



PROGRAMME INSIGHT

Focus Piece

Realising Ambition Programme Insights: Issue 2

Evidence is Confidence

How to create a richer evidential tapestry

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.

Rather than writing a long evaluation report at the end of the five-year programme – which would likely be read by very few people – we are instead producing a series of 12 *Programme Insights* so people can get information about the programme while it is happening in bite-sized pieces. Some issues, like this one, are **Focus Pieces** that describe concepts and share some of our reflections and opinions. Others will be **Findings** pieces, reporting empirical data emerging from the programme and associated evaluation activities. The last type will be **Field Guides**: practical 'how to' guides for a variety of audiences. By sharing ideas, successes, challenges and even some mistakes, we hope to support and inspire others considering, undertaking or commissioning their own replication journey.

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. There you will also find some key reading we have drawn on in the development of this series.

About us: The Realising Ambition programme is supporting and is powered by 22 organisations – large and small – 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. This issue was written by the Dartington Social Research Unit (DSRU), with contributions from all partners in the consortium.



Realising Ambition Programme Insights: Issue 2

About this Programme Insight

The first issue of the Realising Ambition Programme Insight series – [the Secret Life of Innovation](#) – laid the foundations for understanding the replication of services for children and young people: what replication is and how successful replication may be defined. Key learning points from the first issue were:

There are moral, economic and policy arguments in favour of replication: These include providing services for which there is good evidence that children will most likely be better off as a result (when otherwise the impact of many services are unknown), an economic case to achieve value for money, and creating the foundation for scale.

Replication paves the way for innovation: Incremental improvements are the primary source of innovation, and replication creates the foundation for improvement.

There are five key ingredients of successful replication: a tightly defined service; effectively and faithfully delivered to those who can benefit from it; that provides confidence that outcomes have improved; that is cost-beneficial and scalable; and that is delivered by an organisation that uses evidence to learn and adapt, as required.

There are five main stages in the replication journey: prove; design; systemise; pilot; and scale. These stages are not necessarily sequential; replication requires iteration, testing and refinement.

Successful replication requires a degree of careful adaptation: The trick – and greatest challenge – with replication is in knowing what to keep the same and what can change. This requires a good understanding of what is core to the service that is being delivered (the things that make it work) and what is surface (the things that make it fit into a new context and make people want to use it).

Replication is just one of the ways to achieve impact at scale: Scale needs to be considered and built into the design stages right at the outset.

In this issue we build upon these themes and focus on the role of evidence in the replication process. In Part 1, we start by arguing that the process of replication requires a broader definition of evidence, one that includes but also moves beyond just evidence of impact. We present four questions that are central to building a diverse evidence-base in the context of replication and introduce a range of different types of evidence that may be generated to help address these questions. In Part 2 we consider how to take an overview of this broader and more nuanced take on evidence. We introduce the Realising Ambition 'Evidence-Confidence Framework' as a tool for considering the breadth and strength of different types of evidence relevant to the process of replication. We conclude by considering some implications of this approach for the commissioning of evaluations and generation of evidence in pursuit of service improvement.

Part 1: A broader definition of evidence

When most people in children's services think about [evidence](#), the first thing they usually think of is evidence of impact: does a particular service improve outcomes of those receiving it, or, put simply, does it work? This is an important and fundamental question, both for those funding and commissioning services and for those delivering those services. It is a question that at one time was rarely asked in the context of services for children, or was at least rarely answered with robust evidence. Now – particularly in the current climate of austerity and heavy cuts to children's services – a focus on the generation of evidence of impact in children's services is more prevalent than it has ever been.

However, 'does it work?' is not the only question that should be asked when considering if, what and how to replicate. We argue that [replication](#) demands a broader, more nuanced, definition of evidence.

Within the Realising Ambition programme we started by asking three broad questions, which various types of evidence have a role in helping to address:

- 1. What should be replicated?
- 2. How well is a service being replicated?
- 3. What is the impact of a service that is being replicated?

As the programme has developed, we have come to realise that a fourth question is also central to the process of replication:

