors and use of innovative resources to complement the CSC curriculum were reported in surveys and interviews and also observed. These included examples such as tutors tracking each child's progress week by week in order

to meet their individual learning needs, and use of specialist chess software such as ChessBase on the whiteboard as well as other IT resources.

Headteachers who submitted post-intervention survey data were also generally very positive about the Chess in Primary Schools programme. Some took great interest in the chess lessons, as observed by researchers conducting observations who saw headteachers 'dropping into' chess lessons and giving out prizes for class chess competitions. Others relied on class teachers to update them on progress with the chess lessons.

'We are a chess loving/playing school now! We never even spoke about chess before and I wonder if some of the kids knew what it was! It has been an amazing experience, we love it here!' (Headteacher)

The key criticisms of the programme from teachers centred on the tutors' performance. These are detailed in the 'Factors influencing implementation' section below.

Chess tutors—views on responsiveness

Most tutors also considered that the intervention had gone either very well or well in the classes in which they had taught. In general they were very or fairly satisfied with the support they received from headteachers and teachers. Furthermore, most tutors felt well supported by CSC and were very positive about the CSC curriculum. Only one tutor said the lessons had gone badly in one class though many tutors thought there could have been improvements in the level of engagement of the class teacher. This issue, which was the tutor's key criticism, is discussed in the 'Factors influencing implementation' section below.

Factors influencing implementation

The process evaluation data suggests that there were two key factors that inhibited implementation of the Chess in Primary Schools programme in treatment schools: the tutors' teaching skills and the level of involvement of class teachers.

Implementation inhibitor: tutors' teaching and class management skills

The most common criticisms of the programme by teachers and headteachers in interviews and surveys were in relation to tutor teaching and class behaviour management skills. Over half the tutors had limited teaching training but there was no expectation by school staff that tutors should be trained teachers, in fact their status as experts in chess was considered by teachers to be key to gaining respect from the pupils. However, the ability of tutors to engage all the pupils and control the class was a clear area of concern.

'My class is very difficult in behaviour and you need to be firm and strong to maintain order. At times this was lacking [from the CSC tutor] and I had to take over to control the class.' (Class teacher)

Sub themes from teacher data on this issue included the following:

- Tutors stood at the front of the class and talked too much, particularly at the start of the lesson. Pupil data supported this.

- Where little or no technology was used in the teaching this increased the risk of engagement problems—this was supported by the observation data.

Many of the tutors acknowledged, in the survey, that their biggest challenge was accommodation of the wide diversity of pupil academic ability in most classes and many wanted more support with this from the class teacher. Conversely teachers expressed the view that it was important to the success of the lessons for the pupils to perceive the tutor, as opposed to the class teacher, as being in control of the lesson and therefore teachers were reluctant to intervene too readily. One tutor, who had worked for CSC prior to the study, commented on the more challenging nature of the treatment school classes with the suggestion that the range of ability in study school classes presented a particular challenge.

A number of teachers said that tutors needed more training in this aspect of their role.

'Professional development for tutors would be good on how to keep as many of the children engaged at any one time as possible and how to make the sessions as pacy as possible...that is the key to a successful lesson.' (Class teacher)

CSC offered a one-day training course and a number of seminars for tutors to attend. Tutors raised concerns about the usefulness of the one-day training with 5 of the 18 tutors who had attended stating that they were not satisfied with the course. An intention of this training, which deliberately mixes tutors and classroom teachers, is to achieve 'cross fertilisation' of experiences and views. Many tutors enjoyed this aspect of the training. Some however felt that an important focus for them should be on learning key teaching skills and that this could not be achieved while the group was mixed in this way.

For many teachers and headteachers the tutor teaching and class management issues discussed above were expressed as relatively minor concerns and were offered as constructive suggestions for improvement in the future to a programme that they felt had great potential. In a small number of schools, however, classroom management and delivery issues were perceived to be a more serious problem and made any future engagement with the programme by the school questionable.

It is evident that some teachers who had concerns made efforts to find ways to give tutors feedback and suggestions for change. This was clearly not easy when an expectation of such feedback did not appear to have been discussed. In general, where feedback happened, it appeared to be well received by the tutor and led to positive change in tutor performance.

'I emailed [the tutor] and said that there were too many long periods of listening to him talking and suggested that he should break it down with them having more chances to have a go at things. And he responded to this very well and children's engagement improved a lot more.' (Class teacher)

While most tutors liked the CSC curriculum that they used in the classroom, one tutor was overtly critical of it, saying it was old-fashioned and needed to have more 'hooks' that would capture the pupils' attention. The curriculum workbooks were generally liked by the tutors—

‘they facilitate pattern recognition’— but pupil data suggested mixed views on these workbooks.

Implementation inhibitor: class teachers' role in the lessons

From the tutor and CSC head office staff perspective the solution to the issues of delivery and classroom management was a high level of involvement of the class teacher, including attendance at the one-day CSC training course. As stated previously this involvement was less extensive than hoped. For example:

- Only 23 class teachers from 19 treatment schools (31% of teachers) attended the CSC oneday training. The main reasons given by the teachers were that they weren't aware of the training, lack of time or lack of permission from their manager.
- The majority of class teachers provided some help with the chess lessons but few tutors described a formalised team teaching approach with the class teacher in the lessons. Interview, survey and observation data suggests that one reason for this was differing perceptions between teachers and CSC staff about appropriate approaches to leadership of a lesson.
- In the chess lessons, approximately two thirds of the teachers played or learnt to play chess alongside the children. The other third did not take part in playing. CSC tutors suggested this was due to teachers feeling self-conscious about being seen by their pupils in the role of a learner, as opposed to an expert. While in some classes teaching assistants helpfully played when a teacher did not, this was not considered by tutors to have as powerful an effect on pupils as observing their teacher engaging in learning a new skill.

There were some examples given by tutors of classes where there was little or no joint working or support of any kind from teachers.

‘He [the class teacher] would sometimes be there as I arrived then would leave and just come back at the end of the day to dismiss the class.’ (CSC tutor)

The evidence from the process evaluation was that many teachers were not fully aware of these jointworking expectations, even by the end of the intervention. It was clear that in many schools, tutors and class teachers had had little time to liaise, prior to or during, the intervention period.

One aspect of the programme which should have helped with clarifying this role was the one-day CSC training course for teachers and tutors. Uptake of this was low, but those teachers that did attend had mixed views. Most reported enjoying the training, but did not think it was critical to the success of the intervention in their school. Teachers reported that the emphasis on the specifics of chess in the training was too strong (and as non-chess players went over their heads). They wanted a greater focus on how to make the chess lessons as successful as possible in the classroom setting.

Perceptions of programme impact

Teachers were dubious about whether the chess lessons would have impact on the primary outcomes of maths and English attainment. The teachers' survey showed that only about a quarter of teachers thought the chess lessons would have quite a lot of impact on pupils' maths attainment; half thought they would have a little impact and a few thought there would be no impact or were uncertain about impact. For English attainment, teacher views were roughly split between predicting the chess lessons would have a little impact and no impact.

Teachers were, however, overwhelmingly positive that the chess lessons would have impact on pupils':

- thinking/cognitive skills;
- confidence/self-esteem;
- ability to cope with winning/losing;
- concentration; and
- ability to play a game of chess.

'They [pupils] have mostly developed in their ability to slow down and really think about problems.' (Class teacher)

'I think this has been an invaluable experience for the children in my class. It has raised morale, achievements and sportsmanship. It has revealed hidden skills and talents, crossed the barriers between games and education and should, in my personal opinion, be made part of the National Curriculum.' (Class teacher)

While many teachers were also positive about impact of the chess lessons on peer relationships and pupil behaviour, some teachers thought there was no impact (or occasionally negative impact) on these.

Perceived impact on lower achievers

In the survey and in interviews, teachers gave examples of pupils at both ends of the academic ability spectrum that had enjoyed and benefitted from the lessons, with progress by individuals who were at the lower end particularly being selected as examples of positive outcomes of the programme.

'Some of the children that wouldn't have expected to excel have—and others want to pair with them— and this has been really nice for those children.' (Class teacher)

However, there was also considerable concern expressed by teachers and headteachers in some schools that the lower achievers got left behind in the lessons and as a consequence disengaged and often became disruptive. There were also concerns expressed by a few teachers that slower learners might be relatively negatively affected by the loss of a maths lesson (in schools where chess replaced maths).

'In mixed ability groups they [children with additional needs] are not able to access the learning as quickly as others so they become frustrated. The other children become impatient because they want to get on with the game.' (Class teacher)

Impact on the school

The impact of the programme on the school as a whole was mentioned by many schools. The majority of teachers reported on the survey that chess playing in the school had spread beyond the Year 5 chess lessons. Examples given were newly established or reinvigorated chess clubs and chess playing during free time in classrooms and/or the playground.

'Children from Year 2 upwards have benefited from the purchase of a giant chess set in the playground—Year 5 have been able to teach chess to others.' (Headteacher)

Sustainability of the programme

Tutors were fairly optimistic about the potential for chess playing to continue among the pupils they had taught. They reported that in the majority of classes over three-quarters of the children could play a reasonable game of chess by the end of the intervention. They also thought that about two-thirds of schools had a member of staff who was confident enough to teach chess themselves and could carry on doing so within their school.

The teacher survey data also suggested that many teachers intended to continue to incorporate chess in the classroom and/or school. Examples given were setting up a chess club and using some of the maths challenges based around chess. However, teachers were clearly concerned about potential barriers such as their own lack of confidence and time pressures and there was no suggestion that chess lessons of the type delivered by the Chess in Primary Schools programme would continue, unless the school purchased the programme.

