Delivered by

West Yorkshire Combined Authority | Tracy Brabin Mayor of West Yorkshire

West Yorkshire Violence Reduction Partnership

ECORYS | Answering tomorrow's challenges today

Theory of Change, Monitoring, and Evaluation

A toolkit for developing a culture of evaluation and learning

How to navigate this document

Click on the ‘home’ icon to return to the main contents.

Click on the arrows to move backwards and forwards through the document.

Click on these icons to navigate to the contents of the relevant section.

Where you see this icon read the instructions to access the Theory of Change templates.

Where you see this icon read the instructions to access the online presentations.

Click on this icon to go to the ‘draw.io’ website.

Contents

We have created some short presentations explaining the purpose and content of each section of this toolkit

You can watch the presentations by clicking on the relevant image below or clicking on the YouTube icon at the beginning of sections 1-5.

Do this before reading the toolkit section, so that you know what to expect and how Theory of Change, Monitoring, and Evaluation fit together.

Presentation 1
Introduction

Presentation 2
Theories of Change

Presentation 3
Monitoring: Understanding if your project is working

Presentation 4
Evaluation: Laying the foundations

Presentation 5
Evaluation methods

1.0

Introduction to the toolkit

Delivered by

West Yorkshire Combined Authority | Tracy Brabin Mayor of West Yorkshire

West Yorkshire Violence Reduction Partnership

ECORYS | Answering tomorrow's challenges today

1.0 Introduction to the toolkit

This toolkit was funded by the West Yorkshire Violence Reduction Partnership (WYVRP) and has been created by Ecorys, an independent research and evaluation organisation. The toolkit aims to equip those delivering interventions in West Yorkshire to contribute to an evidence base on what works in preventing and reducing violent crime. It gives partners the building blocks to create Theories of Change for projects which, in turn, support the creation of robust monitoring systems and evaluation techniques.

The toolkit was written by George Horton, Maya Hill-Newell, Tave Browett, and Kate Smith at Ecorys. The authors would like to thank the Knowledge Hub team at WYVRP, and the partners who contributed to consultations and discussions in the development process.

Why is monitoring and evaluation important?

Understanding what works well and what doesn't is vital for publicly funded interventions, and particularly so when those interventions are designed to support some of the most vulnerable people in society. Effective monitoring can help organisations to refine their interventions as they are being delivered to get the maximum impact for service users.

It's important that innovative approaches are evaluated so they can help shape future funding approaches and policy, and so we can do more of what works. Having a strong evidence base around what you did well is also really helpful for securing your own future funding and demonstrating your value.

Finally, we understand that resources can be tight for evaluation. As a result, it's really important to make sure that you're focused on what you need to know as early as possible; this toolkit can help you to pinpoint what you need to evidence, and how to collect and use the data effectively.

Using the toolkit

This toolkit contains advice and resources for practical approaches to embedding a culture of learning in your organisation. You'll find most benefit from reading it as a complete document, even if your organisation already has a Theory of Change, or already evaluates projects. It is based around best practice, so may give you some ideas for tweaking or adjusting your approaches. However, if you do need to dip in and out of the document to find guidance on specific aspects of Theory of Change and evaluation, then each section can work in a standalone way.

The document is not just aimed at those carrying out evaluation. It is also helpful for frontline staff (who may not usually be involved in Theory of Change or evaluation) to support learning approaches being embedded in day-to-day delivery.

2.0

Theories of Change:
Understanding your project

Delivered by

West Yorkshire Combined Authority | Tracy Brabin Mayor of West Yorkshire

West Yorkshire Violence Reduction Partnership

ECORYS | Answering tomorrow's challenges today

What is a Theory of Change?

A Theory of Change (ToC) shows how an intervention (or project, programme, or policy) is intended to work. The 'Change' refers to the difference that you wish to make, and the 'Theory' is how you will try to make this happen (in other words, how what you do might create change). It is like a roadmap that indicates how the inputs and activities are expected to produce the desired outcomes and impact.[1]

> *A 'Theory of Change' explains how activities are understood to produce a series of results that contribute to achieving the final intended impacts.*
>
> **UNICEF**

A ToC can help you and your organisation visualise and understand what you are trying to achieve, why, and how. It can serve several important functions, such as helping you to think critically about whether a response will work in your local context, communicate with partners and commissioners about your plan, identify how to maximise the likelihood of meeting the intervention's desired outcomes, and monitor the progress and impact of your work.[2] The different uses of ToCs are discussed in more detail in section 2.2.

It is helpful to think of a ToC as a simplified diagram or narrative that can be built-upon, adapted and re-written over time. ToCs are sometimes called Logic Models or Logical Frameworks and can be visualised in many ways. Whichever approach or terminology you use, it is important that your ToC makes sense to you and your involved stakeholders.