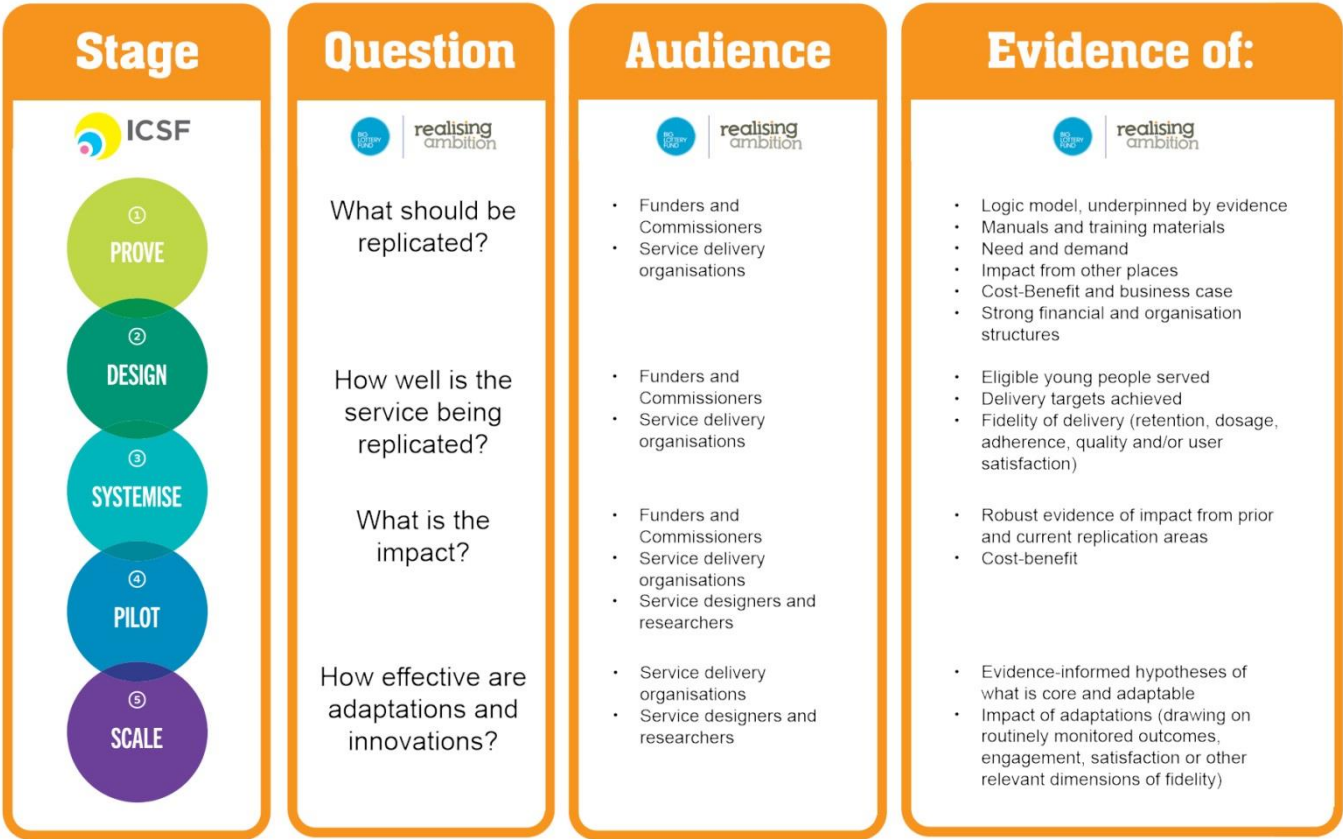
- 4. How effective are adaptations and innovations that emerge from replication?

Each question demands a different type of evidence to address it and some types of evidence contribute to addressing multiple

questions. A key point is that evidence of impact is just one thread in a richer **evidential tapestry**.

In Figure 1 we align our four main questions to the International Centre for Social Franchising's (ICSF) five stages of replication, which we introduced in the first issue of the Programme Insights series. We also acknowledge that different questions and types of evidence will be more or less important to different audiences, including **funders, commissioners**, service delivery managers and practitioners, as well as service designers, scientists and researchers. In Figure 1 we therefore also align the key audiences to each question. The final part of the figure is a list of the different types of evidence relevant to each question, which we consider in the next section.

Figure 1: Stage of replication, questions, audiences and evidence



Types of evidence required to address different questions

In this section we unpack what is summarised in Figure 1. We consider why each question is important and what audience it is important to, along with different types and sources of evidence that may be generated to help answer the question. What will become apparent is how many different types of evidence can potentially be generated, and how variable the strength and quality of that evidence might be. Therefore in Part 2 we introduce the Realising Ambition ‘[Evidence-Confidence Framework](#)’ – a tool we have developed that can be used to help judge the strength and overall balance of different types of evidence for a particular service being replicated, and to identify areas of development and opportunity.

Q1. What should be replicated?

Who and why? This question is particularly important to [funders](#) or [commissioners](#) deciding what services they may wish to replicate. It is also a key question for service delivery organisations considering whether or not to replicate more widely one of their existing services, or looking to replicate an existing service from elsewhere.

Types and sources of evidence required:

Evidence of a tightly defined service: This may be indicated by a strong [logic model](#), underpinned by previous research, and a clearly defined and articulated set of [core activities](#). These core activities are the elements of the service that are believed to lead to improvement in [outcomes](#) – and ideally these beliefs will be supported by evidence.

Manuals, implementation handbooks and training: The availability and quality of these will provide further confidence that the service is tightly defined and can be consistently delivered.

Evidence of need and demand in the planned replication area: There is little point in replicating a particular service if there is not a clearly identified [need](#) and [demand](#) for it. Need may, in part, be supported by a local authority needs analysis or data from public health observatories. Local epidemiological studies, such as the one that underpinned [Action for Children’s successful replication of Functional](#)

[Family Therapy in Renfrewshire](#), can strengthen the evidence of need and help ensure a good fit. Demand may be evidenced through consultations with local commissioners, potential service users or beneficiaries, members of the local community, as well as through an analysis of other competing services.

Evidence of impact on outcomes from other places:

Evidence of impact from other places may come from prior impact evaluation studies. These can vary a great deal in their quality. Very few services for children evaluated by an [experimental evaluation](#) – whereby outcomes of those receiving a service are compared to outcomes of those not receiving the service – yet this generally provides the greatest confidence that a service is or is not responsible for a positive [impact](#). Other evaluation approaches, such as pre- and post- intervention outcome monitoring, have an important role in service improvement and may indicate the direction in which outcomes are moving in, but they cannot be used to attribute changes to the service itself. The Dartington Social Research Unit’s [Standards of Evidence](#), used extensively within Realising Ambition and underpinning many other Standards – such as those of [Project Oracle](#) and [NESTA](#) – provide a tool for assessing evaluation quality and impact. Prior impact evaluation studies may also show varying degrees of impact in terms of size or consistency, so consideration is needed as to whether the evidence of impact is sufficiently compelling.

