



**realising
ambition**

The trials and tribulations of Realising Ambition

Doing randomised controlled trials in the real world

**Realising Ambition
Programme Insights: Issue 6**

About this series: This series of Programme Insights shares reflections, learning and practical implications from Realising Ambition, a £25m Big Lottery Fund programme supporting the replication of evidence-based and promising services designed to improve outcomes for children and young people and prevent them from entering the youth justice system.

Rather than writing a long evaluation report at the end of the five-year programme, we are producing a series of short Programme Insights so people get information about the programme while it is happening.

Some issues, like this one, are **focus pieces** that present ideas and concepts emerging from the programme. Others are **findings pieces**, describing preliminary data and learning from the evaluation activities, and their implications. Our **field guides** are practical 'how to' guides. Throughout each issue, some words are highlighted in blue. For these you will find definitions in the Glossary of Terms box at the end of this piece.

About us: The Realising Ambition programme is supporting and is powered by [22 organisations](#) replicating 25 different services all over the UK. The programme is managed by a consortium of four organisations committed to improving outcomes for children. It is led by [Catch22](#), alongside the [Dartington Social Research Unit](#), [Substance](#) and [The Young Foundation](#).



Realising Ambition Programme Insights: Issue 6

Introduction and recap

Five years ago there had been relatively few [randomised controlled trials](#) (RCTs) of home-grown social interventions for children and young people in the UK. In that context, Big Lottery took the bold step of investing in four real-world RCTs as part of Realising Ambition. Not because we or others think that RCTs are the only or even best method of evaluation in all circumstances: we don't. But because when it comes to testing the impact of an intervention on outcomes they do a good job of helping us to attribute cause by filtering out other possible explanations for any impact observed.

Realising Ambition delivery organisations were invited to apply as part of a competitive process to be part of an RCT. Four projects were selected for the RCTs: two mentoring programmes for young people (provided by [YMCA Scotland](#) and [Chance UK](#) respectively), one therapeutic parenting programme for parents/carers of children with behavioural or emotional difficulties ([Malachi Trust](#)) and one school-based programme to prevent psychologically abusive and controlling behaviour in teenage romantic relationships ([Ariel Trust](#)). Among the factors shaping this decision were the need to focus on promising as opposed to proven programmes, the perceived readiness of the respective organisations for such an evaluation and the desire to have a mixed portfolio in terms of types of intervention and representation from different parts of the UK.

About the RCT projects

Chance UK: Early Intervention Mentoring

Chance UK's mentoring programme is for children aged 5-11 years with challenging behaviour and emotional problems at school and at home. It is designed to prevent future anti-social and criminal behaviour. A volunteer mentor is paired with the child and applies solution-focused techniques in weekly 2-4 hour sessions over one year. Child-centred goals are set for the year and mentors develop a programme of interactive activities that are tailored to the child's interests and needs. A graduation ceremony attended by family and friends marks the end of the year. Optionally, parents/carers can receive support aimed at maintaining positive changes in the child's behaviour and stability for the family once the mentoring ends.

Ariel Trust: Face Up

Face Up is a universal, school-based multi-media prevention programme for children aged 11-16 years, developed by the Ariel Trust. It seeks to promote healthy adolescent dating relationships and prevent psychologically abusive behaviour (particularly of a controlling or coercive nature). The eight one-hour lessons are designed to be delivered in personal, social and health education. They cover knowledge about what is appropriate/abusive behaviour and sources of help, and skills in problem-solving, asking for help and intervening as a bystander. A whole-school element includes documents for school staff about procedures in case of a disclosure and a poster campaign to reinforce messages from the class lessons. Schools also receive a letter to send to parents describing the programme and relevant resources.

About the RCT projects continued...

Malachi Trust: Inspiring Futures

The Malachi Trust's Inspiring Futures programme is for parents of children aged 6-11 years with behavioural and emotional difficulties. It comprises two parts, delivered sequentially: (1) a group-based element for all parents (10-12 weekly 90-minute sessions), and (2) individual one-hour sessions over 6-12 weeks with selected parents. The group sessions first identify parents' early adverse experiences and raise awareness of how these can influence current behaviour, notably parenting. A combination of child development education and solution-focused therapy is used to improve parenting skills. A key underlying mechanism related to parenting skills that the programme attempts to address is the empathy parents feel towards their child.

YMCA Scotland: Plusone Mentoring

YMCA Scotland's Plusone programme works with young people aged 8-14 years. It aims to reduce offending and aggressive and anti-social behaviour, increase engagement with school and realise participants' potential. A trained volunteer mentor from the community develops a positive and trusting relationship over one year with a matched young person. The mentor provides a positive role model, introduces the young person to new activities, encourages them to build confidence and develop new skills, and seeks to help build the young person's understanding of the link between their behaviour, the decisions they make and consequences. Programme managers provide advocacy and broker support with families.

In this Programme Insight we take a frank look at the process of conducting real-world trials – real-world in the sense of how the interventions were developed (not in research centres or universities, but by practitioners and managers), how they are delivered (not by specially-trained researchers or clinicians, but by regular practitioners and volunteers) and how they are evaluated (not in laboratory-like conditions, but in environments hampered by all the constraints of everyday practice). We draw on our own reflections and those of leaders in Chance UK, Malachi, Ariel and YMCA Scotland.¹ We highlight the challenges encountered and how we and the respective organisations sought to address them. In doing so, we seek to highlight lessons for funders, researchers and intervention developers and providers.

We have structured the Programme Insight so that the first part focuses broadly on things to get right before the trial starts and the second part looks at things to get right during the trial (inevitably the line between the two is somewhat blurred). We also present some interim findings related to the process and implementation of the trials. As the trials are still underway, data on outcomes will be published in due course. There are strong echoes throughout of the processes and disciplines we have advocated throughout the lifetime of Realising Ambition, including: whole organisational buy-in; refining and clearly defining the intervention; ensuring fidelity; and taking time to reflect on data and learn and adapt accordingly. These approaches are therefore not exclusive to the RCT projects, although the RCTs have arguably demanded that the participating organisations undertake this activity more robustly and under the glare of the evaluator's spotlight. As will become clear, one of the main messages is that this has helped those organisations to improve not only the RCT interventions but also the quality of their service design and delivery more widely.