'I would definitely encourage schools to teach chess and I would like to become more confident to have a go myself. I don't feel ready to teach it yet but possibly in the future I would.' (Class teacher)

Level of continued chess playing

The pupil survey was carried out approximately seven months after the chess lessons finished. Pupils were asked if they were still playing chess. Table 19 shows the responses to this question.

Table 19: Amount of chess playing by pupils 7 months after the intervention finished

Amount of chess played – 7 months post intervention	% All pupils N=772	% Where not played before Chess in Primary Schools	% Where had played frequently before Chess in Primary Schools
At least once a week	28	17	49
Between 1 and 3 games a month	39	35	39
Not playing any chess at all	34	48	12

- Unsurprisingly, there was a significant difference in continued playing between those who had played chess before the lessons, and those who had not ($p < .01$).
- There was no significant difference in continued playing between children from schools with historically lower achievement or higher proportions of pupils with free school meals status.

Most children who were still playing chess, said they were doing so with family and friends (88%, $n=449$). Some classes were given chess sets to take home by teachers (for example as a Christmas present) to support them in extending their playing into the home. Additionally, 68 children from 14 schools were playing in a school chess club, and a further 56 pupils were playing more informally at school. Twenty-nine children, including seven who had never played chess before the lessons, said they were now playing in chess clubs outside of school.

Continuation with the Chess in Primary Schools programme

Despite the wide scale acceptability of the programme to schools as a free one-year study intervention, when asked in the survey about whether they would pay for the programme during the following year, most teachers were very uncertain. The main barriers mentioned were cost, pressure on curriculum time, concern about the view of Ofsted and potential adverse effects on groups of learners, especially slower learners.

‘It’s a whole afternoon out essentially and...has really eaten into the curriculum—if they had all been completely engaged and excited about it that would be different but there has been this group that has struggled with enthusiasm.’ (Class teacher)

CSC reported that 24 treatment schools paid for their programme to continue, as part of the curriculum, in the school for a second year. Information on which year groups were receiving the programme in these schools was not available. However, no school was allowed to

continue the programme with the pupils who had participated while in Year 5, to keep the intervention dose the same for all participating schools.

Lessons for future implementation

Analysis across sources of process evaluation data provided a range of formative findings to inform future implementation:

1. **A set-up meeting between CSC head office staff and the headteacher** should be part of the 'sign-up' process. This could cover background detail for headteachers on the perceived benefits of learning chess (to help with justification for Ofsted), and expectations of CSC and of the school, and involve clear terms of reference and agreement with staff. Effective communication of the key points from this meeting from headteacher to class teachers is important. In particular headteachers should convey the expectation that class teachers and tutors team-teach and teachers learn to play chess alongside the children.
2. **A set-up meeting/training event on school premises between the chess tutor and relevant teaching staff** should be a part of the programme. This should be paid time for the tutor and would replace the current one-day CSC training day for teachers. The aims of this meeting would be: to establish a good teacher/tutor working relationship; for teachers to learn about the programme; for teachers and tutors to share expectations and requirements of their respective roles in the classroom; for tutors to learn about the classes receiving the lessons and how best to work with them; to teach non-chess-playing teachers the rudiments of chess.
3. The training provided by CSC for tutors, including the one-day introductory course, should **focus on teaching techniques and class management skills**. It should have at least some input from an experienced teacher/teacher trainer who would provide training on teaching techniques, particularly those that assist with differentiation in the classroom. Training for tutors should also focus on minimising risk to pupils who lose chess games frequently or find aspects of the partnering process stressful.
4. **Opportunities for more development of individual tutors' teaching skills** should be created. These could include: opportunities to work with and be mentored by a CSC tutor who excels as a teacher; putting in place systems that require schools to provide termly feedback, including pupil views, on tutor performance; and constructive suggestions for raising this.
5. **Audits to ensure tutor quality** should be conducted by CSC with associated tutor supervision and support where appropriate to improve performance.
6. **The chess lessons should be more interactive** and CSC tutors should be trained and supported in the use of new technologies.
7. **CSC should consider tutors working as a pair** in the classroom as the ideal model, to be achieved wherever possible. Other approaches that support effective differentiation in the classroom should be considered. Positive examples of splitting a

single class or two amalgamated classes, by chess-playing ability once sufficient chess classes have taken place for this to be assessed for each pupil, were reported by teachers. This approach was reliant on a teacher or teaching assistant being able to teach one of the two groups.

8. **The chess lessons should start at the beginning of the academic year** in order that the full 30 lessons can be readily fitted in and should not replace a maths or English lesson.
9. Reverting to the more common CSC practice of **delivering the lessons to a younger year group** would potentially reduce curriculum pressure concerns.

Control group activity

Fifty schools were initially randomised to the control group, but one refused their randomisation status, and booked CSC to deliver chess lessons to their Year 3 students.

Thirty-eight headteachers of schools in the control group completed the pre-randomisation baseline survey. Comparative analysis suggests that there was little difference at baseline between the two arms of the trial in terms of chess-playing activity in their school, with 45% of control schools and 48% of treatment schools saying that this occurred. In both trial arms seven of these chess-playing schools reported only occasional informal chess-playing in the classroom while four schools said there was a lot of chess-playing—including in a chess club and informally in the classroom. Five control schools and six treatment schools said they ran a chess club but did not appear to have any other chessplaying in the school.

The response rate to the follow-up control teachers' survey was extremely low (20%, see Table 13). This was despite concerted efforts by the research team, which included online, telephone and paper-based completion routes. This low response could be the result of their disappointment at their trial arm allocation. Equally though, this could also be an indication of lack of engagement or even knowledge of the programme by class teachers in schools that never received the intervention.

Our survey results, although limited, did not show any sign of compensation rivalry relating to chess initiatives within the control schools. A third of those that responded continued to offer chess clubs and recreational chess; none had external chess activities. We do not have a complete picture of whether extraordinary additional maths or English support was brought into control classes to compensate for the lack of the Chess in Primary Schools lessons. Four of the 15 control teachers who responded to the questionnaire had some additional maths programme for their Year 5 classes. This level, if replicated across the full set of control schools, would not be considered extraordinary. Unfortunately, the poor response means that we cannot be certain.

Conclusion

Key Conclusions

1. There is no evidence that the intervention had a positive impact on mathematics attainment for the children in the trial, as measured by Key Stage 2 scores one year after the intervention ended. The same is true for science and reading.
2. There is no evidence that the intervention had a positive impact on mathematics attainment for the children in the trial, as measured by Key Stage 2 scores one year after the intervention ended. The same is true for science and reading.
3. There is no evidence that the intervention had a positive impact on Key Stage 2 scores for children eligible for free school meals (FSM).
4. Although a current school teacher is allocated to every chess class, it is desirable for the tutors themselves to have strong class management and teaching skills. Without these, it was difficult to ensure that all children were suitably engaged in the chess lessons.
5. For successful implementation, class teachers need to work closely with the tutor and actively contribute to the intervention. It was felt that classes were less effective if the teacher did not actively take part, or was present only at the beginning and end of the class.
6. Half of the pupils who participated in the trial said that they liked the chess lessons a lot, and only 8% reported that they didn't like them. School teachers were very positive about the intervention and its impact on pupils' skills and behaviour.

Limitations

The findings outlined above should be considered within the context of the limitations of this study. The following factors particularly stand out:

1. Focus on academic achievement. The purpose of this trial was to examine the impact of Chess in Primary Schools upon children's academic achievement. Although we find little evidence of any impact, we cannot rule out the possibility that the programme has wider benefits for children. This includes potential impacts upon their well-being, self-confidence and non-cognitive skills.
2. Small 'dose' of the intervention. Children have been exposed to the Chess in Primary Schools intervention for just one academic year. This may be a relatively small 'dose' of the programme. A longer exposure may be needed to have a sustained impact upon educational achievement. Little is currently known about the cumulative impact of playing chess over a sustained period of time.
3. External validity. A strength of this RCT is that we have examined external validity, and considered how well the participants compared to the population eligible to receive the intervention. However, the population of interest was quite specific, and had

different characteristics from children in England as a whole. It therefore remains unknown how far our results generalise to the rest of the country.

Interpretation

The central hypothesis of this study was that teaching primary school children how to play chess would have a positive impact upon their educational achievement (measured one year after the intervention had finished). This RCT provided very little evidence in support of this hypothesis—the estimated effect on reading, science and multiple elements of mathematics after one year was essentially zero. This is in contrast to the only other large-scale RCT of the impact of chess on educational attainment that we are aware of, by Boruch and Romano (2011), who detected a substantial effect of more than 0.3 standard deviations for primary school children in Italy. Our results are also in contrast to another recent quasi-experimental study by Gumede and Rosholm (2015), who found a positive effect of chess on primary school children’s achievement in Denmark (effect size 0.15).