Evidence that the service is cost-beneficial and that there is a strong business case for replicating it:

Evidence of costs may be expressed in terms of [start-up](#) and [unit costs](#), which require a full and realistic estimation, taking into account direct and indirect costs. There are then a variety of ways of estimating whether or not a service is cost-beneficial, ranging from rough estimations of [cost-avoidance](#) through to sophisticated [cost-benefit analysis](#) based on observable impacts and a solid research foundation. (Costs and benefits will be the focus of a forthcoming issue in this Programme Insights series.) A strong [business case](#) builds on an analysis of costs and likely benefits, as well as a good budget, and is also in part related to need and demand: a delivery organisation should have, and be able to evidence, a strong and compelling business case for replicating a particular service.

Evidence of the necessary financial and organisational structures to support replication: Just as important as replicating a strong service is having a strong organisation that is doing the replicating. A mantra within Realising Ambition is that we are supporting strong organisations to replicate strong services: one without the other is a weak combination. Strong financial and organisational structures – including qualified, sufficiently trained and motivated staff – may be evidenced via tools such as the [Young Foundation's Organisational Health Scorecard](#) or external accreditations, kite marks or quality standards.

Q2. How well is a service being replicated?

Who and why? Once a service has been commissioned, the service delivery organisation – including both the managers and practitioners delivering it – will want to know how well it is being implemented to whom, how faithfully it is being replicated, and how the quality of delivery might be improved. Similarly, those funding or commissioning will want to ensure that their resources are being used as intended.

Types and sources of evidence required:

Evidence that eligible young people are being offered and provided the service: There is a good body of research indicating that if a service is delivered to those for whom it was not intended, that outcomes are likely to be diluted (and worse, the service could be harmful). It is therefore important to first of all ensure that there is a good degree of specificity in regards to whom is to be served. Clearly specified inclusion and exclusion criteria may be evidenced in [manuals](#), and training materials. Yet specificity is necessary but not sufficient: it is also important to monitor that these criteria are adhered to. This may be done using simple referral checklists combined with professional judgement, or, better still, in conjunction with standardised and validated screening tools (such as the [Strengths and Difficulties Questionnaire](#), if this were a relevant tool for a specific service).

Evidence that delivery targets are met: [Funders](#) and [commissioners](#) typically pay for and contract a service to be delivered to an agreed number of beneficiaries. Care must be taken at the outset to agree the parameters around these numbers. For example, do initial contacts that then drop

out count? Are different members of a family unit counted separately or as one unit? Care in agreeing these targets at the outset will help ensure that targets are realistic, achievable and accountable.

Evidence that the core of the service is being delivered with fidelity: [Fidelity](#) refers to consistent delivery that is faithful to the planned delivery model. This includes ensuring that [eligible young people](#) are served for a sufficient intensity or duration (referred to as [exposure](#) or [dosage](#)) and that the quality of service delivery is good. Also of central importance is being able to monitor and evidence [adherence to core components](#): those hypothesised key elements of the service that are designed to bring about a change in [outcomes](#). Fidelity incorporates information on retention, dosage and adherence, as well as [user engagement](#) and [user satisfaction](#). This requires that service-specific monitoring processes are developed and integrated into an effective [client management information system](#) that can then be tailored to report on required aspects of fidelity. This will help identify if delivery is going off-course, and if so, help find out why and course-correct.

Q3. What is the impact of the service being replicated?

Who and why? As we have described, a central question in the context of [replication](#) – as well as other forms of service delivery – is 'does it work'? Are children and families better off as a result of the service? This is a particularly important question for funders and commissioners in a time of austerity and heavy cuts to children's services, as they want effective services that represent value for money. It is also key for service delivery organisations who are committed to making a positive difference to those they serve. Yet answering this question, at least with a high degree of confidence, is challenging; it is easy (and common) to over-claim based on poor quality evidence or limited impact, and yet time- and cost-intensive to answer the question with confidence.

Types and sources of evidence required:

Evidence of impact from prior and current replication areas: As described in relation to the question 'what to replicate?', evidence of impact may come from prior impact evaluation studies, of

variable quality. Strong and consistent evidence of impact when a service is delivered elsewhere will provide greater confidence that similar outcomes may be achieved locally, particularly if the context is similar and the service is delivered with fidelity. However, just because an impact is achieved elsewhere does not necessarily mean the same impact will be achieved when replicated. There are examples of rigorous **evidence-based** programmes that have been widely replicated with a consistently positive impact on outcomes. However, there are a growing number of evidence-based programmes that show strong effects in one context that then fail to produce those effects when replicated in a different context.

The implications of this with relevance to **replication**, are as follows: (i) even when strong evidence of impact exists from elsewhere, it is important to build a robust evidence base locally; and (ii) this needs to be approached thoughtfully, considering and testing local adaptations in a **formative** way (i.e. throughout the course of delivery in order to improve practice as you go) and not just in a **summative** way (i.e. waiting until the end of a delivery phase before testing whether or not something worked). We return to these themes shortly.

Evidence that the service is cost beneficial: evidence about impact on outcomes is important, but so is knowing whether or not achieving that impact is worthwhile: do the benefits outweigh the costs of service delivery? We have touched on the types and sources of evidence of costs and benefits previously when considering what evidence is important and when thinking about what to replicate: the same types of evidence apply here.

Q4. How effective are adaptations and innovations that emerge from replication?