Things to get right before the trial starts²

Buy-in

Randomised controlled trials are invariably demanding for all involved, not least because for managers and practitioners they can create extra work and introduce unwanted complexities. Efforts were therefore made to ensure that the delivery organisations were fully equipped with the information they needed to decide if an RCT was right for them. The lesson we learnt is that securing organisational buy-in to a trial requires understanding an organisation's culture and processes, and securing buy-in in the right places. Clearly the person in overall charge needs to agree, but if several people are in charge, or at least have delegated responsibilities, this agreement is necessary but not sufficient. It is essential to complete a site readiness assessment early on, to ensure that all key parties are engaged in the process as early as possible and to produce a terms of reference document that spells out roles and responsibilities for the provider and the research organisations so that everybody is clear about practicalities and expectations.

Getting Buy-in

In the case of Chance UK and Malachi, we presented to the whole staff group on evidence, trials and early plans for the respective RCTs, and gave everyone the opportunity to clarify any issues or concerns they had. This led to open and honest discussions about the benefits and drawbacks of doing a trial and seemed to help allay project staff's fears and build trust, ownership and a sense that we were in it together. Even then, since the recruitment of participants was via schools, there was a need for managers to engage head teachers, who in turn required individual teachers to talk to parents/carers, demonstrating that buy-in extends beyond delivery organisations to the wider professional community.

In the case of YMCA Scotland, the CEO and other senior staff were very supportive and enthusiastic. However, the intervention was based on a social franchise model, so the RCT needed approval from franchise partners, many of whom felt the random allocation of young people to control and treatment groups was deeply unethical. This, combined with issues about the number of participants needed to make the trial meaningful, contributed to the decision not to proceed with a trial. Instead, efforts were concentrated on further refining the intervention logic model and completing a [break-even analysis](#).

“Securing organisational buy-in to a trial requires understanding an organisation's culture, processes and structure.”

Intervention design

In an RCT it is important to be clear about what the intervention comprises, whom it is for, what it is seeking to achieve, and why, in theory, it should have the desired impact. In the standards of evidence that underpin Realising Ambition, this is called **intervention specificity**. Without this it is hard to make an ethical case for doing a trial, let alone to devise a suitable method.

As part of work to improve intervention specificity, we have found that investing significant time in developing or refining the **logic model** pays long-term dividends. It makes the intervention easier to evaluate – because it is clearer what needs to be measured – and it helps the people who deliver it to deliver it better: it is only human that when we know why we are doing something we are more likely to do it.

In each of the Realising Ambition trials, we met with the delivery teams, discussed what their intervention comprised and what it sought to achieve, and consulted the evidence on how relevant problems with child behaviour and emotions develop and what can be done to prevent or address them. We also helped the organisations to sharpen-up aspects of the intervention description – including re-formulating their respective logic models.³ They all reported that this was helpful. This was partly because it helped to develop a better understanding of the intervention in question and highlighted weaknesses that could be addressed. The process also helped to improve intervention materials and training, and yielded wider benefits for the organisations and other interventions in their portfolios.

Intervention refinement

The logic model work with Ariel Trust led to some redesign of the Face Up intervention, including a reformulation of lesson materials to include more emphasis on skills development. For this reason, we undertook a **feasibility study** of the redeveloped version of the intervention. The results of that, and research on what is effective in school-based interventions,⁴ yielded further changes, including the addition of whole school and home-based elements. Rather than proceed to a full trial, we decided to conduct a pilot RCT with an in-built feasibility study. This reflects the fact that if an intervention is still in development it should be tested for the feasibility of its implementation prior to testing in a main trial.⁵

Size of the trial

In order to generate meaningful results, our experience is that trials invariably need to have considerably more participants than the providers expect initially. Psychosocial prevention and early intervention programmes tend to produce small-to-modest effects. Smaller effects require larger sample sizes to detect them.

The size of the sample is largely determined by the minimum **effect size** it is worth detecting, which should therefore be worked out as early as possible. The task of doing this is often based on clinical judgement and a good understanding of the primary measure that will be used to assess change in outcomes. There is a balance to be struck between trying to find an effect so small that it is of little practical value, and one that is so large that, while desirable, it is unlikely to be achieved. With the former there is a danger of recruiting lots of people to no end, while the latter risks setting the intervention up to fail.

It is also necessary to take into account expected **attrition** during the trial. Some **participants** will drop out of the trial as it progresses, meaning that we cannot obtain data from them, either because they refuse to provide it (in some cases formally indicating their desire to stop being part of the trial) or go off radar (ie prove difficult or impossible to contact).⁶ Consequently, the number of participants who sign-up for the RCT must be larger than the actual number needed in the analysis.

The issue of numbers posed a particular challenge for the Realising Ambition trials because decisions about delivery numbers were necessarily taken prior to the programmes in question being selected for the trials, let alone any calculation of suitable sample sizes having been made. In some respects we were trying to retrofit trials into established service provision arrangements rather than planning them prospectively.

Working out trial size

In the case of Chance UK, a [meta-analysis](#) of mentoring programmes found a small average [effect size](#) of 0.2 for young people's behavioural and emotional outcomes.⁷ However, it is plausible that Chance UK's intervention will produce a larger effect size because it demonstrates many of the features of more effective mentoring programmes. For example: it only serves children with an elevated level of need; volunteer mentors are highly trained to deliver a tailored programme of structured activities; a thorough matching process is designed to create successful matches; the sessions take place weekly for one year; and parents/carers are offered support as part of the programme. Additionally, it works with 5-11 year-olds, whose behaviour may be more malleable than that of adolescents (the usual target group for mentoring programmes).

In order to detect a small effect size of 0.2, Chance UK would have needed to recruit 944 eligible children and match 472 with mentors. This was impossible with the time and resource available - their capacity allowed a maximum study size of 132 participants in the intervention arm. We calculated that this permitted the detection of a slightly larger - but plausible - effect size of 0.4, meaning that 264 eligible children needed to be recruited.

