Guidance published February 2026

TIIPE Guidance

Theory-Informed Implementation and Process Evaluation

DARTINGTON SERVICE DESIGN LAB

Centre for Evidence and Implementation

youth futures FOUNDATION

Youth Futures Foundation commissioned the Centre for Evidence and Implementation and the Dartington Service Design Lab to undertake a literature review and develop this Theory-Informed Implementation and Process Evaluation (TIIPE) guidance. This guidance was shaped through iterative review with a Youth Futures Foundation review and advisory group, alongside consultation with external experts. A full list of authors and contributors is provided below.

Dartington Service Design Lab
Ediane Santana de Lima (Research Lead)
Cristina Preece (Research and Design Officer)
Dr Julie Harris (Director of Research)

Centre for Evidence and Implementation
Dr Stephanie Smith (Associate Director)
Dr Gayatri Kembhavi-Tam (Associate Director)
Jane Lewis (Associate Director)

Youth Futures Foundation (Funder and project oversight)
Jane Mackey (Senior Evidence and Evaluation Manager)
Hannah Murphy (Head of Evaluation)

Suggested citation
Santana de Lima, E.*, Smith, S.*, Kembhavi-Tam, G., Preece, C., Lewis, J.†, and Harris, J.† (2026) Theory-informed implementation and process evaluation guidance. Commissioned by Youth Futures Foundation, London.

* Joint first authors.

† Joint senior authors.

See Report of Literature Review: Smith, S.*, Santana de Lima, E.*, Kembhavi-Tam, G., Preece, C., Verdugo, P. Hall, A. Rowland, J., Green, F., Harris, J.†, Lewis J.† (2026). Theory-informed implementation and process evaluation: a rapid review of guidance. Youth Futures Foundation: London

What is the purpose of this guidance?

This guidance is designed to support evaluators undertaking Theory-Informed Implementation and Process Evaluations (TIIPEs) for Youth Futures Foundation (Youth Futures).

TIIPE is an approach that uses programme and implementation theory to systematically explore how programmes are delivered, and how and why outcomes are achieved (or not). By exploring the mechanisms that drive change, and the contexts in which this change occurs, TIIPE also enhances the understanding of causal pathways, and provides actionable insights for improving, replicating and scaling programmes.

The focus of this guidance is to support evaluators in using TIIPE to:

- **Assess and inform programme design and implementation to determine readiness for impact evaluation.**
- **Enhance the learning generated through an impact evaluation, typically a randomised controlled trial (RCT) or quasi-experimental design (QED).**

Impact evaluation and TIIPEs are complementary: impact evaluation explores whether a programme[2] is effective, and TIIPE generates an understanding of how and why a programme works in real-life settings, or why it may not.

As the What Works Centre for Youth Employment, Youth Futures is committed to generating high-quality evidence. This guidance sets out the best practice principles in designing and delivering TIIPE. It has been informed by a rapid evidence review and consultation with Youth Futures' advisory group and panel of evaluators.

2 The term programme is used throughout to mean programme, intervention, innovation or policy.

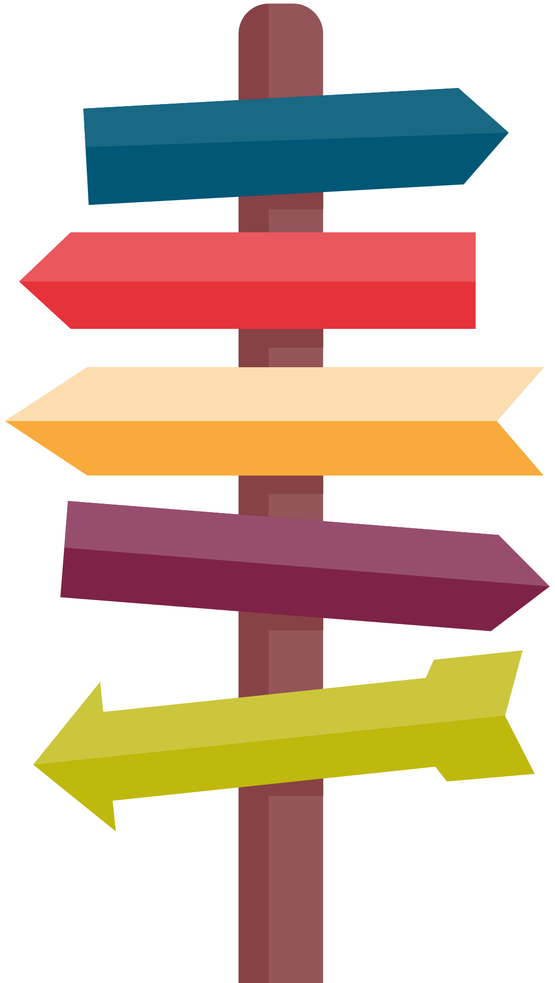
What is the guidance structure?

How can I navigate this guidance?

Start-to-finish
for those new to TIIPE or seeking step-by-step guidance.

Thematically
for those focusing on particular topics, such as programme theory or implementation theory.

By evaluation stage
for those already conducting a TIIPE, focusing on a particular stage of evaluation.



Chapter	
1	Explains the theoretical foundations of Theory-Informed Implementation and Process Evaluation (TIIPE)
2	Focuses on exploring programme theory and implementation theory
3	Focuses on the development, refinement and testing of programme theory and Theories of Change (ToCs)
4	Provides guidance on developing and testing implementation theory
5	Outlines how a TIIPE can be tailored across different evaluation stages.

Chapter 1

Theory-informed implementation and process evaluation

youth futures FOUNDATION

What is Implementation and Process Evaluation (IPE)?

IPE focuses on understanding how programmes are delivered in practice and why they do or do not achieve their intended outcomes.

IPE helps researchers and evaluators "look inside the black box" to understand how and why programmes create change, examining both the practical delivery and theoretical foundations.

Rather than exploring whether an programme works (effectiveness), IPE explores the delivery and dynamics of implementation in real-world settings. It enhances the real-world relevance of evaluation findings by providing policymakers, funders, and practitioners with learning about how a programme needs to be delivered within their context for it to achieve its intended outcomes.

What is Theory-Informed IPE (TIIPE)?

TIIPE builds on IPE by providing a theoretical framework for understanding how and why a programme works. "Theory" in TIIPE refers both to the theory behind the programme itself (programme theory) and the theory behind how it is put into action (implementation theory).

Programme theory

Explains the underlying logic of how a programme is hypothesised to bring about change for its participants. Programme theory is focused on the essence of the programme itself. It includes an understanding of the programme's core activities and components, as well as an understanding of the mechanisms through which these activities are expected to deliver outcomes for participants.

Implementation theory

Explains how a programme is expected to be delivered in practice in order for it to bring about change. It includes an understanding of the factors that can facilitate or hinder implementation, how the context in which the programme is delivered can affect outcomes, and how certain strategies may support effective implementation.

In TIIPE, the integration of programme theory and implementation theory strengthens both the design and interpretation of the evaluation. These two perspectives offer complementary lenses that, when used together, provide a more complete and rigorous understanding of how programmes function in the real world.

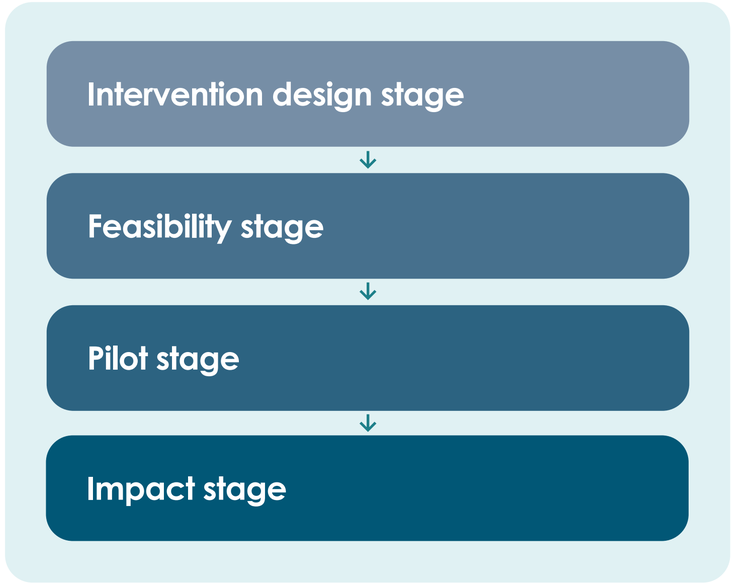
The theoretical underpinnings of programme theory and implementation theory are explored in further detail in Chapter 2. The application of programme and implementation theory is explored in Chapters 3, 4, and 5.

How does TIIPE support programme design, impact evaluation and decision-making?

Youth Futures commissions four distinct stages of programme development and evaluation, aligned to where a programme is in its development journey:

- **Intervention design:** to strengthen and consolidate the programme's design and to build the foundations for future evaluation.
- **Feasibility evaluation:** to assess whether the programme can be delivered as intended, identify any necessary adaptations and determine readiness for impact evaluation.
- **Pilot evaluation:** to test delivery of the programme and the proposed impact evaluation design (or elements of it), refining both where needed.
- **Impact evaluation:** to robustly assess programme effectiveness using a randomised controlled trial (RCT), quasi-experimental design (QED) or other high-quality method.

Figure 1: Youth Futures' programme development and evaluation stages:



TIIPE is used across these evaluation stages to support evidence-informed programme design and iteration, to design and interpret impact evaluations, and to build transferable learning for decision-makers.

In the lead-up to impact evaluation TIIPE is used to:

- **Provide decision-makers with a clear theoretical framework and evidence base for selecting programme components and methods of delivery.**
- **Explore whether a programme has sufficient evidence of promise and implementation stability to be considered ready for impact evaluation.**
- **Support evaluators to identify where adjustments are required to enable a programme to be evaluated through an impact evaluation.**
- **Support evaluators to identify which aspects of programme delivery and its surrounding context are most important to test in an impact evaluation, as well as when and how to test these elements.**

During and post-impact evaluation TIIPE is used to:

- Support the interpretation of impact evaluation findings through generating an understanding of how, why and in what context outcomes were (or were not) achieved.
- Support decision-makers to understand how an effective programme can be sustained or scaled within their context.

Without TIIPE, programmes can become **black boxes**. We may observe outcomes but not understand how or why they occurred. This limits learning and the ability to make informed decisions, increasing the risks of:

- **Superficial understanding** of what drives change, making it difficult to replicate or scale success.
- **Weak or inconsistent decision-making**, driven by intuition or politics rather than evidence.
- **Limited adaptability**, as implementers lack insight into what to modify when contexts shift.
- **Poor use of resources**, since ineffective or poorly understood programmes may continue unchecked.
- **Fragmented knowledge**, as findings remain context-specific and fail to build cumulative scientific understanding.

→
The role of TIIPE across Youth Futures' evaluation stages is explored further in Chapter 5.

Centring equity in TIIPE

Youth Futures is committed to equitable evaluation, and equity should be a key consideration within the design, delivery and reporting of TIIPE.

Key methods and approaches that will be discussed include:

- **Co-producing overarching research questions and theories**
- **Co-designing data collection tools**
- **Co-delivery of data collection**
- **Co-producing the focus for analysis**
- **Co-delivery of analysis**
- **Disaggregation of data in analysis: identifying the key equity groups in focus and how the programme works for them**
- **Promoting participant voice in the dissemination of learnings**
- **Youth advisory panels**

Embedding equity within TIIPE means:

- **Facilitating participation and ownership:** Where possible, TIIPE should be co-designed with a range of stakeholders (including those being supported through the programme). Co-design helps to ensure the approach is responsive to the specific needs and preferences of communities, that their voices are meaningfully included throughout the research process, and that trust and ownership are fostered in both the process and its outcomes. Co-design also helps to ensure the evaluation methods used are not harmful to participants and that the findings and outputs do not cause unintended harm to others.
- **Understanding how a programme works for different groups of young people:** TIIPE should consider differences in programme effectiveness and implementation across contexts and among different groups of young people. For example, an evaluator may want to consider whether the programme is reaching all of its target population, whether the programme is being delivered in the same way for all groups of young people, whether outcomes are being achieved equally for all young people, and whether these outcomes are being achieved in the same ways for all young people.
- **Making recommendations for how a programme could achieve more equitable outcomes:** Through considering how the programme works for different groups of young people, a TIIPE can be used to assess whether there are any unintended negative outcomes for particular groups, make recommendations for how a programme could achieve more equitable outcomes, and make recommendations for how the programme should be adapted across contexts and across different groups.

To meaningfully take equity into account in TIIPE, it must be treated not as a one-time event but as an ongoing process: a series of deliberate choices around methods, approaches and tools that are considered throughout the evaluation lifecycle. This guidance offers practical steps to support the centring of equity throughout the evaluation process.

Equity starts with those who are involved in shaping the work. Participatory approaches ensure that diverse voices and perspectives, especially those with lived experience, can influence how problems are defined, what questions are asked, what tools are used and what outcomes matter. Done well, they improve relevance, rigour and trust in the process.

Participation must be meaningful. If not carefully designed, these approaches can become tokenistic and risk reinforcing existing inequalities. It's essential to ensure that participation genuinely influences decision-making.

Case study

A youth employment training and placement programme

This example is used across the different chapters to explain different constructs.

The programme

A youth employment training and placement programme is launched to help young people (ages 18–24) gain job skills and secure sustainable employment. The programme includes workshops on CV building, interview preparation and workplace readiness, alongside job placement support.

Programme theory

The programme theory assumed that providing skills training and job leads would increase participants' access to sustainable employment.

Initial findings from a feasibility evaluation

After several months, data showed mixed outcomes. While some participants secured jobs, many did not.

Applying a TIIPE approach

To understand variations in outcomes, evaluators examined the programme and implementation theories together. They used interviews, contribution analysis and realist-informed analysis to identify causal mechanisms and contextual differences.

Case study

What the TIIPE revealed

Contextual factors / determinants

Both programme theory and implementation theory identified several contextual conditions shaping delivery and outcomes.

Positive enablers:
Some settings had strong links with schools and colleges, which enabled recruitment, and good links with some employers, which helped with finding job placement opportunities.

Barriers:
Limited staff capacity constrained one-to-one support, variation in staff expertise led to inconsistent delivery, and variation in participant readiness (e.g., confidence, prior experience) affected uptake.

Implementation outcomes

Implementation theory identified that as a result of these contextual factors, the programme was not being delivered as intended.

- Fidelity varied: while group workshops were consistently delivered, the intended one-to-one support was less consistent due to staffing pressures.
- Reach was uneven: young people already closer to the labour market benefitted most, while those with multiple barriers disengaged.
- Acceptability was mixed: participants valued mentorship highly but were less satisfied with standardised workshops.

Mechanism of change

A key mechanism connecting skills training and job leads to employment outcomes was motivation and self-efficacy. However, this was only activated if young people received sufficient one-to-one mentorship support, and the TIIPE demonstrated that implementation was mixed.

Case study

Resulting adaptations and strategies

Adaptations

As a result of these findings, several programme adaptations were implemented:

- Mentors provided personalised support to young people assessed as having additional needs for support.
- Employer engagement activities were introduced earlier in the programme.
- Staff delivered differentiated levels of support tailored to participant readiness.

Implementation strategies

To support the identified mechanisms, enhance enablers and address barriers, several strategies were also deployed or adapted:

- Additional staff training (in life coaching and motivational support) to ensure more consistency in one-to-one support.
- Workload protection was implemented for staff working on the programme.
- A new data system was introduced to reduce administrative burden.
- The leadership team introduced more feedback loops to monitor progress

Observed outcomes

Programme outcomes

Young people who had additional support needs and who took up the offer of a mentor became more motivated, demonstrating more confidence and willingness to consider a variety of employment opportunities, resulting in a higher number of successful job applications.

Implementation outcomes

- Fidelity improved through staff having more time to commit to providing one-to-one support.
- Acceptability improved due to higher satisfaction with the quality of support being provided.
- Reach improved due to the differentiated approach to support provision.
- Feasibility of delivery improved due to reduced caseloads and improved supervision.

Learning and refinement

Insights from the TIIPE led to a revised Theory of Change (ToC) that:

- Differentiated pathways for participants with varying levels of readiness.
- Clarified which mechanisms are central to success.
- Explicitly linked implementation strategies to implementation outcomes.

Chapter 2

Programme and implementation theory

What is programme theory?

Programme theory explains the underlying logic behind a programme: why and how it is expected to work, and who it is meant to help.

It shows how different parts of the programme are linked, based on the assumptions of those developing it. By making assumptions clear, programme theory helps evaluators test and refine these assumptions, supporting improvements to programme design and delivery.

There are different ways to develop and articulate programme theory. These have emerged over time, shaped by ideas from fields like realist evaluation, which highlight the importance of understanding context, how change happens, why change happens and for whom.

Over the past decades, programme theory has developed significantly. Early contributions such as those from Chen (1990) and Weiss (1997) laid the groundwork for theory-based evaluation, whilst further developments, particularly from realist evaluation (Pawson & Tilley, 1997) have introduced a greater focus on context and mechanisms.

Today, programme theory is used in a range of approaches, from traditional logic modelling to complex systems and realist-informed evaluations. As both theory and practice continue to evolve, programme theory remains central to understanding, designing and evaluating effective programmes.