Who and why? In the first issue of the Programme Insights series, we described how rather than being a rigid process, replication can pave the way for incremental innovation: it allows us to think about and test what we might try to do differently to maximise impact, particularly when replicating services in new and different contexts. Yet knowing what not to change – the **core** of the service responsible for changes in outcomes – and where to adapt is the key. Evidence to inform and test adaptations is important for service delivery organisations

looking to improve their work, as well as other innovators and researchers engaged in service design and optimisation. This type of evidence gathering and testing is formative rather than summative; it acknowledges that evidence and service improvement are dynamic and do not necessarily have an end-point. This way of thinking has long been used in technology and commercial product design and innovation. It is now slowly emerging in the field of service design and evaluation.

Types and sources of evidence required:

Evidence to inform hypotheses of what is core and what is adaptable: evidence is required to inform judgements about what aspects of a service should remain fixed and what aspects may be adapted or refined. More often than not, these are hypotheses, and the danger is that they will be no more than guess. Hypotheses will be stronger, however, if existing evidence is drawn upon. Ideally evidence will come from prior research that has empirically tested the impact of similar adaptations in an experimental design (but such research is thin on the ground). At a minimum, strong empirically grounded theory should guide adaptations to be subsequently tested.

Evidence on the impact of adaptations: once planned adaptations are agreed, it is important to test whether they are effective. It follows that effectiveness must therefore be defined: this may be greater **engagement**, **satisfaction** or retention of young people or families, reduced cost of the service or a greater improvement in **outcomes**. As such, the intended impact of the adaptations should be routinely monitored, allowing comparisons either over time or between replication sites. This will not unequivocally confirm that the adaptations were responsible for changes, but it can provide good clues as to what changes appear to have a desired or undesired effect. This may pave the way for more robust summative evaluation. Central to testing adaptations is a strong evaluation framework, supporting service delivery organisations to ask the right questions at the right time, and an associated **client management information system** that can capture relevant data that will be analysed, used and valued by practitioners and managers delivering a service.

In Part 1 we introduced a wide range of different types of evidence to answer four distinct questions: (i) what should be replicated?; (ii) how well is a service being replicated?; (iii) what is the impact of a service being replicated?; and (iv) how effective are adaptations and innovations emerging from replication? In posing these questions, we have demonstrated that evidence of impact is just one part of this [evidential tapestry](#). Also implicit in this overview of different types of evidence is that the breadth and quality of evidence can vary enormously. In Part 2 we explore how to engage with this broader and more nuanced take on evidence in the context of replication.

Part 2: Evidence is confidence

Irrespective of the specific question being asked, too often we rely on evidence to support unequivocal claims of truth. We say things like ‘this works’ and ‘this is cost-beneficial’, yet the truth is often more nuanced than this. In relation to evidence of impact, for example, even with the most robust evaluations, we cannot unequivocally conclude ‘this works and this does not’. Rather, at best, we might be able to say that the evidence suggests that a particular intervention (or practice, or approach) is effective (or ineffective) in improving one or more specified outcomes (assuming certain contextual factors hold). Put another way, we might say “We can be (reasonably) confident that...”. This is not as strong – or confident – as saying “This works” but equally it is stronger – or more confident – than concluding that we can only say that a particular service had a positive impact when delivered by a particular organisation to a particular group of beneficiaries in a particular context in a particular point in time. Such a conclusion would mean that we can only say that something works if it is tested in the specific context, which is neither realistic nor plausible. It is true that evidence is rarely clean-cut, and that it can be messy and should invariably be interpreted with the context in mind, but confidence is a matter of degree.

Karl Popper, the eminent philosopher of science, once said, “The role of evidence is, in the main, to correct our mistakes, our prejudices, our

tentative theories”. Evidence cannot prove a given assertion (like ‘our service works’ or ‘our service is replicable’). What evidence can do, however, is improve the confidence that we – and others – have in our beliefs. It can also challenge our confidence in these beliefs. Evidence is not the whole truth; rather evidence is confidence.

The Evidence-Confidence Framework

The Realising Ambition team has used the concept of ‘evidence is confidence’ to develop a tool to help service delivery organisations, funders and commissioners understand the variety and strength of different types of evidence relevant to replication.

This tool is called the ‘[Evidence-Confidence Framework](#)’. It is structured around our five-part definition of successful replication introduced in the first issue of this Programme Insights series: (i) a tightly defined service; (ii) that is effectively and faithfully delivered to those that need it; (iii) evidence is used to learn and adapt, as required; (iv) there is confidence that outcomes have improved; and (v) the service is cost-beneficial and sustainable.

The tool can provide delivery organisations, funders or commissioners with an overview of the breadth and strength of evidence related to a variety of aspects of replication, not just impact. It can also provide a degree of confidence when considering each of the four questions we have introduced in this Programme Insight, as well as help to identify areas for the refinement of a service and opportunities to strengthen the evidence-base underpinning that service.

The blank template of the framework is presented in Figure 2. The five elements of our definition of successful replication are listed in the first column. To the right of each element of the definition is a range of different types of evidence that may be generated in support of that aspect of the definition. These include each of the types of evidence we outlined in Part 1, as well as a few **wider organisational dimensions** relevant to our broad definition of successful replication.