Preparing for the results

It would make for an easy life if some interventions were unequivocally effective and others equally obviously ineffective. While this is sometimes the case, reality tends to be more complex: some things work for some people but not others, or in some circumstances but not others, or according to some measures but not others. It depends on a host of factors and is often far from clear-cut. It is not helped by the tendency sometimes for evaluators to cherry-pick the positive results and gloss over the others.

In the Realising Ambition trials we want to report faithfully on what we find, and to be open about reasons for the results, whatever they are.

Our starting point, then, is true uncertainty: when the trials started we could confidently say that we didn't know for certain whether the interventions being tested were superior to business as usual. This is referred to as the principle of [equipoise](#).

When trials find null or negative effects, it is very easy to blame the study (eg "We had the wrong measures") or blame programme implementation (eg "It was done badly") rather than accept the result. When results are positive, people tend to worry far less about measures or implementation. We explain away what we want to pretend didn't happen, and accept uncritically what we are glad did happen. The bias in this stance is obvious.

We think that at least three things can be done about this problem. The first is to pre-empt it by articulating before the trial starts what the possible unintended adverse effects of the intervention might be - the so-called dark logic - and how likely these are to occur.⁸ This can inform the intervention design, either in the form of modifying ideas or implementing mitigating actions, and it can shape what is measured (ie it may be appropriate to measure outcomes hypothesised to get worse as well as those predicted to improve).

Second, during the trial - ie ahead of the results - it is worth reflecting on how well things are going. Is implementation good? If not, why not? Can anything be done about it? For example, can drop-out be reduced or fidelity strengthened? Is everyone still happy with the measures and how the evaluation is being executed, or are there problems that need to be resolved, such as [contamination](#) or the timing of measures?

Third, it is helpful to think about what actions will be taken depending on what the results show. What does it mean if the results are positive (eg "We'll recommend it for wider implementation")? What is the message if the results show no effect or demonstrate that the intervention is harmful (eg "Stop doing the intervention and revise or abandon", or "Do further analysis")? What if the results are mixed (eg "Revise and test further" or "We recommend it but only in some circumstances")?

What if...?

Although we didn't refer to it as dark logic, in our logic model workshops with the Realising Ambition RCT projects we discussed potential adverse effects of the interventions and possible negative relationships between components of the model. As the trials have progressed, we have regularly reviewed with all three provider organisations progress on programme implementation and the evaluation and sought to troubleshoot difficulties. Chance UK and Malachi look at their fidelity data regularly and feed back to staff on how implementation can be improved. We also held a workshop with the respective organisations' leaders to consider factors – political, economic, cultural, organisational – that they consider might affect the success of their respective interventions and to scope out different scenarios depending on what the results show.

in a trial may require tightening up or fundamentally changing established [recruitment](#) procedures, which can reduce the flow of potentially eligible participants. Fourth, trials often require double the usual numbers (for a 1:1 random allocation ratio): if there are 100 spaces to fill, 200 participants are needed.

Unsurprisingly, both of these problems raised their heads in the Realising Ambition trials. We sought to use research techniques to help the relevant organisations to address them, with some success, and these approaches have arguably helped the respective organisations to take a more evidence-based approach to recruiting users for their services. Nevertheless, there is no escaping the fact that an incredible amount of hard work was still needed on the ground to engage and retain participants.

Things to get right during the trial

Recruitment and retention

When a trial starts, it is not uncommon to find that the large number of expected participants seems to disappear. Calculations have suggested that the pool of eligible and therefore potential service users exists, and providers have given assurances that they regularly encounter service users fulfilling the relevant target group criteria. Yet once the starting gun is fired, triallists are frequently found fretting about not having enough recruits and the delays this creates. To add to their woes, they find that participants drop out of the study from day one and can be elusive.

The tendency to overestimate actual numbers happens for several reasons. First is the flawed assumption that since there are lots of children who are eligible they will translate into actual cases. The problem is that eligibility does not equal demand: experience suggests that it is hard work to engage eligible participants. The second reason is the failure to appreciate that not all current users meet the target group criteria. Unfortunately, demand does not equal eligibility: actual service users may not have serious enough needs to qualify.⁹ Third, the need for greater specificity

Examples of recruitment and retention strategies

In the case of Malachi, we used a cohort design, meaning that families were recruited in six batches over more than a year. This reduced the number needed at any one point and spread recruitment out. We also estimated the number of children in each school whose parents/carers were eligible for the programme. We did this using a combination of official data on the sizes of schools in the relevant geographical areas and data on the proportion of children in the relevant age group (17%) known nationally to reach the eligibility threshold on the Strengths and Difficulties Questionnaire.¹⁰ Next, based on Malachi's experience and data from earlier cohorts in the trial, including on attrition, we estimated how many of these children's parents/carers would need to attend **recruitment** events. This helped ensure that at least four – the lowest viable parent/carer group size – could be randomly allocated to the intervention.¹¹

We used a traffic light system to distinguish between schools that would be likely to have enough eligible children to run their own parenting group, schools that would need to combine with another nearby school to make a group viable and schools that would struggle to run a group either by themselves or when combined with a nearby school. We also tried to communicate this information in a meaningful way. For example, we produced pen-portraits describing the kind of

children likely to be suitable for the programme and suggested that teachers consider referring the five children in their class whose behaviour and/or emotional well-being they were most worried about (five children out of a class of 30 being about 17%).

In order to boost numbers, Malachi also adapted their recruitment strategy several times. The original strategy involved placing adverts in school newsletters inviting parents/carers to coffee mornings. Parents/carers who attended were then asked to complete a questionnaire to establish whether or not they fitted the eligibility criteria. When it became clear that, due to attrition, this would not secure the number or type of parents needed, Malachi adopted a more targeted approach, asking schools to identify the families, then sending them a letter directly. They then tried phoning parents/carers instead of writing to them; the most successful method.