There are several possible explanations for this difference in results. First, our study was concerned with whether teaching children how to play chess had a *medium-term* impact upon their educational achievement (measured one year after the intervention had finished). In contrast, Boruch and Romano (2011) investigated the immediate impact, straight after the trial had finished. Consequently, their results are more likely to be subject to Hawthorne effects than ours. It is also possible that interventions of this nature have a short-term but not a medium-term impact on academic outcomes, which would explain why an impact was found in the 2011 study but not in this one. Second, our study has used high stakes, external tests as the outcome measure. This is in contrast to Boruch and Romano (2011), and indeed many other RCTs, where the use of low-stakes tests is common. It is possible that the treatment group will be more motivated than the control group when completing such low-stakes tests. Consequently, the study by Boruch and Romano (2011) may have actually been driven by a ‘test motivation’ effect. Third, the studies were conducted in very different settings. Although Boruch and Romano (2011) did not comment upon the external validity of their study, different findings in the UK should not be unexpected. Finally, we note that 7 of the 44 schools that delivered the intervention chose to deliver chess in place of a maths lesson. However, this is a small proportion of all participating schools, and our robustness checks indicate that this is unlikely to have an impact upon our substantive conclusion. **Security of findings**

The Education Endowment Foundation has designed a range of criteria to assess the security of research findings (available from http://educationendowmentfoundation.org.uk/uploads/pdf/Classifying_the_security_of_EEF_findings_FINAL.pdf). Independent peer reviewers are asked to rate each evaluation against five criteria (planned design, power, attrition, balance, and threats to validity). In Table 20 the evaluators present a summary of key pieces of evidence related to these criteria.

Table 20: Evidence regarding the security of research findings

Criteria	Evidence
Planned design	Randomised controlled trial
Power	100 clusters
	Minimum detectable effect from (apriori) power calculation ≈ 0.18
	Actual minimum detectable effect = 0.16
Attrition	0 (0%) of 100 clusters lost due to attrition
	144 (4%) of 4,009 pupils lost due to attrition
Balance	0.05 standard deviation difference in KS1 APS between treatment and control at baseline. Minimal difference in KS1 maths test score distribution.
	% FSM. Control = 36%. Treatment = 33%
Threats to validity	7% of clusters suffer from potential contamination. Robustness of findings tested by conducting a CA-ITT analysis.
	Randomisation, analysis and testing all conducted blind to treatment
	Key Stage 2 tests are high stakes, externally marked and non-specific to the intervention
Other markers	Long-term follow-up built into trial design via NPD
	Protocol published online
	Trial registered with independent organisation
	External validity / representativeness considered
	Randomisation conducted by independent evaluator

Future research and publications

We believe that this study has provided strong evidence that teaching primary school children how to play chess has little lasting impact upon their educational achievement. Future work should therefore concentrate on the potential wider benefits of chess, such as children's well-being and non-cognitive skills. The project team are planning to publish this study as an academic working paper and journal publication in 2016.

References

Aciego, R., García, L. and Betancort, M. (2012) 'The benefits of chess for the intellectual and socialemotional enrichment in schoolchildren'. *The Spanish Journal of Psychology*, 15: 551–9.

Boruch, R. and Romano, B. (2011) 'Does playing chess improve math learning? Promising (and inexpensive) results from Italy.' Available at: http://www.europechesspromotion.org/upload/pagine/doc/SAM_research_synthesis.pdf (accessed 11 December 2015)

Gumede, K. and Rosholm, M. (2015) 'Your move. The effect of chess on mathematics test scores.' IZA DP No. 9370. Available at <http://ftp.iza.org/dp9370.pdf> (accessed 8 October 2015)

Sussman, J. and Hayward, R. (2010) 'Using instrumental variables to adjust for treatment contamination in randomised controlled trials'. *British Medical Journal*, 340: c2073 doi: 10.1136/bmj.c207

Appendix A. School consent form to access the National Pupil Database (NPD)

“Chess in Schools and Communities” (CSC) programme

National Pupil Database (NPD) agreement form

This form is to be returned to Malcolm Pein, Programme Coordinator, by <INSERT DATE>.

As a school taking part in the “CSC” programme you agree to (i) provide some key information on pupils within your school, (ii) provide consent for the evaluation team at the Institute of Education to access pupils school records held on the National Pupil Database (NPD) and (iii) for the Institute of Education to link the test score data to any additional information collected through questionnaires as part of the CSC programme

The independent evaluation carried out by the **Institute of Education** requires this information in order to conduct a statistically robust evaluation of the CSC programme. Pupils’ test scores and any other pupil data will be treated with the strictest confidence. Named data will be matched with the National Pupil Database and shared with the Institute of Education and EEF for research purposes. No individual school or pupil will be identified in any report arising from the research.

I understand and agree that:

- The school consents to the use of National Pupil Database pupil data for purposes of this evaluation.
- That any data collected as part of the evaluation can be matched to individual NPD records, and that this data can be shared with the Institute of Education and Education Endowment Fund for research purposes (at a level of Tier 1 access).
- That the school will complete the attached spreadsheet capturing key information on year 5 pupils and send it (electronically) to Malcolm Pein by <INSERT DATE>.

Headteacher name:			
Headteacher signature:		Date:	

If you have any queries about the evaluation, please contact John Jerrim at the IoE at **J.Jerrim@ioe.ac.uk** or 07590761755.

Any queries relating to the CSC programme can be directed to Malcolm Pein, Programme Coordinator, at <INSERT EMAIL ADDRESS> or <INSERT TELEPHONE>

Appendix B. Questions CSC were asked to estimate costs

Question 1. Please could you provide an estimate of the average cost of the equipment needed to run the programme per school (e.g. Chess sets etc)

Question 2. How many hours, in total, did the regular class teacher have to attend the training in the CSC programme?

Question 3. How many hours training did the CSC tutors complete?

Question 4. What is the average hourly pay of the CSC tutors?

Question 5. What expenses do you pay the CSC tutors?

Question 6. Do you pay your tutors anything for 'preparation time'? If so, how much? And how many hours (on average) do they spend preparing per class?

Question 7. If a school wanted to take part in your programme next academic year, how much would you charge them?

Appendix C: Padlock rating

Figure 1: Summary grid of criteria for rating the security of evaluation findings

Criteria for interim rating			Adjust	Rating
Design	Power	Attrition		
Well conducted experimental design with appropriate analysis	MDES < 0.2	0-10%		5
Fair and clear quasi-experimental design for comparison (e.g. RDD) with appropriate analysis, or experimental design with minor concerns about validity	MDES < 0.3	11-20%	Adjustment for Balance [-]	4
Well-matched comparison (using propensity score matching, or similar) or experimental design with moderate concerns about validity	MDES < 0.4	21-30%	Adjustment for threats to internal validity [-]	3
Weakly matched comparison or experimental design with major flaws	MDES < 0.5	31-40%		2
Comparison group with poor or no matching (E.g. volunteer versus others)	MDES < 0.6	51-50%		1
No comparator	MDES > 0.6	<50%		0

Appendix D: Cost rating

Cost ratings are based on the approximate cost per pupil per year of implementing the intervention over three years. Cost ratings are awarded using the following criteria.

Cost	Description
£	<i>Very low:</i> less than £80 per pupil per year.
£ £	<i>Low:</i> up to about £200 per pupil per year.
£ £ £	<i>Moderate:</i> up to about £700 per pupil per year.
£ £ £ £	<i>High:</i> up to £1,200 per pupil per year.
£ £ £ £ £	<i>Very high:</i> over £1,200 per pupil per year.

APPENDIX 4

CANDIDATE PUBLICATION 4

Evaluation of community-level interventions to increase early initiation of antenatal care in pregnancy: Protocol for The Community REACH study, a cluster randomised controlled trial with integrated process and economic evaluations.

Sawtell M., Sweeney L., Wiggins M., Salisbury C., Eldridge S., Greenberg L., Hunter R., Kaur I., Hatherall B., Morris J., Reading S., Renton A., Adekoya R., Green B., Harvey B., Latham S., Patel K., Vanlessen L., Harden A.

Trials. 2018 Mar 5;19 (1):163.

ABSTRACT

Background

The provision of high-quality maternity services is a priority for reducing inequalities in health outcomes for mothers and infants. Best practice includes women having their initial antenatal appointment within the first trimester of pregnancy, in order to provide screening and support for healthy lifestyles, well-being and self-care in pregnancy. Previous research has identified inequalities in access to antenatal care, yet there is little evidence on interventions to improve early initiation of antenatal care. The Community REACH trial will assess the effectiveness and cost-effectiveness of engaging communities in the co-production and delivery of an intervention that addresses this issue.

Methods/ Design

The study design is a matched cluster randomised controlled trial with integrated process and economic evaluations. The unit of randomisation is electoral ward. The intervention will be delivered in 10 wards; 10 comparator wards will have normal practice. The primary outcome is the proportion of pregnant women attending their antenatal booking appointment by the 12th completed week of pregnancy. This and a number of secondary outcomes will be assessed for cohorts of women (n=approximately 1450 per arm) who give birth 2-7 and 8-13 months after intervention delivery completion in the included wards, using routinely collected maternity data. Eight hospitals commissioned to provide maternity services in 6 NHS trusts in north and east London and Essex have been recruited to the study. These trusts will provide anonymised routine data for randomisation and outcomes analysis. The process evaluation will examine intervention implementation, acceptability, reach and possible causal pathways. The economic evaluation will use a cost-

99

consequences analysis and decision model to evaluate the intervention. Targeted community engagement in the research process was a priority.

Discussion

Community REACH aims to increase early initiation of antenatal care using an intervention that is co-produced and delivered by local communities. This pragmatic cluster randomised controlled trial, with integrated process and economic evaluation, aims to rigorously assess the effectiveness of this public health intervention, which is particularly complex due to the required combination of standardisation with local flexibility. It will also answer questions about scalability and generalisability.

Trial Registration

ISRCTN registry: registration number 63066975.