The personal and practical experiences of key stakeholders (both those delivering and those receiving support through the programme) are crucial in shaping a credible and contextually grounded programme theory. Integrating lived experience with other key stakeholder insights, alongside careful consideration of existing research and evidence, further strengthens a theory's development and ensures its relevance and authenticity.

Programme theories are also informed by ideas from broader social theories to explain how and why a programme is expected to change people's behaviour. These theories help explain how individual actions, societal structures and economic systems interact.

For example, when trying to improve employment outcomes, social theories like Human Capital Theory or Social Exclusion Theory may help us understand the wider influences and contextual factors that determine whether someone finds and keeps a job.

Articulating programme theory with a Theory of Change (ToC)

ToCs are the most common tool used by programme developers and evaluators to articulate and represent programme theories. All Youth Futures' TIIPEs require a ToC.

ToCs are crucial for a TIIPE as they offer a clear framework for understanding how and why a programme is expected to achieve its intended outcomes. A clearly articulated and well-specified ToC is also crucial in supporting the design of a TIIPE evaluation framework or plan.

A ToC should include the programme's rationale, target population, context, assumptions, activities, resources, mechanisms, outcomes, impacts and the pathways between components **(see Table 1).**

Component	Definition
Problems (rationale)	The challenges currently experienced by people, communities, organisations or systems and how the programme is intended to address them. A ToC may also include a discussion of opportunities – these are the ways in which social conditions might be improved through the programme.
Target population	The people, groups, organisations or systems the programme is designed to benefit, support or influence.
Context	The set of circumstances, such as social, cultural, institutional, political, environmental or organisational, that interact with a programme's mechanisms to influence its implementation and outcomes.
Assumptions	The preconceived ideas, concepts and beliefs that underly the thinking about how a programme will achieve its aims.
Resources	The inputs required to deliver the programme (e.g. funding, staff time, equipment and materials).
Activities	The actions that are carried out as part of delivering the programme (e.g. training, mental health support).
Mechanisms	The underlying processes that bring about the desired change.
Short-, medium- and long-term programme outcomes	The short-, medium- and long-term changes expected to result from the programme.

Table 1: Summary of the definitions commonly used in ToC to articulate programme theory

Key programme theory constructs

When thinking about programme theory and its articulation through ToCs, there are four key constructs that are often described in different ways: outcomes, mechanisms, contexts and assumptions. The following sections explore these in more detail.

Programme outcomes

Programme outcomes are the changes the programme is intended to achieve. Within a ToC these changes are usually divided into **short-term, medium-term and long-term effects**. A well-specified ToC should also indicate the level at which these changes are intended to occur, for example, **at the individual, group, organisation or system level.**

A ToC should aim to cover all the different effects the programme is trying to achieve (as far as is possible). During programme implementation it is also important to identify and assess any unintended harmful outcomes, especially those that disproportionally impact marginalised groups and populations.

See more on this in Chapter 3, page 49.

Case study

Intended outcomes of a youth employment training and placement programme at multiple levels

Individual

- Short-term: Young people completing the course show greater motivation and clarity in job searching efforts and better performance in interviews and applications.
- Medium-term: Increased number of young people securing internships, apprenticeships or entry-level jobs in their chosen field.
- Long-term: Sustained employment or progression into higher-skilled roles.

Organisation level

- Short-term: Employers increase the number and quality of placement opportunities.
- Medium-term: Stronger relationships with local employers, training providers and community partners.
- Long-term: Sustained partnerships with employers' influence recruitment practices toward greater inclusivity.

System level

- Short-term: Inclusive recruitment and onboarding practices, informed by programme partnerships, are shared and replicated across local employers.
- Medium-term: Local partners co-ordinate training, employment and support services more effectively to meet young people's needs.
- Long-term: Local employment systems become more joined up and inclusive, supporting improved access and progression for young people.

Mechanisms

Mechanisms are central to TIIPE. They help explain how and why an programme produces change by exploring not just what happens, but why it happens in a particular context.

For the purpose of a Youth Futures' TIIPE, mechanisms should be conceived as the underlying processes triggered when participants engage with a programme.

Mechanisms have core features that make them distinct from activities or outcomes. Understanding these characteristics can help evaluators identify, test and refine mechanisms within a TIIPE.

Characteristic	Explanation
Hidden	**Mechanisms are often not directly visible. They are underlying processes (like motivation, beliefs or social norms) that drive change.**
Sensitive to context	**Mechanisms may work differently in different settings, and in unpredictable ways, i.e., the same mechanism might lead to different outcomes depending on the setting or group.**
Generate outcomes	**Mechanisms are the drivers of change that interact with context to produce observable outcomes. They are distinct from activities (what is delivered) and outcomes (what is observed).**

Table 2: Overview of the characteristics of programme mechanisms

Mechanisms can be described in varying levels of depth, depending on the stage and purpose of the evaluation:

01 Behaviours or participant reactions to programme components

At a basic level, mechanisms can be expressed as how participants respond to specific programme components. This includes changes in motivation, confidence, behaviour or understanding that help explain observed outcomes.

Case study

A youth employment training and placement programme

This example is used across the different chapters to explain different constructs. For background to this example see page 11

Skills and confidence building

Workshops build practical job-seeking skills in a supportive environment. In response, participants begin to **believe in their ability** to apply for jobs and succeed. This **increased confidence** leads to improved applications and performance in interviews.

Job opportunity awareness

Increased exposure to job leads and employer networks clarifies what opportunities are available and how to access them. This provides **greater motivation** and clarity in job searching efforts.

Mechanisms can be described in varying levels of depth, depending on the stage and purpose of the evaluation:

02 Mechanisms as causal accounts ("mini-theories")

In more developed theories of change, mechanisms are described as part of a causal chain integrating participants' reasoning, the processes of change and the context in which they occur. This approach aligns with realist-informed evaluation and often draws on Context–Mechanism–Outcome (CMO) configurations.

Case study

A youth employment training and placement programme

This example is used across the different chapters to explain different constructs. For background to this example **see page11**

Where a realist informed approach is used, the evaluator identifies the key mechanisms that bring about the desired outcomes, and considers the contexts that either support or inhibit them:

Example of a supportive context

C - In the context of small, supportive group workshops which help young people to set achievable goals, and provide them with mentoring support and feedback.

M - Young people increase both their clarity of goals and their sense of self-efficacy in achieving them.

O – Leading to higher quality job applications and performance in interviews.

Example of context as a barrier

C - In the context of large impersonal group sessions which are not tailored to the needs of individuals, where young people receive limited individual support and little direct feedback.

M - Young people do not feel clearer about their own goals and their sense of self efficacy does not increase.

O – Leading to little change in motivation or job search behaviour.

Context

A key part of creating or surfacing a programme theory is understanding how the setting or environment (the context) affects how the programme works.

A well-developed ToC should account for how different contextual factors may support or constrain the changes the programme aims to achieve.

Relevant categories of context to include in programme theory

Programmes operate within political, cultural and economic systems, which shape how well they work. Effectiveness also depends on the organisation delivering it and the people involved.

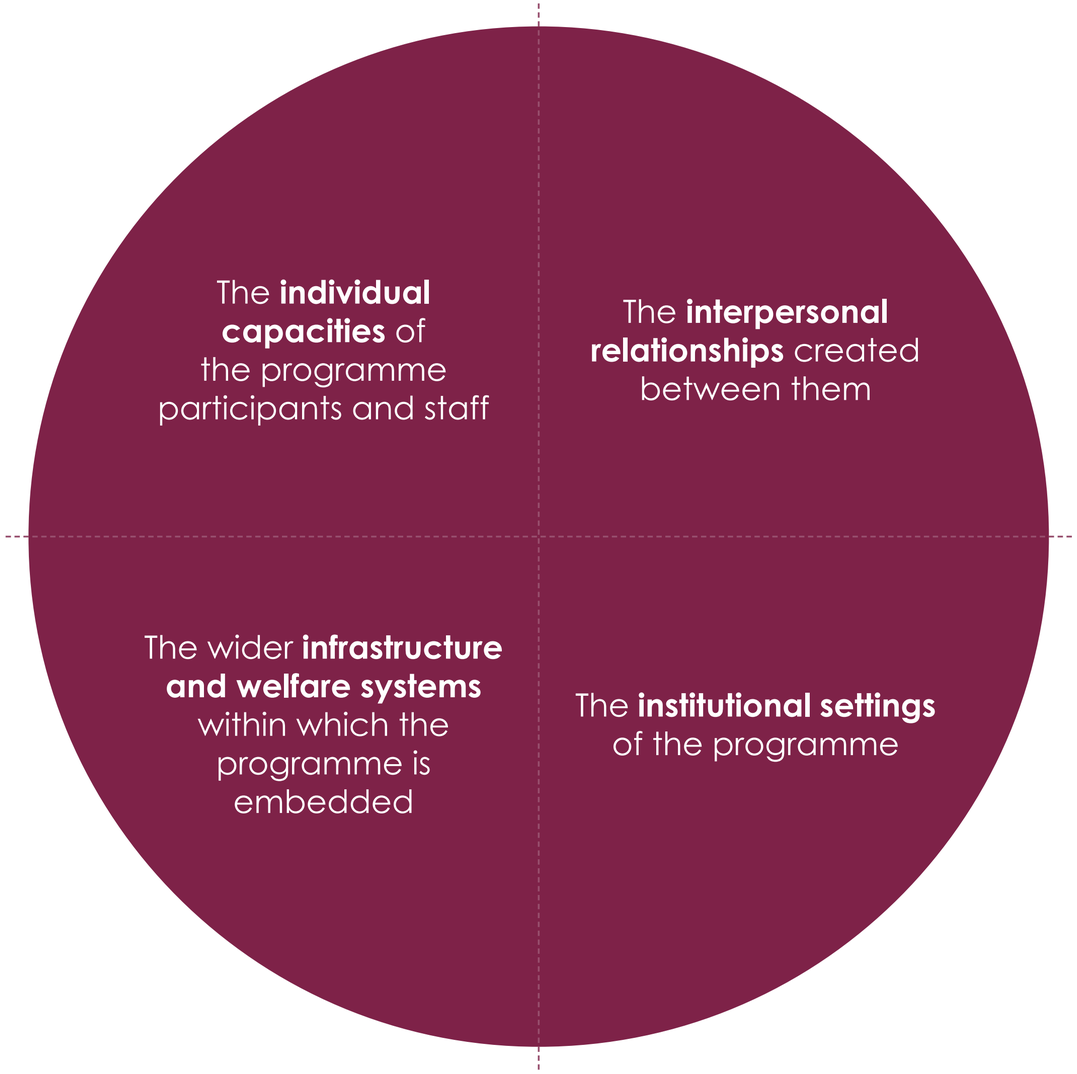


Figure 2: Pawson and Tilley's (2004) "four Is" of categorising aspects of programme context

Case study

Behaviours or participant reactions to programme components

This example is used across the different chapters to explain different constructs. For background to this example **see page 11**

Examples of context:

Individual (participants)
Level of prior education and qualifications, health and wellbeing, family and peer support, previous experience of education or work, housing stability.

Interpersonal (staff and participants' relationships)
Quality of relationships, level of trust, length and nature of time spent together.

Institutional (provider)
Resources and capacity, strategic priorities, partnership networks, staff expertise and core values.

Infra-structural
Local labour market conditions, policies, economic climate and funding stability.

Time is also part of context, and a ToC should consider how the relationship between the programme and its surroundings might change over time.

This is especially important in programmes aiming to change systems, where feedback loops or earlier results create new situations for later stages, sometimes called "ripple effects". These are more flexible or dynamic ways of thinking, where context doesn't just affect the programme but instead changes with it, which can lead to unexpected outcomes.

Case study

A youth employment training and placement programme

This example is used across the different chapters to explain different constructs. For background to this example **see page 11**

Ripple effect:

Participants who found jobs begin sharing tips and resources with peers who didn't take part in the programme.

Employers who hired from the programme start offering similar placements to other local young people.

Community organisations take interest and begin adapting their own youth support services to reflect what's worked.

Local schools and colleges begin to integrate employability skills into their curriculum.

So now, the early outcomes (trained young people getting jobs) become part of the new context by shaping how others in the community act, even though they weren't part of the original programme.

Assumptions

Assumptions are the ideas, concepts and beliefs of designers, implementers, evaluators and other stakeholders involved in creating a ToC that underly the logic of a programme, and the conditions required for it to be effective.

In practice, there are underlying assumptions in relation to every component in a ToC and the links between these components. It is important to make these assumptions explicit so that they can be tested and weaknesses in the programme logic identified.

Assumptions are usually surfaced through group discussion and the development of a shared understanding of how the programme is expected to deliver the desired change. Testing assumptions with different stakeholder groups (such as participants, delivery partners and funders) clarifies expectations of how, for whom and under what conditions a programme will be effective.

Case study

A youth employment training and placement programme

This example is used across the different chapters to explain different constructs. For background to this example **see page 11**

Programme assumptions

Examples of assumptions that programme designers, delivery partners and evaluators might make:

- Participants feel motivated to seek employment and engage with the support offered.
- Job placement opportunities are meaningful and accessible for the intended participants.
- The chosen venue is accepted and accessible for young people.
- Local employers are willing and able to offer entry-level jobs to young people with limited work experience.

→
The next chapter will offer guidance on how to put these constructs into practice when developing and refining a programme's ToC **(see Chapter 3).**

What is implementation theory?

TIIPEs should explore not only the theory of how the programme will lead to its intended outcomes (programme theory) but also the theory of how the programme is expected to be effectively delivered (implementation theory).

Implementation theory is the theory about why implementing the programme in a certain way will mean that it is delivered well, so that it can achieve the programme outcomes.

Implementation refers to the specific intentional activities by which the programme is put into practice.

There are four key concepts within implementation theory:

1. **Implementation outcomes:** what effective delivery "looks like".
2. **Determinants:** factors that support (enablers) or hinder (barriers) implementation.
3. **Strategies:** actions taken to strengthen enablers and mitigate barriers.
4. **Implementation mechanisms** explain how and why implementation strategies work.There are many frameworks that can support the exploration of these concepts. We briefly describe two commonly used implementation frameworks in the next section – the Consolidated Framework for Implementation Research (CFIR) and the Theoretical Domains Framework (TDF).

Using a framework helps ensure that examination of implementation is thorough and reflects established evidence about effective implementation. A framework establishes consistency in the aspects of implementation assessed across evaluations, both in their conceptualisation and the language used.

Implementation outcomes

Implementation outcomes are distinct from programme outcomes.

Whereas programme outcomes are the changes expected to result from the programme, implementation outcomes are the changes (for people or organisations involved in delivery) that indicate the programme is being delivered well. They are essential precursors for the programme to achieve its intended impacts. Implementation outcomes are indicators of whether the implementation is successful and offer insight into the implementation process. If a programme is not implemented well (even if it is grounded in robust evidence) it will not be effective in bringing about the desired change.

Case study

A youth employment training and placement programme

This example is used across the different chapters to explain different constructs. For background to this example **see page 11**

Implementation outcomes

The implementation outcomes might include:

- the programme being taken up by a high proportion of the sites it is offered to.
- the programme reaching the intended target group, particularly those who would most benefit from the support offered.
- the programme being delivered with quality and fidelity, for example, the intended number of sessions, to a pre-established standard.
- the programme fitting well into organisational structures, processes and roles.

These are distinct from programme outcomes, which might include an increased number of young people doing more job-seeking activity, as well as getting and sustaining high quality jobs.

The specific implementation outcomes that are most relevant will depend on the stage and maturity of the programme.

Figure 3 outlines commonly measured implementation outcomes as outlined in Glasgow et al. (1999) and Proctor et al's (2011) seminal papers.

Figure 3: Commonly measured implementation outcomes (Glasgow et al.,1999) and (Proctor et al., 2011)

What are the types of implementation outcomes?

Table 3 describes commonly measured implementation outcomes in more detail and why each outcome is important to consider in a TIIPE. It can be used as a checklist of aspects to consider in an evaluation design.

Table 3: Implementation outcomes overview

Outcome	What is it?	Why does it matter?	What factors may impact this outcome?
Adoption	Whether the programme is taken up by programme staff or an organisation as a whole.	A programme cannot achieve the intended outcomes if few organisations or staff adopt it.	• A programme might be adopted by senior leaders but not by delivery staff – consider shared intent. • Adoption can be affected by alignment with organisational values, motivation and readiness of staff, resources, and how it fits with existing initiatives.
Acceptability	The extent to which the programme is acceptable to those who are involved (young people, delivery staff).	If a programme is not acceptable to young people and/or staff, they are unlikely to participate or put the effort required into delivering it well.	• Acceptability may change over the course of a programme's implementation. • It may vary substantially across dif-ferent groups and contexts.
Reach	The degree to which the intended target group is engaged in the programme.	Understanding reach indicates: • The inclusiveness of the programme, particularly whether it reaches. priority or marginalised groups • The feasibility and potential to scale the programme.	• Participant motivation and the programme's perceived relevance to their goals (e.g. securing a job, building practical skills). • Delivery staff capacity and continuity – skilled, stable staff are critical to initial and sustained engagement. • Accessibility and fit with participants' lives – reach is strongly influenced by whether young people can realistically attend and participate.
Appropriateness	Whether the programme is perceived to be a good fit, relevant to, or compatible with: • the setting, issue or concern being addressed. • the delivery staff or young people who are participating.	• Programmes perceived to be a good fit with the organisation are more likely to be adopted by staff and leaders. • Long-term sustainability and scalability is impacted by perceived appropriateness.	A programme might be evidence-based or appropriate for the delivery context but still perceived as inappropriate if misaligned with organisational or cultural values or norms.