None of this was an easy ride. The need to apply the Strengths and Difficulties Questionnaire (SDQ) meant that established methods of identifying suitable parents/carers, such as at the end of a drama project for an entire year group, were changed. Relationships with schools were also threatened because Malachi had to inform schools that only half of the parents/carers schools identified would actually receive the intervention.

“ Malachi adopted a targeted approach, asking schools to identify the families, then sending them a letter directly. They then tried phoning parents/carers. ”

As indicated earlier, part of the numbers game is about keeping hold of trial participants once they have started (or **retention** in research language). There are few secrets here: in our experience (of these as well as other trials) it requires a lot of hard work, persistence and courtesy. Some data collectors are better than others at engaging people and can provide useful tips, but ultimately there is a strong relational – and therefore personal – element. We have also found that working closely with the provider organisation helps – both Chance UK and Malachi have assisted with our efforts to collect data, whether by encouraging teachers to respond to requests or providing updated contact details. (A significant proportion of participants are fairly mobile and/or change their telephone numbers frequently.) There was some value to keeping data collectors consistent between time points, so that a family would know who was getting in touch with them about meeting. Lastly, in order to help participants – especially those in the control group – feel part of the project, we branded the trials¹² and sent newsletters to parents/carers and birthday cards to children.

A final point in this section is that problems with recruitment and retention can be pre-empted to an extent: pilot studies are useful to test the trial in miniature and check estimated parameters like recruitment and retention rates.¹³ This is what we are doing with the Ariel Trust, which we hope will pave the way for a full trial in the future. For instance, together we discovered that engaging with the Healthy Schools Network in a given city was the most efficient and fruitful means of contacting and signing up schools. We also found that many schools were concerned about the time commitment required by the programme and found it easier to deliver lessons in smaller chunks. This suggests that presenting it as a collection of 15-minute activities might make it easier to recruit schools in the first instance.

Fidelity monitoring

Implementation **fidelity** refers to the extent to which an intervention is delivered as it was originally intended by the programme developers. The word fidelity conveys the sense of being true or faithful to the design.

Understandably, implementers are often too busy implementing an intervention to record what they are doing. They have an intuitive sense of what is happening, and that is enough. It is different in a trial. The evaluator and provider need to know what is actually being delivered – as opposed to what is in the manual – because, without a good sense of this, it is hard to interpret the results. This is particularly the case if the results are disappointing and suggest that the intervention was ineffective. Was it because the intervention wasn't delivered as intended?

A by-product of collecting fidelity data is that it can encourage practitioners to deliver the intervention with greater fidelity – a good example of how the requirements of an RCT can improve intervention delivery. For example, a review of meta-analyses of whole-school anti-bullying programmes and mentoring programmes found that the programmes that monitored fidelity achieved up to three times the level of impact as those that did not.¹⁴

For these reasons, fidelity is being monitored in each of the Realising Ambition trials.¹⁵ A range of methods are being used, including practitioner self-report, in vivo (live) observation, user feedback and the coding of videoed sessions. We have tried to make the fidelity monitoring tools as short as possible and beneficial for practitioners and the respective delivery organisation.

The delivery organisations have acknowledged that this has helped them to develop a better understanding of what gets delivered on the ground and, accordingly, how to make improvements where necessary.

“ A by-product of collecting fidelity data is that it can encourage practitioners to deliver the intervention with greater fidelity. ”

Examples of fidelity monitoring

Chance UK understood that for the RCT they needed to make sure that everybody delivered the non-negotiable building blocks, and that mentors did what they were trained to do, such as spending the right amount of time with young people and using a solution-focused approach. As well as being important for the evaluation, they reasoned that if they discovered that fidelity was weak then they could do something about it.

We worked with Chance UK to design the fidelity measures and work out how to integrate them into employees' day-to-day activities. In order to minimise the burden on staff, we sought to let Chance UK take ownership over what the fidelity measures would look like and involved service delivery staff to ensure that there was proper consideration of how measures would be integrated into regular practice. We made recommendations, but did not force Chance UK to adopt measures that were not useful to them.

Before the trial, Chance UK had less in place to check on fidelity. By their own admission, much of it was done based on gut feeling and trust. They report finding it useful to have numbers to confirm or contradict their preconceptions. Indeed, they have integrated fidelity measures into their wider portfolio of services. They say that they now have better insight into the detailed work that is taking place with beneficiaries, including whether mentors are using the right tools and techniques and how children feel about their mentors.

The Ariel Trust's interventions are different insofar as they are delivered at arms length by regular teachers in schools. This permits large scale replication, particularly since the interventions are software-based (eg interactive websites and DVDs). Before the trial, Ariel knew

that when they presented their products to teachers, the teachers were really keen to use them. They were less clear, however, about what actually happened when the lessons were taught in classrooms.

As part of the feasibility study and pilot RCT, we have conducted extensive fidelity monitoring. We asked teachers to complete data collection forms, observed the teacher training and observed selected classes where Face Up was being taught. Obtaining fidelity data from teachers has been very challenging. Nevertheless, collectively the data showed that teachers were tending to miss out key lesson elements, notably the role play. Instead, teachers would reach the end of the DVD then allow the class to discuss the issues raised. This was a significant omission because this activity enables young people to practise what they have learnt and is therefore considered to be crucial to overall impact. The fidelity monitoring also showed that the problem might lie in the light touch nature of the teacher training.

As a result of this analysis, Ariel redesigned the teacher training for Face Up, making it longer and more evidence-focused. They report that teachers are more on board with the entire programme now as they are told exactly why they have to do each section. That said, getting all the teachers involved to attend the training is difficult if not impossible, and there is a concern that the in-school dissemination of training may be insufficient. In the second cohort of schools in the pilot, therefore, Ariel gave the lead teacher training packs for all teachers, provided contact details for additional support, suggested that they use the training slides to disseminate the training and directed them to the online manual to encourage the comprehensive dissemination of training.

