Ministry of Housing, Communities & Local Government

Evaluation of Community-Based English Language Provision

Process evaluation of the implementation of the Community-Based English Language Randomised Controlled Trial

March 2018
National Learning and Work Institute
Ministry of Housing, Communities and Local Government

This document/publication is also available on our website at www.gov.uk/MHCLG

Any enquiries regarding this document/publication should be sent to us at:

Ministry of Housing, Communities and Local Government
Fry Building
2 Marsham Street
London
SW1P 4DF
Telephone: 030 3444 0000

For all our latest news and updates follow us on Twitter: https://twitter.com/mhclg

March 2018

ISBN: 978-1-4098-5216-2

Abstract

This evaluation was commissioned by MHCLG to review the process of implementing a randomised controlled trial (RCT) on a Community-Based English Language (CBEL) intervention.

The process evaluation sought to:

- explore the implementation of the CBEL intervention, identify any aspects of variation in delivery and make an overall assessment of whether fidelity to the RCT design was achieved;
- identify key lessons, issues and challenges in implementing this social intervention as a randomised controlled trial;
- generate evidence, learning and recommendations to feed into the design and implementation of future RCTs and English language interventions aimed at similar groups.

The RCT was delivered to 527 individuals in Greater Manchester and West Yorkshire to improve basic English skills. Learners in the treatment group began classes in April 2016 and those in the control group began their classes in September 2016 (after completion of the trial). The RCT sought to test the effect of the CBEL intervention on English language proficiency and social integration.

The process evaluation uses data from qualitative interviews with 54 individuals involved in the trial (both learners and those responsible for delivering the intervention and assessments). It also draws on analysis of attendance records and RCT information, in order to add further insight into the process and implementation of the trial. A separate report sets out the findings of the RCT itself.

Conducting the RCT involved strict approaches to recruitment, assessment, course design and delivery. Coordinators found the task of recruiting enough learners to fill the classes a challenge given the short timeframe and restrictive eligibility criteria. The target for recruitment was 600; however, the final number of learners recruited was 527. Learners were assigned to either the treatment or control group on a random basis, clustering learners who signed up with friends or family to attend the course. Learner clusters were more common than first anticipated; however, balanced treatment and control groups were achieved.

We conclude from analysis of the evidence collected that the trial was implemented effectively, with high fidelity to the experimental design and to the design of the course itself. Area leads, teachers and independent assessors reported a good understanding of the aims and objectives of the trial, understood the design of the intervention, and were generally equipped to deliver it effectively.

The main conclusions of this report are that:

- The randomisation of learners was completed successfully. Teachers and delivery partners were sufficiently trained to be aware of the importance of the allocation process and were proactive in preserving the allocations made, despite some participants wishing to change groups.
- There was no evidence of control group participants being exposed to the intervention or receiving course material prematurely (known as control group contamination). The RCT was successfully designed and implemented to avoid participants mixing once they had been allocated to a trial group. In particular, the clustering of learners who were known to each other at the point of randomisation did reduce this risk.
- Accounts of teachers and learners suggest that despite some minor variation in the delivery of the course, in relation to vocabulary covered in specific sessions and strategies for dealing with absence, this variation did not compromise the overall consistency of the intervention, or the trial. Teachers were motivated to adjust materials with the aim of accommodating learners, while still achieving the aims of the specific class.

The report presents recommendations for the delivery and evaluation of future programmes of this kind.

Contents

About the authors

Carmen Hoya was a Researcher at Learning and Work Institute. Carmen managed the process evaluation.

Alex McCallum was formerly a Labour Market Researcher and Economist at Learning and Work Institute. Alex conducted analysis of the learner survey.

Ash Patel is Head of Research (Inclusion) at the Learning and Work Institute. Ash led the quantitative strand of the evaluation, overseeing the participant survey and impact assessment.

Tony Wilson is Director of Policy and Research at Learning and Work Institute. Tony directed the evaluation and had oversight of all both strands and all findings.

Glossary

English for Speakers of other languages (ESOL)	An abbreviation for English for Speakers of Other Languages: used, especially in the UK, to refer to the teaching of English to students whose first language is not English.
Pre-entry level English proficiency	Pre-entry level denotes a very low level of English proficiency. An individual with pre-entry English may be able to answer questions on basic personal information and follow basic instructions but would have very limited (if any) letter and word recognition.
Entry level 1 English proficiency	Entry level 1 denotes a very basic level of English proficiency. An individual at this level may be able to ask and respond to personal information questions (in more than one word answers); give short accounts of activities; and make simple statements of fact. This level equates to standards of literacy and language expected of native speakers aged 5 to 7.
Entry level 2 English proficiency	Entry level 2 denotes a basic level of English proficiency. An individual at this level may be able to answer questions about their daily routine; give short accounts of previous experiences; and ask similar questions with the correct verbs and tense. This level equates to standards of literacy and language expected of native speakers aged 7 to 9.
English proficiency assessment	A proficiency assessment is a test to measure a learner's level of language and ability to use English.
Randomised Controlled Trial	Randomised Controlled Trials (RCTs) are a research method used to establish impact. They involve a control group who does not receive an intervention and effects are compared with those who do. Participants are randomly allocated to each group.
Randomisation	The process by which participants in an RCT are allocated to the treatment or control group (to receive or not receive the intervention). Randomisation is a critical element of an RCT as it ensures there are no systematic selection biases between participants allocated to either group. This helps to remove bias or interference caused by other factors. The result of randomisation will be that the two groups share, on average, very similar characteristics.
Contamination	Contamination occurs where individuals randomised to the treatment group and control groups are exposed to the wrong condition (i.e. people in the control group receive part of the intervention). This can occur inadvertently or intentionally as people discuss their

	experiences or share materials.
Trial participant	An individual eligible to enter the RCT and has provided informed consent to do so.
Treatment group	The group that receives the intervention within an RCT design following randomisation.
Control group	The group that does not receive the intervention within an RCT design following randomisation. They are monitored alongside the group receiving the intervention, and their results are compared to their treatment counterparts to understand what impact the intervention has had, compared to receiving no intervention. Any changes or effects detected within the control group over the course of the RCT can be interpreted as what would have happened anyway.

1 Executive Summary

1.1 About the process evaluation

This evaluation was commissioned by the Ministry of Housing, Communities and Local Government (MHCLG) to review the process of implementing a Community-Based English Language (CBEL) intervention in a randomised controlled trial (RCT) setting.

The process evaluation sought to:

- explore the implementation of the CBEL intervention, identify any aspects of variation in delivery and overall assess whether fidelity to the RCT design was maintained;
- identify key lessons, issues and challenges in implementing a social intervention as a randomised controlled trial;
- generate evidence, learning and recommendations for the design and implementation of future RCTs and English language interventions aimed at similar groups.

The report takes a chronological approach, considering in sequence the steps taken to deliver the CBEL RCT.

1.2 About the intervention and trial

The trial was delivered to 527 individuals across five areas in Greater Manchester and West Yorkshire[1] who had low levels of English language skills.

This report makes use of 54 qualitative interviews with individuals involved in the trial, including 32 interviews with participants from both treatment and control group and 22 interviews with those responsible for delivery. It also draws on analysis of attendance records and trial information (such as the intervention manual, briefing material provided to learners and staff, class records and attendance records), in order to add further insight into the process and implementation of the trial. A separate report sets out the findings of the trial itself.[2]

[1] CBEL was delivered in 23 centres spread across Manchester, Rochdale, Oldham, Kirklees, and Bradford.
[2] https://www.gov.uk/government/publications/community-based-english-language-programme-a-randomised-controlled-trial--2

The intervention being tested was an eleven-week Community-Based English Language course, comprising three sessions per week. Two sessions were class-based, using traditional and interactive methods to learn speaking, listening and writing skills. A third ‘club’ session was also included so that class-based learning could be practiced in a local setting. Club sessions involved trips out to local amenities such as council offices, bus stations, libraries and recreation facilities so that learners could engage with their local area through English and have opportunities to practise what they had learned. There were 33 sessions in total, offering 66 hours of guided learning and support.

The course was designed to address the practicalities of everyday English language usage and was targeted specifically at individuals with low levels of English proficiency.

Manchester Talk English were commissioned by MHCLG to develop and deliver a Community-Based English Language intervention, based on their existing programme, which could be delivered under RCT conditions. To fulfil this, a consortium of partners was established with representation across the five participating local authorities.

Manchester Talk English were responsible for the overall delivery and programme development, which included the design of an intervention manual and course material. Area co-ordinators and administrators were responsible for the engagement and recruitment of trial participants, sourcing community organisations and venues, managing the registration events and ensuring the course was delivered in line with the Talk English Together manual. Participants were identified and recruited locally by the organisations delivering the intervention, who also conducted initial screening to assess participants’ eligibility.

Qualified ESOL teachers delivered the classes, with support from two volunteers in each session. This differed from previous similar provision delivered by the lead delivery organisation, which relied on volunteer teachers.

1.3 Recruitment

The process of delivering the RCT involved strict approaches to recruitment, assessment, course design and delivery. As such, coordinators, teachers, volunteers and learners had a higher level of demand placed upon them than were typical in delivery of such courses.

Local area coordinators were tasked to recruit 24 learners for each class. Half of these would begin receiving the intervention immediately in April 2016 as the 'treatment group', with the remaining learners constituting a 'control group' against whom progress would be compared (and who would go on to receive the intervention from September 2016). The target for recruitment was 600, and the final number of learners recruited was 527.

Inclusion criteria were specified by MHCLG. Trial participants were expected to:

- be equivalent to pre-entry level or entry level 1 on one or more of the assessment areas (speaking and listening, reading or writing) as defined by the ESOL classification criteria);[3]
- be resident in one of the 5 local authority areas the trial was operating in;
- be aged 19 and above;
- have been resident in the UK for more than a year;
- have not received formal support from Talk English in the past (e.g. a Talk English course);
- not be in work or claiming Jobseeker's Allowance or the equivalent in Universal Credit (as this could entitle them to other funded ESOL support);
- consent to participate in the RCT research.

Coordinators said that eligibility criteria set for the trial made recruitment challenging particularly given the relatively short timeframe available to conduct the recruitment. However, a sufficient number of learners were recruited to satisfy the requirements of the RCT.

Learners were randomly allocated to either the treatment or control group. The registration process enabled the identification of learners who were known to each other personally, and these individuals were clustered together for randomisation. This was done to minimise the risk of contamination of the two groups during the trial, as there was a concern that individuals in the treatment group could share their learning and/or course materials with acquaintances in the control group. This would effectively compromise the counterfactual condition (e.g. not receiving the intervention) and likely reduce the effect that would be observed in the RCT. Participant clusters were more common and larger than anticipated.

[3] For more information about the core ESOL curriculum visit: http://www.excellencegateway.org.uk/content/etf2385

The randomisation process resulted in the creation of two groups – a control, and a treatment group - with similar characteristics. These groups' characteristics were assessed on a number of variables, such as age, gender, initial English capabilities and education. On nearly all measures, there were no statistically significant differences between the treatment and control groups at the point of randomisation or at the baseline, indicating that the groups were well balanced. One notable exception was in relation to baseline reading scores, which were higher among participants in the treatment group when compared to the control group. The cause of this difference was not apparent despite investigation. This difference in baseline reading score was subsequently accounted for in the RCT analysis.

1.4 Assessing learners' English Proficiency and integration

The RCT assessed participants' English proficiency in speaking and listening, reading and writing both at the beginning of the course and at its end using a series of bespoke English proficiency assessments developed by the English Speaking Board (ESB). This allowed for comparison of both changes in proficiency over time and overall differences in proficiency between the treatment and control groups.

Qualified assessors from the ESB administered the proficiency assessments with support from CBEL tutors on reading and writing measures. The proficiency tests were themselves developed specifically for this RCT by ESB.

Social integration outcomes (including social interactions and mixing; participation in everyday activities; confidence in engaging public services; local and national belonging; trust in others and attitudes to community integration) were assessed through a paper-based survey, which was also conducted at both the beginning of the course and at its end. Only participants who spoke Arabic, Bengali, Punjabi, Somali or Urdu were required to complete the survey. The survey was administered by researchers from BMG Research fluent in one or more of these languages with the support of additional translators.

The survey was developed by MHCLG with input from the Behavioural Insights Team. The survey was further refined, cognitively tested and piloted by BMG Research. Some of the questions were adapted from existing surveys (including the Citizenship Survey and European Social Survey) while others were created specifically for this research.

Assessment events were held at the beginning and end of the trial. Baseline (or 'pre') measures were collected as close to the commencement of the trial as possible (in most instances during the first week of treatment group classes). Follow-up (or 'post') measures were collected as close to the conclusion of the learning as possible (in most instances during the final week of treatment group classes). Assessments were a challenging aspect of the trial, as they required a number of

different agencies to work together to collect the multiple measures being used to assess the intervention. The need to collect the array of assessments within these sessions added to the challenge.

It was not always possible during assessment events to keep the treatment and control groups separate. However, the process evaluation found no evidence to suggest participants from the treatment and control groups were sharing resources or information with each other during assessments.

Despite the organisational challenge, the vast majority of assessments were completed as required, and the necessary data was collected. One centre was not able to deliver the reading and writing assessments as intended at the baseline point. To ensure trial consistency, reading and writing proficiency data from this centre was excluded from the final analysis. However, the scores for the speaking and listening assessment from this centre were not excluded as these were implemented in line with other centres.

As a precaution, the assessment procedures in all other centres were checked, however no further evidence was found to suggest any other systematic inconsistencies in delivery of assessments.

1.5 Trial fidelity

A real risk for the research was that coordinators, teachers or learners would compromise the RCT by deviating from the intervention or the trial design. However, these concerns did not materialise. Area coordinators briefed teachers in advance of the formal delivery period and teachers reported understanding and could describe the key elements of the trial design.

The evaluation did not identify any evidence that participant allocation to the trial groups had been compromised. The teachers interviewed understood that it was critical that the randomised participants were not allowed to switch classes, even if it was not always fully understood why this was the case. Likewise, teachers were aware that despite poor attendance, participants could not be removed from classes – even where teachers reported wanting to do so. There were very few instances where control group participants attempted to join classes and when this did occur teachers took the correct action and refused.

Evidence around fidelity to the intervention manual in terms of content and approach was more mixed. Overall, the manual and course materials were well received by teachers, who broadly found them useful aids. There was some frustration reported about the limited flexibility in how they delivered their classes and the course content. The manual did not define what level of deviation from set materials was acceptable, which left room for interpretation and (potentially) more adaptation of the content than desirable. Interviews with teachers found some minor adjustments to materials did occur; for example, they spoke about changing vocabulary used or

avoiding concepts that they felt participants could not comprehend. However, there was no evidence of more substantial adaptions such as changing the topics or study modules covered, or discarding teaching approaches all together (apart from one teacher who began using the manual materials several weeks after the start of the course). Overall, while accounts of teachers and learners suggest some variation in delivery of the course, this variation did not compromise the consistency of the intervention as the adaptations were relatively minor and teachers were motivated to adjust materials with the aim of accommodating learners, while still achieving the aims of the specific class.

Some teachers felt that they did not have sufficient training or consultation on the manual. With greater time for reflection it is likely that teachers would have been able to feedback their concerns about the level of prescription in the manual, and an agreed approach (to common issues) could have been developed. Importantly for the trial, teachers were practised in making ad hoc adjustments to materials, seeing this as part of their skill set as a teacher. Adjustments were motivated by making the materials more accessible for participants while still attempting to achieve the aims of the lesson.

The trial benefited from the experience and skill of the delivery staff within the consortium delivering the intervention. The timetable of activities was a challenge for all concerned, but the delivery staff ensured that the intervention was delivered successfully and on time. Without the ability to draw on their existing local networks, facilities and experienced staff the results of the intervention may not have been as positive.

1.6 Practical challenges of delivery

Overall, delivery of the intervention was found to be successful with few issues reported by teachers or participants.

Treatment learners were almost universally positive about the content of the classes. Several reported having achieved personal targets (such as going shopping alone, increasing their confidence, or improving their written English). Participants in the treatment group also reported finding the club sessions particularly useful and a number reported making use of local facilities because of these sessions.

Teachers reported some issues with mixed-gender classes, where female participants were unhappy attending classes if men were present. However, this problem reportedly declined over time as participants became familiar with each other. There was no explicit indication from interviews that this had significantly affected attendance or motivations to attend. Learners commented on enjoying learning in groups and most found working with others, regardless of their background, beneficial.

Importantly for the trial, none of the treatment group participants interviewed identified any aspects of the curriculum or delivery method that did not work for them.

The intensity of the course was an issue that course providers were keen to discuss. Most teachers agreed that three sessions a week was conducive to learning a language but it was also a challenge for some people to attend with such frequency. Many participants were women with children, so issues related to childcare and family illness occurred frequently and were cited as the main reasons for low attendance among some classes.

Another challenge cited frequently by teachers and coordinators was the impact of Ramadan and Eid on attendance (given the high proportion of Muslim participants). Teachers and participants found ways to mitigate the effect of these important religious periods, such as providing homework, or running fewer, but longer sessions in the affected weeks. Nevertheless, attendance was clearly impacted. While the mitigations employed may have moderated the effect of low attendance that occurred during this period, there was no evidence that teachers' responses departed from the intervention to, in any way, jeopardise the trial. Indeed, teachers were able to draw on the manual to share course content that learners would have otherwise missed (due to their absence when this was covered in class). No new material (i.e. not contained in the manual itself) was introduced to learners.

All trial participants were very motivated to learn English. It was clear from interviews conducted with individuals in both the treatment and control groups that most were very grateful for the opportunity to receive formal support with their learning. Participants in the treatment group did not indicate any instances of engagement with others in the control group, or 'cross contamination' which could undermine the overall RCT. However, many control group participants reported preparing for lessons due to start in September 2016 (after the completion of the treatment group's course) by attempting to use English more in everyday situations, watching English television, and talking to their children in English. They remained motivated despite having to wait for their classes to begin, which may explain why they too achieved improvements in their English skills between the baseline and follow-up measurement.

1.7 Conclusions

This evaluation sought to:

- explore the implementation of the intervention, identify any aspects of variation in delivery and assess whether fidelity to the RCT design was maintained;
- identify key lessons, issues and challenges in implementing a social intervention as a randomised controlled trial;

- generate evidence, learning and recommendations for the design and implementation of future RCTs and English language interventions aimed at similar groups.