- Some ToCs can be **project-level** (used for specific projects, such as your intervention), others can be more **high-level**, used for policies or wider programmes of work (such as the whole organisation, or the whole VRU). This toolkit focusses on project-level ToC examples.
- Some are more detailed than others. The underlying principle of ToCs is to show the '**causal pathways**' from your activities to your intended outcomes and impact. This means the chain of results that you anticipate your activities will set in motion, to arrive at specific outcomes. Specifically, causal links show which activities lead to which outputs, which outputs lead to which outcomes, and which outcomes lead to which impacts. Some ToCs include arrows to help visualise these pathways, but basic ToCs do not.
- Some ToCs have more components than others (see components below).

1 College of policing, 2021, Knife crime A problem solving guide, p5. Available at: https://assets.college.police.uk/s3fs-public/2021-11/Knife-crime-a-problem-solving-guide.pdf

2 College of policing, 2021, Knife crime A problem solving guide, p5. Available at: https://assets.college.police.uk/s3fs-public/2021-11/Knife-crime-a-problem-solving-guide.pdf

There are some common pitfalls that people encounter in understanding what a ToC is, or what it is for:

- It should not show operational details of an organisation; it is not a delivery plan exploring how you will deliver your outputs. Rather, it is to help you understand how your outputs (and the activities you deliver to create those outputs) will lead to your intended outcomes and impact.
- A ToC should not focus on a particular project as a snapshot, or at a particular point in time (or in chronological order).
- It should not be a 'tick-box' exercise that you complete and never look at again, but rather a tool that you revisit and adapt as your intervention and priorities evolve. It should be treated as a reference point and 'living document'.[3]

Here is a glossary of key terms, to help you better understand the language around ToCs, which we'll explore in this chapter.

As Figure 1 helps to show, a ToC traces the logic behind an intervention, visualising how and why inputs and activities are expected to lead to the intended outcomes and impact. It also shows what factors need to be in place (assumptions) for the change to be realised and considers any barriers (risks) which may prevent or delay this. Not all ToCs include this, but we believe setting out your assumptions (and the associated risks) is an important step to help you identify the casual pathways between different components of your ToC.

Some ToCs may also include information on the **context** the project is working in (i.e., other programs or funding in place locally, which may have an impact on the extent to which outcomes are met), or possible **unintended outcomes** which may arise following the activities. Also, some ToCs separate outcomes into 'immediate' or 'short-term' and 'medium-term' outcomes where appropriate.

Although there is no fixed way to present a ToC, **it should always cover inputs, activities, outputs, outcomes and impact**.

3 TSIP, Theory of Change: eight common mistakes.
Available at: https://www.tsip.co.uk/blog/2019/7/18/theory-of-change-eight-common-mistakes-9bt5f

Components of a Theory of Change

It is helpful to think of a ToC as a diagram with several, distinct components. The basic ToC presented in this toolkit has six columns, explained in the template below.

Figure 1: Components of a Theory of Change

Rationale	Inputs	Activities	Outputs	Outcomes	Impact
The reason, or problem statement, for the proposed intervention. It should be evident that the intervention focused on addressing an identified need (i.e. links to the VRU SNA and Response Strategy). **Assumptions** Assumptions are a necessity to a ToC. Many assumptions can be captured with casual pathways (arrows between components) but others need to be mentioned separately. Assumptions include the necessary conditions for change, such as the resources that need to be in place for the planned changes to occur. **Risks** What are the key risks to the ToC? Focusing on the assumptions not being met are is a useful start point.	The inputs required to deliver the proposed activities. Financial and non-financial (e.g. time in kind, existing process/ systems) should be included.	The proposed activities to deliver the intended outputs/outcomes. As well as the direct support to clients, supporting activities (e.g. referral/engagement systems, staff training) should also be included.	The immediate (often quantifiable) outputs from activities. Typically, we would be interested in the number of clients engaged and subsequently supported. Other examples include the number of (group) sessions delivered and staff/ organisations trianed.	The intermediate outcomes that (are assumed) to lead to/support the intended impacts. Typically outcomes are observable shortly after the intervention. Examples include positive changes in mental wellbeing, accessing services, and progress towards positive life outcomes (e.g. education/ employment.	Impacts can be considered as the ultimate outcomes of an intervention. In the VRU context, impacts will likely focus on reductions in crime/violence. Depending on the type of intervention and intended target group, impacts can take longer to materialise and be observed.

2.1 Theory of Change worked examples

The following pages provide example ToCs we have created for fictional violence reduction and prevention projects. Figure 2 explores a locality-based youth work project which includes detached outreach and individual, one-to-one advice sessions. The arrows highlight the assumed **casual links** between inputs, activities, outputs, outcomes and impacts; for example, you can see that the existing data-sharing arrangements between partners (the input) facilitates the identification of hotspots (activity) in which to deliver outreach sessions (output). These outreach sessions support the identification of individual young people in need of one-to-one support; once that support has been put in place, outcomes such as improved social and emotional skills can be realised, leading to longer term impacts including reduced criminality.