Table 3: Implementation outcomes overview (cont.)

Outcome	What is it?	Why does it matter?	What factors may impact this outcome?
Fidelity	Whether the programme is implemented as it was originally designed or intended to be delivered. Fidelity includes: • Adherence: are staff able to adhere to programme protocols and deliver as intended? • Dosage: can the required amount or intensity of the programme be delivered (duration, # of sessions, etc.)? • Quality: can the programme be delivered to the quality required, and how is this achieved? • Engagement/responsiveness: to what extent are programme recipients engaging with the programme and responsive to it? • Differentiation: is the programme clearly different from other programmes or activities?	Without consistency and quality of delivery, even an effective, evidence-informed programme can fail to achieve the intended outcomes.	Fidelity is connected to other implementation outcomes. For example, aspects of pro-gramme fidelity may make the programme more or less acceptable or feasible.
Feasibility	The extent to which the programme can be successfully delivered in a particular setting or context with the intended group of young people. Important considerations for feasibility include: • The resources required for delivery and their availability. • Whether staff have the necessary skills and time to deliver it. • Training and other additional supports that staff may require to deliver the programme well. • Other organisational supports and systems that will facilitate the programme implementation (e.g., IT systems, monitoring systems, physical space).	Assessing feasibility is critical for understanding whether a programme is practically deliverable.	• Delivery partner capacity, skills and resourcing – even promising programmes fail if staff do not have adequate training, manageable caseloads or infrastructure, or if the programme does not fit well with existing work processes.
Penetration	The extent to which the programme is integrated into routine operations or other services offered within an organisation.	Penetration is an important consideration when thinking about the potential for sustainment and scalability of programme delivery.	• A higher degree of penetration may con-tribute to the long-term sustainability of a programme. • Penetration is also linked to adoption – if a programme is adopted by an organisa-tion and its staff, it is more likely to be in-tegrated into an organisation's existing delivery.
Implementation cost	The cost of programme implementation involves two elements: 1. How much the programme itself costs i.e., financial and resource costs, and cost of delivery in different settings 2. The cost of the implementation strategies used to implement the programme i.e., training, coaching etc.	Understanding the cost of programme implementation aids decision-making about viability, sustainability and scalability of implementation.	• The more complex the programme and the implementation strategies required, the higher the costs.
Sustainability	The extent to which the programme continues to be delivered over time, beyond the initial delivery period and after initial implementation supports and efforts have ended.	It is not enough for a programme to be delivered during an evaluation – the key question is whether relevant or-ganisations can continue to deliver it well. Sustainability is an indicator of the extent to which a programme has become or could become "institu-tionalised" within a particular setting.	• Ongoing funding and policy alignment – programmes that align with local and national priorities are more likely to attract continued investment • Programmes are more sustainable when they are embedded in host organisations' existing workflows and there is a collective sense of ownership

Determinants

Implementation determinants are the factors – barriers and enablers – that influence implementation and the achievement of desired outcomes.

These factors operate at multiple levels and may relate to the programme itself (e.g. its credibility or complexity), the intended target population (e.g. their needs and preferences), the delivery organisation (e.g. its culture, processes, networks) or the wider social or political context (e.g. community preferences, social norms, policies and funding).

While some determinants may be modifiable (e.g., leadership engagement, readiness of staff to deliver the programme) some are harder to change (e.g., the community context in which the programme is delivered, or the policy environment). It is important to think broadly and comprehensively about determinants (enablers and barriers, modifiable and non-modifiable).

Determinants frameworks

There are several different frameworks that can be used to help identify and conceptualise the determinants that may have an influence on programme implementation activities. Determinant frameworks have been developed based on extensive reviews of evaluations and draw on underlying theories about how change occurs. Frameworks can also be used to inform evaluation questions and organise identified determinants.

Table 4 summarises two of the more commonly used implementation determinant frameworks. Chapter 4 (Developing and testing implementation theory) provides more detail on when you would use each framework.

Characteristic	Consolidated Framework for Implementation Research (CFIR) (Damschroder et al., 2022)	Theoretical Domains Framework (TDF) (Atkins et al., 2017)
Focus	Barriers and enablers that influence implementation outcomes. Provides a comprehensive understanding of the factors influencing implementation success.	Behavioural and psychological factors influencing individuals' behaviour in implementation. Helps to identify specific psychological and behavioural barriers or enablers to implementation.
Level of focus	Individual, organisational and wider contexts.	Primarily focused on the individual and their behaviour.
Key constructs	5 domains: Programme (the attributes of what is being implemented); Outer Setting (external influences on implementation); Inner Setting (organisational context where implementation happens); Individuals (people involved in implementation); Implementation Process (how implementation is planned and executed).	14 theoretical domains (e.g., knowledge, skills, beliefs, motivation, etc.).
Examples of constructs	Relative advantage (of programme), available resources (in organisation), policies and laws.	Knowledge, skills, social influences, beliefs, reinforcement.

Table 4: Overview of two determinant frameworks: Consolidated Framework for Implementation Research (CFIR) and Theoretical Domains Framework (TDF)

Case study

A youth employment training and placement programme

This example is used across the different chapters to explain different constructs. For background to this example **see page 11**

CFIR-organised implementation determinants

CFIR domain	Example CFIR determinants	Determinants applied to a youth employment training & placement programme
Intervention characteristics (Attributes of what's being implemented)	• Relative advantage • Evidence base • Adaptability • Complexity • Cos	• Perceived relevance of this training programme compared to other employment programmes • Participants' and partners' perception of programme effectiveness • Extent to which modules can be tailored for different sectors • Complexity (length of course, admin burden) • Cost to implement
Outer setting (External influences on implementation)	• Local conditions • Policies and laws • Partnerships and connections • Local attitudes	• Local labour market demand and availability of placements • Policy incentives for hiring young people • Referral pathways (e.g., Jobcentre Plus, colleges, youth centres) • Community perceptions (e.g., stigma or support for employment schemes)
Inner setting (Organisational context where implementation happens)	• Available resources • Mission alignment • Relational connections • Communications • Compatibility • IT infrastructure	• Local labour market demand and availability of placements • Policy incentives for hiring young people • Referral pathways (e.g., Jobcentre Plus, colleges, youth centres) • Community perceptions (e.g., stigma or support for employment schemes)
Characteristics of the individuals (People involved in implementation)	• Individuals have competence, knowledge and skills to fulfil their role • Individuals have availability, scope and power to fulfil their role	• Perceived relevance of this training programme compared to other employment programmes • Participants' and partners' perception of programme effectiveness • Extent to which modules can be tailored for different sectors • Complexity (length of course, admin burden) • Cost to implement
Implementation process (How implementation is planned and executed)	• Planning, including participatory approaches • Engaging with opinion leaders, champions • Reflecting and evaluating	• Co-design and stakeholder engagement in tailoring training modules • Fidelity monitoring processes • Post-placement follow-up and sustainment planning

Link between determinants in implementation theory, and context in programme theory.

Implementation determinants and contextual considerations are closely related: both describe the conditions under which a programme is delivered and how these conditions shape delivery and outcomes. In programme theory, context often refers broadly to the structural, organisational and social conditions in which a programme is embedded (e.g., workforce capacity, policy environment, population needs). Implementation determinants (e.g., as described in CFIR) are aspects of context, drilling down further into specific factors that enable or hinder delivery in a particular setting – such as leadership engagement, available resources and participants' readiness for change.

The key difference is that context in the case of implementation theory is delivery-focused, whereas programme theory considers a wider range of contexts which may impact effectiveness, irrespective of whether or not the programme is delivered well.

Although, in practice, the same issue might be relevant as context in both programme theory and in implementation theory.

The two constructs (determinants and context) align in that both help us understand "for whom and under what circumstances" a programme works. Determinants in implementation theory are typically framed as modifiable levers that can be targeted (i.e., through implementation strategies – see next section below). In contrast, contextual factors in programme theory are more often viewed as descriptive or fixed, and as likely to moderate programme effects. However, programme theory also recognises the dynamism of context, as illustrated by the notion of "ripple effects" – where intentional changes to key determinants can trigger wider, sometimes unanticipated impacts across the system, shaped in part by the surrounding context.

Implementation strategies

Implementation strategies are the methods or activities used to achieve and enhance implementation. They should be selected to address identified determinants: mitigating barriers to implementation of a programme or catalysing enablers to implementation.

The ERIC framework (Powell et al., 2015) identified 73 discrete strategies, which were later organised into eight overarching categories (or clusters) based on concept mapping and expert consensus (Waltz et al., 2019). These clusters help researchers and practitioners navigate the wide range of strategies available. Some commonly used implementation strategies are mapped onto the clusters on this page to illustrate how they function in practice.

Use evaluative & iterative strategies

Purpose: to assess and refine implementation processes or outcomes.

Example: audit and feedback – provide regular, structured feedback to staff based on observations of programme delivery to improve quality and fidelity.

Provide interactive assistance

Purpose: to offer personalised hands-on support during implementation.

Example: helplines or help desks, provide on-demand assistance for staff encountering challenges.

Adapt & tailor to context

Purpose: to modify the programme or implementation strategies to fit the local context.

Example: contextual adaptation, modifying training materials to make them culturally relevant.delivery to improve quality and fidelity.

Develop stakeholder interrelationships

Purpose: to build connections and partnerships with key groups.

Example: champions, identify and empower individuals to advocate for the programme.

Train & educate stakeholders

Purpose: to improve stakeholder knowledge skills and confidence.

Example: training workshops, provide in-person or virtual sessions to build competency in delivering the programme.

Support individuals or teams

Purpose: to motivate or reinforce staff and team capacity during implementation.

Example: incentives or rewards, recognise individuals or teams for achieving implementation goals.s

Engage consumers

Purpose: to involve end-users (young people families communities) in the implementation process.

Example: co-design workshops, engage end-users in shaping programme materials or delivery.

Change infrastructure

Purpose: to adjust organisational systems and structures to support implementation.

Example: introduce new IT systems, digitise data tracking to improve programme monitoring.

Case study

A youth employment training and placement programme

This example is used across the different chapters to explain different constructs. For background to this example **see page 11**

CFIR-organised implementation determinants

CFIR domain	Determinants applied to a youth employment training and placement programme	Potential implementation strategies
Intervention characteristics	Perceived relevance of training to participants' career goals	• A higher degree of penetration may contribute to the long-term sustainability of a programme. • Penetration is also linked to adoption – if a programme is adopted by an organisation and its staff, it is more likely to be integrated into an organisation's existing delivery. • The more complex the programme and the implementation strategies required, the higher the costs.
Outer setting	Local labour market demand and availability of placements	• A higher degree of penetration may contribute to the long-term sustainability of a programme. • Penetration is also linked to adoption – if a programme is adopted by an organisation and its staff, it is more likely to be integrated into an organisation's existing delivery. • The more complex the programme and the implementation strategies required, the higher the costs.
Inner setting	Leadership engagement and prioritisation	• A higher degree of penetration may contribute to the long-term sustainability of a programme. • Penetration is also linked to adoption – if a programme is adopted by an organisation and its staff, it is more likely to be integrated into an organisation's existing delivery. • The more complex the programme and the implementation strategies required, the higher the costs.
Characteristics of the individuals	Motivation and readiness of young people	• Provide motivational interviewing at intake • Use peer mentors or alumni ambassadors • Offer milestone-based incentives (e.g., gift cards at course completion)
Implementation process	Fidelity monitoring processes	• Provide motivational interviewing at intake • Use peer mentors or alumni ambassadors • Offer milestone-based incentives (e.g., gift cards at course completion)

Implementation mechanisms

Implementation mechanisms are the causal processes or events through which an implementation strategy produces its effects on implementation outcomes (e.g., on adoption, fidelity, sustainability).

Strategies can be thought of as the "intervention" for implementation. Implementation mechanisms explain how and why they work.

They bridge the gap between:

- **The implementation strategy** (what we do to improve implementation) and
- **The implementation outcome** (what changes as a result, such as increased fidelity or reach)

To illustrate an example of an implementation mechanism within an implementation pathway, here is a simplified causal chain:

Determinant

Staff report low confidence and reluctance to provide honest feedback due to fear of blame.

→

Implementation strategy

Leadership engagement – senior leaders actively model openness, encourage feedback and demonstrate non-punitive responses to mistakes.

→

Implementation mechanism

This approach increases staff psychological safety, creating an environment where practitioners feel comfortable asking for help, raising concerns and learning from errors.

→

Implementation outcome

Higher psychological safety enables staff to adhere more closely to programme guidelines, ask for clarification when needed and share best practices – all of which enhance fidelity of delivery.

→

Programme outcome

Improved fidelity contributes to better participant outcomes.

Researchers have not yet theorised mechanisms for the 70+ ERIC strategies, nor identified more objective measures of implementation mechanisms. To address this gap, they can draw on complementary implementation frameworks (e.g., the TDF, outlined earlier).

Such frameworks can help to unpack potential mechanisms that underpin behaviour change that an implementation strategy might activate. In parallel, the CFIR can help researchers identify relevant determinants (e.g., contextual barriers or enablers) that shape whether those mechanisms are activated in practice. Used together, these frameworks can provide a way to theorise both what an implementation strategy is targeting and how context might influence its effects.

Table 5: Overview of the characteristics of implementation mechanisms

Characteristic	Explanation
Dynamic and context-dependent	Implementation mechanisms can vary over time and work differently across settings, influenced by organisational culture, resources and the stage of implementation.
Multi-level in operation	They function at various levels (e.g., individual, team, organisational, system).
Often latent or inferred	Mechanisms are not always directly observable – they are typically inferred through proxy indicators (e.g., perceived norms, workflow changes, self-efficacy).
Assessed primarily by self-report	Mechanisms are usually measured through surveys, interviews or focus groups, rather than direct observation or objective measures.
Link strategies to outcomes	They explain how a chosen implementation strategy brings about a change in implementation outcomes (e.g., adoption, fidelity).

Considering programme theory and implementation theory

Programme theory (PT) and implementation theory (IT) are often developed in parallel, but evaluators benefit most when they are used as complementary, interconnected frameworks. Programme theory explains what is expected to change for the young people who are programme participants and why, whereas implementation theory explains what is required for effective delivery to happen in practice, under what conditions and with what supports.

Combined, they enable evaluators to understand both what works and what it takes to make it work in practice. When combined, they create a more complete picture that clarifies both the causal chain of outcomes and the conditions for successful delivery.

Determinants (IT) and context (PT)

Both highlight the conditions shaping delivery and the achievement of programme outcomes (e.g., leadership support, staff capacity, local systems). Considering them together helps evaluators judge where the programme is most likely to work and how the programme, or aspects of the delivery context, may need to be adapted.

Activities (PT) and implementation outcomes (IT)

Implementation outcomes (e.g., fidelity, reach, feasibility) provide evidence on whether the intended programme activities were delivered as intended, and whether they should be modified or scaled.

Activities (PT) and implementation strategies (IT)

Implementation strategies inform what needs to be done (e.g., training, supervision and coaching of practitioners) so that programme activities can be delivered.

Programme mechanisms/ outcomes (PT) and implementation strategies (IT)

Evidence on whether programme mechanisms were activated helps identify whether the implementation strategies were sufficient (e.g. whether practitioner training enabled them to coach young people in ways that activated the intended mechanisms, leading to improved employment outcomes).

Map the overlaps early

When developing a ToC, explicitly note, for example, where programme activities depend on implementation strategies, or where programme mechanisms may be constrained or enabled by context/ determinants. Evaluators can develop ToCs that cover both programme theory and implementation theory constructs (see section in **Chapter 4** for more detail).

Frame evaluation questions across both theories

For example, not just "Did self-efficacy improve?" but "Did training staff (strategy) lead to higher fidelity in delivery of a programme component (implementation outcome), which in turn meant that staff's coaching of young people activated young people's self-efficacy (programme mechanism)?"

Align data collection

Ensure that data on implementation outcomes (e.g., fidelity logs, acceptability ratings) and programme outcomes (e.g., participant measures) are collected in ways that allow them to be linked so that you can explore whether implementation outcomes are associated with programme outcomes.

Interpret findings in an integrated way

Use evidence on implementation to interpret programme effects: lack of change may be due to weak implementation, not programme failure.

Table 6: Overlapping areas between programme theory and implementation theory

Dimension (PT/IT)	Overlapping area	What to explore	Why it matters
Context (PT) / determinants (IT)	Shared conditions that shape delivery	Which contextual factors (fixed/dynamic) most influence delivery? Which are modifiable via implementation strategies?	Clarifies where and for whom the programme is most likely to work, and what supports are needed
Activities (PT) / implementation outcomes (IT)	Delivery of programme content	Were programme activities delivered with fidelity, acceptability, feasibility?	Explains whether outcomes (or lack of) reflect delivery quality or programme content.
Activities (PT) / implementation strategies (IT)	Supports for delivery of programme activities	What strategies enabled staff to deliver activities effectively (e.g., did training or coaching of staff enable them to deliver programme sessions well?)	Links what is delivered with how it is supported, guiding optimisation.
Programme mechanisms & outcomes (PT) / implementation strategies (IT)	Testing strategy-mechanism links	hich strategies were essential to activate mechanisms?	Helps refine the implementation support package and scale-up guidance.