Overall, the research conducted as part of this evaluation found:

- The process of randomisation, critical to the success of the RCT, was completed successfully. All parties involved in this process, guided by the briefings that formed part of the trial itself, worked hard to ensure random allocation was not compromised. Learners were recruited, clustered where necessary, randomised and informed of their allocation in sufficient time to begin the intervention.

- The trial was successful in ensuring that contamination (i.e. participants in the control group obtaining course material or receiving elements of the intervention) did not occur. Clustering participants reduced the likelihood of contamination, however the most critical part of the trial that reduced the risk of contamination, was training all delivery staff on the importance of randomisation and assignment in advance of delivering the intervention.

- There was some minor variation in the delivery of the course. Teachers adjusted content and materials, in line with their professional capacity, with the aim of accommodating particular learner needs, while still achieving the aims of the specific classes. As such, it is not believed that these variations present a deviation beyond what would reasonably be expected and therefore do not jeopardise the conclusions of the RCT.

- Overall, the intervention manual was a well-regarded resource. While the ambition was that teachers would be able to deliver the manual in exactly the same way in each location, the reality was that they saw the need to make ad hoc adjustments to the materials and prescribed approaches to accommodate learner needs.

- The manual was not finalised with sufficient time in advance of the trial beginning for all parties to review it and to reflect on the proposed content and structure. On reflection, teachers would have preferred the manual to be less prescriptive and allow them more discretion in deciding what content and approaches were most appropriate for their learners. This could have been achieved (while also maintaining a consistent approach across trial sites) by focusing guidance on the core learning aims of each session. The manual gave no guidance on what constituted a deviation from the intervention and what would be considered a reasonable adjustment to ensure course materials were appropriate for learners. Likewise, there was no guidance on how to treat learners that missed classes (i.e. whether to offer homework). In

hindsight, both issues could have been anticipated and covered in the manual.

1.8 Recommendations for future CBEL trials

- Extend timescales for design and delivery planning. A longer development phase would allow delivery partners to review materials and ask questions that could be beneficial to successful delivery. In particular, this would allow more time for teachers not only to review and comment on the manual, but also to familiarise themselves with the content to ensure its accurate implementation.

- Allowing more preparation time would also ensure sufficient piloting of resources and activities that are likely to prove challenging to delivery. For example, the assessment days showed that with practice (i.e. follow-up compared with baseline assessments) a complex task could become easier. Similarly, more time would also enable more refined cognitive testing and back translation of the survey materials.

- Provide clarity about teachers' roles and responsibilities in lesson planning and delivery. For example, setting out how teachers should deal with regularly occurring issues such as helping participants catch up after an absence, or what elements could be adapted for local circumstances. Furthermore, there should be clear guidance on what the core elements of the intervention and each class are so that true deviations can be identified.

- Consider testing interventions of varying lengths or capturing measures at different points, to test the extent to which intervention length impacts on attainment. Also, taking follow-up measures over a longer period will ensure that longer-term impacts can be observed (i.e. allowing more time for impacts on social integration to occur).

- Review the intensity of provision – for example by conducting 'clubs' every second week, which again may be more compatible with a longer intervention.

- Consider how local areas and partners are supported in the delivery of the trial. It was particularly striking in this case that both Talk English and its partners were well established and able to deliver this, and that MHCLG could provide relatively intensive support. Future trials may be less well placed in this regard and may need further support in terms of administration and resource.

2 Introduction

Learning and Work Institute, in partnership with BMG Research, were commissioned to conduct an impact evaluation of a Community-Based English Language (CBEL) intervention aimed at people with very low levels of functional English proficiency. The intervention was commissioned by the Ministry of Housing, Communities and Local Government (MHCLG) and conducted as a randomised controlled trial (RCT).

This is a process evaluation of the implementation of the trial. The Magenta Book defines a process evaluation as an evaluation which "primarily aims to understand the process of how a policy has been implemented and delivered, and identify factors that have helped or hindered its effectiveness" (HMT, 2011:84)[4]. It is therefore distinct from the impact evaluation of which the trial itself is the principal part.

This process evaluation was designed to:

- explore the implementation of the intervention, identify any aspects of variation in delivery and assess whether fidelity to the RCT design was maintained;
- identify key lessons, issues and challenges in implementing a social intervention as a randomised controlled trial;
- generate evidence, learning and recommendations for the design and implementation of future RCTs and English language interventions aimed at similar groups.

The report makes use of 54 qualitative interviews with individuals involved in the trial – both treatment and control group participants and those responsible for delivery. This has been supplemented by analysis of management information in order to add further insight into the process and implementation of the trial. The report takes a chronological approach, considering in sequence the steps taken to deliver the roll out of the CBEL RCT.

A separate, accompanying report sets out the findings of the RCT itself.[5]

---

[4] Her Majesty's Treasury. (2011). The Magenta Book: Guidance for evaluation. ISBN 978-1-84532-879-5

[5] https://www.gov.uk/government/publications/community-based-english-language-programme-a-randomised-controlled-trial--2

2.1 Background

Overall, within government it is the Department for Education (DfE) that is responsible for funding English for Speakers of Other Languages (ESOL) provision through the Adult Education Budget. Fully funded provision is prioritised for unemployed individuals on benefits, whose poor command of English is a barrier to getting a job. In 2014/15, DfE invested an estimated £104 million on fully and part-funded ESOL courses, supporting 131,000 adult learners.

MHCLG is responsible for policy on integration. Supporting people to learn English has been a core part of MHCLG's approach, which is set out in *Creating the Conditions for Integration published in* 2012.[6]

Evidence shows that poor English proficiency is a barrier to both economic and social integration. The 2011 Census found that over 760,000 adults born outside of the UK and living in England and Wales cannot speak English well or at all.[7]

From 2013/14 to 2015/16, MHCLG funded an £8m Community-Based English Language programme, supporting six projects to deliver English courses to adults with the lowest levels of English. The projects operated in the English language priority areas: broadly in East and North London, East Birmingham, Manchester, towns along the M62 in Yorkshire and Lancashire, Slough, Luton and Bristol. The projects were selected for their innovative teaching and engagement models, which delivered training in community settings or on-line, often using volunteers. Together the projects reached over 39,000 adults - around 80 per cent of whom were women, with over half from Pakistani, Bangladeshi, and Somalian ethnic groups.

To strengthen the evidence base MHCLG commissioned the RCT to test a Community-Based English Language intervention. Projects were invited to submit proposals for delivering an intervention on an RCT basis, and Manchester Talk English's proposal was selected.

2.1.1 Complex interventions in real world settings

[6] For more information the approach published under the 2010 to 2015 Conservative and Liberal Democrat coalition government, see: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/7504/2092103.pdf

7 ONS. (2011). English Language Proficiency by Age by Sex by Country of Birth by Year of Arrival in the UK (Table BD0059). Data available from: http://www.ons.gov.uk/ons/rel/census/2011-census-analysis/social-and-economic-characteristics-by-length-of-residence-of-migrant-populations-in-england-and-wales/rft7---bd0059.xls

The challenge of identifying the impact of an intervention such as this should not be underplayed. The CBEL intervention was considered a complex intervention, as it comprised of a number of interdependent elements (set out below). With such complex interventions, there is a high degree of challenge in attributing impact to a specific element of the intervention and identifying the relationship between these elements.[8]

The elements that form the CBEL intervention include:

- delivery by qualified teachers;
- a high staff to learner ratio;
- use of an intensive, well defined syllabus, using a bespoke intervention manual;
- targeting of highly specified learners;
- delivery within numerous, and different types of community settings (e.g. using local resources and locations).

The RCT tested whether these elements in combination delivered changes to participant outcomes. It was not possible to assess the individual contribution of these elements or their interaction with each other.

It is reasonable to expect that even with a very high level of conformity, interventions delivered in real world settings (i.e. outside of controlled environments such as a school) will have a degree of unquantifiable variation across centres and even individuals. As a result, the intervention may not look identical across the locations in which it was delivered.[9]

As set out in this report, variations occurred between locations because of differences in attendance, the accessibility of local amenities, slight differences in approach by teachers, and different learner abilities. Together these factors led to a slightly different experience for the participants.

This report considers each element of the design of this trial in turn but this should not lead the reader to assume that each element makes an isolated contribution to the success or indeed failure of the intervention.

---

[8] Medical Research Council. (2000). A framework for development and evaluation of RCTs for complex interventions to improve health. Medical Research Council: London. Available from: https://www.mrc.ac.uk/documents/pdf/rcts-for-complex-interventions-to-improve-health/

[9] For a deeper discussion on trailing complex interventions please read Haw, P., Shiell, A., Riley, T. (2004). 'Complex interventions: how "out of control" can a randomised controlled trial be? *British Medical Journal* 328 (26). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC437159/

In considering the overall intervention, this evaluation assessed the extent to which variations in implementation were deemed problematic in terms of the fidelity of the intervention delivered, and the validity of the RCT.

2.1.2 Overview of the delivery model

The intervention tested was an eleven-week English language course, comprising three sessions per week. Two sessions were class-based, using traditional and interactive methods to learn and improve speaking, listening, reading and writing skills. A third club session was also included so that class-based learning was put into practice in a local setting. The course was designed to address the practicalities of everyday English language usage and was targeted at those with very low levels of English proficiency.

Qualified ESOL teachers delivered both the class and club sessions, with support from two volunteers. This differed from previous CBEL provision delivered by Talk English, which relied solely on volunteer teachers. The rationale for using qualified teachers to deliver the intervention was not explicitly set out by the provider, however given the intensity of the support we assume that using paid teaching staff was necessary as volunteers would be less likely to commit to as many classes. We also assume that qualified teachers would improve the quality of teaching.

Club sessions involved trips out to local amenities such as Council offices, bus stations, libraries and recreation facilities so that learners could engage with their local area using their newly acquired English skills and have opportunities to practice what they had learned. There were 33 sessions in total, offering 66 hours of guided learning and support.

Participants were identified and recruited by the organisations delivering the intervention. Potential learners were invited to registration events, which occurred between two and five weeks prior to the commencement of the trial. At these events, partners assessed learners' eligibility and collected registration data and informed consent.

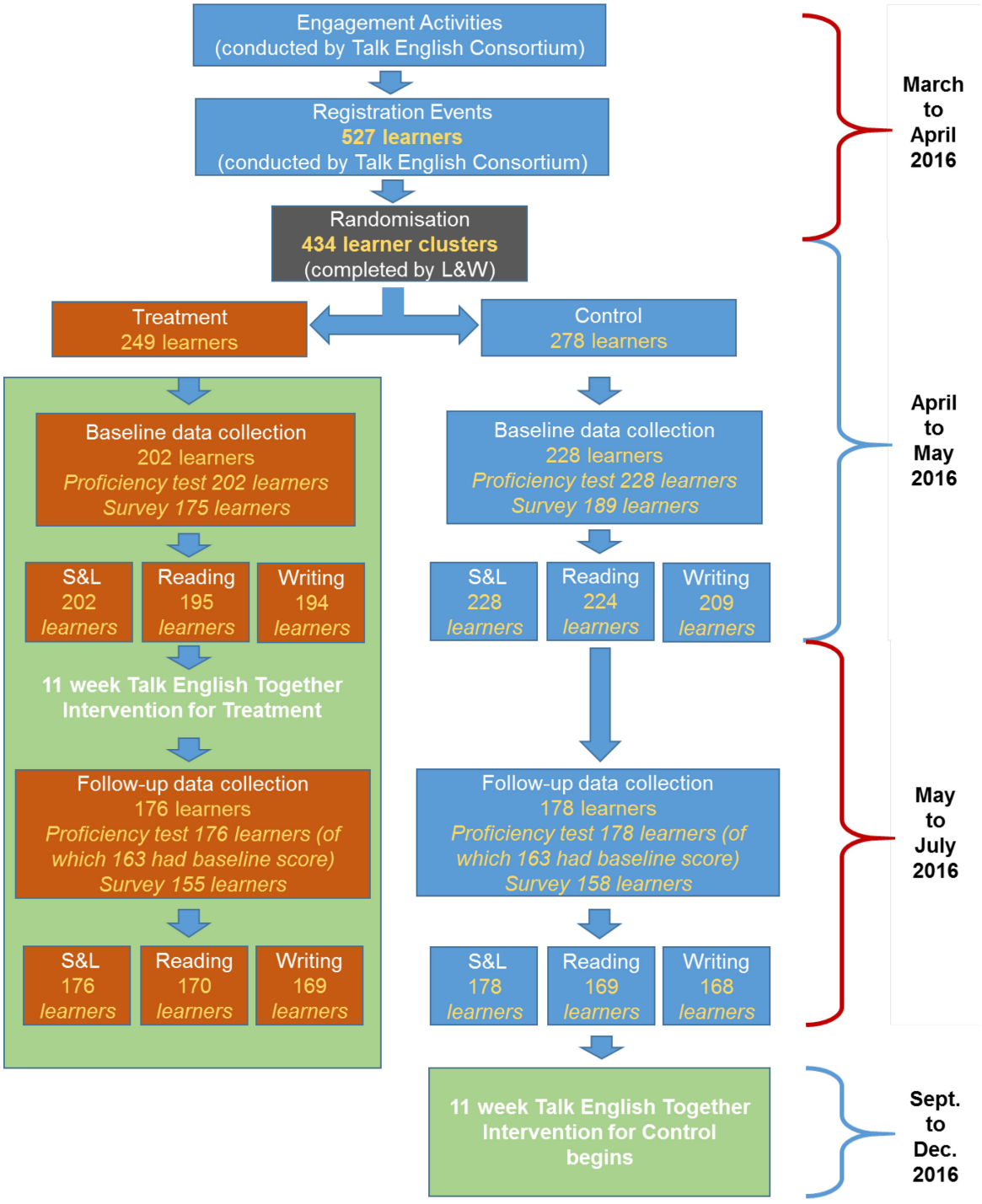
Delivery partners conducted an initial screening of participants to ensure they:

- were equivalent to pre-entry level or entry level 1 on one or more of the assessment areas (speaking and listening, reading or writing) as defined by the ESOL classification criteria);[10]
- were resident in one of the 5 local authority areas the trial was operating in;

[10] For more information about the core ESOL curriculum visit: http://www.excellencegateway.org.uk/content/etf2385

- were aged 19 and above;
- had been resident in the UK for more than a year;
- had not received formal support from Talk English in the past (e.g. a Talk English course);
- were not in work or claiming Jobseeker's Allowance or the equivalent in Universal Credit (as this could entitle them to other funded ESOL support).

Figure 1 Trial summary diagram[11]



[11] This figure excludes 15 learners' reading and writing assessment data at a single centre (further details provided in section 3.9 below). Baseline assessments for a further four learners were excluded from analysis, as their unique identifiers were incorrectly recorded at this stage. Finally, one learner was also excluded from the trial entirely due to registering at two separate centres and being assigned to different trial conditions at each. Individual data on dropout was not collected throughout the trial.

Once collected, the registration data for all participants in a centre were quality assured on site by MHCLG staff. This data were then transferred via a secure online file transfer site to L&W's London office, who then randomised participants into either a treatment or control learner group. A cluster-based approach was used for randomisation, so that clusters of participants who reported that they knew each other (for example family members or friends) would be allocated to a trial group together. This was done to reduce the risks of cross contamination between the two groups. Treatment learners received support between April 2016 and July 2016, while control learners began in September 2016.

The trial ran in 23 community venues across the five areas participating in the trial. A total of 527 individuals (in 434 clusters) participated in the trial; approximately half were assigned to classes beginning in April (treatment) and the remainder were assigned to classes in September (control).

During the intervention, control group participants did not receive anything and were required to wait until September to begin classes. This is known as a waiting list methodology. Its use can be problematic if there is a chance that those who are waiting (control group participants) are able, in some way, to receive the intervention earlier than intended. In this RCT, control group participants had some knowledge of where classes would be held and there was a concern that some within this group would attempt to take part in classes or make use of class materials obtained from treatment learners.

At the start of the treatment group course, all participants – both treatment and control – were invited to attend an assessment event where baseline measures of their English language ability (speaking, listening, reading and writing) were taken. A survey was also used (administered by BMG Research) to assess participants' attitudes, behaviours and their level of integration in their local community. Finally, all participants were invited to a follow-up event (during the final week of treatment group classes) so repeat measurements could be collected.

3 Aims and methodology

This section sets out the aims of the process evaluation, and the methodology followed in its delivery. The overall aims of the main CBEL intervention were to:

- Increase English language proficiency of participants;
- Increase integration among participants.

3.1 Process evaluation aims

This process evaluation sought to:

- explore the implementation of the CBEL provision including any variation in implementation and delivery across trial sites, and the reasons for this, overall assessing whether fidelity to the RCT design was maintained;
- identify key lessons, issues and challenges in implementing a social intervention as a randomised controlled trial – in particular, in implementing an experimental design robustly, and ensuring fidelity across sites and partners;
- generate evidence, learning and recommendations for the design and implementation both of future randomised controlled trials, and of future English language interventions aimed at similar groups.

3.2 Methodology

The primary data collection method for this process evaluation was in-depth qualitative interviews. A total of 54 interviews were conducted, comprising:

- 7 members of the consortium delivering the intervention (regional co-ordinators);
- 9 teachers and 2 volunteers;
- 22 treatment group participants;
- 10 control group participants;
- 4 independent ESOL assessors (employed by ESB), who carried out the English proficiency assessments as part of the trial.

Interviews were conducted using discussion guides with standardised questions designed specifically for each interviewee group to explore their experiences of the trial (see Annex B).

Participant interviews (whether in the treatment or control group) were around half an hour in length and were conducted face to face, and with an interpreter present. Some participants tried to use English in their interviews, however all but one made at least some use of the translation support provided with the majority relying wholly on the interpreter. Permission was sought to digitally record all interviews for post-interview transcription to ensure data was captured exactly, in line with best practice. However, in total ten interviews with learners were not recorded, at the request of the learner. This is a much higher proportion than would ordinarily be expected in this type of research and is likely to be specific to this particular population. In instances where permission to record the interview was not granted, interviewers took detailed written notes.