What the control group gets

In a trial, the services received by the control group affect the results as much as the intervention being tested. There is growing appreciation of this, particularly since the relative strength of provision for the control group has been suggested as a reason for the failure of some imported US programmes to replicate their effects in Europe. It is therefore important to record what the control group receives. This means going beyond saying “business as usual” and instead actually detailing what this means in practice. It is especially important to try to capture three things that effectively cause business as usual to cease being just that: whether control group participants receive any part of the intervention that is being tested (contamination); whether control group participants receive compensation for being in the control group (and, if so, what); and any innovation in standard service delivery – for instance, doing something special to address participants’ needs.

Why are these things important? As indicated already, when trial results are null or negative it is common to say that it is because the control group received a good service. Sometimes this is because the service is good generally. Sometimes it is because the sites had to volunteer to be part of the trial and so are arguably disproportionately motivated to do a good job in the area of interest. Sometimes service providers or members of the control group compensate for their perceived loss (of the intervention being tested). This may happen accidentally – for example, someone who is delivering the programme to the intervention group has some contact with the control group and, without thinking, applies to them some of the same ideas and techniques – but it can also be deliberate – for instance, someone delivering the intervention feels unhappy about an individual being allocated to the control group and subverts the system by providing them with the intervention.

Contamination and control

In the Realising Ambition trials we have tried hard to avoid contamination. For example, in the Ariel trial, where control schools can implement Face Up with another year group, we have specified that teachers who deliver Face Up do not teach children in Year 9 (the control group students). With Malachi we were careful to say that control parents/carers should not receive other Malachi services, because some of them apply the same philosophy and techniques as the parenting programme we are testing.

In order to be able to make sense of the comparisons between the intervention and control arms, we have also sought to find out what the respective control groups actually receive. For instance, we have asked the control schools in the Ariel trial to tell us what they are doing in the area of relationships education, and asked parents/carers in the Chance UK and Malachi trials what other services they and their children have been receiving.

“ In a trial, the services received by the control group affect the results as much as the intervention being tested. ”

Data collection alignment

The aim in an RCT is to compare distance travelled on the outcome(s) of interest in the group getting the intervention and the control group. As such, measures are taken before the intervention commences and after it finishes in the intervention group, and at comparable times in the control group. In an ideal world the following would generally hold: (i) the **baseline** measurement would precede the start of the intervention by the smallest margin possible; (ii) the **follow-up** would be after the intervention has been completed; and (iii) the period between these time points would be the same for each participant (whether they are in the intervention or control group). In order to help achieve the second and third of these, the timing of follow-up data collection points is usually set in relation to when baseline data were collected or the time when participants were randomly allocated to the intervention or control arms of the trial.

Of course, achieving this ideal is easier said than done. First, it is often difficult to collect data on a designated date. This tends to be because the participant or data collector is unavailable. Second, delays at any stage in the process, for example between baseline data collection and the intervention starting, can mean that the scheduled follow-up data collection point falls – for the intervention group – in the middle of them receiving the intervention. Third, where cases are randomly allocated in a block, rather than dynamically,¹⁶ there is usually a delay for some participants between baseline and randomisation, meaning in turn that the time elapsing between baseline and follow-up data collection will vary for each participant. In short, it is hard to align the timing of all data collection elements as specified in the trial **protocol** while still fitting in with operational requirements.

These challenges have certainly been encountered in the Realising Ambition trials. For example, with Face Up, schools have sometimes found it difficult to get all pupils to complete the online surveys in a short space of time because of how lessons are timetabled and the time it takes to get pupils to the IT room and set them up. Similar issues have arisen in the Chance UK and Malachi trials – for instance, because a parent forgets an appointment, or a teacher data collection point for a given child falls in the school holidays.

In response, our approach has been threefold: first, we set out clearly at the outset what we intended to do; second, we have permitted flex within certain boundaries; and third, we have recorded when data are collected so that variations can be taken into account in the analysis. As such, we have sought to strike a balance between rigour and pragmatism. For example, having identified the target date for follow-up data collection based on a set period of time from baseline or randomisation, we have built a data collection window around that date (eg two weeks before, two weeks after). If the target date falls in a school holiday and the source of the data is a teacher, we have allowed teachers to complete data collection before or after the summer holiday.¹⁷ Where there is substantial variation in the time between baseline and randomisation, we have changed the anchor point that determines the follow-up data collection point from baseline to randomisation.¹⁸ Finally, when six months or more elapses between baseline and randomisation, we have re-baselined the participants concerned. This is to ensure that changes in their situation or condition have not affected their eligibility for the intervention.

Interim findings

At the time of writing, we are still collecting data from RCT projects. We plan to analyse and report findings from the trials as soon as possible (in mid-2017). In the meantime, we share some early implementation-related findings.

Chance UK

Chance UK received 326 referrals between May 2014 and December 2015 from schools in Enfield, Hackney, Islington, Lambeth, and Waltham Forest. Following the exclusion of cases that were found to be ineligible (9%) or declined to take part (16%), 246 were randomised: half to the intervention group and half to the control group. Mid-point and post-intervention follow-up on outcomes are scheduled to take place nine months and 16 months after randomisation. This timeline is designed to accommodate up to three months for the matching process and one year of mentoring for the intervention group. Data are collected from parents, teachers, and the child too if they were aged 8-11 years at baseline. Mid-point data collection is due to be completed in December 2016 and the end-point data collection is due to be completed by July 2017. Levels of missing data currently stand at 35% and 22% for teachers and parents respectively at the mid-point, with 4% of data still to be collected. For the end-point there is currently 21% and 14% missing data for teachers and parents respectively, with 38% of data still to be collected. We have sought to minimise attrition by offering an incentive to parents, sending regular newsletters and engaging Chance UK and schools to help us make contact with parents when phone calls and letters from the research team have not been successful. Chance UK also assist in identifying an appropriate school staff member, for example when a child has moved to a new school.