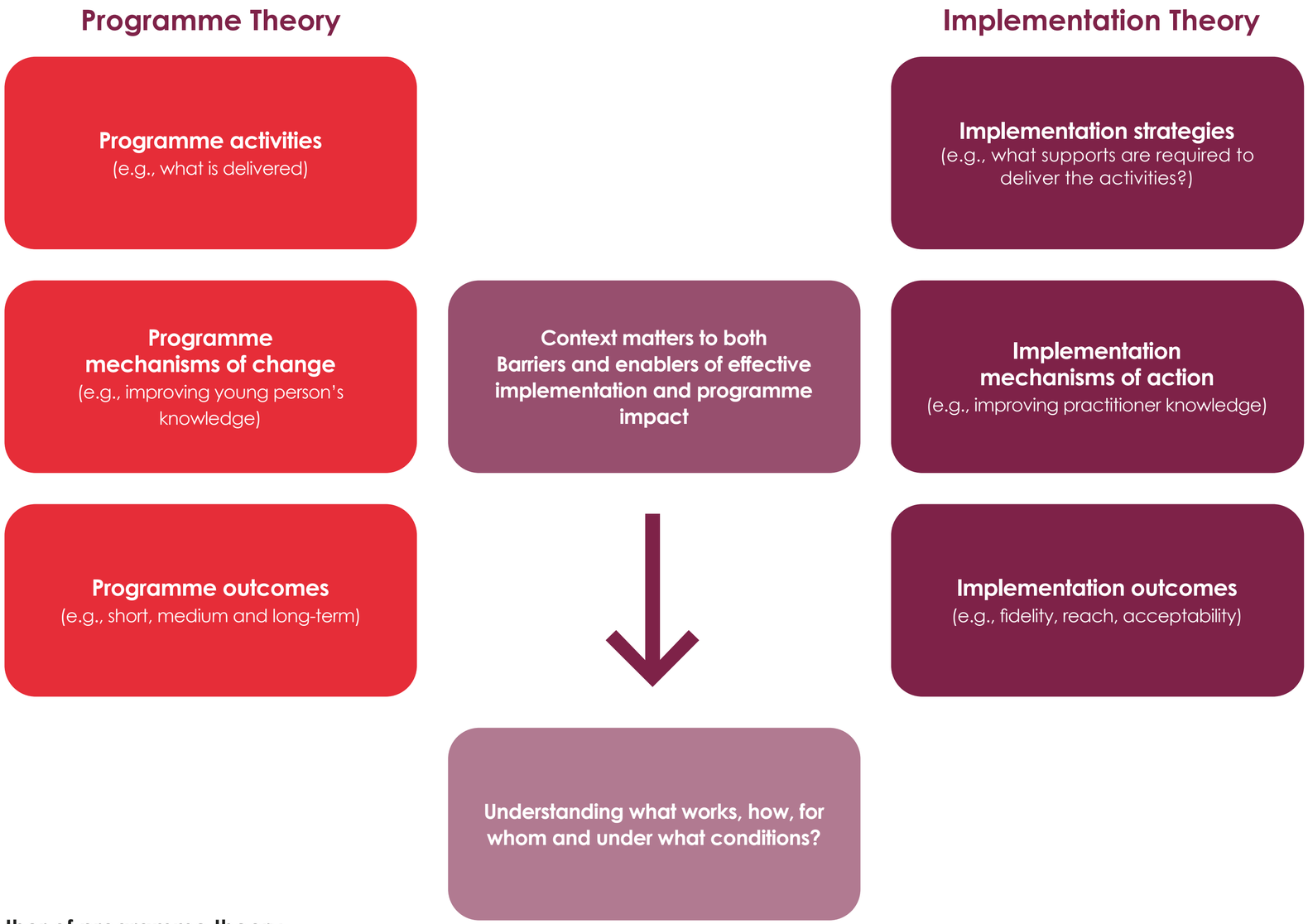


Figure 4: The bringing together of programme theory and implementation theory in a TIIPE

Chapter 3

Developing and refining a ToC

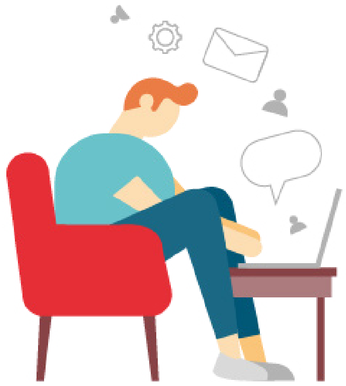
youth futures FOUNDATION

What is programme theory?

This chapter explores how programme theory is developed and tested. Chapter 4 then discusses how implementation theory is developed and tested. Although the focus in this chapter is on programme theory, some references to implementation are included to show that programme and implementation theory need to be understood together.

Articulating and documenting the theory underlying a programme is key to understanding whether and how it works. Without bringing a theoretical lens to evaluation, evaluators are reduced to knowing about change in a single time and place, with the learning difficult to generalise or apply elsewhere. Without a ToC, implementation could become "an expensive version of trial and error" (Nilsen., 2015.)

Features of a good ToC



Explicative

Provide explicit theories to explain how programmes are intended to bring about change.

It should answer the fundamental questions about what a programme aims to achieve, for whom and how.

Detailed

Have sufficient detail and specificity on causal links, assumptions, activities, contextual factors and expected outcomes.

It should include all the ToC components explored in **Table 1,** and a detailed narrative accompanying any visual. Assumptions and contextual factors should be explicit and testable.

Robust

Well-articulated, credible, plausible and logical (Mayne., 2015).

It should clearly indicate the pathways between activities, mechanisms and outcomes. These should be informed by different forms of evidence including lived and practice experience. These sources should be clearly outlined and explained.

Different levels of a ToC

The level at which theory should be developed will be determined by the context in which change efforts are expected to occur. Most of the impact evaluations commissioned by Youth Futures tend to operate at the micro and meso levels.

On complex projects, a series of ToCs may be developed for one project. This can help to capture complexity in different ways (e.g. through creating different ToCs for each sub-theme or objective, or by creating ToCs for different levels of analysis). In this instance, programme-level theory might represent a nested theory within a higher-level organisational (meso) or system (macro) level ToC.

Table 7: Different levels of a ToC

What	Focus	Example
Micro	Programme level	Individual programme: Youth Employability Skills Programme.
Meso	Organisational level	Cross-organisational strategy or approach: Youth Futures' organisational strategy to tackle youth unemployment.
Macro	System Level	Policy systems: national policy to address the root causes of unemployment.

Building a ToC

There are two key ways in which a ToC may be developed:

- **Starting with the problem/opportunity that the programme aims to address: Forward reasoning**

  Starting with the problem/opportunity that the programme aims to address: Forward reasoning In this instance, there may be a complex problem that could be solved in many different ways. Stakeholders will need to agree on what the problem is, the possibilities for addressing the problem, and what can be done given any existing resource constraints.

- **Starting with what success looks like (i.e. the desired outcomes): Backward reasoning**

  In this instance, stakeholders will need to discuss how a new programme can achieve the desired outcomes, or how an existing programme may contribute to the desired outcomes.

In earlier stages of evaluation, particularly the intervention design stage, either approach could be used. In the feasibility, pilot and impact evaluation stages, the second approach is likely to be more appropriate because the programme activities and outcomes are more likely to be defined.

The programme ToC is likely to be developed in the intervention design phase of Youth Future's programmes and refined in later evaluations. In rare cases, the ToC may be developed in later evaluation stages.

When developing (or refining) a ToC there are some key questions to consider that can help define the problem and context, describe the programme, examine assumptions and mechanisms, and clarify expected outcomes, impacts and causal pathways. See Chapter 1, Table 1, page 16 for more detailed definitions of each ToC component described below.

Note: Questions are presented in the present tense; however, when refining and evaluating a ToC, consider some of them in the past tense where appropriate.

Table 8: Questions to consider when developing a ToC

ToC component		Key Questions
Defining problem and context	Understand root causes	• What is(are) the cause(s) of the problem(s) identified?
	Define the problem Describe the context	• What is(are) the cause(s) of the problem(s) identified? • What is the problem, need or opportunity the programme addresses? • How might the broader social and economic conditions determine or shape young people's experience of these issues? • How might the individual capacities of programme participants and staff affect how the programme is received or functions? • How might the interpersonal relationships between programme participants, local organisations, institutions and communities affect how the programme is received or functions? • How might the institutional setting of the programme affect how it is received or functions? • How might the infrastructural systems (e.g. economic conditions, labour market, service availability) and welfare systems within which the programme is embedded affect how the programme is received or functions? • Which aspects of context are key in activating the mechanisms of change? • Which aspects of context are likely to be important for making decisions in relation to transporting the programme to new areas or scaling up (if applicable)? • How may context change over time, and what are the implications of any changes for programme outcomes and sustainability?
Describing programme features	Describe the programme	• What is the programme? • Who is the target population? • What are the activities involved? • What are the resources required?
	Consider assumptions	• What are the assumptions (preconceived ideas, concepts and beliefs) programme deliverers and other stakeholders are making about linear causality between different components of the programme? • What are the underlying hypotheses and preconditions associated with different elements of a ToC? • Note: Surfacing and documenting assumptions is a continuous process and shouldn't just be limited to one initial conversation. By regularly revisiting assumptions throughout the evaluation, you can help ensure they remain relevant and significant to the programme.

Table 8: Questions to consider when developing a ToC (cont.)

ToC component		Key Questions
Describing programme features	Describe mechanisms	• How will the programme achieve its aims? What are the key ingredients of the programme that will bring these changes about? • How are the activities from the programme (or the new resources it brings) and the ways they are implemented by practitioners, departments, organisations and systems expected to trigger changes? • Why are they expected to trigger these changes? • At what level of depth is the description of the mechanism required (e.g. "mechanism as participant response" or "mechanism as causal account"?)
	Identify programme outcomes	• What are the expected short-, medium- and long-term changes for participants or recipients? • What are the expected short-, medium- and long-term changes for practitioners, departments, organisations or systems?
	Identify programme impact	• What is the ultimate goal of the programme? • What change will programme participants ultimately experience in their lives?
	Clarify causal pathways	• What is the ultimate goal of the programme? • What change will programme participants ultimately experience in their lives? • How are different components of the theory connected to each other? • How are programme activities connected to its outcomes? Note: Despite potential similarities, causal pathways show the steps of change, whilst mechanisms explain the reasons behind the change.

It would be impossible for a TIIPE to explore all aspects of context, so evaluators need to decide which are the most important aspects to focus on through engagement with stakeholders and alongside a review of the literature and existing evidence base.

Considerations might include the underlying social theories that a programme theory is based upon (which might emphasise organisational or individual behaviour, for example) and the level at which an evaluator is seeking to determine impact (individual, system or population level).

It is important for TIIPE to focus on the contextual conditions that seem to be most significant in activating the mechanisms of change, and those that will support policy makers to understand how to successfully transport the programme to new areas (scaling up).

Evaluators using TIIPE also need to consider how context may change over different stages of evaluation, and the implications of any changes in context for programme outcomes and sustainability.

Good practice principles in developing a ToC

Youth Futures recommends following the good practice principles when developing a ToC for a TIIPE.

Good practice principle	Description
Co-developed with stakeholders	It should be developed consultatively to reflect the understanding of all relevant stakeholders (see stakeholders list, Table 10). Considerations of equity should be embedded in this stage.
Grounded in evidence	Mechanisms may work differently in different settings, and in unpredictable ways, i.e., the same mechanism might lead to different outcomes depending on the setting or group.
Grounded in evidence	It should be grounded in robust evidence. In addition to stakeholder input, this evidence base should include a review of the existing literature on the topic and available data (such as service and monitoring data, and populational level data). See section below on integrating evidence to support ToC development.

Table 9: TIIPE good practice principles for developing a ToC

Integrating evidence to support

The development and refinement of a ToC should be grounded in existing literature on the topic and available data (such as service and monitoring data and population level data) as well as the experiences of young people and organisations involved in early delivery.

Evaluators should consider the existing evidence to support each component of a ToC. The prompts below can help evaluators to build, or strengthen, a ToC using existing evidence:

Use the evidence collected to justify the programme and to question, test or revise assumptions, causal pathways and outcomes in the ToC.

Understanding the problem, need or opportunity and their root causes:

- How is the problem, need or opportunity described? What does existing research say about the nature, scale and impact of the problem or need? How is the programme positioned in response to this issue?
- What historical, political, social, economic and cultural factors have contributed to the problem, need or opportunity?
- How have these structural and contextual conditions developed over time to contribute to the problem, need or opportunity?

Describing the programme:

- Has this or a similar programme been implemented elsewhere?
- What were the programme **goals** of similar programmes, and how were they operationalised to contribute to outcomes?
- Who was the **target population**, and did the programmes result in positive outcomes or impacts for that group? Why or why not?
- What activities **were** included, and how did they contribute to outcomes?
- What **resources** (e.g., funding, partnerships, expertise) were required, and why were they critical to success?
- In what **context** were the programmes implemented, and how did that context influence delivery and outcomes?

Outcomes and impact:

- What outcomes have been achieved in similar programmes? What conditions supported or hindered these outcomes?
- What longer-term changes or impacts can reasonably be expected, based on findings from other programmes? How does this inform programme expectations and assumptions?

Mechanisms:

- What are the common mechanisms of change described in the literature associated with similar programmes (i.e., the hidden processes by which change occurs)?

Stakeholders involved in ToC development, testing and refinement

It is important to involve a wide range of stakeholders, including those designing, using, delivering and funding the programme, in the development and refinement of a ToC.

Ongoing consultation with diverse groups enhances the ToC's accuracy by ensuring it reflects a fully rounded picture of the programme. It also improves buy-in for evaluation, and increases the possibility that findings are useful to, and used by, professionals, practitioners and funders.

All those involved should be heard and have their views and concerns considered and appropriately reflected in the ToC. **Table 10** illustrates what these may include, but are not limited to.

Table 10: Roles of stakeholder involvement in ToC development, testing and refinement

Characteristic	Explanation
Professionals/ practitioners involved in the design, delivery and implementation of the programme	They have a deep understanding of the programme's components and delivery methods, as well as the context into which it is introduced. Experience across projects provides valuable insight into how these approaches typically work in practice.
Professionals, community members and any other people who might benefit from the programme	They bring diverse lived experience of the issues that a programme aims to address. They also understand first-hand how people experience and interact with different parts of the programme. Consulting these stakeholders will support an understanding of how the programme will work within existing systems.
Policymakers, strategic leaders, funders, researchers, voluntary and community sector representatives, and any other groups who may have an opinion on, are involved in or affected by the programme	They may have knowledge and understanding of the evidence around why and how these programmes usually work to address particular issues. They may possess valuable insights into local strategies, policies and community contexts.

Participatory approaches that encourage collaboration and consensus-building should be prioritised. Below is a non-exhaustive list of potential methods that could be used:

- User journeys
- Service mapping
- Workshops and advisory groups
- Focus groups / group discussions
- Interviews
- Online or in-person consultation tools
- Surveys
- Peer review and review panels

Note: The approach you choose should be determined by what is considered acceptable & feasible within the project timeline and local context

Case study

A youth employment training and placement programme.

This example is used across the different chapters to explain different constructs. For background to this example **see page 11**

Multi-stakeholder structured workshops:

The evaluator brought together programme designers, delivery partners and practitioners with experience of working with young people, and young people from diverse backgrounds with experience of unemployment training and placement programmes or who were looking for jobs, to collectively develop, interrogate and test the ToC.

Multi-stakeholder or separate stakeholder groups review panels:

The evaluator convened diverse stakeholders to critically examine draft ToCs, contributing policy perspectives, contextual knowledge and lived experience.

Storytelling circles:

The evaluator created safe spaces where young people with lived and living experience of such programmes, or groups that were disproportionately affected by a lack of access to them, could share their views and perspectives about how programme elements influence their experiences and outcomes.

Participatory journey mapping:

The evaluator facilitated sessions where young people mapped their employment journeys from unemployment to job placement. Using large visual templates, participants identified key touchpoints, emotional highs and lows, barriers and helpful interventions.

Good practice principles of stakeholder engagement

When working with stakeholders using participatory methodologies, carefully consider safeguarding issues and power imbalances that may emerge when different groups interact

Plan approaches that mitigate or address challenging group dynamics while ensuring safety and meaningful participation, particularly for young people. Effective facilitation depends not only on the activities you choose but how you implement them, in line with good practice principles.

Good practice principle	Description
Inclusivity	Inclusivity supports stakeholders to feel valued, respected and able to contribute in a meaningful, non-tokenistic way, regardless of their background, identity or perspective. This involves creating equitable opportunities that ensure the voices of those most impacted are both represented and centred, while removing or mitigating potential barriers to participation (cultural, communication, physical, and social) through reasonable adjustments.
Responsiveness	Being responsive means actively adapting your approach to the evolving needs and feedback of stakeholders, as well as to other contextual factors that influence engagement. It involves listening closely, recognising when adjustments are needed, and making timely changes to keep the process relevant, inclusive and effective. This can support the building of trust, shows respect for stakeholders' contributions, and helps strengthen engagement throughout the process.
Psychological safety	Psychological safety means creating a collaborative and supportive space where participants can share ideas, ask questions, raise concerns and admit mistakes without fear of judgment or negative consequences. In participatory approaches, it involves facilitators fostering mutual respect, encouraging diverse perspectives and handling conflict constructively. When present, psychological safety enables fuller engagement, constructive challenge and more authentic contributions from stakeholders.
Transparency	Transparency means being open, honest and realistic with stakeholders about the purpose, process and expectations of participation. It involves facilitators clearly communicating how decisions will be made, how input will be used, and what limitations or constraints exist. This helps manage stakeholder expectations, reduces misunderstandings and builds trust by demonstrating that their contributions are valued and handled with integrity.