Most interviews were conducted in person with teachers, coordinators, ESB assessors and volunteers. While interviews with ESB tended to be relatively short, lasting around half an hour, other interviews were generally an hour in length, although some did exceed an hour. Due to time constraints, two teachers and two volunteers were interviewed over the phone; these interviews were conducted in August 2016 after the intervention had been completed. The rest of the teacher interviews were carried out face-to-face, from the mid-point of the intervention. All participants were asked to consent to the recording of the interviews. Where consent was given, the interviews were recorded and transcribed.

Additional data, in the form of English assessment results, attendance records and class notes collated by teachers, have been used to provide further context with which to interpret the qualitative evidence. These data sources enabled evaluators to create a timeline for the course, provided a session-by-session description of content, and offered information on levels of attendance and feedback on the conduct of the class. These sources were compared with the verbal accounts given by interviewees.

3.3 Analysis

The qualitative data collected was analysed using a framework methodology.[12] This is a systematic approach to managing and analysing qualitative data. Using this approach, data from each respondent is classified and organised. This allows for the

[12] Ritchie, J. and Spencer, L. (2002). Qualitative data analysis for applied policy research. In Huberman, A. M. and Miles, M. B. The qualitative researchers' companion. Thousand Oaks: Sage.

structured analysis of data by theme, and enables researchers to draw out the range of experiences and views across respondents.

Emerging themes were initially identified through regular researcher debrief sessions. These were structured around the objectives of the research, and themes were refined and further developed following the review of interview transcriptions and (where no transcripts were available) interview notes. The themes identified through both the debrief sessions and transcript reviews were used to develop a thematic framework – an Excel spreadsheet providing a structure whereby data from transcripts was inputted and interrogated by both the theme of interest (e.g. acceptability of the manual, peer support, etc.) and participant profiles (e.g. centre or gender). This framework evolved in line with the data being reviewed in the debrief sessions.

Thematic frameworks were developed for learners, consortium members, teachers, and volunteers, into which relevant interview data was entered. Given the small number of ESB assessor interviews and relatively consistent approach to the interviews, a less iterative approach was taken with ESB assessors, with analysis framework themes being closely aligned to the broad questions contained in the topic guide.

The analysis of the qualitative data seeks to compare and contrast the experiences of the delivery of Talk English Together and the implementation of the CBEL Randomised Controlled Trial across delivery locations, population groups and in relation to intended delivery.

4 Results

This section presents the results of this research, including an assessment of the implementation of the RCT. It considers the extent to which the intervention was delivered as intended; reviews learners' and teachers' experiences of the intervention; and explores the experience of control group learners who had to delay their learning so the trial could be conducted.

4.1 Programme aims

4.1.1 Understanding of programme aims

Evidence suggests that English language is a key factor in facilitating social integration, as it gives isolated individuals the ability to speak with others who are different from them in a common language, as well as increasing their ability to access services and play a fuller part in society. As such, supporting people to learn English has been a core part of MHCLG's approach to integration. From 2013/14 to 2015/16, MHCLG funded six projects to deliver English tuition to adults with the lowest levels of English.

The trial expanded on the existing Community-Based English Language interventions that had been delivered through MHCLG's integration approach.

Delivery partners reported a good understanding of the aims of the programme. As the design of the intervention was based on previous versions of the Talk English offering, the broad approach was already familiar to many of the partners involved.

Delivery partners also had a working understanding of the experimental research design. Coordinators reported that they had received training on the intervention and the importance of maintaining the randomised participant groups. Their understanding was most evident in the accounts of teachers, who, at times, had to deal with issues that had the potential to impact on the experimental design of the trial. In these instances, it was clear that teachers were referring to senior staff for guidance and that guidance was generally forthcoming and in line with the manual.

4.1.2 Understanding the objectives of the research design

Coordinators were realistic about their understanding of the RCT. For them it was clear that they had to follow the guidance, and they seemed willing to try to understand why, but their understanding was a working knowledge of the process rather than a deep understanding of the complexities of an RCT.

"Well, we want to try and understand the reasons... I'm not a statistician, I'm not a randomised control [sic] trial person, and although it is an academic exercise, clearly, translating the practical from the academic in situations such as this is not easy"
Regional coordinator

The RCT methodology only affected delivery partners when issues arose that were the result of the restrictions it imposed upon them. The most problematic issue for teachers and staff was having to turn potential learners away.

“They don’t feel that we’ve met their needs; but I think it goes a little bit against our missions and vision” ***Regional coordinator***

“The learners’ perception [was] ‘why can’t we do this and why can’t you accommodate us?’ I think that has been really, really difficult...So they can say it’s an unfair thing, but we understand that it’s necessary in order to do what DCLG [sic] is trying to do." **Regional coordinator**

Likewise, learners did not have any choice regarding when they started the intervention as they were randomly allocated to either the treatment group (starting classes in April 2016) or the control group (starting classes in September 2016). This inflexibility was reported as problematic for some people.

“So one of the ladies on the programme, for example, her son’s just been diagnosed as autistic and she really wants to go on the programme. She wanted to come on the programme in September and said could we put her forward for September, ‘No, we can’t’, I’m saying, there’s no flexibility. And even the people coming forward in September, it’s rigid, it’s fixed, and we just don’t know where those people are going to be in September.” **Regional coordinator**

However, this did not mean that the RCT was seen universally as a problem for practitioners.

“I think, to be honest, I’d forgotten about the research going on because we don’t know what exactly your work involves but I think ….it’s going well so we’ve got courses going on and learners attending, community members are happy, tutors are fine, you know, so it’s going well on our side” ***Regional coordinator***

Other challenges also arose. The most commonly cited was in relation to time. Respondents felt that the timescales for implementation had been particularly challenging, which contributed to challenges in recruitment; led in some cases to classes not being filled on time; made it harder to draft and adapt materials and prepare teachers; and contributed to the difficulties noted above in some registration activity. While these issues are not unique to RCTs, the number of steps and processes that were wholly necessary meant that there was a lack of flexibility for delivery partners.

"I just think that [the main difficulty] has probably been the timeframes, you know, where we had a little bit of pressure to get certain things done by specific dates but I think that’s also just the way that we work” **Regional coordinator**

To some extent, these time challenges were an inevitable consequence of the need to take baseline measures in a robust and consistent way. However, there were also

some specific concerns around the additional burden placed on area leads to co-ordinate and distribute materials and on teachers to conduct assessments.

"The admin side of it really, and the demands… I mean, for example, the [ESOL assessment] papers came here and then I went out and distributed them to centres. And just the timings really that we were working against, you know, they're very short deadlines you're expected to meet… It's the amount of manpower really which was needed to be able to make sure the baseline events were effective." **Regional coordinator**

Overall, despite the challenges reported, the interviews conducted found that there was an understanding among those responsible for delivery that the processes in place were necessary.

4.1.3 Design of the Talk English Together course and manual

In order to ensure consistency in the content and delivery of the intervention a detailed course manual was developed by Manchester Talk English. The manual prescribed content for each class. This included an outline of the objectives of the class, issues to cover and class materials for individual sessions, which were available to download from the internet. Nine topics were to be covered over the 11-week course, with modules delivered in a set order and with a prescribed number of classes devoted to each. No flexibility was permitted in the order or time spent covering a specific topic.

Area leads considered the course content to be of high quality. Some were already familiar with the course structure, as they had delivered previous incarnations of Talk English. The final manual was not finalised in time for teachers' briefing sessions so there were no strong objections to the manual in advance of delivery. In general, the assessment of the manual from teachers was that materials were thought to be well suited to the learning outcomes and teachers were happy with the topics that had been included.

The running order of topics was thought to be logical, with the exception of the alphabet warm-up exercise in week six, which felt like a regressive activity that would have been more appropriate for learners towards the beginning of the intervention. Similarly, there was a preference for the democracy module to have been delivered earlier, as this would have enabled learners to apply their understanding to the European Referendum vote that took place shortly before the timetabling of the topic. In both cases however, the classes were run as specified by the manual.

As teachers began to deliver the intervention, some weaknesses in the design process were exposed. For example, there were some examples of minor spelling errors, or grammatical errors. Issues such as these were easily resolved in class.

"We had certain spelling mistakes along the way or, like, grammar issues throughout and then there were, like, kind of the comments that [Teacher] had about certain, you know, like certain activities not being quite appropriate" ***Volunteer***

More fundamentally, teachers also reported having to adapt materials to suit the various needs of their students (as certain aspects were not deemed appropriate for their learners). While this did not pose a practical problem in terms of delivering the class, it led to some minor variation in the content delivered across areas. Area leads and teachers subsequently raised concerns at the high degree of prescription within the course materials and the lack of clarity regarding what constituted a deviation from the intervention.

"It was useful [the manual] but I think it's when you start actually using it that you actually realise what you need to do, if that makes sense. ... But I think that's all to do with the very short time that we have to get the courses… you know, recruit the people and then to train the person who is actually going to deliver it." **Regional coordinator**

"We weren't challenging what was in the manual; it was just a clarification I think, because it was a research trial, on how much we were allowed to make alterations or changes in order to deal with that issue, of meeting the learning needs and meeting the research needs." ***Regional coordinator***

The intervention manual was designed for the purpose of the RCT. The highly prescribed approach to the delivery of classes (set out in the manual) was however at odds with previous provision, where teachers were used to having much more flexibility and discretion in how they taught their classes. This difference was motivated by the need to maintain consistency across locations. Both teachers and coordinators recognised that the course structure was part of the experimental design – to ensure consistency in lesson content and delivery across areas.

In our discussions with teachers, there was consensus on this matter. All those interviewed felt that adjusting materials to meet the needs of their class was part of their skill set as a teacher. For example, one respondent reported changing words in lessons where participants struggled with the underlying concept the word related to, or where the vocabulary was deemed too hard for lower skilled learners in the class.

"If you're teaching a role play, in one colour we would have what the receptionist says, then another colour would be what the patient says…, so they would read, "Receptionist, can I help you?" and the learner might say, "Patient, I need an appointment," rather than understanding that the receptionist and patient are two different characters in a role play, for example. So there were just similar things, things like that, that could have been… that I changed, and used the same language but in a different format." ***Teacher***

Importantly, teachers were not able to identify to what extent the manual could be adjusted before it became problematic for the RCT. The challenge for teachers was

that they had professional experience of responding to learner needs but not within an experimental framework where consistency of delivery had been highlighted as a priority. The adjustments reported by teachers were always undertaken to make sessions more accessible to some or all of their students. The manual did not stipulate what constituted an unacceptable deviation nor how these ad hoc adjustments should be approached.

This issue shows that while the manual was of a good standard, for the purposes of the RCT, further development work was required to give teachers clarity on what should and should not be done to accommodate differing learner needs.

4.2 Experimental design and implementation

The fact that the trial was designed as a randomised controlled trial was a central part of the experience of coordinators, teachers and participants. The requirements of adhering to this methodology made many stages of the process more complicated or time consuming to deliver.

The following section details each stage in turn and sets out some of the challenges that individuals in the trial experienced and their implications for the RCT.

4.2.1 Eligibility criteria and screening

The eligibility criteria were set so that the intervention was targeted at individuals most in need and who may not have accessed English language tuition in the past. The eligibility criteria detailed in the manual was as follows.

The beneficiaries of this intervention:

- had little or no English language skills (Pre-entry - Entry 1 ESOL);
- were not eligible for mainstream ESOL support, as delivered via the Skills Funding Agency (not in employment or actively seeking employment, i.e. JSA claimants);
- were aged 19 and above before beginning classes;
- were resident in the UK for more than 12 months;
- had not received formal support from Talk English in the past (e.g. were not registered on a TE course).

Area leads reported frustrations with what they saw as tight eligibility criteria. The requirements that frustrated regional coordinators and centre leads most were that learners had to have been in the UK for at least 12 months, and that learners from previous Talk English programmes could not take part.

"The criteria was an issue for us as well, people being here for a year because if you go into a community where you know there's a need, lots of the people who are easier to recruit have already been recruited and the new people are going to be new arrivals. And also the evidence shows that if you can get somebody into an ESOL class within the first year of them getting here, they're far more likely to carry on with that and to learn the language. The longer that they go in a community without learning English, then the more they get used to a life that doesn't require English and the more isolated they become, from an English speaking community."
Regional coordinator

From MHCLG's perspective, the tight screening process was seen as necessary to ensure provision was targeted at those who were most in need and least likely to otherwise receive it. In practice, the criteria were seen as another challenge to overcome when recruiting learners to the trial. As such, the criteria did make recruiting sufficient numbers more difficult.

In our interviews with participants, there were indications that some learners had participated in the trial when they may not have been eligible. For example, two participants indicated that they had been in the UK for less than the stipulated 12 months, and another claimed to be in employment. Eligibility was assessed by local partners at registration events (screening tools were provided in the intervention manual); however, learners were not asked for proof of eligibility (i.e. they self-declared their age, residency and employment status). It was not clear in these instances whether learners had intentionally misled recruiters, if their circumstances had changed over time, or if partners assessing eligibility had not applied the criteria correctly. While this suggests that not all participants of the intervention were the intended recipients, eligibility was not a factor that could confound the RCT analysis.

Local partners had, on average, two to three weeks in which to recruit potential learners. The interviews undertaken suggest that it is likely that a longer period for recruitment would improve adherence to the eligibility criteria in the future. It may also be the case that in reality, strict adherence to the criteria would be a challenge in any circumstance unless there is resource to independently check eligibility rather than allow self-reporting.

More importantly for this and future RCTs, the eligibility criteria applied here resulted in competing priorities during recruitment – recruiting sufficient numbers for the RCT analysis to be robust and ensuring the intervention was targeted.

4.2.2 Learner engagement and recruitment

To attract sufficient learners the trial was promoted using posters, leaflets and word of mouth within the communities and geographic areas it was targeting. Centres used their existing networks to engage and recruit learners, and often relied on local partners and community organisations with whom they had existing relationships.

The specific recruitment drives and events differed by location, but the approach was broadly consistent.

"We work closely with our partner organisations, because they've got real clear networks within the local communities. So we did some outreach and we did some looking at people who may be known to the centre but who have not been on a formal programme before. So it's a lot of dialogue with the providers about where we could recruit, and we chose two centres right in the heart of BME communities where we knew there was significant demand, as well as at the centre ourselves."
Regional coordinator

Centres also made use of contact lists of learners who had enrolled on other courses at their venues to promote the trial. This created a problem when in some instances individuals who had accessed other English courses were alerted to the trial, as previous English learners were not eligible to participate in the RCT. Despite this, word of mouth played an important part in engaging learners, and previous learners may have been more likely to recommend the class to friends.

The majority of learners were relaxed and apparently not 'fazed' by the research element of the intervention. However, a few individuals involved in recruitment noted that, in some instances, learners were deterred from registering by the formal nature of the consent form used.

"When I was doing a registration at one venue, so the husband read it in his language and then he said I have to take it away, I'll come back to the next registration session, he was with his wife, because I'm not sure what all these acronyms mean, you know, BMG, ELC or whatever it was, DCLG [sic], you know, I had to google all that and then, I'll explain it all to her and then if she's happy to do that she'll join." ***Regional coordinator***

Care was taken during the design phase to make the consent materials as accessible as possible. All learners were given or read an information sheet detailing what the research entailed. They were then asked to sign a separate form confirming that they had understood what the research was for and that they were happy to participate. It is reasonable to expect that a small proportion of potential learners would always be put off participating in a trial despite every effort to reassure them as to the purpose of any research. Participation in research is voluntary so it is overambitious to hope that everyone will agree to take part.

Importantly for the RCT there was no indication that consent was an issue (beyond what would usually be expected) and despite some 'refusals' recruitment did achieve a sufficient number of learners.

4.2.3 Randomisation procedure and contamination

The randomisation of participants was a fundamental part of the experimental design of the RCT and as such, it was critical that all participants had an equal chance of being selected for the treatment or control group.

The randomisation procedure itself was straightforward. The bigger risk related to adherence to the outcome of the randomisation process (by both individuals and organisations involved in the trial).

In this trial, randomisation used a cluster approach, grouping close friends, relatives or neighbours together, to reduce any potential bias from a member of the treatment group sharing learning with a member of the control group.

Randomisation was also stratified by location to ensure class size was equal across treatment and control, and because it was known that learners would not be willing to travel to locations that were further away. Randomisation only occurred once a centre had registered a sufficient number of participants to run two courses (one immediately for those assigned to the treatment group and one in September for those in the control group). Once achieved, the registration data for all participants in a centre were quality assured on site by MHCLG staff before being transferred via a secure online file transfer site to Learning and Work's London office for randomisation.

Once randomised, lists indicating the allocation of participants to the treatment or control condition were returned to the relevant provider centre. Treatment group participants were identified by an 'April/May' start date while control group participants were identified by a 'September' start date. Providers then contacted participants (by phone and letter) to inform them of their start date and invite them to an assessment session.

4.2.4 The implementation of randomisation

The process of clustering known participants made randomisation more difficult. Maintaining a balance between treatment and control is of course harder if the 'unit' to be randomised varies. Before randomisation it was assumed that there would be only a small number of small clusters; however, in reality some clusters were of a very large size (the biggest being seven learners). This was because some of the areas in which the centres were based had very close-knit communities, making it less likely that learners would not know each other.

Randomisation using a clustering approach appeared to remove the risk of contamination. Throughout the process evaluation, we found no evidence of contamination between treatment learners and control learners. Only one control group participant said they knew someone in the treatment group but advised they

had not seen them for some time. Another participant described knowing treatment group participants "*just a few people but we just say hi to each other".*

Some control group participants did express some frustration at having to wait to begin classes but they did not identify this as resulting from the RCT process.

Most locations appeared to have avoided control group participants attempting to attend classes. Several teachers reported complaints from control learners due to having to wait. Several teachers and coordinators also mentioned that participants were upset when they learned they had to wait for lessons – having been randomised to the control group.

Unsurprisingly, a small number of teachers did report examples where participants in the control group had attempted to attend classes. There were three examples of this reported, however teachers were clear with participants that their selection into groups was fixed and could not be changed.

While it is impossible to be certain that individuals in the control group accessed no course materials (as some learners in the intervention group took materials away as homework etc.), there were no reports from either learners or teachers of requests to share class materials with others.