Figure 2: Evidence-Confidence Framework

A tightly defined service	Supported by a strong logic model	The "core" of the service is well defined	There are clearly specified activities	Delivery supported by manuals and training
That is effectively delivered to those that need it	Eligible individuals in need are served	Realistic delivery targets can be met	The "core" is delivered with fidelity	Service delivered by motivated and qualified staff
Evidence is used to learn and adapt, as required	Outcomes are routinely monitored	Engagement and retention are routinely monitored	Flexible components are identified and adaptations tested	Learning is translated across the delivery organisation
There is confidence that outcomes will improve	Evidence from elsewhere that outcomes improved	Delivery organisation able to effectively gather, analyse and communicate evidence	Evidence from current replication area that outcomes improved	Evidence of wider positive impact
The service is cost-beneficial and sustainable	Analysis of costs and likely financial return on investment	Compelling business case supporting replication	Service fully integrated into core business	Financial and organisational structures sufficiently robust to support replication

In order to provide an overview of areas of evidential strength and areas for potential development, the strength and quality of each type of evidence may be graded. In the framework we use a simple five-point colour grading system: the lightest blue representing the strongest evidence and the darkest blue the weakest (details about how we grade each are provided in the online appendix [here](#)).

In order to illustrate how the framework works in practice, Figures 3 and 4 provide a hypothetical application of the tool to two different organisations replicating two different services. Figure 3 refers to a well-developed **evidence-based programme** that has been widely replicated and rigorously evaluated elsewhere. As can be seen by the light shading in many parts of the table, there is strong evidence that outcomes have been improved when implemented and evaluated elsewhere; there is a strong logic model that is underpinned by evidence and the service is being delivered by a robust organisation.

Despite this, the areas of weakness in this hypothetical replication – indicated by darker shades of blue – are: (a) evidence of ability to engage and serve eligible young people; (b) evidence of ability to meet planned delivery targets; and (c) evidence of identifying and testing adaptations to flexible components. One could conclude from this assessment that while the service itself and the delivery organisation are both strong, there may be an issue with the 'fit' of the replicated service to local need or referral pathways. The framework points towards identifying some aspects of the service for refinement and surface adaptation, and then generating the evidence to test these adaptations in order to better engage and retain young people.

Figure 3: Evidence-Confidence Framework for a hypothetical well-refined service

A tightly defined service	Supported by a strong logic model	The “core” of the service is well defined	There are clearly specified activities	Delivery supported by manuals and training
That is effectively delivered to those that need it	Eligible individuals in need are served	Realistic delivery targets can be met	The “core” is delivered with fidelity	Service delivered by motivated and qualified staff
Evidence is used to learn and adapt, as required	Outcomes are routinely monitored	Engagement and retention are routinely monitored	Flexible components are identified and adaptations tested	Learning is translated across the delivery organisation
There is confidence that outcomes will improve	Evidence from elsewhere that outcomes improved	Delivery organisation able to effectively gather, analyse and communicate evidence	Evidence from current replication area that outcomes improved	Evidence of wider positive impact
The service is cost-beneficial and sustainable	Analysis of costs and likely financial return on investment	Compelling business case supporting replication	Service fully integrated into core business	Financial and organisational structures sufficiently robust to support replication

Figure 4, on the other hand, is an assessment of a hypothetical service that has been developed and refined locally, over a relatively recent period. It has not been subjected to an **experimental evaluation** (hence the relatively dark shading in the row related to confidence in outcomes being improved) but there is promise:

it has a relatively strong **logic model** that could be improved; it is being delivered by dedicated staff in a small but strong organisation committed to learning; the service is a core part of what the organisation does; and the people who deliver it are good at engaging and retaining young people in their service.

Figure 4: Evidence-Confidence Framework for a hypothetical early stage replication

A tightly defined service	Supported by a strong logic model	The “core” of the service is well defined	There are clearly specified activities	Delivery supported by manuals and training
That is effectively delivered to those that need it	Eligible individuals in need are served	Realistic delivery targets can be met	The “core” is delivered with fidelity	Service delivered by motivated and qualified staff
Evidence is used to learn and adapt, as required	Outcomes are routinely monitored	Engagement and retention are routinely monitored	Flexible components are identified and adaptations tested	Learning is translated across the delivery organisation
There is confidence that outcomes will improve	Evidence from elsewhere that outcomes improved	Delivery organisation able to effectively gather, analyse and communicate evidence	Evidence from current replication area that outcomes improved	Evidence of wider positive impact
The service is cost-beneficial and sustainable	Analysis of costs and likely financial return on investment	Compelling business case supporting replication	Service fully integrated into core business	Financial and organisational structures sufficiently robust to support replication

What the framework in Figure 4 suggests is that there is scope to draw upon evidence to better define what is **core** and what is **adaptable** to the service, and to further refine the **logic model** accordingly. It points towards the need to then start building up routinely collected data on outcomes and the engagement of beneficiaries to better understand if their engagement with young people affects outcomes. It also suggests that there is plenty of **formative** evidence that can and should be generated before jumping into rigorous **summative** evaluation of impact.

Using the evidence-confidence framework to answer different questions

As we have described, different types of evidence may be used to address different questions, and sometimes one type of evidence may be used to assess multiple questions. Figure 5 provides an overview of how the different types of evidence described thus far apply to different questions within the **evidence-confidence framework**. The types of evidence shaded in pink are those most relevant to each specific question.

Figure 5: The framework applied to different questions

Q1: What to replicate?

Atightlydefined Service	Supported by a strong logic model	Part of the service evaluated	Parameterised service adaptations	Developed/added by a local team
That is effectively delivered to those that need it	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
Evidence is used to learn and adapt, as required	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
There is confidence that outcomes will improve	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
The service is cost-beneficial and sustainable	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)

Q2: How well is a service being replicated?

Atightlydefined Service	Supported by a strong logic model	Part of the service evaluated	Parameterised service adaptations	Developed/added by a local team
That is effectively delivered to those that need it	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
Evidence is used to learn and adapt, as required	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
There is confidence that outcomes will improve	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
The service is cost-beneficial and sustainable	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)

Q3: What is the impact?