Date of registration: 18th August 2015

Keywords: Access to care; Antenatal care; Cluster randomised controlled trial; Community engagement; Co-production; Maternity

BACKGROUND

Inequalities in maternal and infant mortality and morbidity are a challenge for public health policy and service delivery worldwide. In the UK the provision of high-quality maternity services is a priority for reducing national inequalities in health outcomes throughout pregnancy, birth and the subsequent life course of the mother and infant [1]. Antenatal care is the first step in maternity service provision for the pregnant woman. Antenatal care refers to the package of healthcare services provided throughout pregnancy, from conception to the onset of labour, and includes monitoring the health of the woman and fetus, providing medical and psychosocial support, and health promotion [2]. Under-utilisation of antenatal care is associated with adverse pregnancy outcomes including low birth-weight, neonatal mortality and maternal mortality [3, 4]. Current national guidelines recommend that women have a first contact (with a midwife or GP) followed by an antenatal care booking appointment with a maternity service within the first trimester of pregnancy (and ideally by 10 weeks), in order to fully benefit from the available screening, interventions and support [5]. At the first contact the focus is on provision of antenatal information and screening. The 'booking appointment' involves a health check and medical history, information on nutrition and exercise, information and offers of appropriate screening tests, as well as support for well-being and self-care in pregnancy. The booking appointment is also important for identifying women with social and medical risk factors so that these can be appropriately managed throughout the maternity pathway [6].

The timing of the booking appointment is associated with the quality and availability of health services, and the socio-demographic characteristics of pregnant women [2]. The percentage of women who attend their booking appointment by 12 completed weeks of pregnancy within each NHS maternity service is an indicator used by the UK Department of Health to monitor local and national inequalities in the provision and uptake of antenatal care [7]. Women from minority ethnic groups are less likely to have their booking appointment by the 12th completed week of pregnancy, in comparison to white women [8, 9]. Late booking is also associated with socio-economic deprivation [10]. This delayed initiation of antenatal care for many pregnant women living in marginalised communities is attributed to lower levels of health literacy within these communities (in relation to knowledge regarding the purpose and importance of antenatal care, and understanding of the healthcare system), along with reduced autonomy, resources and support to make use of the healthcare services that are available [11, 12, 13].

A systematic review by Oakley et al. [2] of interventions to increase early access to antenatal care for socially disadvantaged groups of women concluded that there was a lack of good-quality evidence on the effectiveness of such interventions, with no randomised controlled trials identified in this area. Interventions that were noted by the reviewers as potentially applicable to the UK setting and worthy of further development and examination included community-based programmes, where lay

women are trained to provide information on antenatal care and its availability to women within their communities. A review by Hollowell et al. ^[13] that focused on barriers to early initiation of antenatal care by women from minority ethnic communities in the UK, called for the development of interventions that promote the purpose and benefit of early and continued antenatal care in ways that take into account the cultural beliefs and practices of groups at risk of late booking. These authors further emphasised the need for information about antenatal care to be provided in a proactive and accessible format for women who may not be familiar with the UK health care system, or who may not speak English as their first language.

Community engagement to increase early initiation of antenatal care

Community engagement has been broadly defined as the “direct or indirect process of involving communities in decision making and/or in the planning, design, governance and delivery of services, using methods of consultation, collaboration, and/or community control” ^[14, 15]. It is thus an umbrella term for a variety of approaches and methods. In recent years, community engagement has been increasingly recognised as important in national policy and strategy documents for health service delivery, public health promotion and the reduction of health inequalities ^[16, 17]. Theories of community engagement suggest that involving communities as partners in the planning, design and delivery of interventions for health improvement leads to more appropriate and accessible interventions, and increased sense of ownership of the interventions and health outcomes ^[18]. A recent review of community engagement approaches to reducing health inequalities found them to be effective in improving health behaviours, health consequences, participant self-efficacy and perceived social support for disadvantaged groups ^[14]. Community engagement strategies that have targeted disadvantaged pregnant women and new mothers have been found to be most effective when peer delivery or collaborative models are used in interventions ^[19]. Community development is an example of a collaborative model in which communities and other organisations work together to co-produce locally focused activities, by building on existing relationships and assets within communities.

Despite the recent policy focus on utilising local strengths, knowledge and resources of communities to co-produce and deliver interventions for health and well-being, there have been few trials of the effectiveness of such interventions ^[20] and community engagement is not yet routinely embedded in mainstream commissioning and practice. Very few studies have looked at community engagement interventions in relation to antenatal care ^[19].

This paper presents the study protocol (version 2; 22.1.16) for the first cluster randomised controlled trial of a community-centred intervention that seeks to increase early initiation of antenatal care in communities where women are more likely to experience late initiation of antenatal care. The intervention uses community engagement approaches, including community development and peer delivery, as part of a locally focused ‘whole systems’ approach, which will engage and mobilise

community assets (e.g. local health care professionals such as midwives, nurses and GPs, faith groups, local businesses) and enhance local people's capabilities to provide advice and information in relation to health within their own communities.

Development of the Community REACH study

In 2010 members of the study team received UK National Institute for Health Research (NIHR) Programme Development Grant funding (Grant Reference Number RP-PG-1211-20015) to carry out exploratory research that would lead to the development of a new intervention to improve early initiation of antenatal care in urban settings with social disadvantage and ethnic diversity. We worked in an East London borough which has the largest proportion of births to mothers who were not born in the UK, at 76.4% ^[21]. Through epidemiological analysis, socio-demographic and clinical predictors of delayed access to antenatal care in this borough were identified ^[8]. Women identified as most vulnerable to late access included those: from ethnic minority communities; unable to speak English; born outside of the UK; with more than two children. Qualitative research uncovered several barriers to timely antenatal care attendance that corresponded to those identified in the literature outlined above ^[12]. Barriers identified included: difficulties navigating the referral system, especially if women were not already registered with a general practitioner or had limited or no English; lack of understanding regarding the value and benefits of early antenatal care; lack of agency and sense of entitlement to healthcare. As part of a public engagement focussed research process ^[22] a stakeholder workshop was held to plan for intervention development. This brought together maternity service users, maternity service managers, local healthcare commissioners, representatives of community organisations, and the research team. Workshop participants emphasised that the new intervention ought to work collaboratively with women: building on women's networks, empowering women, and harnessing local volunteering.

The Community REACH study was developed as a result of this exploratory work with stakeholders. It will test a local, focused whole systems intervention which aims to a) raise awareness in local communities of the value of antenatal care and its early uptake, and b) support women in how and when to access care, with the longer term aim to change local social norms which will sustain any increase in women's early access of antenatal care. The intervention uses a co-production process to engage local communities in: identifying their perceptions/views on the issues and solutions to increase early booking for antenatal care; tailoring the design of the intervention and form and content of key intervention messages; and facilitating the communication of the intervention messages through community self-help and local social networks. Our community engagement team will work with a co-host community organisation already established in each site, to support and implement the intervention at the local level. Peer volunteers, who are women from the local target community, will be recruited and trained for the role of 'antenatal care champions' to deliver the intervention messages through engagement with women

and wider family members, and local community groups and organisations (ranging from faith groups to pharmacies). A particular focus will be on reaching women from the groups identified through our previous epidemiological analysis to be most vulnerable to late initiation of antenatal care. The theoretical framework for the intervention is informed primarily by the concepts of community engagement and health literacy; the latter is defined as *“the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health”* [23]. Critically, the concept of health literacy goes beyond the ability to read health information and navigate health services, by increasing access to, and the ability and motivation to act on, health information. Figure 1 displays the theory of change in the form of a logic model for the Community REACH intervention.

The Community REACH study is one component of a wider programme of research, the REACH (Research for Equitable Antenatal Care and Health) Pregnancy Programme. With high priority given to public and practitioner involvement [22], the Programme is focused on improving access to, and experience of, antenatal care (Hayes 2012). The University of East London (UEL) is the lead academic partner. Organisations working with UEL on the Community REACH study include: University College London (UCL) Institute of Education providing trials expertise and management; the Pragmatic Clinical Trials Unit (PCTU) at Queen Mary University of London providing data management, statistical support and quality assurance; UCL conducting the economic evaluation. Uscreates, a design agency with a social focus, supported the initial co-design and communication strategy of the intervention. The study runs from 01.04.15 (when NHS Research Ethics Committee approval was received) to 11.10.19 (the end date of the REACH Pregnancy Programme).

This paper was written after funding and approvals were received for the Community REACH study, participating NHS Trusts were enrolled, randomisation of study sites and co-design workshops were completed, but prior to the intervention set-up and collection of any trial data.

METHODS/DESIGN

Study design

The study design is a two-armed matched cluster randomised controlled trial, with integral process and economic evaluations, see Figures 2 and 3 and Additional file 1, which present the study flow chart, Schedule of enrolment, interventions and assessments and Spirit Checklist as per Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) [24]. This design was selected because the interventions are delivered at the community rather than individual level. The unit (cluster) of randomisation and intervention delivery is electoral ward⁷. All outcomes

⁷ An electoral ward in England is a subdivision of a local authority area, typically used for electoral purposes.

are measured using anonymised routinely collected maternity data. The trial includes 20 electoral wards with high delayed rates of initiation of antenatal care, reflecting high rates of inequality in the ward populations. Ten wards were randomised to intervention and 10 to control, matched by these initiation rates and by pattern of use of hospitals in each ward.