During the randomisation process, a single data error occurred when a learner registered at two locations (this was not immediately spotted as the participant's name had been misspelt at one site). Importantly, in terms of the management of the trial, in this and all other instances of problems arising, delivery partners sought advice from the trial administrators about how to proceed. As such, this was dealt with in advance of the learner attending classes at either venue.

Clustering learners in the randomisation process did however generate a different class composition. Some teachers noted that due to clustering there was less diversity of learners within each class. Having classes with a high concentration of existing friendship groups affected negatively on the potential for integration within classes as it exposed learners to fewer new people and fewer people from different backgrounds. Furthermore, it made it more likely that learners E communicate in their own language during lessons. Despite this, interviews did not indicate any further impact of having large clusters on teaching.

4.2.5 Learner motivations

It was apparent from the interviews with both groups of learners that the motivations for learning English were consistent. Almost all learners stated that being able to speak English was very important so they could communicate with other UK residents, and most were enthusiastic about learning and improving their capabilities.

“[It is] just something you should be able to do…you should always be able to speak the language where you live” ***Learner***

“To live in the country you need to learn, and be able to speak in English, you need to go out, make appointments in the doctors. Everywhere you go you need to speak in English.” ***Learner***

Being independent was a frequently mentioned motivation. Other important learner motivations included using English to seek employment, to communicate with local services or travel independently (including by learning to drive).

“I want to be able to call the doctor on my own, not have to wait for a family member to come with me on the tram”. ***Learner***

"Now the children are at school I don't have anyone to help me…if people come to the house, workmen, I can't tell them what I want” ***Learner***

“I can now speak with people if they come to the house, if builders come I’ll be able to tell them what I want.” ***Learner***

“*I want to learn English so that I can eventually get a job”* ***Learner***

Several participants saw the classes as a first step to more learning opportunities – either more English at a higher level, or to help with access to further skills provision such as computer courses.

A small number of learners stated that they were simply there to have company. One elderly learner who was retired was there to meet people and improve their written English. Other learners felt that they already had some grasp of English – generally gained from education back home – but needed help in building their own confidence using it. In these examples, the problem was not English language but UK accents.

“I can speak English but I used to live in America, and the accent is very different here.” ***Learner***

This particular learner was not confident to attempt to speak English at all due to their problems understanding others.

Overall, learners we spoke with did show a great deal of motivation to learn English; however, this had not translated into results prior to the intervention. Participants stated that barriers such as childcare commitments were often the reason for failing to find and sign up to a class.

4.2.6 Challenges of the waiting list approach for control group learners

The RCT employed a waiting list design, meaning control group participants were required to wait around three months to begin English classes. This type of approach can generate a number of problems. Theoretically, control group participants may be able to receive an intervention earlier than intended (i.e. by joining a course

elsewhere). Likewise, maintaining engagement for control groups can also be an issue.

Involvement in the trial also had the potential to change control group participants' behaviour. For example, it could be argued that by being on the waiting list control group participants had greater motivation or encouragement to prepare for the classes in September. This behaviour has the potential to impact on the results of the RCT by reducing the difference between the progress achieved by learners in the treatment group and those waiting to learn.

An alternative approach to the RCT would have been to use a control group that did not receive the intervention at all. This way, the control group would no longer feel the need to prepare for anything, which would make any progress in English proficiency more reflective of a natural progression (without being steered by the trial).

A waiting list design was adopted for a number of reasons. Firstly, delivery partners (who were using their existing networks to recruit learners) were not comfortable with the suggestion that only some of the participants registering would be offered some type of English provision. Secondly, the prospect of receiving the intervention (albeit not immediately) offered an incentive to individuals in the control group to remain engaged in the RCT. This operated on the expectation that the incentive (i.e. receiving English lessons in September) would boost the number of control learners who attended assessment sessions, which was critical to the success of the trial.

Understanding the waiting list approach as a potential source of bias is useful for future RCT design. However, it is important to note that this approach did not affect the validity of the results detected.

A number of the Talk English team reported that it was challenging to keep the control group engaged during the period before their courses started. To maintain engagement, staff made efforts to ensure the assessment events were as enjoyable for control group participants as possible. This included providing food and drink, and giving learners 'goody bags' with Talk English pens and mugs.

"I think under the circumstances we have been pretty successful, quite successful; but my concerns are just with [keeping them engaged in] the activity group... But we just hope that they will come back in September and join the course..." **Regional coordinator**

4.3 Development and application of English language assessments and survey

Assessment events were held at all sites, during which baseline and follow-up measures of English proficiency and social integration were collected.

The English language measures were administered by independent ESOL assessors, with support from Talk English teachers on the reading and writing assessment. BMG Researchers conducted the survey with the assistance of translators. Both data collection tools were designed specifically for the trial.

4.3.1 Development of the English assessment tool

A series of bespoke English proficiency assessments were developed by the English Speaking Board (ESB) to fit the exact needs of the trial. ESB have a staff of trained assessors and regularly produce tests for English capability assessments. The assessment scale and tests were developed based on ESB's expert knowledge of assessment methods and were consistent with the ESOL core curriculum.[13]

These assessments were particularly important to the trial, as the results were used to identify where learners had improved their English abilities.

Assessments covered three domains of English proficiency; speaking and listening, reading, and writing; each was assessed against a nine-point scale. Within each of these domains, three levels of assessments were developed: Pre-entry, Entry Level 1 and Entry Level 2 (nine assessments in total).

"So our materials development team put together the set and the marking criteria and then the people who were going to be involved in the testing we all met together and went through it and altered, arranged, put details in, took details out and came up with the final product." ***ESB assessor***

ESB conducted training with its assessors on the design of the new test so that they could be implemented in a consistent way. As part of the process of ensuring consistency, assessment results were also moderated after they were collected.

"..We had a training session when we went through the materials and how to go through them with the candidates just checking that we all had the same understanding and timing and sort of scheme for marking and all that sort of thing, yeah." ***ESB assessor***

The English proficiency assessment was developed and delivered without any significant problems. However, an important limitation was identified during the course of the trial. The assessments measured ability from Pre-Entry to Entry Level 2 only. It was not possible for participants to achieve a score of greater than nine (indicative of 'established skills' at Entry Level 2). Effectively, this meant that if a participant progressed past this point in one of the three domains, the extra progress

[13] For more information about the core ESOL curriculum visit: http://www.excellencegateway.org.uk/content/etf2385

was not identified or recorded. Given the desired demographic of learners outlined in the trial's eligibility criteria, it was not initially anticipated that any participant would progress above this level but after recruitment it became clear that some participants were achieving higher than anticipated proficiency at baseline (i.e. they were already at the initial stages of Level 2).

Participants with a very high score at baseline would have had limited scope to evidence any improvement in proficiency (because the assessment was artificially stopped at Level 2). While this issue highlighted the failure to screen out participants with higher than expected capability in English, it also meant that some participants could have improved over and above the level assessed by these tools.

The implication for the RCT is that there may have been some progress for more advanced learners that was not captured in the data. However, given the small number of participants affected by this issue, we conclude that this does not present an issue for the findings of the RCT, as any unobserved results would only improve on the overall impact identified.

4.3.2 Survey development

The social integration survey tool was developed by MHCLG with input from the Behavioural Insights Team. Some of the questions were adapted from existing surveys (including the Citizenship Survey and European Social Survey) while others were created specifically for the trial. The survey was further refined, cognitively tested and piloted by BMG Research. The main challenge in delivering the final survey was ensuring there was consistency in the way questions were asked in different languages.

The survey was first developed in English. It was then translated into five core languages (to reflect the language abilities of the majority of learners): Arabic, Bengali, Punjabi, Somali and Urdu. The translated surveys were all cognitively tested with native speakers of those languages and refined accordingly. Finally, the translated surveys were piloted again and final adjustments were made. Cognitive testing in this way helps to ensure that questions are being understood as intended. Translation can change the intended meaning of a specific question and at times, it may be difficult to translate questions that relate to concepts that may be unfamiliar in other languages. A full briefing session was undertaken with the survey test team prior to the cognitive interviews being undertaken so that all involved understood the background to the project, the cognitive interview process, the survey itself and what was required.

The testing explored participants':

- understanding of the survey questions;
- understanding of the response options/scales;

- preference of response options/scales;
- views on clarity, length and terminology used in the questions and response options/scales;
- suggestions to improve the questions and response options/scales.

The survey development process initially included plans to undertake back translation in the five different languages. This is where translated surveys are translated back into English and matched against the original English version to check for consistency. The compressed timetable for delivery of this trial meant that this was not possible prior to beginning the baseline events.

In practice, there were issues with the Punjabi translation of the baseline survey. One question was incorrectly translated; meaning that the data for this question was not comparable across language groups. The Punjabi survey asked 'Apart from your English class, how many people did you speak to last week from a different country or religion to you in English?' The reference to 'in English' was an error and not present in the other language versions. This error did not become apparent until after the baseline survey had been completed by 51 learners. To ensure consistency at baseline and post-test stage, the error in the question was kept the same for the post-test survey. However, the responses of Punjabi learners were excluded from the main analysis and analysed separately.

Some further issues with comprehension became apparent on assessment days. This was often due to the differing dialects used by learners and their own language capabilities. Many learners were not literate in their own languages. To overcome these sorts of barriers researchers attempted to explain the concepts behind the question.

BMG fully briefed researchers who implemented the survey in the field. This was designed to ensure the survey was delivered consistently as possible.

Overall, the survey tool was well developed and delivered but would have benefited from some more time in which to test and refine it.

4.4 Baseline and follow-up data collection (assessment) events

The objective of the baseline assessment events was to measure the English language capabilities of participants (in both groups) for speaking and listening, reading and writing; and to administer the social integration survey. Both measures were critical to the delivery of the RCT.

ESB assessors were tasked with measuring the English capabilities of learners for all speaking and listening assessments, as well as Pre-entry reading and writing

tests. Alongside this, class tutors administered the Entry 1 & 2 reading and writing assessments in group settings. The research team from BMG, with the support of interpreters, completed the survey with learners (see Table 1).

The intervention was designed to improve functional English language proficiency. As such, the primary success measure for the RCT was the impact on English language proficiency with speaking and listening considered the principal outcome measure (though reading and writing assessments were also explored).

Table 1 Assessment and assessor summary table

	Speaking & Listening Assessments	Reading Assessments	Writing Assessments
Entry Level	ESB Assessor	ESB Assessor	ESB Assessor
Level 1	ESB Assessor	TET Teacher	TET Teacher
Level 2	ESB Assessor	TET Teacher	TET Teacher

**** All assessments were marked and moderated by ESB assessors.***

The assessment events were intended to be held in the first week of classes for the treatment group. Participants assigned to the control group were also required to attend the assessment sessions during the same period (in most instances on the same day). Delivery staff attempted to keep treatment and control learners separate during assessment events, so that the risk of cross contamination was minimised.

The assessment events were challenging to deliver to the desired specification. These challenges stemmed both from the timescales within which events were planned and organised, and the number of parties involved in the process. The research materials (the ESB assessments and the social integration survey) were completed days before the first assessment event. The dates of assessment events were also confirmed at short notice (as dates were based on the combination of assessor and room availability at the community venues). Both factors added to pressure in terms of organising assessment events, inviting learners and booking the interpreter support required.

ESB assessors were initially expected to carry out all of the English proficiency assessments. However, during the trial preparation stage it became apparent this would not be feasible as the complete assessment process, on a one to one basis, could take up to an hour. In order to reduce the time burden, responsibility for administering the Entry 1 and 2 reading and writing assessments in group settings

was given to CBEL tutors. The change was made primarily to avoid learners leaving assessment sessions without sitting all the necessary assessments. ESB assessors completed all the Pre-entry level assessments and all the speaking and listening tasks. All centres were informed of the change in expectations and instructions on how to administer the assessments were provided so that it was implemented consistently. Importantly, although the deviation was not desirable, ESB assessors marked all assessments.

The change in assessing responsibilities was not actioned in one centre and ESB assessors attempted to complete all the assessments (including the Entry 1 and 2 reading and writing assessments). Two learners were not given the chance to complete all the assessments, as the assessments took too long. However, the level of attendance at this centre was relatively high and there was no indication that learner groups were treated differently (i.e. more or less likely to have sat all the assessments) so their data was retained for analysis in the RCT.

At another centre the duration of assessments also caused problems and staff reported failing to administer any Entry 1 or 2 reading and writing assessments to control group participants at baseline. All learners were invited back to a 'mop-up' session so that assessments could be completed, but it was clear that few returned. This issue may have introduced systematic bias to the data resulting from those learners who were put off attending a second session, or the delay in data collection that occurred.[14] As such, the reading and writing proficiency data from this centre was excluded from the final analysis. Speaking and listening data was retained, as all measures were collected by ESB in line with trial expectations.

In contrast to the baseline events, the follow-up collection sessions were far less of a challenge to coordinate and deliver. There were no examples of errors in the collection of data, or confusion regarding who should be conducting assessments. This was due to the experience gained at baseline events and the longer timescale in which to prepare for the events.

Despite the challenges faced by delivery staff at the baseline events, we conclude that the measures were collected with a good level of accuracy. Future trials will benefit from the lessons learned here, the most prominent of which is that practice assessment events should be used to iron out challenges and confusion around how the events should be run.

It is important to note that these issues were identified using the processes and data collected as part of the trial. As such, we are confident that the data was correctly

[14] It is plausible that control learners would have more time to practice their language skills in preparation for the assessments, particularly if they observed others sitting the assessment.

collected. More generally, this aspect of the trial highlights the need to allow time for practice.

4.4.1 Data quality assessment

To ensure that no other systematic biases occurred during the collection of assessment data, further enquiries outside of the dedicated fieldwork period were undertaken. Evaluators revisited the data, project documentation and communications from delivery partners to assess whether any other issues had been overlooked. Enquiries were made with Talk English Together consortium partners and ESB assessors to investigate if other similar issues had occurred in other centres, and whether assessments had been implemented consistently. This further exploration found no additional evidence of systematic biases in the implementation of assessments at any other centre.

During these enquiries, a minor issue with the recording of learner IDs was identified, resulting in assessment scores for four learners being excluded. Eight learner reference numbers were incorrectly recorded at baseline. In four cases, it was possible to retrospectively match assessments to the correct learner. However, for the remaining four learners it was not possible to do this with any confidence. As such, these measures were excluded from the final analysis. This error was likely the result of input error on the assessment day.

Overall, the evaluators concluded that the data collected was robust. The issues detected represent a very small number given the total number of events that were run.

All the data issues described above were accounted for in the RCT analysis by excluding affected learners/centres. While this approach had a marginal effect on results (i.e. the size of the unadjusted means and model estimates), sensitivity analysis found that the results of the RCT did not change fundamentally because of these exclusions. Therefore, we concluded that there was no evidence that the measurement issues affected the treatment and control groups unduly.

4.4.2 Delivery partners' experience of assessments

Despite the issues described above, overall the assessment days were successful.

"I suppose I mean in organisation you could be picky and say well, you know, it would have been better to have rooms, separate rooms, for each sort of activity so you completely cut out any distraction but I'm not sure that was actually feasible at some of the centres, they're quite small places really. I think it went as smoothly as it could have done really, yeah." ***ESB assessor***

"The centres were smashing, the support staff were really good and kind and helpful. The candidates were there, they were ready, started on time, finished on time, in my experience, and I had no issues with anybody or anything." ***ESB assessor***

In some locations, baseline assessments took longer than anticipated, or were conducted later (in the first or second week of the course) which affected learning time. Those teachers that reported this were able to catch up on the curriculum with little difficulty. This was usually achieved by compressing the club sessions to provide time for other materials to be completed.

Not all participants turned up for the assessments, so mop up events were needed. These were organised by location to ensure as many participants as possible were assessed.

"They were organised within the community centre where the classes were going to be held. They were conducted on the Thursday, the day that they have the activities…We just thought it would be easier to have in there; to conduct the assessment at the locality convenient for the learners." ***Teacher***

Interviews with ESB assessors found that the delivery of assessments was consistent across centres and participants, with no particular contextual issues or differences in the conditions under which learners were assessed that could have affected their scores.

"We standardised and we moderated them, everything was done according to what you would imagine and what you would hope for so I don't think there's any question about it being… there being any problems. From my point of view, I thought it was fine." ***ESOL assessor***

Despite the challenges encountered, there was no evidence to suggest contamination between the two groups because of assessment events. Most learners attended with friends or family so this may have helped to control this risk by keeping learners occupied while they waited. Likewise, in our interviews with learners, none stated that they were in contact with individuals from another group (though it is possible they were unwilling to admit to any interaction in interviews). Although contamination remained a risk, fully controlling for this may not have been a realistic ambition without additional time and funding to achieve it. As a community-based intervention, the catchment area for participants was purposefully small; this meant that participants were likely to encounter one another outside of formal intervention events. However, our exploration of these issues did not uncover any explicit cases of contamination occurring during the assessment events.

4.4.3 Impact of assessment and evaluation on learner experience

The registration events and assessments did not appear to cause any major issues for participants, with some stating that they enjoyed taking the tests and found them rewarding.

"It's a good idea to have these tests to figure out if we have actually improved our English level." ***Learner***

"One thing I will say, actually, about the survey and with the speaking and the listening with the ESB…my observation was that weirdly they kind of enjoyed that….people being interested in them and wanting their opinion…" ***Regional coordinator***

Assessments were initially confusing for some participants but most were content to take them when their purpose was explained. This confusion or concern may have been, at times, related to a general mistrust of officials. This mistrust was apparent during research interviews. For example, a relatively high proportion of participants refused to have their conversations recorded (only 10 of the 54 interviewed agreed to have their conversation recorded). While participants did not explicitly state their mistrust of the research, some teachers suggested this could be the problem.

A small number of participants reported that they found the assessments difficult and challenging.

"The English tests were quite stressful; especially the last question, which needed a bit more explaining and more thinking. The rest of the questions were pretty straightforward." ***Learner***

Assessors and teachers who were involved in the baseline events indicated there were instances where more explanation of certain questions and tasks for participants was required. This seems to reflect the low levels of education, or literacy, among some participants rather than an issue with the process of proficiency testing.[15] However, because of the briefing sessions held in advance of the assessment events all assessors and tutors were confident in providing these explanations.