Malachi Trust

Malachi received 329 referrals between October 2014 and September 2016 from parents of school-aged children across Birmingham and Somerset. After withdrawals due to parents no longer wanting to take part and ineligible referrals, 265 parents were randomised (134 to the intervention arm and 131 to the control arm). Participants are spread across six cohorts, ranging in size from 25 to 74). In each cohort, data are collected at three time points: baseline, 16 weeks post-randomisation

(T2), and 32 weeks post-randomisation (T3). At the time of writing, a preliminary figure for attrition for the two respective time points can be provided for the first three cohorts: cohort 1 (24% T2, 48% T3), cohort 2 (44%, 38%), and cohort 3 (22%, 24%). Data collection is ongoing for Cohorts 4, 5 and 6 and will finish in May 2017. We have attempted to reduce attrition significantly in Cohorts 4-6 through targeted data collection and working more closely with Malachi to connect the data collectors with parents. Malachi also coordinate with the schools to ensure that contact information is correct and to help with contacting parents who are hard to reach.

Ariel Trust

An initial feasibility study was conducted in Summer 2015 to understand teachers' and young people's respective experience of delivering and receiving Face Up. At the time, Face Up comprised three modules spread over nine lessons, which addressed healthy/controlling relationships, cyber bullying and advanced forms of psychological abuse. The study included interviews with teachers in eight schools, four focus groups with pupils, four classroom observations, a review of implementation materials (ie manuals, lesson plans and training slides) and analysis of teacher-completed lesson records. The main recommendations of this study were incorporated in the revised version of the programme. These included:

- adding a complementary whole-school element and information for parents
- having a stronger focus on skill development
- reducing programme length (now eight lessons over two modules)
- diversifying examples (eg addition of female perpetrators and different types of relationships)
- calibrating activities to engage pupils with diverse ability levels
- providing more relevant information about sources of help
- indicating in lesson plans the desired length of constituent activities
- ensuring that teachers understand the logic model and how each aspect of the programme contributes to it
- consolidating training materials

Final reflections

The process of undertaking the Realising Ambition trials has not been easy. Although some of the difficulties were arguably exacerbated by how the trials originated,¹⁹ we do not think that they are exceptional – most real-world trials face similar challenges. We hope that being open about them and how we and the respective provider organisations have sought to address them offers some useful learning to others who have embarked on or who are thinking about embarking on such ventures. Here, we offer a final few reflections on the experience.

First, for providers, being involved in an RCT is by no means just having someone come along and do an evaluation to them. It demands significant capacity and extensive collaboration with the evaluator. In one trial, for example, we had weekly check-in calls on recruitment targets and to troubleshoot problems with individual cases, and a member of the evaluation team spent considerable time in the relevant provider organisation's offices. As well as building trust and a sense of ownership, which are essential for running a trial smoothly, particularly when project staff are needed to assist with data collection, these processes helped with making timely decisions.

Second, evaluators need to make good use of data already held by the provider organisation. In one trial, for instance, data on young people served previously helped with refining eligibility criteria and estimating potential movement on the outcome measure. In another trial, existing data assisted with making estimates about demand and drop-out.

Third, it is necessary for all concerned to think on their feet. This is because, despite the best

planning, unexpected things crop up. As far as possible it is important to refer back to the trial [protocol](#). Since this won't cover every eventuality, it will be necessary to consider carefully problems that arise and work out solutions in collaboration (ie potentially involving evaluators, providers, statisticians and the ethics committee) and update the protocol accordingly.

Fourth, and this point is pertinent to commissioners and funders as well as providers and evaluators, the method of evaluation needs to be aligned to the stage of intervention development.²⁰ In the portfolio of Realising Ambition trials, one has involved a feasibility study followed by a pilot trial because the significant changes to programme design meant that a full trial was deemed to be premature. Failure to achieve a suitable alignment can mean a lot of wasted time and also risk setting up fledgling interventions to fail.

Finally, the benefits of doing an RCT, at least with social interventions, may lie as much in the refinements made to the intervention as in the end results. Put another way, trials are about improving as much as proving. There are other evaluation methods more obviously associated with intervention improvement, such as [rapid cycle testing](#), and certainly we are not suggesting that RCTs be undertaken unless a robust test of impact is warranted. But as a by-product of addressing some of the requirements for a trial, interventions can be improved significantly. As this Programme Insight demonstrates, there is the potential for particular gains in terms of intervention design, fidelity monitoring and participant engagement, many of which will last well beyond the end of the trial. Moreover, presentationally, an emphasis on improving as well as proving can help to increase buy-in from all provider staff.

“ All involved in an RCT need to think on their feet. It will be necessary to consider unexpected problems, work out solutions in collaboration and update the trial protocol accordingly. ”

To reinforce this last point, we leave the final word to leaders of the three programmes who have been travelling with us on this journey.

“I think it’s helped the programme reach its potential. It has professionalised the service, so that ultimately we’re delivering a better service for our children and families. By developing a better understanding of and creating fidelity measures, logic models and outcome measures we feel that our work is better thought out and stands on firmer, evidence-based grounds. Some of these processes we did before, but now we understand exactly what we’re doing and why we’re doing it. We have been able to use some of what we’ve learned in other programmes. Ultimately we’ve created a better service for beneficiaries. Rather than just assuming that what we’re doing is having a positive impact, we are now better organised, more focused on what we’re actually trying to achieve and better placed to achieve it.”

**Caroline Hopkins,
Senior Programme Manager, Chance UK**

“I’m proud of what we’ve done. People think that getting an RCT is just a gold stamp, a pat on the back, but it’s completely the opposite. It’s a warts-and-all look at every detail of your programme. Most licensed parenting programmes come from the US, and they often have limited impact in the UK. We want to be one of the first home-grown UK-based programmes to be licensed as an effective programme. Getting sufficient numbers was a crucial part in moving towards this goal, both in terms of reaching a large number of parents and being a rigorously evaluated organisation. We’ve been brave to make ourselves vulnerable to this kind of exposure and now I can stand up at the end of the day and say I know where our programme’s strong, I know where it’s weak and here’s what we’re going to do about it. It’s about saying I believe in what we’re doing, but I also believe in making it the best it can possibly be.”