Randomised trial design has traditionally required standardisation (or fidelity) of intervention delivery across experimental sites, but the effectiveness of community-based interventions is likely to depend upon whether they are responsive to local needs and contexts ^[25]. The Community REACH trial is one of a small, but growing, number of cluster-randomised controlled trials of community-led complex public health interventions, where the function and general approach of the interventions are standardised across study sites, but the exact nature of each intervention is adapted at the community level to suit the local context ^[20].

Study population

The study population is women in the selected electoral wards who give birth, at a hospital enrolled in the study, over a 12 month period. The timing of this period will be determined by the delivery of the intervention; all women in control and intervention sites will be included if they give birth 2-13 months after the intervention delivery has ended in intervention sites. There will be no active recruitment of participants to this trial as outcomes will be measured using anonymised routinely collected maternity services data. Individuals will be recruited to the integral process evaluation and this is described below.

Inclusion and exclusion criteria

The unit of randomisation is electoral ward. Inclusion criteria for electoral wards (clusters) are:

Where the proportion of women who have their first appointment for antenatal care with the maternity service by the 12th completed week of pregnancy is below the NHS national target of 90%, using data collected prior to randomisation over a six month period within each hospital.

Where historically the majority of pregnant women have chosen to access the maternity services of the commissioned maternity care provider for the area in which they live (i.e. the service participating in the study), as opposed to services provided by another NHS trust.

Exclusion criteria are:

Where the proportion of women who have their first appointment for antenatal care with the maternity service by the 12th completed week of pregnancy is 90% or above.

Where it is common for pregnant women to access maternity services other than those provided by the hospitals participating in the study.

Recruitment and selection of clusters for randomisation

For both substantive and pragmatic reasons NHS trusts in north and east London and Essex were targeted for recruitment to the study. The inclusion of an out of London area (Essex) is intended to enhance the generalisability of findings. Six NHS trusts agreed to take part, from which eight hospitals providing NHS maternity care (three providers from one large trust) enrolled in the study. NHS trusts were required to commit to providing routine maternity data for randomisation and assessment of outcomes as part of the participation agreement. Any trust unable to commit to this was not able to participate.

The research team identified 20 electoral wards eligible to include in the study, distributed across the geographical areas served by the participating hospitals. This identification process required informatics staff, already responsible for managing routine maternity data in each participating hospital, to extract retrospective 'gestation at booking' data for women using the maternity services of these hospitals during a defined recent six-month period. Routinely this maternity data is organised by postcode. In order to render it non-identifiable for study purposes, the informatics staff converted the postcode data to its corresponding electoral ward name using a postcode-to-ward database sourced from Local Authorities and provided to the informatics staff by the research team. Informatics staff transferred the data to the research team in a dedicated study database sent in an encrypted form via email. A pilot of this data transfer procedure was completed in one hospital, in order to ensure that all processes could be operationalised and audited.

From the databases of electoral wards, we selected 10 pairs of wards that met the eligibility criteria, ensuring that no wards neighboured one another to minimise intervention/control site contamination. Each pair of wards was matched on the hospitals accessed and baseline rate of antenatal booking by the 12th completed week of pregnancy, categorised as very low ($\leq 70\%$) or low (71-89%). A matched pair consisted of either two 'very low' wards or two 'low' wards, with similar patterns of hospital usage by women seeking antenatal care. An independent member of the research team, who was not involved in the recruitment of the trusts, oversaw the decision on the final list of 20 sites; to ensure the risk of selection bias was minimised. Each selected site was allocated a unique ID code, which was used in place of electoral ward names for randomisation purposes.

Randomisation, blinding and retention

Matched randomisation, of the 10 pairs of electoral wards with 1:1 ratio was undertaken remotely by the Pragmatic Clinical Trials Unit (PCTU) at Queen Mary's University London (QMUL) using Stata software (version 12; StataCorp, College Station, TX, USA). The research team were informed of the results of the allocation by secure email using a password protected file.

As with most social intervention trials, those involved in delivering the intervention in the clusters (i.e. the electoral wards) cannot be 'blinded' to allocation status.

Informatics staff in the participating services were given the names of the selected wards in their area but were not actively informed of the results of the randomisation.

Given the nature of this trial, sites cannot drop out of the trial, unless an NHS Trust withdraws from the study. We aim to work closely with local site principal investigators (PI) to ensure this does not occur. The use of routine data for assessment of outcomes means

limited missing data can be expected. Data sets provided for the site selection process (described above) confirmed this to be the case; as such no consideration of how to deal with missing data is indicated.

The intervention

The Community REACH intervention will be delivered in the 10 intervention sites (electoral wards), with each component of the intervention tailored to the local community thereby tapping into local assets and addressing local cultural beliefs and motivational barriers, and community perceived needs and solutions.

Phase 1: Mapping, community engagement and co-design

In developing the intervention plans the research team prepared a profile/map for each intervention site, in relation to current referral pathways to antenatal care, demographic information and community assets. Midwives working in the area provided local knowledge on barriers to accessing services. Staff from the design agency Uscreates and members of the UEL community engagement team then spent 2-3 days engaging with local people in each intervention site (through speaking to people at local community facilities, marketplaces, and other areas of local footfall). Local women and other family members were asked about experiences of antenatal care, perceived importance of antenatal care and their thoughts and opinions on the local area. A co-design workshop was held in each intervention site, facilitated by Uscreates. Local women who had registered their interest in the project during the street engagement, outlined above, were invited to attend, along with representatives from local community organisations and midwives working locally. Attendees participated in exercises to stimulate creativity and worked collaboratively on developing ideas for key messages, materials and events to improve early uptake of antenatal care in the local area.

Workshop participants highlighted the need for greater information about referrals to antenatal care, the services that are available, and the purpose and benefits of antenatal care. They felt that this information ought to be provided through local connections, networks and languages, i.e. women from the community engaging with other local women about antenatal care. In response to workshop outputs, a co-produced community-based intervention will be prepared and implemented in each site, as outlined below:

Phase 2: Set-up and training (3 months)

The intervention will be centrally co-ordinated by a community engagement team at UEL. A co-host community organisation will be recruited within each intervention site to support the local delivery of the intervention. Co-host organisations must meet certain criteria in order to be involved, for example: experience of managing, supporting and developing outreach teams; demonstrable experience of working with vulnerable groups including black and ethnic minority communities: strong links to local community groups and organisations and statutory health services.

Six to eight local people will be recruited in each site as voluntary antenatal care champions, to engage directly and indirectly with women and families, and locally specific groups and organisations from their community, to raise awareness of the value and benefits of early antenatal care and how and when to access care. The antenatal care champions will receive training for this role (which will involve input from a midwife) on the antenatal care system, including referral pathways and the purpose and benefits of antenatal care. Champions will also be trained in presentation and communication skills, ongoing practice reflection, adult and children safe-guarding, and health and well-being coaching.

Co-host organisations and antenatal care champions in each site will work collaboratively with the UEL community engagement team and the research team to build on the detailed profiles and mapping of community assets for each intervention site, and to further develop their local outreach plans for intervention implementation, which will engage the whole local system within each ward (e.g. schools, children's centres, GPs, pharmacies, faith groups).

A communications strategy and materials for the intervention messages, based on the outputs from the co-design workshops will be developed. The co-host organisations and antenatal care champions will also be part of the decision-making about whether and how the intervention messages ought to be tailored for each site.

Phase 3: Implementation (6 months)

Implementation of the intervention will take place, in each of the sites, immediately after the development phase. As this will be co-produced and locally tailored, the outreach and engagement plans in each site may vary.

Antenatal care champions are likely to engage with their local communities about antenatal care, particularly women who are vulnerable to later service access, through: presenting and discussing information with groups (e.g. at community events, evening classes, faith groups); one-to-one sessions, where antenatal care champions will engage with local people directly and indirectly in places of high footfall (e.g. GP surgeries, pharmacies, shopping centres); informal, opportunistic outreach, building on existing networks and relationships within the community.

We will use a staggered approach, with implementation taking place first in three 'pathfinder' sites before starting in the remaining seven sites, in order that intervention development and delivery can benefit from some initial learning.

The comparator

Normal maternity care promotion and practice will continue in the 10 electoral wards randomised to the control arm. All participating providers of maternity services are free to engage in additional actions to enhance early booking of antenatal care that affect populations in the electoral wards in either arm of the study.

Outcome measures

Based on the hypothesis that the intervention will promote earlier booking of care, particularly by those most at risk of late booking (as per the pathways shown in Figure 1), the primary outcome is the proportion of pregnant women in each ward who have attended their antenatal booking appointment by the end of the 12th completed week of their pregnancy. Secondary outcome measures are: the proportion of women who have attended their antenatal booking appointment by 10 weeks and 0 days of pregnancy; antenatal admissions; emergency caesarean rates; gestation and weight at delivery; maternal and infant death; APGAR score at five minutes; smoking at booking appointment and at birth; feeding method at discharge. Outcomes will be measured using anonymised routine hospital data. Outcomes data, covering periods of six months, will be extracted by trust informatics staff at 3 time points. These time points are as follows: time point 1 (baseline) - women giving birth who were at least 13 weeks pregnant at the point that any intervention related activities commenced; time point 2 (first follow up) - women giving birth 2-7 months after the end of intervention delivery; time point 3 (second follow up) - women giving birth 8-13 months after the end of intervention delivery. Further data on age, ethnicity, housing tenure, parity and deprivation will be provided along with the outcomes data.