4.4.4 Self-reported progress in English proficiency and integration during treatment period

Control group learners were interviewed as part of the process evaluation. This was to explore whether they had taken any steps over the course of the trial to independently improve their English language ability. As expected, and given that all learners had volunteered to take part in the course, most showed a high level of motivation to learn English and some described engaging in activities that would help them prepare and make the most of the classes starting in September. Much like the treatment learners, many stated they were spending more time listening to others speak English, talking with children or trying to practise the little that they may already know.

[15] Around a third of learners (30 per cent) had no formal education; another quarter had not progressed past primary equation (25 per cent).

“I’ve been trying to learn English by looking at [TV] with my grandchildren, also speaking to them, watching TV in English. I have really been trying to learn and pick up English; trying to talk to people... I think that what has improved the most is my writing.” ***Learner***

Behaviours such as these could explain some of the positive increases in English proficiency recorded among the control group participants. Another explanation could be greater confidence among control learners in the second assessment. In the second assessment they would know what to expect, and may be able to perform better. This improvement in performance, known as practice effect, is cancelled out when the two learner groups are compared but would still show as an improvement when assessing control group scores in isolation.

“The tests were good. I think I did better in the second test, the first time I didn't know what to expect and what to do… I wasn't sure what was going on” ***Learner***

Importantly, no participants in the control group reported seeing materials from the classes conducted between April and July.

4.5 Programme implementation and delivery

For the results of the RCT to be robust and valid, it was important for the intervention to be delivered consistently across all the sites involved. Pre-launch briefing and training sessions were used to communicate this to delivery staff and the intervention manual was developed to be used by teachers so that a high degree of consistency could be achieved.

4.5.1 Preparation and suitability of material

Almost all teachers and volunteers interviewed had received at least some training or induction on the intervention manual. However, the satisfaction with the training varied. Despite some complaints from teachers around the prescriptive nature of the manual, they were clear that as part of the training they had been told that it was critical to stick to the manual, as this was fundamental to the trial itself.

“It [the training] was just to tell us about the [course]… we’ll be getting the manual and we had to stick to the manual and not deviate in any way. It had to be… you know, lesson one, lesson two and this is the way you’re going to do it.” ***Teacher***

As has been noted in section 4.1.3, some teachers suggested that the training was rushed or had covered a lot in a short time. While this was not a problem for most of the teachers, it did cause confusion for some. One teacher, having overlooked the pre-prepared lesson material provided alongside the manual, put together their own materials based on the topic guidance for the first three weeks, after which point they started to use the correct manual tools once they became aware of them. Once discovered, they saw the materials as being well set out and very helpful but only in

the sense that it saved her from having to prepare classes herself. She did not suggest that the materials were different from what she would have prepared.

"I realised when I started… that I was still a bit lost in terms of paperwork. I think I didn't realise how closely I had to work with the manual. That was the problem. I kind of just saw it as something that you were given and that you put aside, but really it was a core part of everything and all the lesson plans and everything were in there but I didn't realise that at the time. I was already a few weeks into the programme when I realised." ***Teacher***

In this instance, the teacher had used the information within the manual to deliver what she understood to be part of the course. The original lesson plans did not elaborate on this, but data gathered through the qualitative interviews suggested that she delivered the same class, albeit with different materials than the manual specified. Participants from the class were not aware of the difference between the materials provided in the manual and those developed by the teacher. Analysis of the quantitative data did not suggest the class performed significantly differently to others. It is reasonable to suggest fidelity may have been partially compromised for this class, but not in a manner that posed a threat to the overall trial. This example highlights the need for greater clarity regarding what each session should include to be compliant with the intervention.

Teachers reported mixed views about the course materials. In the instance above, the teacher was not fully aware that the manual contained everything she needed so worked using it as a guide for, rather than the source of, her materials. Once discovered, this teacher gave the materials a glowing review. Likewise, other teachers felt the structure was sensible and the materials were useful and of a good quality. Most teachers emphasised, without prompting, the importance of following the manual and course materials and this was consistent with the descriptions of classes offered by learners. However, at the same time, teachers also described making modifications to classes or sessions to accommodate learner needs or to add more interesting content to particular sessions. In the view of these teachers, the changes they made were not substantive changes, but they were necessary adjustments to make the classes fit with the needs and capabilities of their learners.

"*I just made notes of what worked, what hasn't worked and how I've changed it, or if I felt that they need more reinforcing, so sometimes I've not finished the lesson… and I've had to carry on the next day because the learners haven't grasped the topic."* ***Teacher***

"I've had to change because some learners found it difficult… I have to personally say I had to adapt quite a few of my lessons because of the learners that I had…especially the reading one, I found a lot of the reading material was too demanding." ***Teacher***

In one instance, a teacher felt that the class on information technology was not pitched appropriately, suggesting it assumed that participants had some prior knowledge or understanding of how to use computers, and had email addresses.

"We have followed the course manual as we can to the letter, with the exception of the [IT lesson] … To get onto the IT, actually log them into the Talk English website, they have to have an email address. We had one person who had an email address but they didn't know what it was or how to use it, and the rest of the people had no technical knowledge about how to use a computer at all. So before we could actually get them onto the Talk English website to use the online training, you would actually have to run a training course for them. And for us to… create an email for them, create a log in for them, and then send them away to do something at home, was ridiculous." ***Teacher***

There were also a number of examples of teachers describing having to adapt materials so the content was easier to deliver. In one case, a teacher saw the materials as being too advanced for the abilities of their class.

Based on these accounts, it is difficult to quantify the extent of the deviations from the manual. However, the comments of teachers were consistent; their adjustments were what they expected to do as a qualified teacher. Moreover, there were no examples given where teachers chose not to teach a specific subject area. However, as is noted in the quote above the IT class did cause an issue for one site.[16]

"Personally, I think it was too prescriptive. It doesn't really allow for variation of the different abilities in the class enough. … Any qualified teacher would probably deliver those materials in a different way, depending on the individual needs of the learners in the class". ***Teacher***

One of the innovative aspects of the CBEL intervention was that every week there would be a group activity club session outside of the classroom. Some teachers and volunteers expressed frustration with the fact that, due to the structured course content, these community visits could not be scheduled more flexibly.

"It's quite structured… For example, if we have plans to go say to a chemist or to the café and it's raining and the ladies have no coats, we can't say, 'Well, we'll do this tomorrow,' because we can't because we're onto another lesson and there's no flexibility in it in that way. If we had a bit more flexibility, we could change maybe some of the weeks around." ***Volunteer***

As part of the trial monitoring process, teachers recorded any incidents in a lesson record. A total of 76 incidents were logged over the course of the trial.

[16] The data yielded from the fieldwork does not specify whether the IT class was subsequently omitted from the programme or adapted to suit learner needs.

The information included in lesson records was of relatively low quality for the purposes of analysis. The descriptions provided by teachers were limited (notes were generally a single sentence) and did not offer sufficient detail to fully understand the nature of the incident or its implications. Analysis of the data, presented in Table 2 below, attempts to group incidents by type.

Table 2 Class incident type and frequency

Incident type	Count
Issues with the intervention manual	24
Issues with class dynamics	13
Potential deviation from expected implementation	12
Slower than expected progression	8
Learner attendance	6
Other	13

Source: Lesson records

The most commonly recorded incidents (24 in total) related to the course materials and how they were used. Examples of this included materials being unavailable because teachers had not been able to print or download them from the online resource where they were held, or where segments of lessons were deemed too confusing for learners and certain vocabulary was subsequently changed.

A small number of these incidents (12) may have resulted in a deviation from the manual, but the information included in the lesson records was not sufficiently clear to be concrete for all examples. There were two examples in this category of teachers reporting that tasks did not fit well with the locally sourced community materials being used (for example role-play activities being changed to a group activity, or where there was no paper for the class to make posters). Five incidents were also recorded where the lesson was a poor fit for the learners (for example, one teacher said not all learners wanted to bring in pictures to share as part of the lesson).

The remaining issues logged may have affected the delivery of classes but did not appear to represent a deviation from the manual or specific materials. These issues related to the slow progress made by learners, disruptions due to attendance or in-class issues (for example, one incident related to some learners chatting while another was reading to the class). The ’other’ category relates to incidents that did not appear to be relevant to delivery, for example where a camera did not work in a class; where a fire drill disrupted a class; and where the teacher felt the class repeated too much from the previous class.

These records provide further indication that the manual had no clear direction about what constituted an unacceptable deviation. The remedy for this issue may be clearer guidance on what aspects of lessons must be delivered to meet the requirements of the intervention – whether it is part of a RCT or not – such as a core learning section at the beginning of each lesson plan. This could include topics, vocabulary or teaching methods, or indeed whatever is felt to underpin the approach of the intervention.

Overall, it is apparent that while the manual was generally delivered to a high standard, variation in implementation across sites did occur and had the potential to impact on trial fidelity. However, it does not appear that this variation was problematic for the trial as the qualified teachers, recruited specifically as part of the trial, were guided by the content of the manual to make changes that both helped participants learn while still attempting to achieve the lesson aims set out in the manual.

4.5.2 Class configuration and delivery

A maximum of twelve learners to one teacher was considered the optimal balance when designing the trial. However, in reality class size varied a great deal across different sessions and centres with an average class size of ten. One teacher confirmed that the lowest turnout encountered was a single learner attending one session. Another teacher stated that, apart from initial dropouts, their class had remained stable. Table 3 below shows that as classes progressed, the average attendance declined.

Table 3 Average session attendance

Session	Average attendance
1 to 11	7.4
12 to 22	6.4
23 to end	4.3

Source: Class attendance records

Attendance records indicate attendance varied considerably and that having a full class was rare. As a result, when talking about the challenges of delivery, teachers focused on the problems arising from inconsistent attendance rather than high (or low) numbers of learners per teacher.

Teachers delivered class and club sessions with support from two volunteers. Some teachers made use of their volunteers to support struggling learners. Overall, volunteers appeared to be providing effective support to teachers.

“They did lots … I just used them as I saw fit really. It was quite good really.”
Teacher

“I help with the resources; I also sit with the lower level learners to help them. We go out sometimes and I will, you know, stay with the lower ones again to make sure they’re okay and everything’s all right.” ***Volunteer***

However, one teacher reported that their volunteers were of limited help as neither were as experienced as the teacher had expected.

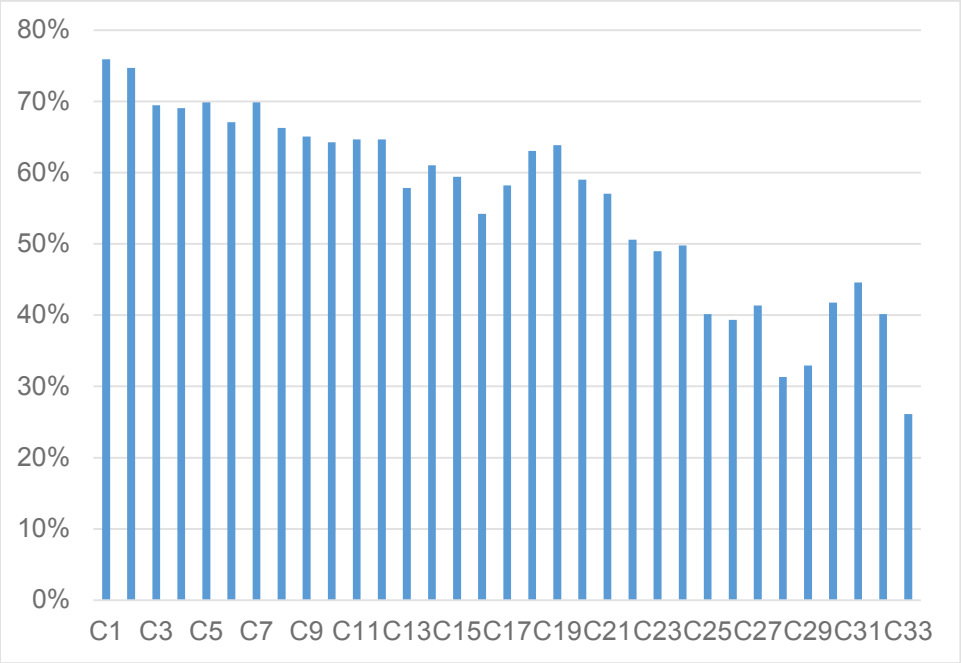
Separately, one volunteer did confirm that they had facilitated a whole class due to the absence of the teacher, but analysis of management information and interviews with area leads suggest that this was an isolated occurrence.

4.5.3 Attendance and course attrition

It was important for the trial to achieve a good level of attendance, as with low attendance it would be difficult to attribute change in outcomes to the intervention. Teachers and coordinators were asked to promote good attendance and were instructed to contact participants by phone and letter if they did not attend classes. Some teachers reported passing these tasks to coordinators, while others made calls themselves.

Average attendance across all sites and weeks was 56 per cent, however this varied considerably over the course of the trial. Figure 2 below sets out the number of learners attending each of the 33 sessions delivered. In the first week, attendance was around 75 per cent (indicating a drop off from registration of 25 per cent). A little over a quarter of learners attended the last class (26 per cent). Overall, there was a steady decline in participant numbers over time. The impact of Eid (sessions 28 and 29) is also noticeable.

Figure 2 Session attendance rates (all treatment learners)



Source: Compiled class attendance records

Attendance was affected by common issues – in particular health, childcare, and travel abroad to visit family. The attendance notes recorded two incidents of learners that were deported during the intervention but no further details were provided of the circumstances.

Course attrition was not a universal issue however, and some teachers reported consistent numbers over the eleven weeks. Reviewing individual site attendance records show that three sites had particularly good attendance. However, with such low numbers per class it is not possible to suggest that one or more specific factors contributed to this observation (a much larger number of learners per class would be needed to make statistically robust claims). It is plausible that the variation seen across sites was purely the result of chance.

"My class has actually been quite good. The ladies, if they don't attend they actually do ring me. If they don't ring, you ring them and find out whether they've had a hospital appointment or their child's been sick, but, generally, I've had good attendance." ***Teacher***

All classes had a high proportion of Muslim participants (overall 94 per cent of learners were Muslim). As a result, there was a clear impact from the start of Ramadan and from Eid. In most centres, there was a visible dip in attendance at Eid in particular.

"Attendance was spiky and this was a problem. We had Ramadan for a whole month, right in the middle of everything. It really slightly fell apart then. I noticed a big difference and some people never recovered, in the sense of the concentration that they lost, so I had people who were really, really clued in to begin with, really

enthusiastic, very good workers and they fell away and they never got that enthusiasm back or were able to reconnect with the material." ***Teacher***

Teachers and participants had different perspectives on the intensity of the course. Most teachers reported that having three classes in a week was beneficial – leading to participants absorbing and consolidating learning. However, a number reported that having three classes a week was difficult to commit to, often due to family responsibilities.

Some teachers reported that certain participants had decided to attend only some of the classes each week. One centre had a group that had to attend twice in one day, and some learners chose instead to attend only one of these two sessions.

"They did not like coming back in the afternoon. That was a big problem and we would have just two people or one person coming back in the afternoon. [Coordinator] had to send in somebody to give them a pep talk and I had to encourage them all the time and we never saw more than three people attending. They did not like it." ***Teacher***

Some participants reported that attending three sessions a week was a lot to fit in due to other commitments. In one example, a participant said they would selectively miss some of the club sessions. This should not be interpreted as participants having a dislike for club sessions. Many felt that these sessions were particularly beneficial to them, as they often related to learning more about their locality.

Teachers and coordinators described their attempts to try to maintain a good level of attendance but the indication was that, in most instances the absenteeism was unavoidable.

It was encouraging, however, that while learners commented on the particular usefulness of some sessions and topics, there was no clear indication that they chose to avoid sessions or topics. Likewise, teachers did not suggest that attendance was worse than would be expected for particular classes.

4.5.4 Assessing delivery

Despite the efforts of all parties, there was some variation in delivery of the intervention across sites, as described above. This was often as a result of adjustments teachers made for learners to accommodate differing abilities and to minimise the impact of poor attendance. Centres adopted a variety of measures to help learners catch up, which may have affected the amount of support different learners received across centres. Some were sent home with class materials to complete as homework. In one centre, the teacher reported using the first half of their club sessions to allow learners to catch up before the planned excursion.

This type of variation can be problematic for a RCT as it is likely to increase the variability seen in the results. Variation in the delivery of the intervention can also

affect the amount of the intervention received (known as the ‘dose’). However, thinking about the intervention in this way may not be helpful in this trial. As a complex intervention, it is not easy to deconstruct the intervention into constituent parts. Each class contributed to the overall objective of improving English proficiency but it is not the case that each class only contributed to one element of the learning process.

Analysis of the relationship between the number of sessions attended and changes in speaking and listening scores indicated that the relationship was non-linear (i.e. a higher score was not directly associated with attending more sessions). While it would be illogical to suggest that we would expect no difference in the amount of learning absorbed by those who attended one class and those who attended all 33, other factors may affect the relationship.

To try and understand this further, the process evaluation sought to explore whether this non-linear relationship could stem from different participant approaches to learning. For example, high attendance may indicate an individual was struggling with the material and attended more sessions to compensate; likewise, low attendance could suggest that an individual was happy with their progress and felt comfortable missing classes. The interviews undertaken with learners did not offer any evidence of this happening. Rather, attendance appeared to be impacted most by childcare responsibilities and illness. The data collected as part of the trial does not allow for further investigation to answer this question.

To be certain that the dose was administered consistently a more stringent approach to monitoring delivery of the trial would be required. Examples of this could include using detailed checklists to confirm the precise content of classes; taking audio recordings of lessons or employing class observers in all sessions (so adherence to the intervention specification could be verified). Such approaches would be more rigorous, but also very costly and difficult to implement (requiring consent to record lessons for example would likely have been an issue in the CBEL trial).

However, the purpose of conducting this RCT was to test the effect of a community-based intervention delivered in a real world setting with the aim of using the findings to inform future programme delivery. As such, adopting such stringent measures would seem to be at odds with permitting wider roll out.

Overall, it is not possible to say categorically that the intervention was delivered with perfect fidelity. There were clearly some issues with materials and due to variation in learner ability across sites the scope and scale of adjustments made by teachers varied – learner to learner and class to class. However, given that all teachers were trained professionals, and were motivated to adjust materials so that learners could progress, we conclude that the range of adjustments reported are not problematic for the RCT and do not undermine the overall results.