Figure 3 provides an example for a Hospital Navigator project, and Figure 4 focuses on a project designed to divert perpetrators of violence against women and girls VAWG crimes from reoffending.

Those items presented with a dotted line represent components which are vital to the project but which we assume will not lead to impact alone, without the other components of the ToC. It is not necessary to include this level of detail when creating your own ToC, but we have included this to show the extent to which you could think through your causal links.

Figure 2: Local based interventions Theory of Change

Rationale

Young people are more likely to engage within a space of their choosing. Aim to positively influence their lives by providing support and guidance and engaging them in a programme of informal social education.

Assumptions

Identifiability of gang hotspots and individuals therein. Prompt availability (and take-up) of detached sessions. Availability and take-up of tailored follow-up support. Good relationships with police and community support services.

Risks

Low take-up, difficulty getting access to areas and individuals, difficulty getting data sharing agreements in place.

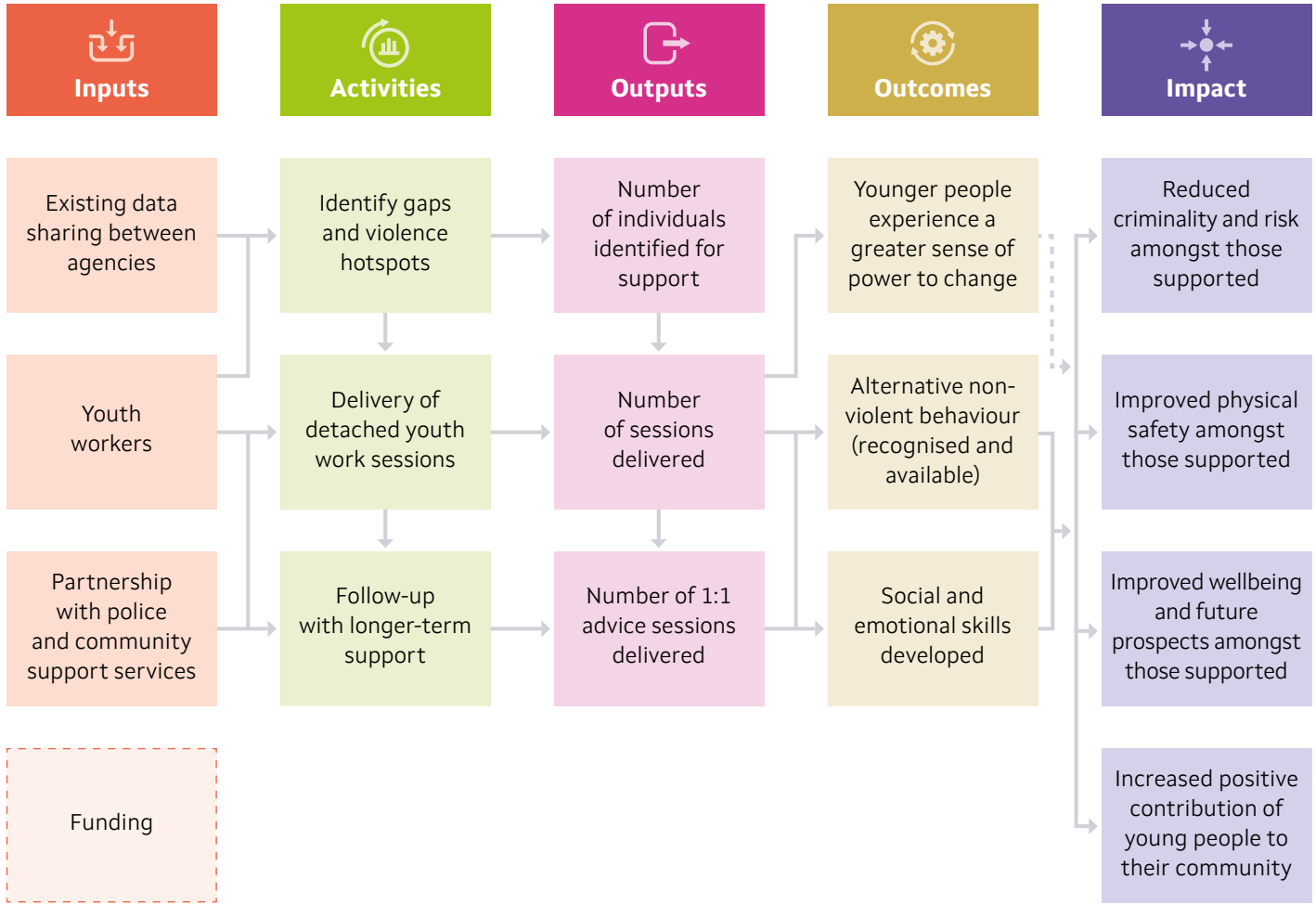


Figure 3: Hospital navigators Theory of Change

Rationale

Provide alternative, non-violent, life choices to victims of knife crime at a time when they are likely most amenable to change.