The study will not require names, NHS numbers, dates of birth, addresses or postcodes. In order to ensure complete anonymisation, trust informatics staff will re-code potentially identifying individual level data, for example age and ethnicity, into broad categories, ensuring this categorisation does not compromise any proposed analyses. The name of each electoral ward will be replaced by a unique ward specific ID code that is not known to staff in the PCTU who will be analysing the data. In addition, an unblinded member of the PCTU data management team will liaise with informatics staff in each trust to ensure secure data transfer methods are understood and used when transferring data from the trust to the PCTU. They will then make the data accessible to their data management colleagues who will remain blind to study site. The data management team will ensure the secure storage of study datasets for statisticians to conduct their analysis. A pilot of the outcomes data transfer process will be conducted across the study sites to ensure procedures are deliverable by informatics staff and fit for purpose. A collaborative approach involving the research

team, informatics staff and site PIs, to finalising the exact list of outcomes and the data transfer processes will be taken.

Power and sample size

In our Programme Development Grant research, we analysed routine data from the maternity service at Newham University Hospital, which contained the corresponding postcode for each pregnancy in Newham from the period April 2007-January 2011. We calculated by ward the variation in cluster size and the proportion accessing antenatal care by 12th completed week of pregnancy as the basis for estimation of intra-cluster correlation coefficient and our sample size calculation. To detect an increase in antenatal booking by 12 weeks gestation from 75% to 82%, with 90% power at the 5% significance level requires at least 751 individuals in each trial arm. To account for clustering (ICC = 0.005, mean cluster size 145, matching correlation = 0.1) requires nine clusters in the intervention group and in the control group, which equates to 1450 women per trial arm. Because this sample size is relatively small, and to guard against a substantial loss of power if a cluster is lost for any reason, we added one cluster to each group. As described, pairs of clusters were matched by the hospital used by those in the ward seeking maternity care and by baseline rate of antenatal booking by 12 weeks. This has been measured per ward and collected pre-randomisation.

Process evaluation

Ongoing formative evaluation work will be conducted alongside Phases 1 and 2 of the intervention, in order to inform the development of the intervention structure, activities and materials. Research team members will conduct observations and interviews to document and analyse the activities in each intervention site during the stages of community engagement, co-design, intervention set-up and training of antenatal care champions. The UEL community engagement team will provide records on intervention set-up in each site, and co-host organisations will provide records on recruitment of antenatal care champions.

A process evaluation will be conducted for Phase 3 of the intervention, to explore its implementation. This integral process evaluation will run alongside the impact evaluation

Process data will be used to examine intervention implementation, acceptability and reach, issues of context, as well as hypothesise possible causal pathways, in order to facilitate interpretation of outcome data. In line with recent Medical Research Council (MRC) guidance on process evaluations for complex interventions^[26], this component of the trial will also enable refinement of the intervention logic model (see figure 1).

The process evaluation has four specific components. These are:

(i) Documentation and analysis of the local social contexts in intervention and control sites

The mapping process described above in Phase 1 of the intervention will be expanded by the research team to document and analyse the local social context of the intervention and control sites and their main maternity providers, in order to identify events and influences that may hinder or support intervention implementation, or affect study outcomes. Information will also be recorded on all previous and current interventions/ innovations regarding antenatal care that have been developed by maternity services, local authorities or community organisations for people living in the study sites.

(ii) Documentation and analysis of intervention activities

Intervention activities will be documented and analysed for each site. Members of the research team will observe up to two purposively selected events in each of the intervention sites, where antenatal care champions engage with their local community about antenatal care. The UEL community engagement team and the co-host organisations will provide detailed records of intervention delivery and progress.

(iii) Interviews with antenatal care champions, co-host representatives and other stakeholders

The experiences of those taking part in the co-production and delivery of the intervention, and their perceptions regarding its impact on their communities, will be explored through qualitative research interviews. Up to 40 people, across 3-4 intervention 'case study' sites (final numbers dependent on data saturation), will be purposively sampled for individual interviews. Participants will be interviewed at the beginning and again towards the end of the intervention delivery phase. In the first interview, interviewees will be asked about their motivations for getting involved with the intervention, their expectations for how the intervention may be received within their local communities, their experiences and perspectives on the training process, and their early experiences with delivering the intervention messages around antenatal care. The second follow-up interview will be conducted with each participant approximately 6 months following their first interview, to address experiences with delivering the intervention and perceived intervention acceptability within the community. We will also aim to conduct interviews with those who cease their involvement with the intervention delivery while it is ongoing, in order to explore reasons for stopping. Interviews will also be conducted with local community midwives to gain insight into local maternity service provision and local barriers/enablers around access to antenatal care within the intervention site. Interviewees will receive a £10 voucher as reimbursement for their time. The qualitative research interviews will explore the social contexts within which the intervention is implemented, building a richer picture of intervention delivery and mechanisms of impact. The number of interviews to conduct in the remaining intervention sites, and who to interview, will be decided pragmatically and judiciously based on the findings from earlier sites.

(iv) Survey to assess exposure to the intervention and its influence

A sample of 400 women across the 20 trial sites will be surveyed to assess reach, exposure to and acceptability of the intervention. The survey will be conducted with women attending appointments at antenatal booking clinics three months following the start of the implementation of the intervention. We will aim to conduct this with an approximate ratio of 3:1 intervention to control women surveyed. Hospital staff will identify eligible women, when they attend the clinic, based on their postcode. Women who agree to participate, will be provided with the survey in written, self-administered form, with the offer of researcher or bilingual health advocate support where required. The survey will be translated into other languages that are common to women using that particular maternity service. The survey will not contain sensitive or personal questions regarding the woman's pregnancy, but rather will focus on the woman's first point of contact in her antenatal care pathway, whether she had heard of the intervention and whether it had any effect on her decision-making about the timing of antenatal care initiation. Participants will receive a £5 voucher as a thank you for their participation.

Women who would like additional time to make a decision about participation in the survey will be provided with a link to an online version that they can complete at a future date. On completion they will be invited to send their address by email to the research team in order that a £5 voucher can be posted to them.

Economic evaluation

A cost-consequences economic evaluation, will evaluate the effectiveness of the community-based intervention compared to current practice. Information collected during the process evaluation will be used to calculate the cost of the intervention for each ward including development and implementation costs. Health care resource use, collected for both trial arms, will include: antenatal bookings and appointments; antenatal admissions; mode of delivery with a focus on emergency cesareans; costs associated with pre term births and low birth weight, maternal and infant deaths.

ANALYSIS

Trial outcomes A cluster-level analysis, appropriate to the analysis of matched cluster randomised trials, will be used with the maternity care providers for the NHS trusts participating in the study as fixed effects. We will use intention to treat principles. We will include individual level prognostic covariates if appropriate and ward-level estimates of baseline levels of outcomes as covariates. These will be chosen in advance of any analyses being conducted and documented in a full analysis plan. In additional analyses we will explore the use of instrumental variable techniques to incorporate some process measures as mediators of effect. Our primary analyses will consider primary and secondary outcomes data pertaining to births in the period 2 – 7 months post intervention start. Further secondary analyses will be conducted using data for births in the period 8-13 months post intervention completion to explore the maintenance of any effects and trends. A small number of sub group analyses may

also be conducted using Stata software (version 14; StataCorp, College Station, TX, USA) .

Process evaluation

All qualitative interview data will be managed and coded using QSR International's NVivo 11 qualitative data analysis software. Interview data will be subjected to thematic analysis. Codes will be applied to transcripts, to identify key themes and how these inter-relate in order to develop an analytical framework. Each transcript will also be coded to indicate the type of participant and electoral ward allowing analytical themes to be explored in relation to different groups' experiences and to compare processes across intervention areas. Drawing on methods associated with 'grounded theory', constant comparisons will be made and deviant cases examined to refine the analysis.

Observation data will be recorded on a semi-structured proforma. Thematic analysis of this data will also be undertaken.

Data from the surveys will be analysed using the current version of IBM SPSS statistical software. Descriptive analysis will be conducted to assess the key themes relating to awareness and level of involvement with the intervention, attitudes towards it, and views on antenatal care.

Economic evaluation

Costs and health care resource will be reported alongside primary and secondary outcomes for each trial arm. Missing data will be assumed to be missing at random and available case analysis used following the principles set out in the statistical analysis plan. We will report 95% confidence intervals calculated using bootstrapping.

A decision analytic model will also be developed to extrapolate the outcomes collected, antenatal admissions and emergency caesarean, their impact on costs and health outcomes and costs published in the literature for pre-term and low weight births. We will use a simplified decision analytical model to look at the benefits and disadvantages of the intervention, including the impact on the 12th completed week of pregnancy target, to assist NHS Trusts with decisions about implementation; synthesising information from other sources where at all possible.

SAFETY AND TRIAL CONDUCT

There are no anticipated risks to study participants or to those involved with the intervention and the first outcomes data collection will occur after the intervention delivery is complete. Therefore, there will be no Data Monitoring Committee for this trial. This decision, made by the Chief Investigator (AH) and members of the Trial Steering Committee (TSC; see below), is due to the reliance on routine monitoring data that is provided at source in non-identifiable form. However, as in all interventions, there may be unanticipated risks and harms will be assessed through examination of outcomes at the two time points.