However, as a learning point for future trials of this nature, the intervention manual should be revised to more clearly stipulate which elements of a class must be delivered to constitute a full dose (and indicate compliance with the intervention). This will help overcome any gaps in knowledge around delivery, both with respect to recording or monitoring ad hoc adjustments, and identifying deviations.

4.5.5 In-class challenges and facilitators

Most participants reported being happy with the way that classes were set up and designed. Very few learners appeared to want to be critical of the classes or teachers as they were grateful for the opportunity to learn. Likewise, having mixed ability classes was not seen as an issue for most learners, with a number reporting finding it helpful that others were asking questions that they could listen to.

Working in groups was another positive factor that was often mentioned by learners. Having others there to ask questions was considered helpful, and there was little concern about where other learners came from.

Teachers commented on the impact of having similar cohorts to teach. For example, where learners all spoke the same native language (more common with large clusters at a site, but not exclusively so) it was easy for them to help each other if a topic proved difficult. Conversely, some teachers pointed out that where classes were more mixed the commonality among participants was their English language; forcing more use of English in the class.

It was also noticeable that, for some learners at least, the teacher's approach had an impact. Several learners in one class described how friendly and patient a teacher was.

"*Back in [home country] teachers shout; but I think it is better to learn when you are not afraid."* ***Learner***

Some teachers reported issues with female participants being reluctant to attend classes with males. Generally, this complaint reduced over time. This issue was not identified explicitly by any of the learners interviewed for the research.

As has been noted, childcare was an issue for many female participants. In some centres, there were crèche facilities but this was rare. It should be noted that having children was also a help for some participants – in particular those with older children (those children already in school) who could help them with their learning and practice. For example, one participant said that she had been watching children's programmes with her children to improve her understanding of English.

These accounts were also supported by the quantitative analysis in the RCT report where participants with children achieved better results.

5 Conclusions

This evaluation sought to:

- explore implementation of the CBEL intervention, identifying any variation in delivery and overall assess whether fidelity to the RCT design was maintained;
- identify key lessons, issues and challenges in implementing a social intervention as a randomised controlled trial; and
- generate evidence, learning and recommendations for the design and implementation of future RCTs and English language interventions aimed at similar groups.

Overall, the RCT was successfully implemented. Despite a number of challenges, an adequate number of participants were recruited; learners were taught sufficiently well, such that they improved their English proficiency; and the RCT was able to identify that effect from the data collected.

However, it is not the case that each element of the RCT was delivered with absolute fidelity to the original RCT design. We will now summarise the overall findings in relation to the randomisation, contamination and the design and delivery of the intervention.

5.1 Randomisation

The process of randomisation, critical to the success of the RCT, was delivered successfully. All parties involved in this process worked hard to achieve this critical step, guided by the briefings that formed part of the trial itself. Learners were recruited, clustered where necessary, randomised and informed of their allocation in sufficient time to begin the intervention.

While a small number of issues were identified, none of these jeopardised the validity of the RCT.

5.2 Cross contamination

The trial was also successful in ensuring that cross contamination did not occur. Clustering of participants reduced the possibility of participants sharing class materials with those who were in different groups to them (i.e. treatment to control). Likewise, the recruitment and assessment events did not appear to contribute to the risk of cross contamination. The most critical part of the trial that influenced the risk of contamination, however, was training all delivery staff on the importance of randomisation and assignment in advance of delivering the intervention. While many

teachers discussed the problems that the RCT rigidity imposed upon them (namely not being able to drop poor attendees and replacing them with other more motivated learners), almost all stated that they were fully aware that learner assignment could not be altered in any way.

5.3 Intervention design and delivery

The expectation was that teachers would be able to deliver the intervention (as directed in the manual) in exactly the same way in each location. The reality however, was that teachers saw the need to make ad hoc adjustments to the materials and their teaching approach in order to accommodate learner needs.

The manual was not finalised with sufficient time in advance of the trial beginning for all parties to review it and to reflect on the proposed content and structure. Minor errors in the manual (such as spelling and grammar) could be amended quite simply. The more complex issues described to us, however, caused variations across lessons and sites that were difficult to quantify.

The manual was also viewed as too prescriptive and outlined the content to be delivered but did not give clear guidance on the core learning aims. Effectively, the manual asked teachers to deliver content rather than teach learners.

In practice, teachers felt capable to make ad hoc adjustments to accommodate learner needs. However, to what extent these adjustments constituted a deviation from the intervention is not clear, as there was no guidance to check against. Likewise, there was no guidance on how to treat learners that missed classes (i.e. whether to offer homework). On reflection, both issues should have been anticipated and covered in the manual. Together these omissions indicate an oversight in the design phase. The manual did not address these very likely risks to trial fidelity.

Therefore, some minor variation in delivery of the course was identified. However, teachers adjusted content and materials, in line with their professional capacity and with the aim of accommodating particular learner needs. Importantly, none of the lesson amendments were found by this evaluation to deviate from the original aims of the specific classes. As such, it is not believed that these variations present a deviation beyond what would reasonably be expected and do not jeopardise the conclusions in the RCT.

A number of other issues were identified within the trial's design and implementation. These related to the requirements of delivering an RCT and issues around the timescales and implementation of this intervention. In particular:

- The eligibility criteria may not have been met by all learners. Some learners interviewed made statements that contradicted the criteria. It is important to remember that eligibility was self-reported. These examples may be due to

misunderstandings (in either our interviews or the understanding of participants while answering the criteria questions).

- Timescales for implementation were generally short at all stages – in programme design, evaluation design, material development, recruitment and delivery. As some respondents pointed out, to some extent this is inevitable in programmes of this sort and challenges were largely overcome. However, in some cases short timescales affected the ability to deliver some elements of the trial and research to the highest possible standard (such as recruitment of learners, delivery of assessments and back translation of the survey). While steps were taken to address any issues raised during assessments, more time would have enabled a more thorough monitoring of the trial and more attention to detail, particularly in terms of refining the manual and research tools.

- In some cases the short timescales may have contributed to some of the challenges in delivering the intervention and evaluation such as allowing for sufficient training, and reflecting on the manual in advance of the trial.

These factors do not appear to have prejudiced the validity of the trial, but they could provide useful learning for future interventions.

Overall, the process evaluation findings indicate that the RCT was implemented in a robust manner, thereby lending weight to the conclusions of the impact assessment, namely that the CBEL intervention led to improvements in participants' English speaking and listening, reading and writing skills, as well as improvements on a number of measures of social integration.

5.4 Recommendations for future CBEL trials

- Extend timescales for design and delivery planning and evaluation. A longer development phase would allow delivery partners to review materials and ask questions that could be beneficial to successful delivery. In particular, this would allow more time for teachers not only to review and comment on the manual, but also to familiarise themselves with the content to ensure its accurate implementation.

- Allowing more preparation time would also ensure sufficient piloting of resources and activities that are likely to prove challenging to delivery. For example, the assessment days showed that with practice (i.e. follow-up compared with baseline assessments) a complex task could become easier. Similarly, more time would also enable more refined cognitive testing and back translation of the survey materials. Finally, taking follow-up measures

over a longer period (e.g. at 6 and 12 months) would enable longer-term impacts to be observed.

- Provide greater clarity for teachers about their expected role in lesson planning and delivery. This could include setting out how teachers should deal with regularly occurring issues such as how to treat learners with poor attendance, or what elements can be adapted for local circumstances. It would also be useful for teachers to have well-defined learning objectives that they must achieve for their class. Whether these objectives should stipulate specific vocabulary, learning methods or concepts is for ESOL practitioners to define.

- More clarity should be provided on exactly what constitutes an acceptable 'deviation' from manual content. Guidance should be included on what aspects of lessons must be delivered to meet the requirements of the intervention.

- Review the intensity of provision – for example by conducting 'clubs' every second week, which again may be more compatible with a longer intervention, or reducing the number of classes.

- Consider how local areas and partners are supported in the delivery of the trial. It was particularly striking in this case that both Talk English and its partners were well established and able to deliver this intervention; and that MHCLG could provide relatively intensive support. Future trials may be less prepared in this regard and may need further support in terms of administration and resource.

Annex A

Qualitative summary of CBEL Outcomes

Background

The Ministry of Housing, Communities and Local Government (MHCLG) commissioned the delivery and evaluation of a Randomised Controlled Trial (RCT) of a Community-Based English Language (CBEL) intervention. The RCT evaluation provides an objective assessment of the impact of the CBEL intervention on English proficiency and social integration.

In order to better understand the implementation of the trial, a supporting process evaluation was also commissioned. This comprised 54 qualitative interviews with individuals involved in the trial, including learners and those responsible for delivery. These interviews explored how the RCT was delivered – focusing on the challenges of delivering the intervention under trial conditions in a real world setting and looking at whether (and if so why) there was variation in implementation and delivery across trial sites.

The qualitative interviews also explored perceived progress, benefits and outcomes achieved by learners, which were attributed to participation in the trial. This paper presents summary analysis of the qualitative outcome data as an addition to the RCT evidence in order to offer a richer description of the range and importance of the outcomes reported by learners and other stakeholders that were attributed to the intervention. Data collected from learners, teachers and volunteers, were analysed in the same way as has been described in the main body of the process evaluation.

Intended outcomes

The CBEL intervention was designed with the aim of improving learners' English proficiency and social integration. By attending intensive English classes with others from their local area, learners were expected to gain skills and confidence using English with other people. Club sessions were designed to provide learners with the opportunity to put their skills into practice and to familiarise them with services and activities within their local community.

Outcomes achieved

Analysis of qualitative data revealed a number of outcomes achieved by learners. These outcomes clustered around five interrelated themes.

- Self-reported progress
- English acquisition
- Confidence
- Experiential learning

- Independence
- Social integration

Each theme is explored in detail below.

Self-reported progress

The progress reported was very consistent across all learners, and often related to achieving the personal goals they had set themselves.

"When I came here I had no English at all I couldn't even ask for directions if I got lost. But now I can go out knowing that if I get lost I can ask for help". (Learner)

While most learners reported having seen progression in their own language skills, some did have ambitions over and above language alone. For example, one learner who had a good grasp of spoken English had hoped to improve his writing abilities, which he felt had happened. Other learners described more intangible changes – namely improving their confidence when communicating.

Another learner, who had studied English in their home country, also wanted to improve their confidence most of all. Their ambition was to move onto other classes; the CBEL course was for them an interim step to further learning, which they had managed to organise.

English acquisition

The qualitative data provided strong evidence of improvement in English proficiency. This was consistent with the RCT, which found that learners improved their English capabilities across the range of domains. Learners reported a range of situations in which they used these new skills, such as speaking with strangers, speaking to businesses and services, and improvements in their interactions with others more generally.

Importantly, their newly acquired English skills were highly valued by individual learners, and appeared to have led to substantive improvements in their lives. For some, this meant being able to directly communicate with family members who had been born and raised in the UK.

"I can speak with my grandchildren now." (Learner)

The expectations of delivery partners regarding the potential for improvements in reading ability were relatively low at the outset of the trial, in part because many learners were not literate in their first language. However, teachers who were interviewed reported improvements in reading skills acquired by some learners (a finding substantiated by the results in the RCT).

"There are two in particular who've done really well and have improved a lot more than I expected. One lady couldn't recognise any letters of the alphabet when she started, there was another one who was really struggling with CVC[17] words and reading and she's come on a lot as well." (Teacher)

Improved English proficiency underpinned numerous other benefits and was evident in all of the other themes identified through this research.

Confidence

Teachers and volunteers strongly believed that confidence in using a language was one of the greatest obstacles to learners' independence and integration.

Becoming more confident in using English was a key objective mentioned by learners, who reported being worried *"about being spoken to in the street"* or even talking to known individuals, like a neighbour.

The CBEL intervention provided learners with a space to practice their English without feeling self-conscious or hesitant about making mistakes:

"I do try to speak English outside of the lessons, but I don't know if I'm saying things wrong; here I can be corrected." (Learner)

The course appeared successful in helping learners achieve this aim with strong evidence of changes and improvements in learners' personal confidence with English over the course of the intervention. The reported increase in confidence is also reflected in the survey findings as part of the impact evaluation.

As well as the increased confidence in being able to speak English, learners also reported being more confident in its everyday usage:

"I feel much more confident… going to the clinic and shopping. I even speak to people in the park now." (Learner)

Many learners mentioned the practical implications of increased confidence. This mainly related to being able to undertake normal activities independently. For some this meant being able to leave the house without the need for support from husbands, children and other family members. Learners also reported being *"happier"* to visit doctors or a dentist *"on my own"* as well as carrying out other activities independently:

[17] CVC words are three letter words that include a consonant, vowel and consonant.

"I'm shopping on my own, I don't have to wait for my husband or mother-in-law. I'm even speaking with others in English outside now too." (Learner)

The progress made by learners was also observed by the teachers, many of whom spoke of improving confidence as key to the continued progress of learners outside of the classroom.

In exploring the relationship between what was taught on the course and improvements in confidence, one teacher explained:

"She [a learner] *comes into class and says, "Hello. Good afternoon, how are you? It's a nice day." This confidence then goes into the outside world which means they tend to start interacting more when they're out. Because we increased that confidence they tend to start using* [English] *more and hopefully it starts a cycle and most of the ladies seem to be following this same sort of cycle. As their confidence is increasing, they are tending to use more and more English outside." (Teacher)*

Aside from improved English proficiency, learners also reported two additional lessons they learnt from the programme which appeared to improve their confidence. First, that it was okay to make mistakes when speaking English outside of the course; second, that they could ask people to repeat a question or speak slower to help them understand. Prior to the course, learners were reluctant to do so. As one teacher noted:

"They would have somebody to help them… make a phone call… they're not as scared as they used to be before, they've got more confidence, they say they will speak even though it's not correct, but they can get their information across." (Teacher)

Overall, learners reported feeling more confident about their English proficiency and their use of English outside of the classroom because of the intervention. Teachers and volunteers echoed this view:

"I've been surprised with their progress, particularly confidence; I've seen an improvement in confidence." (Teacher)

Experiential learning

Learners were asked about their experience of the classes and what sessions stood out most for them in terms of practical usefulness or those of most interest. In total there were 11 different topics covered by the course; however 'health', 'shopping' and 'transport' were the sessions recalled most often and positively in the accounts of learners.

“I liked the trips into town, to see the tram stop, to go to the town hall and head to Asda to practice shopping. The most helpful thing for me was the transport classes and the doctors.” (Learner)

It appeared that the practical application and relevance of these topics and visits were important. As mentioned in the previous section, for many learners, being able to go to the shops without other family members to help them communicate was the first (and almost immediate) manifestation of the impact of the intervention.

Despite the popularity of these particular modules, there was no evidence, either qualitative or quantitative (e.g. though the attendance records) suggesting selective attendance depending on the class topic. Similarly, analysis of the attendance records did not suggest that club sessions were better attended than class based sessions, despite appearing more popular with learners during the qualitative interviews. It seems that while learners were able to draw a distinction between the more traditional class based sessions and the more applied club sessions in terms of content, both aspects were seen as equally important and valuable.

Teachers viewed the value derived from ‘club sessions’ as being able to put what was learnt in class sessions into practice. While learners had made use of local facilities and attended medical appointments with relatives, they had not done so on their own. Teachers reported that the combination of familiar activities but unfamiliar tasks (e.g. learners needing to listen, speak and read within these situations) appeared to have helped some of them to develop greater independence with these everyday events.

Overall, it was apparent the English skills acquired by learners provided the potential to fundamentally impact on their lives, for example by enabling them to talk to their doctor, help with their children’s homework, or simply join a gym or library:

“I learned very useful things such as how to fill in the forms at the library.” (Learner)

“As part of visits they've gone to Tesco, the gym, the library, the park… It's been really helpful. She felt more encouraged to attend the gym; she has always wanted to go but always felt scared to go inside. She has gone to the gym now.” (Teacher)

Independence

As the previous sections have indicated, participation in the CBEL intervention led to improvements in English proficiency, which in turn led to greater confidence using English.

In describing the impact of these outcomes, learners and teachers gave examples of learners independently interacting with people, services and everyday activities (in some instances for the first time).

Other examples of how learners were gaining independence were more subtle. One learner was now 'checking' with her husband on how she would say things on the telephone and then making calls herself rather than getting him to make calls on her behalf.

Another learner reported being *"very happy"* that he was no longer reliant on his son for assistance at his business:

"*Because I had no English I've had to rely on my son in some business situations; negotiating with customers. Now I know that I've done the best I can, so I'll get better results on negotiations." (Learner)*

Teachers and volunteers saw other examples of how learners were gaining independence as a result of their improvements in English ability and confidence combined with their other skills and experiences:

"[The learners] *had to find pictures that they would put into their leaflet, and then they could write a sentence, and then I saw one of the ladies using her internet to actually find out a little bit of information, leisure centre - she used the internet to find out what activities they have in the leisure centre, and she looked at me thinking, "Am I not allowed to do it?" and I was, like, "No, that's fine, because you're researching, that's absolutely fine, you can do that." So that was quite interesting" (Volunteer)*

These examples suggest that there is the potential for sustained improvements because of the intervention. As learners become more independent (and more confident with their language skills), the opportunities for self-initiated learning opportunities are likely to increase. Language skills are improved with accumulated practice.

However, increased independence was not equally shared across groups, with older learners appearing to be more reliant on existing social networks:

"They're the learners, especially my older women learners, they are still I feel very reliant on family members and not quite confident yet to be able to do things by themselves." (Teacher)

Indeed, interviews with older learners themselves suggested that they felt they had made less progress than younger one:

"Maybe my age. If I was younger, I could have absorbed even more… Obviously, when you're younger you learn more" (Learner)

Social integration

For the purposes of the RCT and process evaluation, social integration was measured in terms of social interactions with, and attitudes towards, people from different backgrounds and participation in wider society (including confidence around engaging with professionals).

Teachers were acutely aware of challenges learners experienced in their everyday lives. They reported the practical barriers that learners faced such as having caring responsibilities, limited access to other communities and low levels of confidence that stopped them using local services. These issues affected their opportunities to interact with other people in their communities. Likewise, a lack of language skills added to social isolation.