Atightlydefined Service	Supported by a strong logic model	Part of the service evaluated	Parameterised service adaptations	Developed/added by a local team
That is effectively delivered to those that need it	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
Evidence is used to learn and adapt, as required	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
There is confidence that outcomes will improve	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
The service is cost-beneficial and sustainable	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)

Q4: How effective are adaptations?

Atightlydefined Service	Supported by a strong logic model	Part of the service evaluated	Parameterised service adaptations	Developed/added by a local team
That is effectively delivered to those that need it	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
Evidence is used to learn and adapt, as required	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
There is confidence that outcomes will improve	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)
The service is cost-beneficial and sustainable	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)	Core evidence (pink)

The generation of evidence is not sequential

Our experience with Realising Ambition is that the process of replication is not sequential: in

particular, there are multiple feedback loops between the stages of design, systemising, piloting and scaling. It follows that the

generation of evidence should not be sequential and nor is it ever really complete.

Timely questions should drive the generation of appropriate evidence, yet too often we see evidence commissioned or generated without enough consideration of whether the question is timely or appropriate. Evidence of impact is usually first and foremost in people's mind (especially funders and commissioners of evidence). However, as we have explored in this issue, different questions require different forms of evidence to answer them, and different questions will be more or less relevant at different stages in the replication journey.

For a service in the early stages of its replication journey, it will be more prudent to ask questions about how well it is being replicated before jumping into assessing impact. For a service that is 'bedded down' with a good track record of delivery of good quality implementation, it will be more appropriate to explore questions related to the scope for testing adaptations or assessing impact.

There is also a good chance that questions will be revisited as the process of replication develops. One might start with asking how effective replication has been to date, which may in turn prompt evidence for the identification and testing of adaptations, in turn promoting an assessment of impact, which in turn prompts the quest for further refinement and innovation – and so on.

As illustrated in the [evidence-confidence framework](#), the [evidential tapestry](#), at best, represents a snapshot in time in a specific place or context. The generation of evidence in the context of replication is not a rigidly hierarchical or stepped process. Evidence evolves.

The generation of evidence should be proportionate

The generation of evidence should also be proportionate. We do not necessarily mean proportionate to the size of an organisation. Indeed, we are challenging this assumption within Realising Ambition by undertaking three randomised controlled trials with small and medium-sized third sector organisations. Rather, the generation of new evidence should be proportionate to the breadth and depth of existing evidence. So while we have argued that the generation of evidence is not rigidly

sequential, there are some foundational forms of evidence in the context of replication – such as a strong logic model and tools to support consistent delivery – that should arguably come before other forms of evidence – such as evidence about fidelity of delivery or impact on outcomes. But this does not mean that once this evidential foundation is built, that it should not be strengthened if the weight of other evidence exposes cracks in those foundations or highlights opportunities to make them stronger.

We hope this broader, more nuanced and proportional view of evidence will contribute to a smarter generation of evidence to inform service improvement and ultimately increase the likelihood that outcomes will be improved for children and young people in receipt of services that are being replicated.

How we plan to develop and use the Evidence-Confidence Framework within Realising Ambition

- The [Evidence-Confidence Framework](#) is a tool that we have developed this year to help provide a more nuanced and balanced overview of different types of evidence relevant to the process of [replication](#).
- We are using the framework to help determine how confident we can be that each project has been successful in replicating their specific service as part of the Realising Ambition programme. This will also help inform the success of the programme overall.
- The application of the framework to each project is being applied and agreed in collaboration with each project as part of an exit interview process for those reaching the end of their Realising Ambition grant. We will be reporting the application of the framework for each project as part of a published case study series.
- We will be refining the tool over the coming year and will update and disseminate as it develops. [The Dartington Social Research Unit](#) will draw upon the application of the framework as part of the wider programme-level evaluation of Realising Ambition: findings from which will be reported in due course as part of this Programme Insights series.

Key Learning Points

- **Replication requires a broader definition of evidence.** Evidence of impact is important, but there are many other forms of evidence that also play an important role in replication. Evidence of impact is just one thread in a richer evidential tapestry.
- **Different questions require different types of evidence.** There are four questions central to replication: (i) What to replicate? (ii) How well is something being replicated? (iii) What is the impact? and (iv) How effective is adaption and innovation emerging from replication? Each of these questions requires different types of evidence.
- **The generation of evidence is not sequential:** Different questions – each requiring different types of evidence – are more or less relevant at different stages in the replication journey. Once an evidential foundation is built it should be revisited and strengthened if the weight of other evidence exposes cracks in those foundations. It follows that the generation of evidence should not be sequential and nor is it ever really complete.
- **The generation of evidence should be proportionate:** Not necessarily to the size of the delivery organisation, but rather to the breadth and depth of evidence that already exists.
- **Evidence is confidence, not the whole truth:** Evidence cannot prove a given assertion (like ‘our service works’ or ‘our service is replicable’). What evidence can do, however, is improve the confidence that we – and others – have in our beliefs.
- **The evidence-confidence framework is a useful tool to prompt reflection about the depth and breadth of existing evidence:** It allows service delivery organisations to reflect a more nuanced way about areas of strength and areas in which evidence may be strengthened, in response to questions that are relevant and timely to their particular stage of replication.