Table 11: Good practice principles of stakeholder engagement

Case study

Behaviours or participant reactions to programme components

This example is used across the different chapters to explain different constructs. For background to this example **see page 11**

Practical facilitation examples for each of the four good practice principles for engaging diverse groups of stakeholders:

Inclusivity

Facilitators ensured that group activities used a mix of verbal discussion, written input and visual tools so that all young people, regardless of confidence or communication style, could participate fully.

Responsiveness

During the first evaluation session, facilitators noticed that some participants were more engaged in small groups than in whole-room discussions. They adapted future sessions to include more breakout activities and an online feedback form, ensuring richer and more representative input from all voices.

Psychological safety

At the start of each evaluation session, the facilitator co-created group agreements with the young people, establishing norms around respectful listening, no interruptions and confidentiality.

Transparency:

The facilitator explained how evaluation findings would be used, what decisions they may influence, and any limitations, such as budget constraints. After the sessions, participants received a summary of feedback and planned next steps, showing how their input was shaping the programme.

Equity considerations in developing a ToC

To centre equity, ensure the ToC reflects how different groups experience the problem, programme and outcomes. Consider:

- How might different groups and populations experience this problem/issue/opportunity?
- Who benefits from the programme goals?
- Who might not benefit but still be impacted by the problem, issue or opportunity identified?
- Are there barriers and facilitators that vary across different groups?
- How is the programme accessible, engaging and beneficial to historically underserved populations and young people?
- How is the programme being adapted to meet the needs of different groups?
- How might the outcomes, impact and experiences vary for different groups?

However, it's important not just to ask these questions, but to actively use them to inform the ToC and, where appropriate, the implementation strategies. This helps ensure that both the ToC and the programme are grounded in clear logic, explicit assumptions and aligned with the evaluation focus.

Visualising a ToC

There is no "correct" way to visualise a ToC. Instead, you need to think about what a visualised ToC contains and communicates.

Youth Futures' minimum requirements for a ToC visual:

- **Include context, assumptions, resources, activities, mechanisms, outcomes and impacts.**
- **Illustrate the causal pathways through which the programme achieves its outcomes.**

The visual representation of the ToC should be accompanied by a clear narrative explanation. This should clarify the logic behind the visual, ensure shared understanding amongst stakeholders, and highlight the rationale for key pathways. The narrative may also be a more appropriate place to discuss the programme's rationale and target population.

ToC diagrams are different to logic models (see examples below), which capture the inputs alongside the different stages of the programme or policy in a logical stepwise fashion and do not necessarily include all the elements of a ToC.

Note: Logic models have been criticised for a limited and linear conceptualisation of causality and a failure to adequately link the activities of a programme to its causal mechanisms and outcomes. However, there is significant variation in the presentation of logic models, with some including contextual factors alongside programme mechanisms.

See examples below:

Youth Employment Initiative – Impact Evaluation

Department for Work and Pensions and Government Social Research Profession (2022)

View Youth Employment Initiative Impact Evaluation

A logic model for the Freeports Programme

HM Government (2022)

View Logic Model for the Freeports Programme

Youth Investment Fund: Theory of Change user guide

The National Lottery Community Fund (2021)

View A Theory of Change for Open Access Youth Provision

Evaluation of the Durham Works Futures programme

Youth Futures Foundation (2025) (p.28)

View Durham Works Futures: Theory of Change

Strengthening ToCs: testing, validating and refining

Programme theories and the ToCs that articulate them can diverge from what happens in reality due to external disruptions to delivery, misalignment with participant needs, or weak evidential foundations

Refining ToCs at different stages of evaluation helps ensure programme models remain contextually relevant and provide an evidence-informed view of how a programme works. This enables evaluators to examine programme context interactions, surface unintended consequences, and assess the credibility and plausibility of the theory based on emerging evidence.

A ToC is most useful for evaluators when it serves as a practical framework that can be tested and refined to accurately reflect how and why a programme works in the real world. The steps outlined in this section help guide this process, highlighting the key questions to consider, how to prioritise what to test, and the methods commonly used to test and adapt the theory.

Key questions in testing and refinement of ToCs

Table 12 lists the key questions that should be answered in the process of testing and refining a ToC. For more details, across different stages of evaluation, see **Chapter 4.**

Table 12: Good practice principles of stakeholder engagement

Good practice principle	Description
Assumptions	Are the initial assumptions underpinning the ToC still valid?
Context	Have external factors (e.g., policy changes, economic conditions, crises) or a changed context for the delivery partner or recipients influenced the programme implementation or effectiveness?
Activities	Are the programme activities being delivered as originally designed? If not, what modifications have been made and why? (Further information in this chapter.)
Mechanisms	Are the expected mechanisms of change occurring as anticipated? If not, what alternative mechanisms are emerging?
Outcomes	Are the expected short-, medium- and long-term outcomes being achieved?

Prioritising what to test

Why it matters:

To effectively evaluate a programme, it is crucial to first prioritise which elements of the ToC to test. Evaluation resources are often limited, so it is rarely feasible to explore all components of the ToC. Below is a key list of factors to consider when engaging in prioritisation:

Not all programme elements are equally important.

Identify which components seem essential to achieving outcomes, and those which may be adaptable, depending on the context.

Some programme elements may be untested or have limited evidence.

Focus on components that are uncertain or novel, especially those not previously tested or those being implemented in a new context. Consider whether these are core elements as described above.

Outcomes may differ across groups.

Prioritise elements that are particularly relevant to the key groups your programme is designed to benefit, as effectiveness may vary.

Look beyond outcomes and consider how the programme is implemented.

Include implementation factors such as reach, adoption, and acceptability, which can have important implications for equity and sustainability. This is covered in detail in **Chapter 3**.

Consider the feasibility of data collection and measurement.

Reflect on how time, budget and data constraints may affect what and how you can evaluate. Assess which aspects of the ToC can be measured with the available resources, and identify suitable and proportionate methods for data collection that make best use of those resources.

Priorities will vary depending on the programme's evaluation stage.

See recommendations around the different elements to test across the different stages of evaluation in **Chapter 4**, to determine priorities.

Monitoring unintended and adverse outcomes and consequences

Unintended outcomes, both positive and negative, can emerge during programme implementation. Evaluators should plan from the outset how these effects will be identified, monitored and addressed. This is an important part of both prioritising what to test and using the ToC to guide measurement.

When developing and refining a ToC, teams should:

Anticipate risks:
Identify risks as part of the ToC. These may include reinforcing inequalities, excluding certain groups. Regularly review risks with key stakeholders to understand how they might be changing or emerging as the programme develops.

Anticipate opportunities:
Unexpected positive benefits can also emerge during an evaluation. Evaluators should build in approaches to capture this learning and update the ToC accordingly. Recording positive unintended outcomes supports evaluators to ensure that the right outcomes are being measured at the impact evaluation stage.

Plan how to detect and track them:
Define indicators, feedback mechanisms or participatory sensemaking activities that help surface emerging issues throughout delivery.

Pre-agree escalation routes for addressing adverse outcomes:
Set out clear processes for documenting, reporting and responding to concerns, including governance channels and thresholds for when adaptation or pausing of activities might be required.

Keep the process participatory: Involve a diverse range of stakeholders, particularly young people and those most affected, in monitoring and interpreting unintended outcomes, to ensure different perspectives on potential harms or benefits are recognised.

Monitoring unintended effects is not only a safeguard against harm; it also helps learning and improvement. When evaluators or delivery partners identify adverse impacts, these should be logged and discussed through agreed governance routes.

Depending on the scale and seriousness of the issue, responsaes might include making programme adjustments, adding mitigation activities or, in extreme cases, temporarily pausing or stopping delivery.

Using a ToC to guide measurement and data collection

Each component of a ToC, as well as the overall theory, can be developed, tested and refined using a range of methods. Table 13 explains how different methods can support evaluators in the testing and refinement of theory, and highlights which specific components it may be particularly suited to explore.

Although the methods below are presented separately, they can be combined in practice. Using multiple methods together can strengthen your ability to identify, understand and assess different components of the ToC. In many cases, a single method may also help you explore several ToC components at the same time. In the table, we include some references to implementation to show how this is integrated with programme theory testing in TIIPE, but we discuss developing and testing implementation theory fully in **Chapter 4.**

Key principles for testing programme theory:

1. **Operationalise each construct**
   Translate programme theory constructs into measurable indicators and identify the data source for measuring these indicators **(see Table 13).**

2. **Use mixed methods**
   Combine quantitative data (e.g., employment, education and training status) with qualitative insights (e.g., young people's perceptions of support). This enables evaluators to triangulate findings and explain variation.

3. **Adjust approaches by stage**
   In the intervention stage, a component can be included if it is plausible based on theory, existing evidence and stakeholder input. At the feasibility evaluation and pilot evaluation stages, validation from further testing using the methods described in **Table 13** is required. At the impact evaluation stage, programme outcomes should be supported by causal evidence, and the evaluator should consider whether any other elements of the ToC can be explored empirically (although this is not a requirement).

4. **Pay attention to variation**
   To understand what works for whom, and in what contexts evaluators need to explore variation between participant groups and contexts.

5. **Use the ToC to determine when programme constructs should be measured**
   Programme constructs should be measured at the time point, and in the order, they are expected to emerge. For example, if a mechanism is expected to emerge before an outcome this should inform the phasing of data collection for mechanisms and outcomes.

Table 13: Methods to support testing and refinement of a ToC

Component	Potential data sources / approaches	Considerations
Problem (rationale)	• Evidence reviews • Problem tree analysis	• Evaluators should explore the problem during the intervention design stage.
Target population	• Evidence reviews • Local admin data	• Evaluators should explore the target population during the intervention design stage.
Context	• Evidence reviews • Observations • Moderator analysi	• Evidence reviews should be used in the intervention design stage. • Observations and moderator analysis can be used during feasibility, pilot and impact evaluation stages.
Assumptions	• Participatory methods • Interviews with young people, programme delivery staff, local stakeholders • Focus groups with young people, programme delivery staff, local stakeholders	• Evaluators should revisit assumptions throughout the evaluation cycle and log changes to the ToC. • All methods are appropriate across evaluation stages.
Activities	• Management Information data • Fidelity and dosage logs • Surveys • Interviews with young people, programme delivery staff, local stakeholders • Focus groups with young people, programme delivery staff, local stakeholders	• Interviews and/or focus groups should typically be used in the intervention design stage. • A mixed method approach combining either MI data or survey data with qualitative data from interviews and/or focus groups is appropriate during feasibility, pilot and impact evaluation stages. • Fidelity and dosage logs can also be used in the pilot and impact evaluation stages.

Table 13: Methods to support testing and refinement of a ToC (cont.)

Component	Potential data sources / approaches	Considerations
Mechanisms	• Evidence reviews • Interviews with young people, programme delivery staff, local stakeholders • Focus groups with young people, programme delivery staff, local stakeholders • Observations • Mediator analysis	• Evidence reviews, interviews and/or focus groups should typically be used in the intervention design stage. • Interviews, focus groups, observations, mediator analysis can also be used to further test mechanisms in the feasibility, pilot and impact evaluation stages.
Outcomes	• Longitudinal administrative data	• Evidence reviews, interviews and/or focus groups can be used in the intervention design stage. • A mixed method approach combining either MI data or survey data with qualitative data from interviews and/or focus groups should be used during the feasibility evaluation stage. • In the pilot and impact evaluation stage the measurement of primary outcomes does not fall in the remit of the TIIPE, but surveys, MI data, interviews and focus groups may be used to explore other outcomes, or to support the interpretation of the primary outcome analysis. • Ripple effects mapping can be used to surface unintended outcomes.
Impact	• Evidence reviews • Local admin data	• Impact cannot be measured within the intervention design, feasibility, pilot or impact evaluation stages, but longitudinal administrative data can facilitate subsequent analysis into programme impact.
Causal pathways	• Causal mapping	• Causal mapping can help evaluators map out the causal pathways at all evaluation stages.

Case study

A youth employment training and placement programme

This example is used across the different chapters to explain different constructs. For background to this example **see page 11**

ToC components, examples

- **Activities:** Skills training, job search workshops, mentorship, employer placements.
- **Mechanisms**: Increased skills, self-efficacy, goal clarity, motivation, psychological safety, access to networks, employer engagement.
- **Outcomes:** Employment or education placement, sustained work, career progression.

Examples of survey questions

(Focus on enabling comparisons by making them specific and setting out clear measurement timescales)

- **Activities:** Which programme activities have you taken part in? (Training, mentoring, work placement, job search workshops, other)
- **Mechanisms:** Since starting the programme, how confident do you feel in your ability to apply for jobs? How much has your understanding of what employers are looking for improved? (A scale [1-5] could be used to make answers comparable.)

Interviews and focus groups

(Focus on enabling more open and in-depth exploration)

- **Activities:** Can you describe the parts of the programme you've been involved in? Which activities were most useful, and why?
- **Mechanisms:** How has the programme changed the way you approach looking for work? What helped you feel able to apply for jobs?
- **Outcomes:** What's changed for you since you joined the programme (skills, job search activity, interviews, offers)? Have you seen any unexpected benefits or changes as a result of joining?

Evidence for revising a ToC

Throughout TIIPE, it is important to consider the ToC as an evolving framework that will require revision as insights emerge. Evidence from a range of sources can prompt a revision to a ToC. As described in Table 13, an evaluator should consider whether there is evidence to support each of the components of a ToC. A revision may occur if:

- **Assumptions:** Participants engage in ways that challenge the assumed pathways of change.
- **Context:** Changes in systems, policies or community conditions alter the feasibility or relevance of the ToC.
- **Activities:** Activities are delivered differently than expected.
- **Mechanisms and outcomes:** Key outcomes are not achieved or emerge through different mechanisms than those originally proposed.
- **Unintended consequences:** Unexpected effects (positive or negative) reveal new causal dynamics or risks that were not previously recognised or anticipated.
- **Causal pathways:** Evidence emerges that questions the sequencing, logic or relevance of the ToC components

As a general rule, evaluators should follow these guidelines when making revisions to ToCs:

- If there is contradictory evidence (from multiple sources), revise or remove the ToC component.
- If evidence for a ToC component is mixed or weak, further testing is needed. Evaluators should define what evidence is required and set a timeline for this testing.
- If there is consistent evidence (from multiple sources) supporting a ToC component, retain the component.

It is important to log changes made, including the rationale, evidence used and decision date, for transparency.

It is important to log changes made, including the rationale, evidence used and decision date, for transparency.

The strength of evidence required at each stage of evaluation may vary. In earlier stages, such as the intervention design, feasibility and pilot evaluation stages, lower or more limited levels of evidence may be required than at the impact evaluation stage (which requires programme outcomes to be tested empirically). However, although lower levels of evidence may be acceptable during earlier stages of evaluation, a TIIPE should build sufficient evidence to justify progression to an impact evaluation. A primary purpose of early-stage evaluation is to de-risk the likelihood of trial failure. If there is insufficient evidence to support the plausibility of the ToC, then an impact evaluation should not be carried out.

The evidence to support the plausibility of the ToC, and therefore to justify progression to an impact evaluation, builds from evaluation stage to evaluation stage. To move from the intervention design stage to the feasibility evaluation stage there must be reasonable evidence to conclude that the ToC is plausible. To move from the feasibility evaluation stage to the pilot evaluation and impact evaluation stages there must be sufficient evidence of promise to warrant testing the effectiveness of the programme further.

Table 14: Overview of theory-based approaches that can complement TIIPE

Approach	Definition	How it complements a TIIPE (by stage)	Considerations
Realist evaluation	Explores programme theory by examining Context-Mechanism-Outcome (CMO) relationships to explain how and why outcomes occur for different groups in specific contexts.	What works, for whom, in what circumstances and why?	**Intervention design:** identify target population needs, clarify causal pathways and expected CMOs. Feasibility/Pilot evaluation: test core components and mechanisms; identify the "ideal" context for delivery. **Impact evaluation:** explain observed effects using CMOs; highlight subgroup differences. **Scale-up:** assess how context and mechanisms might affect replication or roll-out.
Contribution analysis	Builds a credible contribution story linking activities, mechanisms and outcomes, testing the ToC against alternative explanations using multiple evidence sources.	To what extent did the intervention contribute to observed outcomes (accounting for other factors)?	**Intervention design**: develop an initial "contribution story" linking activities to outcomes; identify causal "hotspots" where evidence is weaker. **Feasibility/Pilot evaluation:** gather evidence on mechanisms; assess likelihood of contribution versus alternative explanations. **Impact evaluation:** Assess likelihood that the proposed mechanisms contributed to any outcomes achieved; help to explain why outcomes may have failed to emerge.
Process tracing	Uses tests to weigh evidence for and against hypothesised links in a causal chain by (i) identifying potential explanations for outcomes, (ii) disregarding explanations the evidence contradicts, (iii) assessing the strength of the evidence for multiple causal explanations.	How strong is the evidence that an programme has contributed to observed changes?	**Feasibility/Pilot evaluation:** test plausibility of causal chains in the ToC. **Impact evaluation:** test whether achieved outcomes plausibly resulted from the programme's mechanisms.