**Laura Evans,
CEO, Malachi Trust**

“There is a tension with evaluation between proving and improving. Organisations like Ariel have to adapt, learn and improve very quickly if we want to stay afloat, particularly in such an austere funding environment. RCTs cannot keep up with this pace of change; since the RCT has begun, the product it is evaluating is not even among the top three or four products that we focus on. The benefits from being part of the [RCT] process have largely been through the significant changes we had to make to the organisation in order to be a part of it. The word that I would use is ‘transformational’. Both personally and professionally, both for me and the Ariel Trust, it’s been a great learning experience.”

**Paul Ainsworth,
Director, Ariel Trust**

Key Learning Points

- **The benefits of doing an RCT may lie as much in the refinements made to the intervention as in the end results.**
Trials, at least of social interventions, are often about improving as much as proving.
- **For providers, being involved in an RCT is not just having someone come along and do an evaluation to them.**
It demands significant capacity and extensive collaboration with the evaluator.
- **All involved in an RCT need to think on their feet.**
It will be necessary to consider unexpected problems, work out solutions in collaboration and update the trial protocol accordingly.
- **Securing organisational buy-in to a trial requires understanding an organisation's culture, processes and structure.**
The person in overall charge needs to agree, but if several people have responsibilities, they also need to buy in. Other factors that contribute to readiness for an RCT are the quality of the intervention, the organisation's capacity to recruit and deliver, and a commitment to learn from and not twist the findings.
- **Evaluators need to make good use of data already held by the provider organisation.**
- **The method of evaluation needs to be aligned to the stage of intervention development.**
Failure to achieve a suitable alignment can mean a lot of wasted time and also risk setting up fledgling interventions to fail.
- **The size of the sample is largely determined by the minimum effect size worth detecting, which should therefore be worked out as early as possible.**
- **Research techniques can usefully be employed to help recruit the required number of eligible participants.**
This includes using a combination of existing data and educated assumptions to determine the size of the eligible population and the numbers who need to be contacted or referred in order to arrive at the required number, taking into account factors such as lack of interest and ineligibility of some referrals/contacts.
- **Problems with recruitment and retention can be pre-empted to an extent.**
Pilot studies are useful to test the trial in miniature and check these and other factors.
- **Significant development or refinement of the logic model pays long-term dividends because it makes it clearer what needs to be measured.**
This makes the intervention easier to evaluate and can improve the quality of delivery.
- **In a trial, the services received by the control group affect the results as much as the intervention being tested.**
It is important to record what the control group receives, going beyond saying "business as usual" and instead detailing what this means in practice.
- **A by-product of collecting fidelity data is that it can encourage practitioners to deliver the intervention with greater fidelity.**
- **During the trial, it is worth reflecting on how well things are going and thinking about what actions will be taken when the results emerge.**
What does it mean if the results are positive, show no effect, suggest the intervention is harmful or are mixed?

Glossary of Terms

■ Attrition

The loss of participants from the study, typically defined as the number or proportion of participants who drop out.

■ Baseline

A measurement of participant characteristics and outcomes taken at the beginning of a study. In the case of an impact evaluation, this is done before the intervention is implemented.

■ Break-even analysis

An analysis that calculates a break-even point at which a profit begins to be made per unit. In the context of cost-benefit analysis, this point is shown in terms of the size of an effect on outcomes that would yield sufficient monetary benefits to break-even after accounting for unit costs.

■ Contamination

The spillover of effects from one study group to another. It occurs when participants allocated to the intervention and control groups interact, resulting in control group participants receiving part or all of the intervention.

■ Effect size

The size of the effect of an intervention, ie the amount of change that can be attributed to the intervention. It can be expressed in different ways, one of which is referred to as 'Cohen's d' (named after Jacob Cohen). A rule of thumb for Cohen's d is that 0.2 is a small effect, 0.50 a medium effect and 0.8 a large effect (not to be confused with the importance of the effect).

■ Equipoise

In medicine, clinical equipoise means that there is genuine uncertainty in the expert medical community over whether a treatment will be beneficial.

■ Feasibility study

Examines the practicality of an intervention with a view to refining it. It looks at the acceptability of and engagement with the intervention as well as adherence in delivery and viability of implementation.

■ Fidelity

Whether an intervention is delivered as intended/designed, covering adherence (delivery of core components), exposure (delivery of the specified dose), quality (eg provider's preparation, attitude and engagement of participants), responsiveness (engagement of participants in the activities) and reach (in terms of the target group).

■ Follow-up

A measure of participant outcomes taken after the end of an intervention, whether immediately or soon after or at some subsequent point.

Glossary of Terms

- **Intervention specificity**

Relates to the design of an intervention and whether it is focused, practical, logical and based on the best available evidence.

- **Logic model**

Explains how an intervention is designed to work and why it could achieve the desired outcomes – it is often represented in a diagram with a supporting narrative.

- **Meta-analysis**

A statistical method for combining the results from two or more studies.

- **Participants**

In the context of research, participants are individuals who agree (provide voluntary consent) to take part in a study, and should be distinguished from service users – in a trial, some but not necessarily all users of an intervention will be participants, and consenting individuals who do not receive the intervention because they are in the control group are also participants.

- **Protocol**

A detailed blueprint that sets out what needs to happen and when in a study, such as a trial.

- **Randomised controlled trial**

An experimental study in which participants are allocated to the study conditions (eg intervention and control) at random (ie by chance alone).

- **Rapid cycle testing**

Iterative testing of changes to an intervention with a view to improving the intervention.

- **Recruitment**

The process of getting initial involvement and sign-up from participants in the study.

- **Retention**

The task of keeping participants involved in a study to complete assessments and procedures as outlined in the study protocol.

Further Reading

We have drawn on many sources in the production of this Programme Insight. Our top picks for further reading on the themes discussed are listed below.