Glossary of Terms

- **Adaptable**
Those aspects of a service that may be altered, refined or adapted in order to foster greater engagement, retention or satisfaction of those in receipt of a service (yet do not disrupt the underlying core mechanisms of the service or intervention).
- **Adherence**
A dimension of fidelity. Refers to whether the core components of a programme are delivered as designed, to those who are eligible for the service, by appropriately trained staff, with the right protocols, techniques and materials and in the prescribed locations or contexts.
- **Business case**
A business case provides justification for a proposed project or programme. Ideally it includes an analysis of costs and likely benefits, as well as a detailed budget, and also evidence of the need and demand for the service.
- **Client management information system**
A database that allows projects to view their real time data on outcomes, fidelity

Glossary of Terms

monitoring, quality assurance processes and other delivery data such as costs and staffing. High quality systems will typically allow users to view data in a visual format (graphs, charts etc) and enable data to be analysed and presented in a variety of ways (by delivery year, project type, outcome etc). These systems are useful for monitoring children's outcomes as they progress through a programme, monitoring the quality of delivery across multiple sites, and testing the results of adaptations to programme components.

- **Commissioner**

Responsible for the strategic allocation of public funds to projects, programmes or services that best address the needs of children, young people and families in their geographical and service area (for example Children's Services, Health, Education, Youth Justice etc). The priorities of commissioners are to engage services that represent good value for money as well as quality delivery and increasing the likelihood of positive impact.

- **Core components**

The key activities that make a service work. Put another way, the specific aspects or mechanisms of a service that lead to the desired change in outcomes. For a service to be replicated successfully, providers need to be clear about what the immutable core the service is.

- **Cost-avoidance**

Refers to actions taken to reduce future costs. Cost-avoidance as a value is the difference between what is actually spent and what would have been spent had no avoidance measures been implemented.

- **Cost-benefit analysis**

The estimation of financial returns on an investment or service. Returns are typically estimated for individual recipients of a service, agencies providing the service and the state. Cost-benefit analyses rely upon accurate cost information and robust evidence of impact (ideally from experimental evaluations). Cost-benefit analysis may produce a calculation of net cost (benefits minus cost) or the ratio of costs and benefits.

- **Demand**

In the context of social interventions the number of individuals who (a) match the particular target group within a given population and (b) actually want to participate in the programme.

- **Eligible young people**

Those young people who fit the target criteria for a specific service or programme. This could be based upon factors such as their age or gender, or relate to the difficulties they may be experiencing such as homelessness, conduct disorder, or educational problems. Those young people who are eligible for a service or programme should be the same young people who are likely to benefit most from receiving it.

- **Evaluation**

Various aspects of a programme can be evaluated, including the process of delivery, user satisfaction and impact. Here evaluation refers to the use of social research procedures to investigate systematically the effectiveness of programmes or services in terms of improving children's health and development.

Glossary of Terms

▪ **Evidence**

Generally speaking evidence is information that acts in support of a conclusion, statement or belief. In children's services this tends to be information indicating that the service works, i.e. is achieving the intended change in outcomes. We take a broader view in that evidence may support or challenge other aspects of service delivery, such as quality of implementation, reach and value for money.

▪ **Evidence-based programmes**

A discrete, organised package of practices or services – often accompanied by implementation manuals, training and technical support – that has been tested through rigorous experimental evaluation, comparing the outcomes of those receiving the service with those who do not, and found to be effective, i.e. it has a clear positive effect on child outcomes. In the Standards of Evidence developed by the Dartington Social Research Unit, used by Project Oracle, NESTA and others, this relates to 'at least Level 3' on the Standards.

▪ **Evidence-Confidence Framework**

The Realising Ambition 'Evidence-Confidence Framework' is a tool that can be used to help judge the strength and overall balance of different types of evidence for a particular service being replicated, and to identify areas of development and opportunity. It is structured around a five-part definition of successful replication: (i) a tightly defined service; (ii) that is effectively and faithfully delivered to those that need it; (iii) evidence is used to learn and adapt, as required; (iv) there is confidence that outcomes have improved; and (v) the service is cost-beneficial and sustainable. A simple five-point colour grading system is used to grade the strength and quality of each type of evidence: the lightest blue representing the strongest evidence and the darkest blue the weakest.

▪ **Evidential tapestry**

Replication requires a range of evidence to support both its justification, and to maintain high quality delivery. For example, not only is evidence of impact important for understanding the outcome of a service, but it is also useful in justifying the replication of a service in a new area. Alongside this can be evidence of the need for the service and demand for it in a local area. Evidence can also relate to delivery quality and fidelity to the model. Different types of evidence, all varying in quality and utility can provide answers to a range of questions helpful to practitioners and managers delivering services for children and families. When viewed holistically together, this overview of the breadth, depth, and quality forms an 'evidential tapestry'.

▪ **Experimental evaluation**

An evaluation that compares the outcomes of children and young people who receive a service to those of a control group of similar children and young people who do not. The control group may be identified by randomly allocating children and young people who meet the target group criteria – a randomised controlled trial or RCT –, or by identifying a comparable group of children and young people in receipt of similar service – a quasi-experimental design or QED.

▪ **Exposure / Dosage**

Refers to the "amount" of programme or service a person receives. This could be the number of total sessions attended, the length of those sessions, or how frequently they took place.

▪ **Fidelity / Faithful delivery**

The faithfulness to the original design and core components of a service. This can be assessed by fidelity monitoring tools, checklists or observations.

Glossary of Terms

- **Formative evaluation**

An evaluation that takes place before or during the implementation of a programme or service to improve the quality of its design and delivery. This type of evaluation is useful for providing on-going information and feedback to staff, and can also be useful in observing changes that take place after adaptations or modifications to a programme have been made (see also summative evaluation).