Chapter 4

Developing an implementation theory

youth futures FOUNDATION

Developing an implementation theory

Developing an implementation theory involves specifying how and why particular approaches to implementing a programme are expected to lead to it being successfully delivered in practice.

Implementation theory complements the programme theory by articulating the pathways that link implementation strategies, mechanisms and outcomes. While a programme theory explains how programme activities lead to changes in participants, implementation theory focuses on understanding the conditions needed for effective and sustainable implementation.

The processes involved in developing implementation theory mirror those involved in developing programme theory (as described in **Chapter 2** and summarised below in **Table 15).** As with the development of programme theory, it involves drawing on insights from delivery partners; practice-based knowledge and lived experience of those involved in early delivery; evidence about the design and evaluation of similar programmes, and from any early evaluations of the programme; and existing theory about implementation (e.g., implementation frameworks or wider theory about organisational change or individual behavioural change). As with the development of programme theory, meaningful stakeholder engagement and co-design is very important.

Implementation theory should be developed alongside **programme theory** in the **intervention design** stage.

Good principles in developing an implementation theory

Explicative

Clearly explains how and why an implementation strategy is expected to produce change in implementation outcomes.

It should specify the mechanism(s) through which the strategy works (e.g., training > increases self-efficacy > improves fidelity) and under what conditions.

Detailed

It includes sufficient detail on implementation determinants, strategies, mechanisms and outcomes.

It should outline causal pathways linking strategies to mechanisms and implementation outcomes, with all assumptions made explicit and testable.

Robust

Well-articulated, credible and grounded in relevant theory and evidence (e.g., using frameworks such as CFIR and TDR, and/or refer to prior implementation studies).

It should clearly show the logic from determinants > strategy > mechanism > outcomes, and cite empirical evidence, theory or practice-based / lived experience insights supporting each link.

Key steps for developing an implementation theory

01. Start with determinants
Identify the contextual and organisational factors likely to influence delivery (e.g., leadership, staff capacity, participant engagement). Use a framework such as CFIR to structure this.

02. Map strategies to determinants
Specify which implementation strategies (e.g., training, supervision, incentives, stakeholder engagement) are designed to address which barriers or enablers.

03. Identify expected mechanisms
Describe the processes that explain how each strategy is expected to produce change (e.g., coaching [strategy] → increased understanding of the programme + increased staff confidence [mechanisms]→ improved fidelity [outcome]).

04. Define implementation outcomes
Define what high-quality implementation would look like based on selected proximal implementation outcomes (e.g., fidelity, reach, acceptability, feasibility, adoption and sustainability).

05. Align with programme theory
Ensure each implementation component links clearly to the programme's causal chain, showing how successful implementation supports intended participant outcomes.

The resulting implementation theory should form a testable logic model for delivery – clarifying which supports matter most, under what conditions and why.

As noted in **Chapter 2**, there are different ways of reflecting this within or alongside the programme ToC. In more complex evaluations intended to drill down into implementation in more detail, tools such as an **Implementation Research Logic Model (IRLM)** can help visualise and structure this theory, linking implementation determinants, strategies, mechanisms and outcomes in a single visual.

Identifying implementation determinants

Identifying implementation determinants at the intervention design stage means forming hypotheses about the contextual factors that might act as barriers to or enablers of implementation (see Chapter 1). Implementation determinants may change across evaluation stages, and evaluators should update and refine an implementation theory, where relevant, in response to this.

Using a determinants framework is helpful to ensure this is done comprehensively There are no rules about which of the many determinant frameworks to use – this depends on how well they fit with the programme and the aims and focus of the evaluation.

- CFIR is useful when the focus is on **broader contextual variables** influencing implementation, for example, organisational culture.
- TDF is useful where there is interest in the **individual level variables** influencing implementation outcomes, for example, behaviour change.

Frameworks can also be used in combination. For example, CFIR and TDF complement each other through their focus on different influencing variables that can impact implementation, but other frameworks might be more appropriate to your study.

There are free resources to support evaluators in choosing which determinant framework to use to help surface relevant barriers and enablers.

For example, the T-CaST tool is an online checklist tool that can be used to compare implementation theories, models and frameworks (TMFs) and help researchers to compare TMFs based on their applicability, usability, testability and acceptability. Tools such as T-CaST support transparent TMF selection and minimise superficial use of TMFs. If researchers do not have a set of candidate TMFs in mind, they can refer to the D&I Models webtool, which houses an online repository of hundreds of potential TMFs that can be considered for a TIIPE.

Identifying implementation strategies and mechanisms

A central aim of TIIPE is to test the theory of how implementation is expected to work. This means surfacing the assumptions behind the implementation approach and assessing whether those assumptions hold true in practice. At the heart of this is a focus on implementation strategies – the deliberate methods or techniques used to support adoption, delivery and sustainability of an programme.

Implementation strategies are not neutral activities; they are theory-informed actions designed to overcome specific barriers or enhance enablers in a particular setting. So, the starting point in identifying implementation strategies is to decide which strategies are likely to be most successful in addressing the barriers and enablers that have been identified. This work is aided by using a framework of implementation strategies such as ERIC (Powell et al., 2015) as well as theories of organisational or individual change. It is also important to think about what implementation strategies are likely to be feasible and not to be onerous for the programme delivery team and for settings.

In formulating strategies, it is also important to think about the mechanisms by which a strategy is intended to address barriers or enablers and achieve implementation outcomes. As with programme theory, mechanisms here are the underlying processes through which occurs. In the example we used earlier, coaching (strategy) might be planned to overcome practitioners' unfamiliarity with the programme (barrier), with the assumption that it will help them to understand the critical elements of the programme and to be more confident in their delivery of them (mechanism).

The degree to which implementation strategies can be introduced and adjusted depends on the evaluation stage and programme maturity. In the intervention design stage, there is high flexibility to develop and modify strategies. In the feasibility evaluation stage there is moderate flexibility to modify strategies based on evidence from the intervention design stage. In the pilot evaluation stage, there is low flexibility to alter strategies, and these should only be minor refinements. In the impact evaluation stage, implementation strategies should be stable and pre-specified, so they can be evaluated in a stable condition. The role of the evaluator mirrors these changing dynamics – in earlier stages the evaluator can play an advisory role in the development of strategies, but by the impact evaluation stage their role is purely to evaluate the effectiveness of the strategies. For more information on the focus of different evaluation stages see **Chapter 5.**

Identifying implementation outcomes

Identifying implementation outcomes involves forming a view about what effective implementation would look like, in relation to selected implementation outcomes. For example:

The relevant **implementation outcomes** will vary at each **evaluation stage.**

Reach

What proportion of settings are expected or intended to take up the programme? What proportion of young people offered the programme are intended to take it up, and how does this vary between key sub-groups?

Acceptability

Whose acceptability of the programme is relevant (e.g., organisation leaders, practitioners, participants, referring agencies etc.)? What level or evidence of acceptability would be expected?

Fidelity

What are the key aspects of delivery that are essential to programme effectiveness, based on the programme theory? What level of fidelity is expected for each?

Integrating implementation into ToC development

It is important for an implementation theory to be developed alongside a programme theory. A ToC is not only a map of how programme activities are expected to lead to outcomes – it can also be adapted to capture how the programme will be implemented, under what conditions and with what supports.

Rather than treating implementation as an afterthought, evaluators can use the ToC as the central organising framework to align programme theory with implementation determinants, strategies, mechanisms and outcomes. This makes explicit what needs to be in place for delivery to succeed and guides what the evaluation should measure and test.

Approaches for embedding implementation in ToCs

There are different ways in which evaluators can embed implementation theory into ToC development, including:

01. Add dedicated implementation columns within the ToC framework

Extend traditional ToC templates by including columns that explicitly map:

- **Implementation determinants:** factors that hinder or support implementation (e.g., leadership buy-in, staff capacity, participant readiness). Note: these might overlap with contextual factors in the programme ToC.
- **Implementation strategies:** actions and supports required to overcome or activate those determinants (e.g., training, supervision, co-design, stakeholder engagement).
- **Implementation mechanisms:** processes by which strategies produce desired effects on implementation outcomes.
- **Implementation outcomes:** measurable markers of how well the programme is being implemented (e.g., fidelity, reach, acceptability, adoption, feasibility).

These additions encourage evaluators and programme teams to think in advance about what needs to happen for successful delivery and sustainability, and to define success not just in terms of outcomes for participants, but also outcomes for implementation.

When to use: This approach is most appropriate for small-to-medium-scale evaluations where the delivery system is relatively simple (e.g., one delivery provider, a limited number of sites).

02. Include an implementation narrative alongside the ToC diagram

Alongside the ToC diagram, evaluators add a short narrative section describing:

- The delivery system and key actors (who does what, where).
- Key determinants shaping implementation.
- How targeted implementation strategies are expected to influence determinants.
- The mechanisms by which strategies are expected to produce desired effects on implementation outcomes.
- How implementation outcomes are monitored and interpreted

When to use: This approach can be useful when space in the ToC diagram is limited and/or where contextual nuance is important (e.g., multi-agency partnerships, varied local settings).

Approaches for embedding implementation in ToCs

03. Use a supplementary implementation ToC

Implementation scientists have developed an implementation ToC, known as an Implementation Research Logic Model (IRLM) which can be constructed alongside a programme ToC.

An IRLM is a structured tool that connects

- Determinants of implementation.
- Chosen implementation strategies.
- Implementation mechanisms.
- Implementation outcomes.
- Programme outcomes (short-, medium-, long-term

When to use: An IRLM is particularly valuable for more complex programmes, or multi-site evaluations, where there is a need to compare how implementation factors operate across different contexts.

The IRLM helps make assumptions about implementation more explicit and testable. It can also support clear alignment between theory, practice and data collection – helping evaluators trace whether outcomes are not achieved because the programme was ineffective, or because it was never fully or faithfully delivered. Evaluators can access the free IRLM building tool and further resources **View Implementation Logic Model**

Testing implementation theory

Light touch testing of an implementation theory should start in the intervention design stage. However, the testing of implementation theory becomes increasingly important as evaluation stages progress from the feasibility and pilot evaluation stages to the impact evaluation stage.

Testing implementation determinants

This involves assessing whether the implementation determinants identified were indeed those that were relevant in programme delivery, assessing which were the most significant barriers and enablers and, where possible, measuring their prevalence (e.g. whether they were experienced by all settings or all practitioners) and extent or scale (e.g. the level of staff capacity gaps). Many of the tools and the frameworks that we recommend to identify implementation determinants also include advice on how to measure or assess determinants.

Testing implementation outcomes

It is not necessary nor feasible to measure all the implementation outcomes introduced in this guidance, and the relevant implementation outcomes of interest will also vary according to the evaluation stage (more details provided in Chapter 5.

Typically, outcomes such as adoption, reach, acceptability and feasibility are key outcomes to assess in earlier stages of evaluation.

It is also important to continue to assess them once delivery has stabilised, for example, as the programme is offered to a wider range of settings more distant from the original context, and with a focus on equity considerations. Fidelity is also a key consideration once it is established that the programme is, in principle, feasible. In the later stages of evaluation, implementation outcomes such as cost, penetration and sustainment are more likely to come into focus.

There are a range of validated measures available to support the quantitative assessment of implementation outcomes. Evaluators can use the free, online Implementation Outcome Repository to search for suitable validated, quantitative measures for a variety of implementation outcomes.

Unless they have already been developed, evaluators will need to design bespoke fidelity criteria and measures. This should be done with delivery partners so that the criteria selected are linked with the programme theory. Measurable indicators for each will then need to be developed.

As an example, if a youth employment programme's ToC specifies that coaching of young people builds self-efficacy, then fidelity measures might include: whether an identified minimum number of required coaching sessions were delivered (dosage); whether the prescribed coaching techniques were used well (adherence, quality of delivery); and whether participants actively engaged in coaching (participant responsiveness).

Balancing fidelity and programme adaptation

In practice, fidelity in programme delivery and adaptation to the programme need to be balanced. High fidelity means that the programme is being delivered in full and as intended, but some adaptation is inevitable – and indeed necessary – for the programme to fit well with all relevant settings.

As programme development is refined, it will be important to identify and support the adaptation that is "in scope" with the programme theory and acceptable, and to try to minimise adaptation that is out of scope and not acceptable. One way of approaching this is for evaluators to work with delivery partners to distinguish between core elements that must be delivered as designed, and flexible elements that can adapted or dropped without undermining the programme. For example, it might be determined that a minimum of eight hours of coaching is needed and that a set of core topics are covered (core elements), but staff can decide whether or not to offer additional coaching sessions (flexible element). Acceptable adaptations often include changes to surface structures (e.g., language, examples used in materials or training, timing) to improve fit and equity. Deeper adaptations, for example, to content or to the intended function of a component, would need to be monitored closely with an eye on whether they are consistent with the programme ToC.

The balance between fidelity and adaptation also depends on the stage of evaluation. For example, in the intervention design stage there is scope for higher flexibility – core components are hypothesised but not yet fixed, and adaptation is encouraged in order to learn what works for whom and why. In feasibility and pilot evaluation stages, there is moderate scope for flexibility – this is where you begin specifying which components are essential and adaptations are closely tracked to identify those which, for example, improve reach, equity or delivery quality. In an impact evaluation, core components must be delivered consistently and any adaptations should be minimal, pre-agreed and justified, as the purpose here is to test effectiveness under defined conditions.

Testing implementation strategies and mechanisms

Testing whether the proposed strategies worked as intended is crucial to understanding both how and why an programme was (or wasn't) delivered successfully.

The focus for testing strategies will need to change across evaluation stages. In the intervention design stage, evaluators may explore a range of strategies that may be effective in addressing barriers and emphasising enablers. In the feasibility evaluation stage, testing should be more focused on understanding whether there is plausible evidence to suggest that specific strategies enhance implementation or whether adaptations need to be made. In the pilot and impact evaluation stages, testing should be focused on evaluating a set of pre-determined strategies.

We've broken down 5 key considerations over the next pages.

1. What implementation strategies were used?

- In the intervention design stage this may involve capturing what strategies were used. In the feasibility, pilot and impact evaluation stages it involves capturing whether the intended strategies were used and whether settings used any additional implementation strategies (e.g., reorganising of roles even if not explicitly requested, or revisions to workflows). Real-world implementation rarely follows a fixed plan. Testing implementation theory involves documenting any adaptations and examining their effects on delivery and outcomes, appropriate to the evaluation stage:
- What was changed, added or dropped?
- Were these changes reactive (in response to challenges) or proactive (to improve fit)?

2. Did the strategies successfully address the right barriers and enablers?

- When testing an implementation strategy, evaluators should revisit their diagnostic work – what barriers and facilitators were identified and what strategies were used to respond to them? Across all evaluation stages, a robust TIIPE will examine whether the strategies:
- Addressed all the key barriers or enablers.
- Successfully reduced or removed key barriers (e.g., time constraints, lack of confidence, unclear procedures).
- Amplified key enablers (e.g., strong leadership, alignment with organisational goals, positive practitioner attitudes).
- This requires collecting data not only on what implementation strategies were used but also on how they were used and how stakeholders experienced them, for example: Were co-design processes genuinely inclusive? Were training sessions perceived as useful? Did supervision feel supportive?

3. Testing of implementation strategies is aligned with programme theory.

The choice of which implementation strategies to test in feasibility, pilot and evaluation stages should be guided by their role in activating the programme mechanisms of change set out in the programme ToC. For instance, if the ToC specifies that practitioner confidence is a prerequisite for effective delivery, then coaching, supervision or peer-learning strategies aimed at building skills and boosting confidence should be prioritised in both measurement and analysis. This framing allows evaluators to ask focused validation questions such as:

- Did the implementation strategies activate the intended mechanisms?
- Were any intended mechanisms not activated by implementation strategies?

Note: a TIIPE does not usually attribute causality to specific strategies but, in some cases, this may be the primary aim of an evaluation (see next page).

4. Did the strategies achieve the intended implementation outcomes?

Each strategy should contribute to at least one implementation outcome – such as fidelity, reach, acceptability, feasibility or adoption. Testing implementation theory means checking these pathways:

- Did the strategies to improve fidelity (e.g., a structured coaching model) actually result in more consistent delivery?
- Did efforts to improve acceptability (e.g., participatory design workshops) increase buy-in among staff and service users?

5. Which strategies, or combinations, were most effective?

In many TIIPEs, multiple strategies may need to be evaluated simultaneously. Where possible, an evaluator should assess the contribution of each strategy or bundle:

- Was a combination of training and peer support more effective than training alone?
- Did layering strategies (e.g., leadership engagement plus audit and feedback) create synergistic effects?

Exploring patterns in the delivery of implementation strategies, mechanism activation and implementation outcomes can generate useful insights about which combinations of strategies seem to drive the most success, and under what conditions. By taking these steps, evaluators can move beyond simply describing what was done to critically testing whether the implementation theory holds in practice. This strengthens the explanatory power of the evaluation and supports learning that can guide future adaptation, scale-up or replication. It also helps in ensuring that the implementation strategies required are proportionate, by checking that all the intended strategies were necessary to achieve high quality implementation.