The UK MRC Guidelines on Good Clinical Practice in Clinical Trials ^[27] will be followed. The University of East London, the employer of AH, will act as the sponsor of this trial. The trial will be overseen by a TSC. This group will meet face to face once a year and will be responsible for overseeing the trial, ensuring scientific quality and clinical relevance, and adherence to ethics and research governance. All key collaborators on the trial will attend the TSC, as well as a range of experts who are not directly involved in the trial, including a chair with relevant expertise, a statistician and an economist. There will also be a maternity service user representative on the TSC. Bi-monthly trial management meetings, with the PCTU, will be held and study team meetings, involving AH, MW, MS, LS and CS, will take place once a month to oversee day-to-day progress.

DISSEMINATION

The findings of the trial will be presented at national and international conferences (e.g. Royal Colleges of Midwives annual conference, the International Confederation of Midwives Congress and relevant national public health conferences). They will also be published in peer reviewed academic journals and in professional and practitioner journals. Findings will also be made available on the study website and in newsletters. Briefing papers to healthcare commissioners and managers and to service users via Maternity Voices Partnerships, will be prepared. We will use links with the Reproductive and Childbirth topic network to further disseminate throughout the NHS.

DISCUSSION

The Community REACH trial is one of a growing number of randomised controlled trials of public health interventions. The antenatal intervention being tested in the trial builds on evidence of effectiveness of lay or peer-delivered interventions when using community engagement strategies to provide health interventions to vulnerable or disadvantaged populations. A number of elements of the study will aid generalisability and scalability if effectiveness is shown. These elements include the integrated process and economic evaluations, the range of participating providers of maternity care in the trial, the flexibility of the intervention and the central involvement of local community members and community organisations. The application of a cluster randomised controlled trial design to the testing of an intervention that combines standardisation of overall approach with adaption to the local context will make a valuable contribution to the existing body of work on study design, as will the use of routine hospital data for outcomes analysis and the collaborative approach to research processes. If the intervention is shown to be effective it will be of benefit to those who received it and to society generally, in terms of improved health and associated reduction in cost to society and the NHS.

TRIAL STATUS

The study protocol reported here is version 2 (22.1.16). There will be no active recruitment of participants to this trial as outcomes will be measured using

anonymised routinely collected maternity services data. Individuals will be recruited to the integral process evaluation; recruitment began in June 2017 and is expected to finish in June 2018.

LIST OF ABBREVIATIONS

CLAHRC - Collaboration for Leadership in Applied Health Research and Care

IT -Information Technology

MRC - Medical Research Council

NIHR - National Institute for Health Research

PCTU – Pragmatic Clinical Trials Unit

PGfAR - Programme Grants for Applied Research

PI - Principal investigators

SPIRIT - Standard Protocol Items: Recommendations for Interventional Trials

QMUL - Queen Mary's University London

TSC - Trial Steering Committee

UCL – University College London

UEL - University of East London

DECLARATIONS

Ethics approval and consent to participate

The study has been approved by the NHS Health Research Authority National Research Ethics Committee North East-York (27/03/2015, ref.15/NE/0106) and the appropriate permissions for all the participating sites secured. All information collected during the trial will be kept confidential and adhere to the 1998 Data Protection Act.

There is no individual consent process for the outcomes evaluation due to the reliance on anonymous routine data. For the observation of community co-design workshops, participants were informed before the workshop began, that the researcher would not record any personal or identifiable data concerning anyone in attendance. Remaining in the workshop was interpreted as participants giving consent to the researcher making non-identifiable notes of the workshop processes. All potential survey and interview participants will be taken through information and consent procedures with verbal and written information provided before written consent is requested. Potential participants will be given adequate time to consider whether or not they want to participate – with contact details given to the research team should they want more time for this. Bilingual research assistants or healthcare

interpreters (in antenatal clinics) will be included in the informed consent and data collection processes to provide language support where necessary.

Consent for publication

Not applicable

Availability of data and material

Audio recordings and consent forms will be destroyed when the final report of the study has been completed (October 2019). When the study has ended, data that is suitable for open sharing will be stored (along with relevant metadata and documentation) in the data repository of UEL (the sponsor organisation for the study)- data.uel- without any restrictions (Open Data). As this data is anonymised/ not personal no security will be required. In line with the Research Data Management policy of UEL (sponsor organisation for the study) any archived research data will be reviewed (for removal) every five years, and if appropriate destroyed. Data that is not suitable for sharing will be securely stored in UEL's Arkivum data archive. Data is encrypted and only project personnel and UEL administrative staff will have access to it. This security will be handled by UEL IT (information technology) by restricting folder access. There will be the same 5 year review which will look at whether the data should be retained. In both cases, the destruction would involve secure erasing of the data in consultation with UEL IT. The appraisal of the data after five years will be undertaken by the Chief Investigator on the study (Angela Harden).

Competing Interests

None

Funding

This paper presents independent research funded by the National Institute for Health Research (NIHR) Programme Grants for Applied Research (PGfAR) Programme (Grant Reference Number RP-PG-1211-20015). The funding body influenced the study design in the initial application stage, by recommending that the study was expanded to a wider geographical area than originally planned. This involved the recruitment of several additional trusts to the study. The funding body will not be involved in the collection, analysis and interpretation of data or in the writing of the manuscript.

Author's contributions

The chief and co-investigators on the REACH pregnancy programme contributed to conceptualising and designing the Community REACH trial (AH, AR, MW, SE, RH, JM, SR, CM, BH and RA) with input from LS, MS, IK and GF. All of these individuals also contributed to the protocol. The wider REACH pregnancy programme team also contributed when additional informed views were required about aspects of Community REACH. The local principal investigators from NHS Trusts beyond Barts Health NHS Trust contributed to decisions about outcomes measures and to the design of the procedures for transferring routine NHS data (BG, BH, IK, SL, KP, and

LV). CS is carrying out a PhD as part of the process evaluation of the study and has been particularly involved in this aspect of the protocol. All authors have carried out the work on the study so far. MS and LS drafted the manuscript; all other authors read, edited, and approved it.

Acknowledgments

The authors would like to thank the staff in the NHS trust informatics teams who provided routine maternity data to support site randomisation and all the practitioners and members of the public who have assisted with the trial so far. They would also like to thank staff at Uscreates and Kevin Sheridan and Asmat Syed at UEL for their community engagement, co-design and communication strategy input and members of the data management team at Queen Mary University of London PCTU who advised on all data management matters. Angela Harden was (in part) supported by the NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) North Thames at Bart's Health NHS Trust. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

References

- [1] Marmot M, Allen J, Goldblatt P et al. Fair society, healthy lives: strategic review of health inequalities in England post 2010. London: Marmot Review Team; 2010.
- [2] Oakley L, Gray R, Kurinczuk JJ, Brocklehurst P, Hollowell J. A systematic review of the effectiveness of interventions to increase the early initiation of antenatal care in socially disadvantaged and vulnerable women. Oxford: National Perinatal Epidemiology Unit; 2009.
- [3] Partridge S, Balayla J, Holcroft CA, Abenheim HA. Inadequate prenatal care utilization and risks of infant mortality and poor birth outcome: a retrospective analysis of 28,729,765 U.S. deliveries over 8 years. *Am J Perinatol*. 2012;29:787–794.
- [4] Raatikainen K, Heiskanen N, Heinonen S. Under-attending free antenatal care is associated with adverse pregnancy outcomes. *BMC Public Health*. 2007; doi:10.1186/1471-2458-7-268
- [5] National Institute for Health and Clinical Excellence. Antenatal Care for Uncomplicated Pregnancies: NICE guidelines [CG62]. Manchester: National Institute for Health and Clinical Excellence; 2008.
- [6] Dhange P, Breeze ACG, Kean LH. Routine antenatal management at the booking clinic. *Obstet Gynaecol Reprod Med*. 2013;23(2):45-52.
- [7] National Institute for Health and Clinical Excellence. Antenatal Care: Quality Standard [QS22]. Manchester: National Institute for Health and Clinical Excellence; 2012.

- [8] Cresswell J, Yu G, Hatherall B, Morris J, Jamal F, Harden A, Renton A. Predictors of the timing of initiation of antenatal care in an ethnically diverse urban cohort in the UK. *BMC Pregnancy Childbirth*. 2013; doi: 10.1186/1471-2393-13-103
- [9] Henderson J, Gao H, Redshaw M. Experiencing maternity care: the care received and perceptions of women from different ethnic groups. *BMC Pregnancy Childbirth*. 2013; doi: 10.1186/1471-2393-13-196.
- [10] Lindquist A, Kurinczuk JJ, Redshaw M, Knight M. Experiences, utilisation and outcomes of maternity care in England among women from different socio-economic groups: findings from the 2010 National Maternity Survey. *BJOG*. 2015;122(12):1610-1617.
- [11] Downe S, Finlayson K, Walsh D, Lavender T. 'Weighing up and balancing out': a meta-synthesis of barriers to antenatal care for marginalised women in high-income countries. *BJOG*. 2009;116:518–529.
- [12] Hatherall B, Morris J, Jamal F, Sweeney L, Wiggins M, Kaur I, et al. Timing of the initiation of antenatal care: An exploratory qualitative study of women and service providers in East London. *Midwifery*. 2016;36:1-7.
- [13] Hollowell J, Oakley L, Vigurs C, Barnett-Page E, Kavanagh J, Oliver S. Increasing the early initiation of antenatal care by black and minority ethnic women in the United Kingdom: a systematic review and mixed methods synthesis of women's views and the literature on intervention effectiveness. Oxford: National Perinatal Epidemiology Unit; 2012.
- [14] O'Mara-Eves A, Brunton G, McDaid D, Oliver S, Kavanagh J, Jamal F, et al. Community engagement to reduce inequalities in health: a systematic review, meta-analysis and economic analysis. *Public Health Res*. 2013;1(4).
- [15] Popay J, Whitehead M, Carr-Hill R, Dibben C, Dixon P, Halliday E, et al. The impact on health inequalities of approaches to community engagement in the New Deal for Communities regeneration initiative: a mixed-methods evaluation. *Public Health Res*. 2015;3(12)
- [16] National Institute for Health and Clinical Excellence. *Community Engagement: NICE Guidelines [PH9]*. Manchester: National Institute for Health and Clinical Excellence; 2008.
- [17] Public Health England. *A guide to community-centred approaches for health and wellbeing: Full report*. London: Crown Copyright; 2015
- [18] Popay J, Attree P, Hornby D, Milton B, Whitehead M, French B, et al. *Community engagement in initiatives addressing the wider social determinants of health. A rapid review of evidence on impact, experience and process*. Lancaster: University of Lancaster; 2007.