"The other thing for the learners is some of them have never visited the library or leisure centre. They got to do that on the course. They had never taken the tram, none of them" (Teacher)

Learners described having more opportunities to integrate as a result of finding out about local services and amenities as part of the course. For example, one learner, who subsequently joined a local gym of her own accord reported:

"The trip to the leisure centre and gym were great. I have the information I need now to go on my own. I didn't realise until we all went together that they have women only classes." (Learner)

The integration benefits of the intervention, in terms of greater participation in the local community also spread beyond the individual learner to include their families. One teacher noted that as a result of work on museums and local events carried out as part of the course, a couple of learners with young children had acquired leaflets of local events being held in the city centre which they intended to take their children to during the school holidays; neither of the women had previously taken their children to such activities.

There was some evidence to suggest that the increased integration appeared to be self-sustaining with some expanding their involvement in the wider community:

"A couple of them have already, off their own back, come out and joined little clubs in the library, local libraries and joined the leisure centres, things like that." (Teacher)

In theory, participation in CBEL classes provided the opportunity for increased social interaction (as learners were exposed to new people). However, the extent to which learners did indeed interact with others from different backgrounds was varied, as many classes were relatively homogenous. This was due to a host of factors, including the clustered randomisation process, the demography of the local areas

where interventions were hosted, as well as the eligibility criteria. Indeed, interviews with learners did not suggest that the classes had provided increased contact with individuals from different backgrounds. Despite this, there were still some instances reported by teachers where students had difficult interactions with their peers. Learners did not necessarily all get along from the outset.

The issue most often cited was where women felt uncomfortable learning with men in the class; however, there were at times other frictions between women. Teachers reported, however, that most issues eventually disappeared with time. Learners overcame their initial caution of others in their class and there were no examples given by teachers of learners who chose to no longer attend because of these issues.

"I think they had more cultural awareness by the end of the course and more tolerance for each other because, like I said at the beginning, people were refusing to sit next to people. By the end, that wasn't happening at all." (Teacher)

Further, teachers and volunteers observed that learners had developed friendship groups through attendance of the course:

"They go out and they're in their little group and they all walk down together whereas before they all walked separately and now they all walk together so they've got that little friendship going and they're more confident when they come in." (Teacher)

In another example, it was reported that a group of learners drawn from different backgrounds had arranged to meet on a weekly basis at a café that they visited as part of the course to catch-up, socialise, and practice their English.

This qualitative evidence supports the results of the RCT that social integration was positively impacted by the intervention. The indication from these accounts is that mixing within the classes themselves was an important factor in achieving this result; learners were brought together and were motivated to maintain attendance due to a shared ambition (i.e. learning English) which in turn allowed time for more constructive interactions between learners of different backgrounds to occur.

Further research would be helpful in exploring just how impactful such social interactions within classes are on language proficiency outcomes. It would be useful to test the hypothesis that facilitated interactions between different groups (i.e. formal classes with mixed groups of learners) is more effective at improving social integration measures than English language classes alone (with homogenous groups of learners).

Conclusion

The qualitative evidence presented here suggests that the CBEL intervention improved English proficiency outcomes, increased confidence and independence and supported wider integration. Language use that is anchored in real world situations provided opportunities to develop in all of theseareas.

The qualitative evidence appears to suggest that, by increasing learners' confidence and baseline English skills, they are quickly able to make use of their new learning. It seems likely that the community approach supports this by providing scenarios encountered in everyday life through which learners can cement their skills and confidence.

The community approach also helps with developing new networks that have the potential to be sustained after classes have ended. Likewise, the 'hand holding' approach used to explore the local services on offer does appear to have opened up further opportunities for learners to continue their integration into the wider community through engaging and participating in such services.

Furthermore, the qualitative evidence presented in this annex directly reflects the positive quantitative findings described in the RCT. The trial found that learners in the intervention group had significantly higher proficiency scores across all measures, and in line with this, learners also affirmed their own progress in the interviews. The intervention has measurably improved the English language capabilities of learners and this has resulted in some tangible results for many interviewed. It has also affected their attitudes to others. Learners have reported improved confidence in both the interviews as well as the RCT survey, which may have helped to increase and improve their interactions with others, and given some of them the tools to begin progressing independently simply by having more opportunities to practice and expand on what they have already learned.

Annex B

Topic guides

The following section contains the full complement of guides used in the course of the process evaluation. These are in the following order.

- Centre and area leads
- Assessors
- Teachers
- Integrated teachers and volunteers
- Treatment learners
- Control learners

Regional coordinators and Centre Leads – Topic Guide

Introduction
Learning and Work Institute is an independent research organisation. We have been commissioned by MHCLG to evaluate the process of delivering the Talk English randomised controlled trial currently being delivered in Greater Manchester and West Yorkshire.

Purpose of the interview
The purpose of the interview is to explore your experience of the trial so far, what has worked well, and what could have been improved.

Confidentiality and consent
Anything that you tell us will be treated in the strictest of confidence, and your comments will remain anonymous. We would like to record the interview so that we can make sure we record what you say exactly, but you do not have to be recorded if you would prefer not to. Recordings will be kept securely, not accessible to anyone outside the research team and will be deleted at the end of the project.
The interview will last around 60 minutes. You are under no obligation to answer any questions, if you do not feel you can comment on specific questions do not worry. There are no wrong or right answers.

Check if participant is happy to proceed. Gain consent to record

Regional coordinators and centre leads
Note for the interviewer: These stakeholders to the trial will give a strategic view of the intervention and trial in the main. Centre leads may also provide insight from the teachers and other front line staff as well as a hands-on understanding of how the engagement and recruitment processes have worked so far.

Introduction and warm up
ASK ALL COORDINATORS AND CENTRE LEADS
Can you tell me what your role is and how long you have been involved in ESOL activities?
What is your role in the Community-Based English Language / CBEL trial?

What does this involve?

How long have you been working in this particular area (geographic)?
Do you have a view on which communities/individuals most need ESOL support in the area?

Tools/guidance (including briefing, manual, other support)

ASK ALL COORDINATORS AND CENTRE LEADS
Note for interviewee if needed Briefing materials were prepared to provide an overview of how the intervention worked and to make it clear to those delivering the intervention what would be required of them while participating in a RCT.
What briefing did you receive in advance of the trial (either in writing or verbally)?

How useful was the briefing information provided?
a. Are there any areas you would have liked more information on?
b. How did this impact (facilitate or hinder) your ability to deliver the trial?

Have there been sufficient opportunities for you to raise questions and receive clarification? If 'No', why not?

Were your colleagues aware of the Talk English delivery manual prior to courses starting?

If yes:
Did they have sufficient time to read it and ask questions; did they receive any training on how to implement the course using the manual; is the manual useful, how could it have been improved?

If no:
What are the reasons for this? What materials are they using in the absence of the manual?

Recruitment and delivery
Note for interviewee if needed specific eligibility criteria has been set as part of the CBEL trial. This differs slightly from the previous approach (with the inclusion of a new minimum 12-month residency period
Can you describe the approach you used to recruit learners to the trial?
How did you and your colleagues reach out to the target communities? *PROBE – leafleting, using existing waiting lists, word of mouth?*
Did you receive any help to recruit learners from any local groups or community partners (e.g., other community organisations, religious establishments, and links with schools/libraries)?
What worked well/less well?
ASK COORDINATORS ONLY
How has recruitment varied across locations / centre in your area?
a. Have any sites stood out/ (e.g. due to very high or very low levels or recruitment)?
ASK CENTRE LEADS ONLY
How did the recruitment process work at your location?
ASK ALL

Were you able to recruit the minimum amount of learners required per centre (18)?
If yes:
And what about the targeted number per centre (24)?
What do you think the reason for this was?
What challenges did you encounter? How were these overcome?

If not enough recruits:
What were the main barriers or issues you encountered with recruitment?
PROBE – length of time had to recruit; overlap with Easter holidays; staffing

What should be done in the future to make sure enough learners are recruited?
Are there any organisations/ agencies that could support you to recruit sufficient numbers in the future?
Are there any changes you would make to the recruitment process that would support you to recruit sufficient numbers in the future? (Ask for specific examples)
Is there anything (else) you would do differently next time?

How did you monitor whether the eligibility criteria were being met?

How would you describe the current attendance rates (attending the course) of learners in your area? Was it generally high or low? Why?

And what about the rate at which Learners have dropped out? Higher or lower than would be expected?

Why do you think participants dropped out after registering?
How could this have been avoided/could be avoided in future similar trials?

Design and targeting of participants
Do you know what the eligibility criteria is for the CBEL trial?
ADVISE OF CRITERIA (BELOW) IF NOT SURE.
have little or no English language skills (Pre-entry - Entry 1 ESOL)
not be eligible for mainstream ESOL support, as delivered via the Skills Funding Agency (therefore not in employment or actively seeking employment, i.e. not on JSA)
be aged 19 and above
have been resident in the UK for more than 12 months

Do you think the criteria are appropriate given the objectives of CBEL? Why/ why not?

How would you adjust the criteria, if at all, to better fit the objectives of the trial, and why?

How do these criteria differ from the criteria you would normally use in this area/ programme?

Do you know of any instances where you or one of you colleagues had to turn potential learners away?

If yes:

Why this happened / what where the circumstances?
What would you do, if anything, in the future to avoid this?
What do you think was the impact of having to turn learners away?
PROBE – impact on reputation/ trust with target communities?

Were there any specific parts of the criteria that participants did not commonly meet / satisfy?

In your opinion, how useful has the eligibility criteria been?
What issues have you or your colleagues had in applying the criteria?
How does the eligibility criteria applied here contrast with your experience (if any) of using other criteria for participation in other ESOL support?
Were there any target groups you had hoped to recruit more of, but were hard-to-reach in the community

Were there any learners that were eligible but subsequently didn't take up the offer of CBEL?
What were the reasons for this?
How do you think this could be best addressed in the future?

Were there any other issues related to CBEL that you believe may have put potential learners off from joining the trial?
Prompt – timing of the classes, duration, having to share data, being assessed, having to give consent to be part of the trial, gender/age balance of the class, etc.
Can you briefly describe an example or two of where this occurred?
How could this best be addressed?

Were there any common motivations for participation that learners mentioned?
Did any learners mention any unexpected motivations?
Were there any other reasons that people wanted to join the course?

Registration events
Were you involved in the registration events in any way?

If yes:
How do you think the registration event went?
Where there any organisational issues at the day? How were they resolved? Would you do anything differently in the future?
If no:
What feedback did you hear from your colleagues or learners about the registration events?

Baseline events

Were you involved in the assessment process in any way?

How do you think the assessment process went? What were the views of participants?
PROBE – overall format of the day; staffing; coordination with ESB and BMG; learner experience

Do you know whether participants had any problems understanding the questions being asked to them in the survey being carried out by BMG? If yes, ask why
a. Were there any questions that caused particular problems for participants?
Did this affect any communities/language groups more than others?

Were there any other issues you or your colleagues were aware of with regards to how the survey was delivered?

Are you aware of any feedback participants gave after their English (ESB) assessment?
If yes:
How did the ESB assessments differ or align with the English language assessments used previously by Talk English?
Do you think that the levels presented were appropriate?
Are you aware if any learners found it too easy/too difficult?
Did this affect any communities/language groups more than others?
Did [name of specific centre(s)] have enough time to assess all learners (including giving them a chance to take the Level 1 and/or 2 tests if applicable)?

How would you describe participant attendance at the baseline sessions?
Was there a difference between April/May and September learners?

What activities did you and your colleagues do to encourage attendance?
How effective was this activity?

Thinking about future assessment events, what do you think can be done to encourage participation and retention during assessment days?
What support could be provided to you and your colleagues to encourage participation?

Overall, how do you think the assessment process could be improved (if at all)?
What were the main challenges that you encountered during the event?
How do you think these could be avoided in future?
Treatment and control – cross contamination risks?
ASK CENTRE LEADS ONLY
Note for the interviewee if needed: A key part of the RCT is to have two groups of people, treatment and control, so that the results seen in those participants receiving English classes can be compared with a similar group of people who are not receiving classes. It is always important in an RCT to make sure that the groups chosen are not changed in any way.
How were the assessment events organised in your location?
How many rooms did you have available?
How many staff members did you use as part of the baseline event?
Were the two groups (those starting in April/May and those starting in September) seen separately or together?

What (if anything) have you done to ensure that treatment and control groups remain separate?

Did many participants at your centre come in groups/ with friends to learn together?
Did they end up in the same group for classes?
Were there any learners who wanted to attend the same class as their friends but weren't selected together?
Did this cause any difficulties?

Since starting classes have the correct participants (i.e. those assigned to the correct centres) turned up to sessions?
Have you had to turn anyone away from lessons because they were not supposed to attend that session?
Have staff allowed any of the 'wrong' participants to attend their class (whether it was a learner assigned to control attending classes now or a learner attending classes in a centre different to the one they were assigned)?
Do you think participants are sharing their learning with others?

What challenges have you had to overcome, if any, to make sure that 'treatment' and 'control' groups have remained separate?

To what extent do you think you have been successful in this?

Overall reflections and close
What do you think the benefits of learning English are for the learners (if any)?
Prompt – Additional benefits like increased employability, integration in the community, etc.

Do you think the right communities/individuals have been targeted?
If yes – why? If not, why not?

Are there any particular issues in this local area that you think are important to consider when assessing the effectiveness of CBEL during this trial stage?
Prompt – the type of participants who have attended, attendance rates, the way the classes have been delivered, etc.
This can be two questions: first of all issues in relation to organising English classes in general – and secondly, issues that relate to how the trial has mapped out in the area.

Overall, how successful do you think CBEL was at engaging those who most need the trial?
What makes you say this?
What do you think is the best way to engage them in CBEL?

Which parts of the intervention do you think works well?

Which parts do you think do not work well?
What would you change to improve the intervention?

What has been the key difficulties in running the intervention as part of an RCT? What could have been done to make this easier?

Thank and close.

ESB assessors – Topic Guide

Introduction
Learning and Work Institute is an independent research organisation. We have been commissioned by MHCLG to undertake an evaluation of the Community-Based English Language trials.

Purpose of the interview
The purpose of the interview is to hear about your experiences in delivering the English proficiency test, what has gone well and possible areas for improvement.

Confidentiality and consent
Anything that you tell us will be treated in the strictest of confidence, and your comments will remain anonymous. We would like to record the interview so that we can make sure we record what you say exactly, but you do not have to be recorded if you would prefer not to. Recordings will be kept securely, not accessible to anyone outside the research team and will be deleted at the end of the project.
The interview will take less than half an hour. You are under no obligation to answer any questions, if you don't feel you can comment on specific questions don't worry. There are no wrong or right answers.
Check if participant is happy to proceed. Gain consent to record

Introduction and warm-up
Can you tell me what your job role is and how long you have been in it?
How long overall have you been involved in the delivery of ESOL activities [either testing or teaching]?

Development of the tool
Were you involved in the design and development of the tool?
If yes:
How were the tests developed?
Probe: theory behind the design and the methods used for the development.

Did you receive any type of training on how to use the new assessment tools? How was it?
Assessment of learners
How did you find the assessment day?
Probe: Any organisational issues; were you clear on what you needed to do when; did the day overall go to plan?

Are you aware of any differences in the way the English tests were delivered across different centres? And by different assessors?
Probe: comparison between ESB assessors and TE assessors

Are you aware of any contextual issues that should be taken into account when interpreting the learners' scores?
Probe: explore possible reasons for the higher English level of learners than expected.

Do you think that the conditions under which learners' were assessed may have affected their responses in any way?
Probe: facilitators and barriers to administering the test.
What went particularly well?
What could have been improved?

Overall would you say that learners were satisfied with the testing process?
If not, explore reasons why not.

The tool
How did you find the delivery of:
a. The pre-entry tests (reading, writing, speaking and listening)?
b. The Speaking and Listening tests for entry levels 1 and 2?

Do you think that the test was suitable for the type of learners being tested?
a. Was it able to capture potential differences in English level learners had across different areas (for instance, different Speaking and Reading levels)?
Prompt – in terms of level, structure of the test, content, across Speaking/Listening/Reading/Writing

If no: why not? How could this be improved in the future?

Were there any specific questions or parts of the test that you found harder to deliver?
Probe for any specific content that was problematic, what were the reasons for this, how could it have been improved

Are you aware if any specific groups of learners struggled more/less with the test?
Prompt – from a specific language group, age group, gender, ethnic background, etc.

What is your opinion on having Talk English tutors deliver the Reading and Writing tests for entry levels 1 and 2?
Prompt – explore again any differences with ESB delivery, any organisational issues, etc.

Do you feel that the score learners achieved matches their actual English level and ability?
Prompt – explore appropriateness of the content being tested and whether the majority of learners were pre-entry/entry1/entry 2.

Would you revise the test in any way or make any changes if it were to be used again for similar learners?

Thanks and close.

Teachers – Topic Guide (mid-point)

Introduction
Learning and Work Institute is an independent research organisation. We have been commissioned by MHCLG to undertake an evaluation of the Community-Based English Language course currently being trialled in Manchester and the North West.

Purpose of the interview
The interview will be used to better understand how the trial has worked and whether there are any areas that could be improved. The purpose of the interview is to enable us to learn about how TALK ENGLISH TOGETHER is being delivered in practice and identify any deviations from the intended delivery model as well as across sites. Drawing on the experiences for front line staff provides rich data on what is being delivered, what is and is not working, factors affecting delivery and views on improvements. It will also explore barriers and facilitators faced by individual participant groups

Confidentiality and consent
Anything that you tell us will be treated in the strictest of confidence, and your comments will remain anonymous. We would like to record the interview so that we can make sure we record what you say exactly, but you do not have to be recorded if you would prefer not to. Recordings will be kept securely, not accessible to anyone outside the research team and will be deleted at the end of the project.
The interview will last around 45 minutes. You are under no obligation to answer any questions, if you don't feel you can comment on specific questions don't worry. There are no wrong or right answers.

Check if participant is happy to proceed. Gain consent to record

Introduction and warm-up
Can you tell me what your job role is and how long you have been in it?