Methods for testing implementation theory

Testing implementation theory means collecting data that can assess whether the hypothesised determinants, strategies, mechanisms and outcomes operate as expected. The goal is to understand not just whether the programme was implemented, but how and why implementation quality varied across sites or contexts. The methods and data sources set out in Chapter 3 in relation to testing programme theory are relevant for testing implementation theory too.

Evaluation designs for assessing impact of implementation strategies

Although a TIIPE explores how strategies work without causal attribution, evaluators can use evaluation designs such as the Multiphase Optimisation Strategy (MOST) or factorial experiments to help isolate the added value of specific strategies, by systematically varying which strategies are present and measuring downstream effects on fidelity, reach or participant outcomes. Small-scale trials that focus on comparing the implementation outcomes achieved through different implementation support packages (e.g., coaching only vs. coaching plus fidelity feedback) can also be run, to provide evidence on what is essential and what may not add sufficient value to be necessary.

Key principles for testing implementation theory:

1. **Operationalise each construct –** Translate determinants, strategies, mechanisms and outcomes into measurable indicators and identify the data source for measuring these indicators **(see Table 13).**

2. **Use mixed methods –** Combine quantitative data (e.g., fidelity scores, participation rates) with qualitative insights (e.g., perceptions of feasibility, experiences of enablers/ barriers). This enables evaluators to triangulate findings and explain variation.

3. **Pay attention to variation –** between sites, between stakeholder groups (e.g., organisational leaders vs. practitioners) and between participant groups (with a particular eye on equity). An evaluator should pre-specify priority subgroups.

4. **Collect data longitudinally –** Implementation evolves. Gathering data at multiple time points (e.g., pre-, mid-, post-delivery) captures changes in determinants and strategies over time.

5. **Link data to programme outcomes –** A key analytic step to support the integration of programme theory and implementation theory is connecting implementation variation to outcome variation. For instance, do sites with higher fidelity or stronger leadership show greater improvements in participant outcomes?

6. **Test mechanisms empirically –** As per the testing of programme mechanisms, researchers can utilise methods such as mediator analysis to examine whether hypothesised mechanisms (e.g., improved staff confidence) explain links between strategies and outcomes.

7. **Use structured frameworks –** Tools like CFIR (for determinants), ERIC (for strategies) and Proctor's outcomes framework provide shared language and validated constructs that improve rigour and comparability.

8. **Adjust approaches to stages –** Testing implementation theory should evolve across evaluation stages. Early on, evaluators may explore a wide range of implementation strategies to address barriers and enablers, while focusing on outcomes like acceptability, reach and feasibility. As delivery stabilises, testing should shift toward fidelity and whether selected strategies work consistently across settings. In later stages, outcomes such as cost, penetration and sustainment become increasingly important, and the focus may shift to evaluating a defined set of strategies based on prior evidence.

9. **Iterate and refine –** Across all stages, iteration is key. Revisiting the implementation theory allows teams to, for example, clarify which assumptions about barriers and facilitators held true, which strategies effectively addressed them, and which intended mechanisms were not activated by implementation strategies?

In practice

A strong implementation theory and measurement plan allow evaluators to identify what delivery supports are essential, what can be adapted, and why some contexts succeed while others struggle. This creates evidence that is not only evaluative but also actionable, guiding scale-up, workforce planning and resource allocation in real-world systems.

Table 15 shows the key components of implementation theory as set out in this guidance and some example data sources and approaches.

Component	Examples	Potential data sources/ approaches	Why it matters
Determinants – factors that enable or hinder implementation	• Leadership engagement • Staff capacity/turnover • Resource availability • Participant motivation/readiness • Partner relationships	• Staff or leadership surveys • Interviews or focus groups • Readiness assessments (more commonly used in earlier evaluation stages – intervention design and feasibility evaluations) • MI data • Asset mapping	Understanding which conditions may support or hinder delivery allows evaluators to interpret variation in outcomes and suggest adaptations.
Implementation strategies – specific actions or supports used to enable implementation or address barriers	• Training and coaching • Supervision and feedback loops • Stakeholder engagement • Incentives or resource supports	• Attendance registers • Delivery MI data • Observation notes • Surveys • Adaption logs • Interviews on perceived usefulness of strate-gies • RCTs comparing implementation support packages (impact evaluation stage only)	Clarifies how implementation is being supported and which strategies are essential to ensure key outcomes (e.g., fidelity, acceptability, sustainability).
Implementation mechanisms – underlying changes triggered by strategies, leading to outcomes	• Increased staff confidence/self-efficacy • Improved team communication • Stronger sense of ownership • Normalisation of new practices	• Staff pre/post training surveys • Focus groups with practitioners • Reflection diaries kept by practitioners • Observation of behaviour change in delivery	Helps explain why strategies work (or don't) and whether mechanisms were activated as hypothesised in the implementation theory.
Implementation outcomes – proximal indicators of implementation success	• Fidelity • Reach • Acceptability • Feasibility • Adoption • Sustainability	• Fidelity checklists • MI data on recruitment, participation and retention • Acceptability/feasibility measures used in surveys • Stakeholder interviews	Provides evidence on the quality and extent of delivery. Enables links between implementation variation and programme outcomes.

Table 15: Components of implementation theory and example measurement approaches

Evidence for validating or revising implementation theory

Developing a clear implementation theory and accompanying measurement plan provides the foundation for meaningful analysis of implementation. Once data have been collected on determinants, strategies, mechanisms and outcomes, evaluators need to examine how these elements interact — identifying which implementation supports were most influential, under what conditions and why.

In other words, analysis moves from describing what happened in the implementation of the programme, to testing whether the implementation theory holds.

This is where implementation theory becomes practical: by analysing relationships between variables (e.g., between leadership engagement and fidelity, or between staff confidence and delivery quality) evaluators can assess which elements of the implementation system are critical to programme success.

Analysing and interpreting implementation data

Analysis of implementation data should not sit apart from the programme evaluation – it provides the interpretive bridge between process and programme outcomes. Ideally, evaluators should:

- Use the implementation theory (and a ToC) to structure both data collection and analysis.
- Test the alignment between the intended and actual delivery of the programme.
- Examine how variation in implementation relates to variation in outcomes.

Implementation data can be broadly structured around three levels of analysis, and the findings from each of these levels of analysis will support an evaluator to understand whether the implementation theory holds, or needs to be revised:

Descriptive analysis
To understand what was delivered and how consistently. Through understanding what was delivered evaluators can assess which elements of the implementation system are critical. If there is evidence that suggests a programme was not delivered consistently this may also prompt a revision to the implementation theory.

Exploratory/explanatory analysis
To identify why implementation varied and what influenced it. Through understanding what drives variation, evaluators can assess whether the barriers and enablers detailed in the implementation theory are crucial determinants of implementation. Evidence to suggest that barriers or enablers are less important than initially believed, or evidence to support the presence of different, important barriers or enablers, may prompt a revision to the implementation theory.

Relational analysis
To examine whether implementation quality is associated with outcomes and to explore if hypothesised mechanisms of change were activated. Where outcomes are not achieved this analysis can support evaluators to distinguish between weakness in the programme theory from weakness in the quality of the implementation.

Once analysis has been conducted, evaluators should consider how the findings impact the validity of the implementation theory. As a general rule, evaluators should follow these guidelines when revising an implementation theory:

- Contradictory evidence (from multiple sources): revise or remove the implementation theory component.
- Mixed/weak evidence to support an implementation theory component: further testing is required and evaluators should define the evidence requirements and timeline for undertaking this further testing.
- Consistent evidence (from multiple sources) supporting an implementation theory component: retain the component.

It is important to log changes made, including the rationale, evidence cited and decision date for transparency.

The strength of evidence required can vary by evaluation stage. Earlier stages (intervention design, feasibility evaluation and pilot evaluation) can tolerate more limited evidence but must still build a sufficiently strong case to justify progression to an impact evaluation. One of the primary purposes of early-stage evaluation is to reduce the risk of trial failure. If there is insufficient evidence to suggest a programme can be delivered to the required standard, then an impact evaluation should not proceed at that time.

Evidence for the plausibility of the implementation theory should accumulate across stages. Progression from the intervention design stage to the feasibility evaluation stage requires reasonable evidence that the programme is stable and can be delivered as intended. Progression from the feasibility to pilot to impact evaluation stages requires evidence to suggest that the implementation approach is clear, supports the achievement of programme outcomes and can be delivered under conditions suitable for a rigorous trial.

Chapter 5

Applying TIIPE throughout evaluation

youth futures FOUNDATION

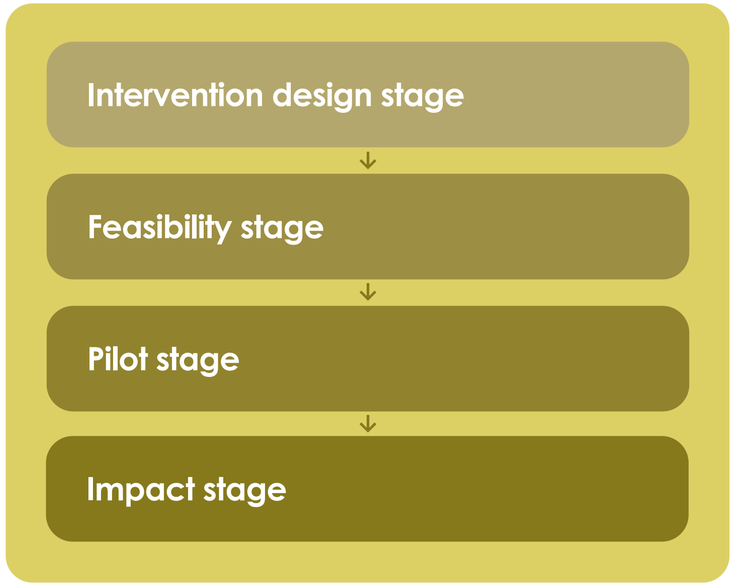
Incorporating TIIPE across evaluation stages

The following section outlines how TIIPE can be tailored across different evaluation stages, highlights the aspects of theory that are most relevant at each point, outlines the equity considerations at each stage, and sets out how TIIPE is used to inform stage-specific decisions about evaluation and programme design.

What are evaluation stages?[2]

- **Intervention design:** to strengthen and consolidate the programme's design and to build the foundations for future evaluation.
- **Feasibility evaluation:** to assess whether the programme can be delivered as intended, identify any necessary adaptations and determine readiness for impact evaluation.
- **Pilot evaluation:** to test delivery of the programme and the proposed impact evaluation design (or elements of it), refining both where needed.
- **Impact evaluation:** to robustly assess programme effectiveness using a randomised controlled trial (RCT), quasi-experimental design (QED) or other high-quality method.

Figure 5: Youth Futures' programme development and evaluation stages



TIIPE is used across the intervention design stage, and the feasibility and pilot evaluation stages to support an understanding of a programme and its implementation. During these early stages, TIIPE is also used to assess whether a programme is suitable for impact evaluation and, where appropriate, to prepare a programme for impact evaluation and support the design of the impact evaluation. During the impact evaluation stage, TIIPE is used to strengthen the interpretation and learning generated from the impact evaluation. If positive outcomes emerge, TIIPE helps to understand how the programme contributes to these outcomes, and what it takes to achieve them. If positive impacts are not found, TIIPE is essential in understanding whether this represents programme failure or implementation failure. TIIPE is also used during the impact evaluation stage to inform decisions about programme sustainment and scaling.

2 There are overlaps between these stages. Sometimes, two stages of an evaluation might happen in parallel, or the "end" of one stage will overlap with the beginning of another. For example, a feasibility evaluation may be conducted in the latter stages of programme design, or a pilot evaluation may take place at the start of an impact evaluation.

Evidence for validating or revising implementation theory

Programme theory and implementation theory are important at every stage of evaluation, but the relative emphasis on each can shift in each stage. The development of programme theory may have greater focus in earlier stages, with implementation theory gaining more prominence in later stages as the ToC becomes more stable.

There are no fixed rules about this, and some research questions may be relevant at several stages. It is not unusual to have overlap between programme theory and implementation theory considerations.

When thinking about evaluation stages, it is also critical to consider issues of equity, and how equity can be embedded in the evaluation activities, as well as the resources that are required to conduct a high quality TIIPE at each stage.

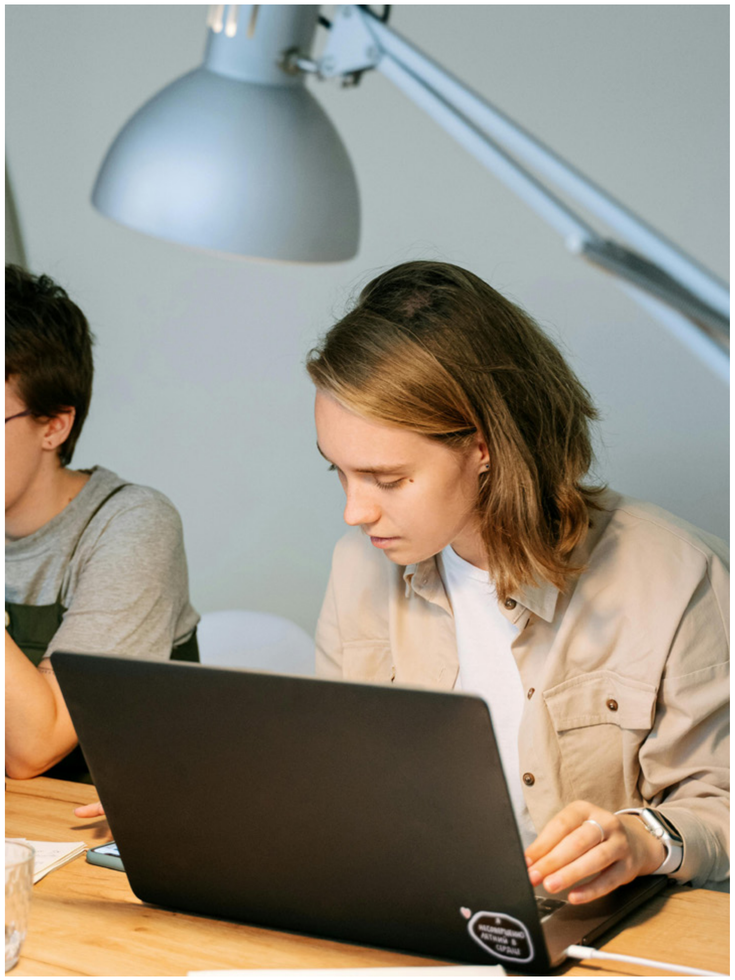
The next section outlines the role that programme theory and implementation theory can have in each evaluation stage, examples of evaluation questions that can be posed in each stage, and equity considerations.

Intervention design stage

What is it?

The development of both programme theory and implementation theory is a core focus of the intervention design stage. This stage is also when an organisation's capacity and infrastructure to conduct evaluations in the future will be built.

TIIPE activities in the intervention design stage are formative – the purpose is to support the process of design, theory development, or to refine implementation approaches.



What are the roles of programme theory and implementation theory in the intervention design stage?

Table 16: Focus of programme theory and implementation theory at the intervention design stage

Theory	Role of theory at this stage
Programme theory	To identify and refine the programme theory of how the programme is intended to reach its desired outcomes.
Implementation theory	To examine how the programme will be implemented in a way that supports the ToC. This includes identifying potential barriers and enablers to implementation, the intended implementation outcomes, and what strategies are needed to achieve the outcomes.

What questions could a TIIPE explore in the intervention design stage?

Questions that can guide the intervention design process are provided in Table 17. Those marked with an asterisk (*) indicate where a co-creation process or engagement with young people would be particularly valuable.

What are the equity considerations in the intervention design stage?

In the intervention design stage, equity considerations centre on ensuring the programme design is informed by a diverse group of young people. This means including their voices in the co-design of the programme and the ToC and seeking their inputs when, for example, understanding the determinants that relate to young people, or how best to reach them through the programme recruitment. Suggestions for how best to include the voices of young people in the design of the programme are described in **Chapter 2** – good practice principles for stakeholder engagement.

Theory	Role of theory at this stage	Role of theory at this stage
How is the programme intended to create impact?	Programme theory (PT)	Assumptions, activities, programme outcomes, mechanisms, causal pathways,
What are the barriers or enablers to implementation?	Implementation theory (IT)	Determinants
How should this programme be delivered?	PT and IT	Activities, assumptions Strategies
Is the programme appropriate for the young people it is aimed to reach?*	IT	Implementation outcomes: appropriateness
Does the programme address a problem that young people view as important?*	PT and IT	Problems (rationale), Determinants
Is the proposed approach to programme delivery aligned with the experiences of young people?*	PT and IT	Context Implementation outcomes: acceptability, appropriateness
Do any elements of the programme need refinement or adaptation before further evaluation?	PT and IT	Activities Strategies, implementation outcomes: feasibility
What elements of programme theory and implementation theory need further testing?	PT and IT	All
What data sources are available to measure programme and implementation theory?	PT and IT	All

Table 17: Testing programme theory and implementation theory at the intervention design stage

Feasibility evaluation stage

What is it?

The feasibility evaluation stage is when TIIPE is used to assess whether the programme can be delivered as intended, has a credible, plausible and logical ToC, and is ready to be evaluated through an impact evaluation. During this stage, TIIPE is also used to identify whether any changes or adaptations to the programme are needed ahead of pilot and impact evaluation.

What are the roles of programme theory and implementation theory in the feasibility evaluation stage?