- Dynarski, M. & Del Grosso, P. (2008) Random assignment in programme assignment and intervention research: questions and answers, *Journal of Children's Services* 3 (1), 9-13.
- Feeley, N. & Cossette, S. (2015) Testing the waters: piloting a complex intervention. In Richards, D. A. & Hallberg, I. R. (Eds) *Complex Interventions in Health: An Overview of Research Methods*. London: Routledge.
- Funnell, S. C. & Rogers, P.J. (2011) *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco: Jossey Bass.
- Giangregorio, L. M. & Thabane, L. (2015) Pilot studies and feasibility studies for complex interventions. In Richards, D. A. & Hallberg, I. R. (Eds) *Complex Interventions in Health: An Overview of Research Methods*, London, Routledge.
- Glennerster, R. & Takavarahsa, K. (2013) *Running Randomized Evaluations: A Practical Guide*. Princeton: Princeton University Press.
- Haynes, L., Service, O., Goldacre, B. & Torgerson, D. (2013) *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials*. London: Cabinet Office.
- Jadad, A. R., & Enkin, M. W. (2007) *Randomized Controlled Trials: Questions, Answers and Musings (Second Edition)*. Oxford: Blackwell Publishing.
- Torgerson, D. J. and Torgerson, C. J. (2008) *Designing Randomised Trials in Health, Education and the Social Sciences: An Introduction*. Basingstoke: Palgrave Macmillan.

This Programme Insight has been written by Nick Axford, Gretchen Bjornstad, Finlay Green, Tim Hobbs, Anam Raja, Shreya Sonthalia, Luke Timmons, Laura Whybra and Zoe Wrigley (all based at Dartington Social Research Unit (DSRU), with the exception of Anam Raja, formerly DSRU and now at the Department of Psychiatry, University of Oxford, and Zoe Wrigley, also formerly DSRU and now at CASCADE/ School of Social Sciences, Cardiff University).

We gratefully acknowledge the helpful comments on earlier drafts of this Programme Insight from various colleagues: Paul Ainsworth, Vashti Berry, Sarah Blower, Laura Evans, Frances Flaxington, Julius Hinks, Caroline Hopkins, Colin MacFarlane, Louise Morpeth, Neil Watson and Shaun Whelan.

Endnotes

- ¹ Some of the material comes from telephone interviews conducted with Paul Ainsworth (Director, Ariel Trust), Laura Evans (CEO, Malachi Trust), Caroline Hopkins (Senior Programme Manager, Chance UK) and Colin MacFarlane (Plusone National Programme Manager, YMCA Scotland).
- ² All programmes in Realising Ambition, including the four discussed here, received support with organisational readiness and intervention refinement as part of the broader Realising Ambition project. Some but by no means all aspects of this work are reflected in this Programme Insight.
- ³ Some of this work preceded the selection of RCT projects, but in all cases further refinement took place after this decision. This reflected the additional resource that was available for working with RCT projects, which enabled us to go into greater depth.
- ⁴ Langford, R., Bonell, C. P., Jones, H. E., Poulidou, T., Murphy, S. M., Waters, E., Komro, K. A., Gibbs, L. F., Magnus, D. & Campbell, R. (2014) *The WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement*. Cochrane Database of Systematic Reviews 2014, Issue 4.
- ⁵ <https://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/>
- ⁶ There is a distinction between dropping out of the trial and dropping out of the intervention. Participants may drop out of the trial in the sense of refusing to provide any data for the evaluation but, if they are in the intervention arm, continue to receive the intervention. Equally, if they are in the intervention arm of the trial they may drop out of the intervention but continue to provide data for the evaluation.
- ⁷ DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002) Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology* 30 (2), 157-197; DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011) How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest* 12 (2), 57-91.
- ⁸ Bonell, C., Jamal, F., Melendez-Torres, G. J. & Cummins, S. (2015) 'Dark logic': theorising the harmful consequences of public health interventions. *Journal of Epidemiology and Community Health* 69, 95-98.
- ⁹ Axford, N., Lehtonen, M., Tobin, K., Kaoukji, D. & Berry, V. (2012) Engaging parents in parenting programs: lessons from research and practice. *Children and Youth Services Review* 34 (10), 2061-2071.
- ¹⁰ <http://www.sdqinfo.com>
- ¹¹ The programme involves groups of parents meeting over successive weeks, and random allocation was done in blocks of at least 8 to guarantee at least 4 parents per group.
- ¹² Chance UK – Evidence for Children's Outcomes (ECHO); Malachi Trust – Helping Others Parent Empathically (HOPE); and Ariel Trust – Healthy Adolescent Relationships Training Study (HEARTS).
- ¹³ While this adds to time and cost in the short-term, it can improve efficiency longer-term if lessons are learnt that either (a) preclude a full trial because of anticipated difficulties or (b) improve the design of a full trial.
- ¹⁴ Durlak, J. A., & DuPre, E. P. (2008) Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology* 41, 327-350.
- ¹⁵ Fidelity is also being monitored in the other Realising Ambition projects, although not necessarily in the same depth.
- ¹⁶ For example, rather than randomly allocate cases as they come in (dynamic) they are 'stored' until a block of the required size is obtained, at which point they are randomly allocated.
- ¹⁷ Preferably before, because we think their recollection is likely to be better, and it can create additional logistical problems if their contact with the pupil is reduced or non-existent in the new school year.
- ¹⁸ Although, as indicated, this then runs the danger of randomisation – and, in turn, the start of the intervention – taking place some time after baseline for some participants.
- ¹⁹ This refers to the predetermined constraints under which we were operating, notably regarding numbers of participants and time before results were due.
- ²⁰ Axford, N., Berry, V., Blower, S., Little, M., Hobbs, T. and Sodha, S. (2013) *Design & Refine: Developing Effective Interventions for Children and Young People*. Dartington: The Social Research Unit.

Find out more

**catch
22**

realisingambition@catch-22.org.uk
catch-22.org.uk/realising-ambition

substance.

neil.watson@substance.net
substance.net/case-studies/realisingambition



**DARTINGTON
SOCIAL
RESEARCH
UNIT**

tim.hobbs@dartington.org.uk
dartington.org.uk

**THE
YOUNG
FOUNDATION**

james.teasdale@youngfoundation.org
youngfoundation.org