- [19] Brunton G, O'Mara-Eaves A, Thomas J. The 'active ingredients' for successful community engagement with disadvantaged expectant and new mothers: a qualitative comparative analysis. *Journal of Advanced Nursing*. 2014;70(12): 2847-2860.
- [20] Phillips G, Bottomley C, Schmidt E, Tobi P, Lais S, Yu G, et al. Well London Phase 1: Results among adults of a cluster randomised trial of a community engagement approach to improving health behaviours and mental well-being in deprived inner city neighbourhoods. *J Epidemiol Community Health*. 2014; doi:10.1136/jech-2013-202505.
- [21] Office for National Statistics. Statistical bulletin: Parents' Country of Birth, England and Wales: 2015. London: Office for National Statistics; 2016.
- [22] Hayes H., Buckland S., Tarpey M. INVOLVE Briefing notes for researchers: Public involvement in NHS, public health and social care research. Eastleigh: INVOLVE Coordinating Centre; 2012; http://www.invo.org.uk/wp-content/uploads/2014/11/9938_INVOLVE_Briefing_Notes_WEB.pdf
- [23] World Health Organization. Health Promotion Glossary. Geneva: WHO; 1998.
- [24] Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200-207.
- [25] Hawe P, Shiell A, Riley T. Complex interventions: how "out of control" can a randomised controlled trial be? *BMJ*. 2004;328:1561-1563.
- [26] Moore G, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: UK Medical Research Council (MRC) guidance. London: MRC Population Health Science Research Network; 2014.
- [27] MRC Guidelines for Good Clinical Practice in Clinical Trials. London: Medical Research Council; 1998.

APPENDIX 5

CANDIDATE BIBLIOGRAPHY

Sawtell, M., Sweeney, L., Wiggins, M., Salisbury, C., Eldridge, S., Greenberg, L., et. al. (2018). Evaluation of community-level interventions to increase early initiation of antenatal care in pregnancy: protocol for the Community REACH study, a cluster randomised controlled trial with integrated process and economic evaluations. *Trials*, 19 (1), 163. doi:10.1186/s13063-018-2526-6

Wiggins, M., Sawtell, M., Wiseman, O., McCourt, C., Greenberg, L., Hunter, R., Eldridge, S., Hoara, P., Kaur, I., Harden, A. (2018). Testing the effectiveness of REACH Pregnancy Circles group antenatal care: protocol for a randomised controlled pilot trial. *Pilot and Feasibility Studies* 4:169. doi.org/10.1186/s40814-018-0361-x

Jerrim, J, Macmillan, L., Micklewright, J., Sawtell, M., Wiggins, M. (2017). Does teaching children how to play cognitively demanding games improve their educational attainment? Evidence from a Randomised Controlled Trial of chess instruction in England. *Journal of Human Resources* 53 (4) pp 993-1021

Christie, D., Thompson, R., Sawtell, M., Allen, E., Cairns, J., Smith, F., . . . Viner, R. M. (2016). Effectiveness of a structured educational intervention using psychological delivery methods in children and adolescents with poorly controlled type 1 diabetes: a cluster randomised controlled trial of the CASCADE intervention. *BMJ Open Diabetes Research & Care*.

Mary Sawtell, Liz Jamieson, Meg Wiggins, Felicity Smith, Anne Ingold, Katrina Hargreaves, Meena Khatwa, Brooks, L., Thompson, R., & Christie, D. (2015). Implementing a structured education program for children with diabetes: Lessons learnt from an integrated process evaluation. *BMJ Open Diabetes Research and Care*, 3 (1). doi:10.1136/bmjdr-2014-000065

Schucan Bird, K. L., Newman, M., Hargreaves, K., & Sawtell, M. (2015). Workplace-based learning for undergraduate and pre-registration healthcare professionals: A systematic map of the UK research literature 2003-2013. London: UCL Institute of Education.

Christie, D., Thompson, R., Sawtell, M., Allen, E., Cairns, J., Hargreaves, K. et al. (2013). Maximising engagement, motivation and long term change in a Structured Intensive Education Programme in Diabetes for children, young people and their families: Child and Adolescent Structured Competencies Approach to Diabetes Education (CASCADE): A Multi Centre Cluster Randomised Control Trial with integral process and economic evaluation.

Christie, D., Thompson, R., Sawtell, M., Allen, E., Cairns, J., Smith, F., . . . Viner, R. (2013). Structured intensive education maximising engagement, motivation and long term change for children and young people with diabetes: : cluster randomised controlled trial with integral process and economic evaluation : The CASCADE study. *Health Technology Assessment Monographs*.

Oakley, A., Wiggins, M., Strange, V., Sawtell, M., & Austerberry, H. (2011). Becoming a Mother: continuities and discontinuities over three decades. In E. F, J. M, R. MPM, H. J (Eds.), *Birth Rites and Rights* (pp. 9-27). Oxford: Hart Publishing.

Sawtell, M., Ingold, A., Wiggins, M., & Austerberry, H. (2010). Innovative services to improve the health and well-being of vulnerable young people.

Sawtell, M., Austerberry, H., Ingold, A., Strange, V., Stevens, M., & Wiggins, M. (2009). Evaluation of the Teenage Health Demonstration Sites Programme: Final Report.

Wiggins, M., Bonell, C., Sawtell, M., Austerberry, H., Burchett, H., Allen, E., & Strange, V. (2009). Health outcomes of youth development programme in England: Prospective matched comparison study. *British Medical Journal*, 339.

Austerberry, H., Sawtell, M., Ingold, A., Wiggins, M., Arai, L., & Strange, V. (2008). Evaluation of the Teenage Health Demonstration Sites Programme: 1st Annual Report 2007.

Sawtell, M., Austerberry, H., Ingold, A., Wiggins, M., Strange, V., Arai, L., & Aicken, C. (2008). NHS Teen LifeCheck Evaluation: Final Report.

Wiggins, M., Bonell, C., Burchett, H., Sawtell, M., Austerberry, H., Allen, E., & Strange, V. (2008). Young People's Development Programme Evaluation: Final Report.

Wiggins, M., Bonell, C., Burchett, H., Sawtell, M., Austerberry, H., Allen, E., & Strange, V. (2008). Young People's Development Programme Evaluation: Executive Summary. London: Social Science Research Unit, Institute of Education, University of London.

Sawtell, M., & Strange, V. (2007). An evaluation of Let's Leave it Till Later' training for youth workers.

Austerberry, H., Wiggins, M., Rosato, M., Sawtell, M., & Oliver, S. (2007). Supporting pregnant teenagers and teenage parents: The Sure Start Plus initiative. In P. Baker, K. Guthrie, C. Hutchinson, R. Kane, K. Wellings (Eds.), *Teenage Pregnancy and Reproductive Health* (pp. 189-198). London: RCOG Press.

Wiggins, M., Sawtell, M., Austerberry, H., Burchett, H., & Strange, V. (2006). Evaluation of the Young People's Development Programme: first interim report.

Wiggins, M., Sawtell, M., Austerberry, H., Burchett, H., & Strange, V. (2006). *Evaluation of the Young People's Development Programme: first interim report.*

Sawtell, M., Wiggins, M., Austerberry, H., Rosato, M., & Oliver, S. (2005). *Reaching out to pregnant teenagers and teenage parents? Innovative practice from Sure Start Plus programmes.* London: Social Science Research Unit, Institute of Education, University of London.

Wiggins, M., Rosato, M., Austerberry, H., Sawtell, M., & Oliver, S. (2005). *Supporting Teenagers who are pregnant or parents. Sure Start Plus National Evaluation: Executive Summary.*

Wiggins, M., Rosato, M., Austerberry, H., Sawtell, M., & Oliver, S. (2005). *Sure Start Plus National Evaluation: Final Report.*

Wiggins, M., Oakley, A., Sawtell, M., Austerberry, H., Clemens, F., & Elbourne, D. (2005). *Teenage Parenthood and Social Exclusion: A multi-method study.*

Sawtell, M., Wiggins, M., Austerberry, H., Oakley, A., Clemens, F., & Elbourne, D. (2003). *Involving users in research: Report of focus group discussions with current and previous teenage mothers.*

Wiggins, M., Austerberry, H., Rosato, M., Sawtell, M., & Oliver, S. (2003). *Sure Start Plus National Evaluation Service Delivery Study: Interim Findings.*

Sawtell, M. (2002). *Lives on Hold: homeless families in temporary accommodation.* London: Maternity Alliance.

Sawtell, M. (2002). *Time to listen: an account of the role of support health visitors.* *Community Practitioner*, 75 (2), 461-463.