Table 18: Focus of programme theory and implementation theory at the feasibility evaluation stage

Theory	Role of theory at this stage
Programme theory	• Testing the assumptions in the programme's ToC • Evaluating the mechanisms through which change occurs and the context in which these changes may occur • Developing or testing assumptions about which elements of the programme are core, and which are adaptive • Identifying the elements of the ToC to test further in an impact evaluation, and how this should be done
Implementation theory	• Evaluating if the programme can be delivered as intended • Assessing enablers and barriers to programme delivery • Examining the strategies that will lead to effective programme delivery • Examining key implementation outcomes and mechanisms of change • Identifying the elements of implementation theory to test in an impact evaluation, and how this should be done

What questions could TIIPE explore in the feasibility evaluation stage?

Table 19: Testing programme theory and implementation theory at the feasibility evaluation stage

Sample question	Relevant theory	Construct
Is there evidence the programme is contributing to its intended outcomes, as articulated in the ToC?	PT	Programme outcomes, mechanisms
Are the hypothesised mechanisms of change occurring as anticipated? If not, what alternative mechanisms are emerging?	PT	Programme outcomes, mechanisms, causal pathways
What is/are the context(s) in which the programme leads to change?	PT	Context, assumptions
Which components of the programme are essential for achieving intended outcomes, and which can be adapted without compromising effectiveness?	PT and IT	Activities, programme outcomes, implementation outcomes: fidelity
What aspects of context are enabling programme delivery or creating barriers to delivery?	PT	Determinants
Is the programme acceptable to young people and how are different groups experiencing it?	IT	Implementation outcomes: acceptability
Is the programme feasible for staff to deliver?	IT	Implementation outcomes: feasibility
What does it look like for the programme to be delivered with fidelity? What does high quality programme delivery look like?	IT	Implementation outcomes: fidelity
Does the programme reach the young people who are intended to benefit from it, especially young people in marginalised communities?	IT	Implementation outcomes: reach

↓

Table 19: Testing programme theory and implementation theory at the feasibility evaluation stage (contd)

Sample question	Relevant theory	Construct
Do any elements of the programme need refinement or adaptation before further evaluation? Does it need to be tailored to meet the needs of different young people ahead of the pilot evaluation stage? **(If yes – consider how changes can be made and if another feasibility study should be conducted before moving to a pilot or impact evaluation)**	PT and IT	Activities Strategies, implementation outcomes: feasibility
What programme outcomes should be measured in the impact evaluation, and when, and how should these be measured?	PT	Programme outcomes
What is a suitable business-as-usual condition for the control group to receive, and how should this be measured?	PT	Context, assumptions
What elements of programme theory need to be further tested at the impact evaluation stage, and when, and how should these be measured?	PT	All
What elements of implementation theory need to be further tested at the impact evaluation stage, and when, and how should these be measured?	IT	All

What are the equity considerations in the feasibility evaluation stage?

In the feasibility evaluation stage, it is important to assess whether the programme is feasible for and acceptable to the young people it is intended to support (particularly those who face the most barriers to accessing the programme). The feasibility evaluation stage is a chance to look more deeply into the barriers that specific groups of young people may face in participating in the programme and/or the evaluation. This can provide useful information for delivery partners on how elements of the programme or its delivery can be adjusted to facilitate inclusion, and also provides useful information to evaluators about where it may be most important to disaggregate data in the impact evaluation analysis. At this stage, young people can be also involved in co-designing the impact evaluation processes (e.g., how to introduce the trial, how to collect informed consent, the appropriate data collection methods and tools to use in the trial, etc.) and the research approach for the TIIPE in the impact evaluation stage (e.g., co-producing research questions, co-designing data collection tools).

Pilot evaluation stage

What is it?

Pilot evaluations, or pilot studies, are commissioned by Youth Futures when a programme is well-established, and prior evaluation has demonstrated that the ToC is credible, plausible, logical and stable, and that the programme can be delivered in a way that would be necessary in an impact evaluation. A pilot evaluation takes place when the process of designing the impact evaluation has started or has been completed but when there are still some questions about the best way to design and deliver the impact evaluation.

The main purpose of a pilot evaluation is to test elements of a proposed impact evaluation design, to determine if any changes need to be made before it is conducted.

What are the roles of programme theory and implementation theory in the pilot stage?

What are the roles of programme theory and implementation theory in the pilot stage?
Although the focus of a pilot evaluation is on testing the design of the impact evaluation, it is important to continue testing elements of the programme theory and implementation theory through the TIIPE.
conducted.

Table 20: Focus of programme theory and implementation theory in the pilot evaluation stage

Theory	Role of theory at this stage
Programme theory	• Pilot testing the approach to measuring components of the ToC that will be used in the impact evaluation (particularly programme outcomes) • If the duration of the pilot evaluation allows, testing the plausibility of the ToC
Implementation theory	• Pilot testing approaches to measuring aspects of the implementation (e.g., fidelity of programme delivery, implementation strategies) • Pilot testing approaches to measurement of implementation outcome constructs (e.g., acceptability, reach)

What questions could a TIIPE explore in the pilot evaluation stage?

Sample question	Relevant theory	Construct
Is it likely that the referral targets for the impact evaluation can be met?	N/A	N/A
Is the referral process feasible, acceptable and appropriate?	IT	Implementation outcomes: acceptability; appropriateness
Is the randomisation approach feasible, acceptable and appropriate?	IT	Implementation outcomes: acceptability; appropriateness
Are the "right" short-, medium- and long-term outcomes being measured?	PT and IT	Activities, programme outcomes, implementation outcomes: fidelity
Is the chosen approach to measuring outcomes feasible to deliver?	PT	Programme outcomes
Are intended changes occurring through the mechanisms and processes set out in the ToC?	PT	Programme outcomes, mechanisms, causal pathways
Is the chosen approach to measuring business-as-usual support feasible to deliver?	N/A	
Is the chosen approach to measuring fidelity appropriate?	IT	Implementation outcomes: fidelity
Are the fidelity criteria that have been selected the key ones that lead to desired outcomes?	IT	Implementation outcomes: fidelity

Table 21: Testing programme theory and implementation theory at the pilot evaluation stage

↓

Sample question	Relevant theory	Construct
Do the implementation strategies need to be refined to improve or maintain outcomes? • Do the implementation strategies lead to different (e.g., less desirable) outcomes in different contexts? • Are different implementation strategies required with different groups of young people?	IT	Strategies
Is the programme acceptable to a wider range of young people who are engaged with the programme?	IT	Implementation outcomes: acceptability
Has the programme been adopted at the delivery sites and by delivery staff?	IT	Implementation outcomes: adoption
Has the programme achieved penetration at the delivery sites (has it been integrated with other programmes that are offered)?	IT	Implementation outcomes: penetration
Is the programme reaching the most marginalised young people or those facing the most disadvantage?	IT	Implementation outcomes: reach
What is your approach to measuring cost or financial viability of the programme and is this an appropriate method for measuring cost	IT	Implementation outcomes: cost

Table 21: Testing programme theory and implementation theory at the pilot evaluation stage (contd)

What are the equity considerations in the pilot evaluation stage?

The pilot evaluation is an important stage in which to assess if:

- Equity monitoring works in practice and that key equity groups remain in focus.
- The programme and impact evaluation design is acceptable to them.
- Any different implementation strategies are required to reach them and deliver the programme well.

The pilot evaluation also provides valuable opportunities for meaningful engagement with young people. In the feasibility evaluation stage, young people can be invited to co-design the impact evaluation. This includes collaborating in decision-making about the programme outcomes that are measured, and how these will be measured (what tools/ approaches). These co-designed elements can then be tested and refined during the pilot evaluation stage, where the focus shifts to trialling the evaluation procedures, assessing the feasibility of data collection, and making adjustments based on young people's feedback and real-world delivery conditions.

Impact evaluation stage

What is it?

The impact evaluation stage is commissioned for well-established programmes when implementation has been stabilised and they are ready for a rigorous evaluation that uses high-quality methods, such as a randomised controlled trial, a quasi-experimental design or an alternative high-quality method. In addition to using rigorous methods, the impact evaluation stage looks more deeply at the context of findings, seeking to understand not just if a programme works, but how it works, for whom and in which circumstances – a TIIPE plays a crucial role in the latter.

Another important purpose of the impact evaluation stage is to understand why a programme required change in order to be delivered at the scale needed for the impact evaluation, how it might need to be changed for delivery at a larger scale, and the effect of these changes on desired outcomes – areas where a TIIPE can also support learning.

What is the role of programme theory and implementation theory in the impact evaluation stage?

Table 22: Focus of programme theory and implementation theory at the impact evaluation stage

Theory	Role of theory at this stage
Programme theory	• Continuing to test the ToC • Examining how and why the outcomes that have been achieved have occurred. This will inform decisions about how (or if) a programme can be sustained or scaled
Implementation theory	• Pilot testing approaches to measuring aspects of the implementation (e.g., fidelity of programme delivery, implementation strategies) • Pilot testing approaches to measurement of implementation outcome constructs (e.g., acceptability, reach)

What questions could a TIIPE explore in the impact evaluation stage?

The overarching question of an impact evaluation is: Was the programme more effective at (insert outcome of interest) in (specify which young people/context) than (specify the control condition)?

Table 23: Testing programme theory and implementation theory at the pilot evaluation stage

Sample question	Relevant theory	Construct
Was there evidence to support the ToC? • **Does the ToC or programme require further refinement in light of the findings?** • **Does the evidence support the key mechanisms of change previously identified?**	PT	Programme outcomes, mechanisms, causal pathways
Which groups of young people achieved the intended outcomes?	PT	Programme outcomes
Was the programme delivered with high quality and fidelity? • **Where did this vary and why?** • **What implications did this have for the outcomes achieved?**	IT	Implementation outcomes: fidelity
What combinations of strategies lead to improved implementation and/or programme outcomes?	PT and IT	Programme outcomes Strategies; implementation outcomes, e.g., acceptability
What business-as-usual support did the control group receive?	N/A	N/A
What challenges and supports are encountered when scaling the programme?	IT	Implementation outcomes: scaling; determinants

Table 23: Testing programme theory and implementation theory at the pilot evaluation stage (contd)

Sample question	Relevant theory	Construct
What was the cost of programme delivery?	IT	Implementation outcomes: reach
Which group(s) of young people were reached / not reached by the programme?	IT	Implementation outcomes: fidelity
Are the fidelity criteria that have been selected the key ones that lead to desired outcomes?	IT	Implementation outcomes: fidelity
What was the cost of programme delivery?	IT	Implementation outcomes: reach
Which group(s) of young people were reached / not reached by the programme?	IT	Implementation outcomes: fidelity
Do the findings of the impact evaluation suggest that the programme should be embedded into the organisation's existing delivery and if not, why not?	IT	Implementation outcomes: penetration; determinants

What are the equity considerations in the impact evaluation stage?

As in the pilot evaluation stage, the key equity consideration in the impact evaluation stage is whether those facing the biggest barriers to engagement were reached and retained. If not, it is important to explore why not and what would be required to increase engagement with these groups. At the impact evaluation stage, there should also be an exploration of whether there were any differences in programme effectiveness and implementation among different groups of young people. Through considering how the programme works for different groups of young people, a TIIPE can make recommendations for how a programme could achieve more equitable outcomes and make recommendations for how the programme should be tailored to specific groups of young people.

At this stage, young people could be invited to join the research team to engage in sense-making around data analysis and interpretation. This is particularly important to ensure that their insights and experiences inform decision-making about sustainment and scalability of the programme based on findings of the impact evaluation.

Chapter 6

Appendix

youth futures FOUNDATION

List of key references

Atkins, L., Francis, J., Islam, R., O'Connor, D., Patey, A., Ivers, N., Foy, R., Duncan, E.M., Colquhoun, H., Grimshaw, J.M. and Michie, S. (2017) A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implementation Science, 12, 77. https://doi.org/10.1186/s13012-017-0605-9

Belcher, B.M., Claus, R., Davel, R. and Ramirez, L.F. (2020) A refined method for theory-based evaluation of the societal impacts of research. Humanities and Social Sciences Communications, 7, 153. https://doi.org/10.1057/s41599-018-0139-z

Collier, D. (2011) Understanding process tracing. Political Science Department, University of California, Berkeley. Available at: https://polisci.berkeley.edu/sites/default/files/people/u3827/Understanding%20Process%20Tracing.pdf

Damschroder, L.J., Reardon, C.M., Widerquist, M.A.O., Lowery, J. and Aron, D.C. (2022) The updated Consolidated Framework for Implementation Research based on user feedback. Implementation Science, 17, 75. https://doi.org/10.1186/s13012-022-01245-0

De Silva, M.J., Breuer, E., Lee, L., Asher, L., Chowdhary, N., Lund, C. and Patel, V. (2014) Theory of Change: A theory-driven approach to enhance the Medical Research Council's framework for complex interventions. Trials, 15(1), 267. https://doi.org/10.1186/1745-6215-15-267

Fletcher, A., Jamal, F., Moore, G., Evans, R.E., Murphy, S. and Bonell, C. (2016) Realist complex intervention science: Applying realist principles across all phases of the Medical Research Council framework for developing and evaluating complex interventions. Evaluation, 22(3), pp. 286–303. https://doi.org/10.1177/1356389016652743

Glasgow, R.E., Vogt, T.M. and Boles, S.M. (1999) Evaluating the public health impact of health promotion interventions: The RE-AIM framework. American Journal of Public Health, 89(9), pp. 1322–1327. https://doi.org/10.2105/ajph.89.9.1322

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R. and Kerr, K. (2016) Implementation and process evaluation for interventions in educational settings: An introductory handbook. Available at: https://www.researchgate.net/publication/328761292

Lewis, C.C., Boyd, M.R., Walsh-Bailey, C., Lyon, A.R., Beidas, R., Mittman, B., Aarons, G.A., Weiner, B.J. and Chambers, D.A. (2020) A systematic review of empirical studies examining mechanisms of implementation in health. Implementation Science, 15, 21. https://doi.org/10.1186/s13012-020-00983-3

Moore, G.F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D. and Baird, J. (2015)

Process evaluation of complex interventions: Medical Research Council guidance. BMJ, 350, h1258. https://doi.org/10.1136/bmj.h1258

Moullin, J.C., Dickson, K.S., Stadnick, N.A., Albers, B., Nilsen, P., Broder-Fingert, S., Mukasa, B. and Aarons, G.A. (2020) Ten recommendations for using implementation frameworks in research and practice. Implementation Science Communications, 1(1), 42. Available at: https://implementationsciencecomms.biomedcentral.com/articles/10.1186/s43058-020-00023-7

Nilsen, P. (2015) Making sense of implementation theories, models and frameworks. Implementation Science, 10, 53. https://doi.org/10.1186/s13012-015-0242-0

Pawson, R. and Tilley, N. (1997) An introduction to scientific realist evaluation. In: Chelimsky, E. and Shadish, W.R. (eds.) Evaluation for the 21st century: A handbook. Thousand Oaks, CA: Sage Publications, pp. 405–418. https://doi.org/10.4135/9781483348896.n29

Peters, D.H., Tran, N.T. and Adam, T. (2020) Implementation research in health: A practical guide. BMC Medical Research Methodology, 20, 240. https://doi.org/10.1186/s12874-020-01164-y

Powell, B.J., Waltz, T.J., Chinman, M.J., Damschroder, L.J., Smith, J.L., Matthieu, M.M., Proctor, E.K. and Kirchner, J.E. (2015) A refined compilation of implementation strategies: Results from the Expert Recommendations for Implementing Change (ERIC) project. Implementation Science, 10, 21. https://doi.org/10.1186/s13012-015-0209-1

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R. and Hensley, M. (2011) Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. Administration and Policy in Mental Health and Mental Health Services Research, 38(2), pp. 65–76. https://doi.org/10.1007/s10488-010-0319-7

Proctor, E.K., Bunger, A.C., Lengnick-Hall, R., Gerke, D.R., Martin, J., Phillips, R. and Swanson, J.C. (2023) Ten years of implementation outcomes research: A scoping review. Implementation Science, 18(1). https://doi.org/10.1186/s13012-023-01286-z

Skivington, K., Matthews, L., Simpson, S.A., Craig, P., Baird, J., Blazeby, J.M., Boyd, K.A., Craig, N., French, D.P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M. and Moore, L. (2021) A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance. BMJ, 374, n2061. Available at: https://www.bmj.com/content/374/bmj.n2061

Smith, J.D., Li, D.H. and Rafferty, M.R. (2020) The Implementation Research Logic Model: A method for planning, executing, reporting and synthesising implementation projects. Implementation Science, 15, 84. https://doi.org/10.1186/s13012-020-01041-8

Treasury Board of Canada Secretariat (2012) Theory-based approaches to evaluation: Concepts and practices. Ottawa: Government of Canada. Available at: https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/theory-based-approaches-evaluation-concepts-practices.html

Waltz, T.J., Powell, B.J., Matthieu, M.M., Damschroder, L.J., Chinman, M.J., Smith, J.L., Proctor, E.K. and Kirchner, J.E. (2015) Use of concept mapping to characterize relationships among implementation strategies and assess their feasibility and importance: Results from the Expert Recommendations for Implementing Change (ERIC) study. Implementation Science, 10, 109. https://doi.org/10.1186/s13012-015-0295-0

youthfuturesfoundation.org

youth futures FOUNDATION