- **Funder**

Typically an organisation – foundation, charitable trust, or other philanthropic entity – that seeks to support social change through the funding of programmes, projects or services aimed at addressing “social problems”. Usually these organisations are focused on particular outcomes – such as reducing inequality and homelessness, tackling the causes of gang violence, improving mental health support etc.

- **Impact**

The impact (positive or negative) of a programme or service on relevant outcomes (ideally according to one or more robust impact evaluations).

- **Logic model**

A typically graphical depiction of the logical connections between resources, activities, outputs, and outcomes of a service. Ideally these connections will have some research underpinning them. Some logic models also include assumptions about the way the service will work.

- **Manual**

A document that covers all the things about a programme or service that are relevant wherever – and whenever it is being implemented. This includes the research base for the programme, the desired outcomes, the logical connection between activities and these outcomes, the target group and all of the relevant training or delivery materials (see also ‘Implementation handbook’).

- **Need**

In relation to services for children and families, this refers to how many individuals in a specified population match the target group for the programme.

- **Outcomes**

Outcomes refer to the ‘impact’ or change that is brought about, such as a change in behaviour or physical or mental health. In Realising Ambition all services seek to improve outcomes associated with a reduced likelihood of involvement in the criminal justice system.

- **Replication**

Delivering a service into new geographical areas or to new or different audiences. Replication is distinct from scaling-up in that replication is just one way of scaling ‘wide’ – i.e. reaching a greater number of beneficiaries in new places. (See definition of ‘scale’).

- **Service designer**

Within the context of services for children and families, any individual or organisation responsible for conceiving, planning and constructing a service or programme aimed at preventing or ameliorating the difficulties or potential difficulties of children and families.

Glossary of Terms

Ideally service designers balance science and knowledge of 'what works' alongside expertise in user engagement and co-production.

- **Start-up costs**

The total cost of setting up a project, programme or service in a new area. Start-up costs typically include capital costs such as IT equipment, planning and training costs, consultancy, recruitment, licensing and legal costs.

- **Summative evaluation**

An evaluation carried out typically at the end of a delivery cycle in order to establish the outcomes of a programme against its original objectives, how effective adaptations may have been, and to inform decisions around whether a programme should continue to be delivered or whether further adaptations should be made (see also 'formative evaluation').

- **Surface adaptations**

Aspects of the service that can be adapted to fit local contexts. These are peripheral components that do not directly alter the core aspects of the service that make it work. Surface adaptations may allow providers in other areas to make the service 'their own' and better serve the needs of local populations.

- **Unit costs**

The cost of everything required to deliver a programme to a participant or a family. A unit cost is normally expressed as an average cost per child or family, but can also be expressed as a range (for example, unit costs ranging for "high need" to "low need" cases).

- **User engagement**

A dimension of fidelity. This refers to the extent to which the children, parents or families receiving a programme are engaged by and involved in its activities and content. How consistently do participants stick with the programme? Do they attend? Do they like it? Do they get involved? Without high levels of user engagement, it is unlikely that programmes will achieve their desired impact.

- **User satisfaction**

Refers to whether children and families in receipt of a particular service are satisfied with the delivery and outcomes of that service. Did they feel they received enough sessions, that they established a good relationship with practitioners? Did they feel like the programme helped to deal with the difficulties they were facing, or prevented the occurrence of others? User satisfaction is typically captured upon completion of a service or programme.

A more expansive glossary of key terms related to Realising Ambition may be found at the Realising Ambition website: catch-22.org.uk/realising-ambition. This will grow as the series of Programme Insights develop.

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.

- Bamberger, M., Rugh, J., & Mabry, L. (2012). *RealWorld evaluation: Working under budget, time, data, and political constraints*. Thousand Oaks, USA: SAGE.
- Blasé, K., & Fixsen, D. (2013). *Core intervention components: Identifying and operationalizing what makes programs work*. Washington, D.C: Office of the Assistant Secretary for Planning and Evaluation, Office of Human Services Policy, U.S. Department of Health and Human Services.
- Dartington Social Research Unit. (2013). *Design and Refine: Developing effective interventions for children and young people*. Dartington, England.
- 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.
- Little, M., & Edovald, T. (2012). Return on Investment. The Evaluation of Costs and Benefits of Evidence-Based Programs. *Psychosocial Intervention*, 21, 2, 215-221.
- LUMA Institute (2012). *Innovating for people: handbook of human-centered design methods*. Pittsburgh, Pennsylvania: LUMA Institute, LLC.
- Paulsell, D., Del, G. P., & Supplee, L. (2014). Supporting replication and scale-up of evidence-based home visiting programs: assessing the implementation knowledge base. *American Journal of Public Health*, 104, 9, 1624-32.
- Puddy, R. W., Wilkins, N., Centers for Disease Control and Prevention (U.S.), & National Center for Injury Prevention and Control (U.S.). (2011). *Understanding evidence: A guide to the continuum of evidence of effectiveness*.
- Spoth, R., et al. (2013). Addressing core challenges for the next generation of type 2 translation research and systems: the translation science to population impact (TSci Impact) framework. *Prevention Science: the Official Journal of the Society for Prevention Science.*, 14, 4, 319-51.
- Stern, E. (2015). *Impact evaluation: A guide for commissioners and managers*. Bond for International Development. London, England.

You can find a full list of additional resources we have drawn on at the Realising Ambition website: catch-22.org.uk/realising-ambition. This will grow as the series of Programme Insights develop.

Find out more



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



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



thobbs@dartington.org.uk
dartington.org.uk



shaun.whelan@youngfoundation.org
youngfoundation.org