Need for ESOL
How long have you been working in this particular area (geographic)?
From your experience, what can you tell me about the need for community-based tuition? *Probe relative levels of need between population/community groups.*

Understanding of the Talk English Together Trial
Note for interviewee if needed: Briefing materials were circulated in advance of the trial starting; they were intended to give an overview of how the intervention worked and also to make it clear to those delivering the intervention what would be required of them while participating in a Randomised Controlled Trial (RCT).
What did you know about the Talk English Together Trial before the classes actually started?
Where did this information come from?
How useful was the briefing material?

Are there any contextual issues in this local area that you think affected the delivery of Talk English Together?

Prompt – the type of participants who have attended, attendance rates, the way the classes have been delivered, etc.

Baseline assessments

Note for the interviewee if needed: some baseline measures were collected while classes were taking place, while in other cases they were not. This section of the topic guide wants to understand for the centres that did hold lessons while collecting baseline measures, whether this affected the normal development of the class and what disruptions it caused.

Were any assessment measures collected while you were teaching a class?
IF NO: move to next section
IF YES: ask following questions.

How were the assessment events organised in your location?
How many rooms did you have available?
Were the two groups (those starting in April/May and those starting in September) seen separately or together?
How many staff members did you use as part of the baseline event?
What was your role, if any, in carrying out the assessments? Did you also carry out any Level 1/2 assessments? *Probe around observation of learners' English levels and appropriateness of the test.*
How did you decide which papers to send for marking by ESB?

Did this affect your ability to run the class normally?
If yes, how?

How do you think the assessment events affected the learners'?

Could anything have been done differently to support you and your learners through the baseline measures?
Previous preparation
What training or instructions did you receive on how to deliver the programme prior to the start of the trial?
And on how to use the manual?

Delivery

Overall, how appropriate is the level of structure to the Talk English Together course? *Explore why.*
What implications did the use of the manual have for you? *Probe whether the structure was helpful or not.*
And for the learner? *Explore whether more flexibility in the manual would have been better.*

To what extent have you followed the course manual?
Prompt: looking at the topics and order of modules covered.

Are there any elements of the manual that you have had to adapt or change?
Probe – to what extent they followed topics, ordering, and timings

Probe – can include unplanned adaptations such as classes overrunning, or including additional topics at the learners' request

If yes: what are the reasons behind this?

Could you please tell me a bit about the supplementary materials/resources you have used to teach your classes?

Explore the role of the teacher and their assistant volunteer(s).

What has been your average class size and composition? *Explore size, teacher/volunteer to student ratio.*
How do you think this has affected your ability to teach the class? And in terms of learning outcomes for the learners?
What have the demographics of the class been like? *Explore implications for class dynamics and mixing of learners.*

What has the attendance rate been like?
Why?
Is there any noticeable pattern in attendance/absence? What has the impact of Ramadan been?
What do you think are the facilitators and barriers to attendance? What helps maintain engagement?

Have any learners swapped classes; have they allowed any September or new learners to join? Have they brought family members or friends to classes?

What are your views on the length and intensity of the course?
Probe: explore any feedback received from learners on the duration and frequency of classes.

What teaching methods have you used?
Explore the split between individual learning, small-group learning, whole-class activities, repetition, role-play, homework; the interaction between the tutor and learners, and whether the tutor only spoke in English during classes.

Are there any parts of the class/topics that learners found particularly easy/difficult? And any that they found particularly useful?

How have the conversation club sessions been delivered?
Probe: explore the differences between classes and conversation clubs, how they benefit the learner, the role of the volunteer in delivery.

Have you invited in any guest speakers for particular sessions? And have you organised community visits? *Explore where they have taken learners, what they have done, the perceived impact.*

What were the main challenges or barriers experienced when teaching these classes? *Explore how they have been overcome.*

Have you found any differences between speaking, reading and writing abilities for learners?

Impact for learners

Has the impact on English proficiency been as you had anticipated?

Are you aware on ay wider impacts on learners as a result of being involved in the TET programme?
Prompt: on day-to-day activities, on their confidence/self-worth, social engagement, employability, etc.

Have there been negative or unexpected impacts on learners that you are aware of (or slower progress on some outcomes than expected)? *Probe for examples.*

Future delivery and closing comments

Which parts of the course do you think work well?

Which parts do you think do not work so well?
What would you change to improve the course?

Is there anything you would do differently if you could start classes all over again?

Ask if there's anything else they would like to add.

Thank and close.

Integrated Teacher and Volunteer Topic Guide

Introduction
Learning and Work Institute is an independent research organisation. We have been commissioned by MHCLG to undertake an evaluation of the Community-Based English Language course currently being trialled in Manchester and the North West.

Purpose of the interview
The interview will be used to better understand how the trial has worked and whether there are any areas that could be improved. The purpose of the interview is to enable us to learn about how TALK ENGLISH TOGETHER is being delivered in practice and identify any deviations from the intended delivery model as well as across sites. Drawing on the experiences for front line staff provides rich data on what is being delivered, what is and is not working, factors affecting delivery and views on improvements. It will also explore barriers and facilitators faced by individual participant groups

Confidentiality and consent
Anything that you tell us will be treated in the strictest of confidence, and your comments will remain anonymous. We would like to record the interview so that we can make sure we record what you say exactly, but you do not have to be recorded if you would prefer not to. Recordings will be kept securely, not accessible to anyone outside the research team and will be deleted at the end of the project.

The interview will last around 45 minutes. You are under no obligation to answer any questions, if you don't feel you can comment on specific questions don't worry. There are no wrong or right answers.

Check if participant is happy to proceed. Gain consent to record

Introduction and warm-up
Can you tell me a bit about your involvement in the TET trial?

Need for ESOL
How long have you been working in this particular area (geographic)?
From your experience, what can you tell me about the need for community-based tuition? *Probe relative levels of need between population/community groups.*

Understanding of the Talk English Together Trial
Note for interviewee if needed: Briefing materials were circulated in advance of the trial starting; they were intended to give an overview of how the intervention worked and also to make it clear to those delivering the intervention what would be required of them while participating in a Randomised Controlled Trial (RCT).
What did you know about the Talk English Together Trial before the classes actually started?
Where did this information come from?
How useful was the briefing material?

Are there any contextual issues in this local area that you think have affected the delivery of Talk English Together?

Prompt – the type of participants who have attended, attendance rates, the way the classes have been delivered, etc.

Baseline assessments

Note for the interviewee if needed: some baseline measures were collected while classes were taking place, while in other cases they were not. This section of the topic guide wants to understand for the centres that did hold lessons while collecting baseline measures, whether this affected the normal development of the class and what disruptions it caused.

Were any assessment measures collected while you were helping teach a class?
IF NO: move to next section
IF YES: ask following questions.

How were the assessment events organised in your location?
How many rooms did the centre have available?
Were the two groups (those starting in April/May and those starting in September) seen separately or together?
How many staff members did they use as part of the baseline event?
What was your role, if any, in carrying out the assessments? Did you also carry out any Level 1/2 assessments? *Probe around observation of learners' English levels and appropriateness of the test.*
How did you decide which papers to send for marking by ESB?

Did this affect the ability to run the class normally?
If yes, how?

How do you think the assessment events affected the learners'?

Could anything have been done differently to support you and your learners through the baseline measures?
Previous preparation
What training or instructions did you receive on how to assist the classes prior to the start of the trial?
And on how to use the manual?
Delivery

Overall, how appropriate do you think the level of structure to the Talk English Together course is? *Explore why.*
What implications did the use of the manual have for you? *Probe whether the structure was helpful or not.*
And for the learner? *Explore whether more flexibility in the manual would have been better.*
To what extent have the classes followed the course manual?
Prompt: looking at the topics and order of modules covered.

Are there any elements of the manual that have had to be adapted or changed?
Probe – to what extent they followed topics, ordering, and timings
Probe – can include unplanned adaptations such as classes overrunning, or including additional topics at the learners' request

If yes: what are the reasons behind this?

Could you please tell me a bit about the supplementary materials/resources you have used to teach your classes?

What has been your role in the classes?

What has been your average class size and composition? *Explore size, teacher/volunteer to student ratio.*
How do you think this has affected your ability to deliver the class? And in terms of learning outcomes for the learners?
What have the demographics of the class been like? *Explore implications for class dynamics and mixing of learners.*

What has the attendance rate been like?
Why?
Is there any noticeable pattern in attendance/absence? What has the impact of Ramadan been?
What do you think are the facilitators and barriers to attendance? What helps maintain engagement?

Have any learners swapped classes; have they allowed any September or new learners to join? Have they brought family members or friends to classes?

What are your views on the length and intensity of the course?
Probe: explore any feedback received from learners on the duration and frequency of classes.

What teaching methods have been used in the class?
Explore the split between individual learning, small-group learning, whole-class activities, repetition, role-play, homework; the interaction between the tutor and learners, and whether the tutor only spoke in English during classes.

Are there any parts of the class/topics that learners found particularly easy/difficult? And any that they found particularly useful?

How have the conversation club sessions been delivered?
Probe: explore the differences between classes and conversation clubs, how they benefit the learner, the role of the volunteer in delivery.

Have there been any guest speakers invited for particular sessions? And have you organised community visits? *Explore where they have taken learners, what they have done, the perceived impact.*

What were the main challenges or barriers experienced when teaching these classes? *Explore how they have been overcome.*

Have you found any differences between speaking, reading and writing abilities for learners?

Impact for learners

Has the impact on English proficiency been as you had anticipated?

Are you aware on ay wider impacts on learners as a result of being involved in the TET programme?
Prompt: on day-to-day activities, on their confidence/self-worth, social engagement, employability, etc.

Have there been negative or unexpected impacts on learners that you are aware of (or slower progress on some outcomes than expected)? *Probe for examples.*

Future delivery and closing comments

Which parts of the course do you think work well?

Which parts do you think do not work so well?
What would you change to improve the course?

Is there anything you would do differently if you could start classes all over again?

Ask if there's anything else they would like to add.

Thank and close.

Learners (treatment group) – Topic Guide

Introduction
Learning and Work Institute is an independent research organisation. We have been commissioned by MHCLG to undertake an evaluation of the Community-Based English Language course currently being trialled in Manchester and the North West.

Purpose of the interview
The interview will be used to better understand how the trial has worked and whether there are any areas that could be improved. The purpose of the interview is to enable us to learn about how CBEL is being delivered in practice and identify any deviations from the intended delivery model as well as across sites. Drawing on the experiences for learners provides rich data on what is being delivered, what is and is not working, factors affecting learning and views on improvements.

Confidentiality and consent
Anything that you tell us will be treated in the strictest of confidence, and your comments will remain anonymous. We would like to record the interview so that we can make sure we record what you say exactly, but you do not have to be recorded if you would prefer not to. Recordings will be kept securely, not accessible to anyone outside the research team and will be deleted at the end of the project.
The interview will last around 50 minutes. You are under no obligation to answer any questions, if you don't feel you can comment on specific questions don't worry. There are no wrong or right answers.
Check if participant is happy to proceed. Gain consent to record

Introduction and warm-up
Tell me a bit about yourself? *Where do you live? Who lives with you? How long have you lived in the UK? (If relevant - Where did live before coming to the UK?)*
Enrolment and registration
How did you hear about the Talk English course?
Why did you join the Talk English course? *What did you want to get from it? Explore motivations for learning English*

How did you feel about the registration events and tests? *Explore whether it changed their opinion of the course; was overly burdensome; affected their first lesson etc.*

What do you think about the English tests and survey? *Explore what was easy or difficult to understand; whether the conditions/timing/waiting may have affected their responses*

Did you sign up with any friends or family?
If yes: did you attend the same classes together?
If no: did you share materials with friends waiting for the September classes?

About the learning course
How often did you attend a session?
If not that often: ask why they missed sessions, and what would have made attendance easier (e.g. travel, childcare, timing, impact of Ramadan)

What do you think about the length of the course?
Probe: Frequency of sessions (3 per week) and overall length (11 weeks); would they have liked to spend longer on certain topics?
Would you suggest any changes? Why?

What activities did you do in your sessions with the tutor?
What activities did you find most useful? Why?
Explore:
General experience of the course (what was useful/interesting)
Classes (formal tuition)
Conversation clubs
Community visits – practising skills in community; learning about local facilities

What topics did you find most interesting and helpful?

Probe: visiting the GP, shopping, democracy
Suggest probing in reference to topics specified in manual (e.g., visiting the GP; talking to teachers; shopping; etc.)
Was there anything you would have liked to learn about, that wasn't covered?

What did you find least interesting? Why?

What was the size of your class? *Explore whether they would have preferred to be in a bigger or smaller class*
Were people in your class similar or different to you? How did that affect your experience on the course? *Explore: for men, how they felt being in the minority; for women, how they felt if a man was present in the course; how the learners interacted and related to each other.*
Can you tell me a bit about your teacher? *Explore how they found working with the teachers, volunteers, other learners.*
What did you think about the place/venue/centre you had your classes in? *Explore learner opinion on the pros/cons of the venue, e.g., church, school, library, children's centre.*
How far did you have to travel to attend the course? *Explore mode of transport (walked; drove by family member; bus) and what impact this had*.

Impact on participant

What did you learn?
What difference has being on this course made to you?
Prompts: gained skills, feel more/less confident in skills, made friends, increased/decreased self-confidence, feel more motivated to learn, knowledge of local area and services, knowledge of further training.
How has the course changed how you talk to people in English?

Do you talk to more people outside of your home? Who/How often?
Are you more or less confident when talking to people outside of your home?
Do you speak to people in English more often?
How confident are you when speaking to people in English?

Has this changed from before you started the course?
What difference, has being on this course made to how you talk to or mix with your neighbours and other people that live near you?
Do you feel more/less involved in your community? e.g. Do you do any community activities now which you didn't do before?
Do you use local services (e.g. library, community centre) more/less?
What impacts have these changes had for your family or friends?
Prompt: more engaged in children's education; encouraged others to join an English class; able to help others more; reduced burden on family due to increased independence…
Did you do any other activities that were not part of the course to help you improve your English? What?
Probe: asking family/children who can speak English for help with homework, practice speaking English with family/friends, meeting up with learners outside class etc.
Overall reflections and close
Thinking about all the changes we've just talked about, which is the most important to you?
a. Which changes have made the biggest difference to your life?
b. Did you experience any difficulties because of the course? *Probe: Explore any negative outcomes experienced, e.g., did they feel overwhelmed, anxious, etc. due to any part of the course; any difficulties juggling attendance with family life*
c. Explore anything they hoped to learn or achieve but didn't.
d. Was there any activity or help you would have liked to have received from the course but didn't? *Probe: Explore changes they'd make to the course; explore potential benefits of other offers for the learner, such as childcare, travel, access to materials, etc.*
What are your next steps after learning English here?
Prompt: gained new work-related skills, increased confidence to look for a job, know more about how to behave in the workplace etc.; enrol in further English language courses; sign-up to other activities (computer skills; textiles; social clubs); volunteering
If offered to you, would you take part again in a similar English course? Would you recommend the course to others / what would you say to someone thinking about doing the course
Is there anything else you would like to say about the course?

Thank and close.

Learners (treatment group) – Topic Guide (VERSION 2)

Introduction
Learning and Work Institute is an independent research organisation. We have been commissioned by MHCLG to undertake an evaluation of the Community-Based English Language course currently being trialled in Manchester and the North West.

Purpose of the interview
The interview will be used to better understand how the trial has worked and whether there are any areas that could be improved. The purpose of the interview is to enable us to learn about how CBEL is being delivered in practice and identify any deviations from the intended delivery model as well as across sites. Drawing on the experiences for learners provides rich data on what is being delivered, what is and is not working, factors affecting learning and views on improvements.

Confidentiality and consent
Anything that you tell us will be treated in the strictest of confidence, and your comments will remain anonymous. We would like to record the interview so that we can make sure we record what you say exactly, but you do not have to be recorded if you would prefer not to. Recordings will be kept securely, not accessible to anyone outside the research team and will be deleted at the end of the project.
The interview will last around 50 minutes. You are under no obligation to answer any questions, if you don't feel you can comment on specific questions don't worry. There are no wrong or right answers.

Check if participant is happy to proceed. Gain consent to record

About the Learner

Aim: Warm-up interviewee by asking background questions and explore motivations for wanting to learn English

Can you tell me a bit about yourself?

Probe for personal and household circumstances, e.g.:

Where Learner lives / with who

What do they usually do on a day-to-day basis

How long in the country and local area

Who (i.e. friends or family) are you attending course with?

Why are you interested to learn English? Probe:

What does the Learner want to achieve by learning English

Identify any support / encouragement they received outside of the course (inc. influence of other Learners known to interviewee)

Identify any barriers discouraging the Learner from learning English in the past or currently, and how overcome

Learner´s perceived performance

Aim: Understand how far the Leaner has come in their learning journey, identify what has contributed to this and what more the Learner thinks could have been done to achieve their goals

Thinking about why you wanted to learn English, what progress have you made since joining the course. Probe:

Examples of improvement in English e.g. using English more frequently, more confidence in using English in particular contexts (talking to children, teachers, GP´s)

What have you not yet achieved. Probe:

Reasons / factors for not achieving goals *("Why do you think you've not achieved this" / "What would you need to do to achieve this"?)*

Getting the Learner to reflect on everything they have achieved, how much of this is accountable to the Talk English Course. Probe:

Identify other influences if not completely attributable to TET

About the course

Aim: Understand Learners experience and understanding of the course and satisfaction with each stage of support

Can you describe the course to me? Probe:

How did they hear about it and their recollection of their first contact (recruitment events)

The classes and activities they have done (*"What sort of things have you done as part of the course?")*

Things done outside of the course, but connected to it e.g. social activities outside of course with staff and other Learners

Course duration and number of classes per week (with reference to Learner's own attendance)

What is the Learner experience of each element *(noted above).* Probe:

What did they find useful

What was less useful

What would they have liked to do more of

How would you describe the other people on the course? *Probe*:

Relationship with of other Learners, including who Learner is attending course with (*"How did you get on with the other people in your class" / "Were there people in the class that you were more or less comfortable to work with; why"?*)

Can you describe the people who were teaching and helping you on the course?
Probe:
Learner views about teaching and volunteer staff
What did each staff member do with the Learner

Overall view of the course
Aim: Summary of Learners view of the Talk English Together course
Explore satisfaction with the course
Explore satisfaction with what Learner has achieved on the course
Any other views Learner wishes to add

Annex C

Delivery model

Talk English Together Structure