Assumptions

Substantial overlap between victims and offenders, identifiability of knife crime victims (who may also be offenders), prompt availability (and take-up) of credible providers of initial interventions and availability and take-up of follow-up support tailored to victim/offender needs.

Risks

Potential unintended/negative outcomes/impact from labelling victim as a potential offender and/or increased need to self protect/reintegrate with peers owing to recognition of vulnerability.

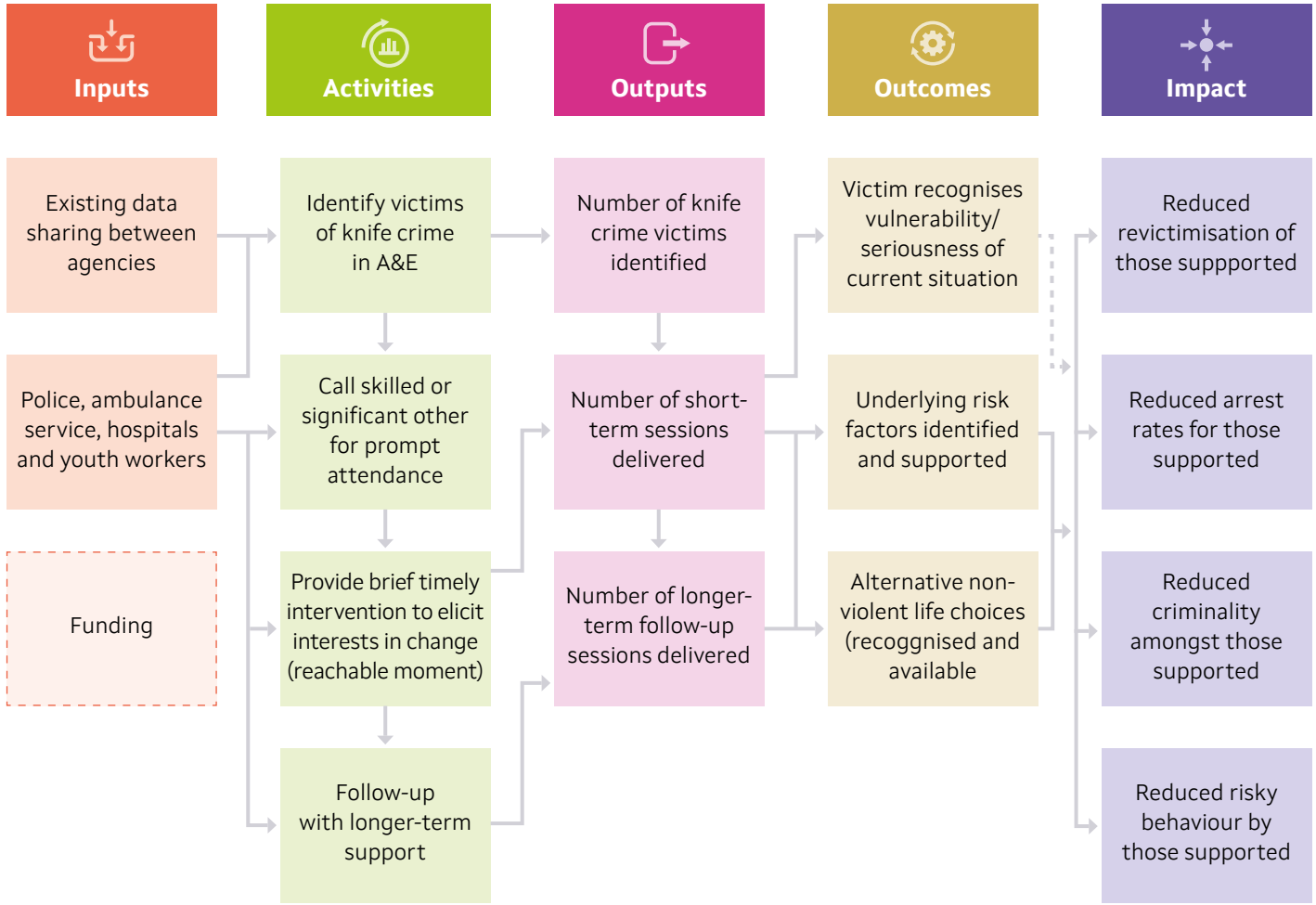


Figure 4: VAWG/criminal justice Theory of Change

Rationale

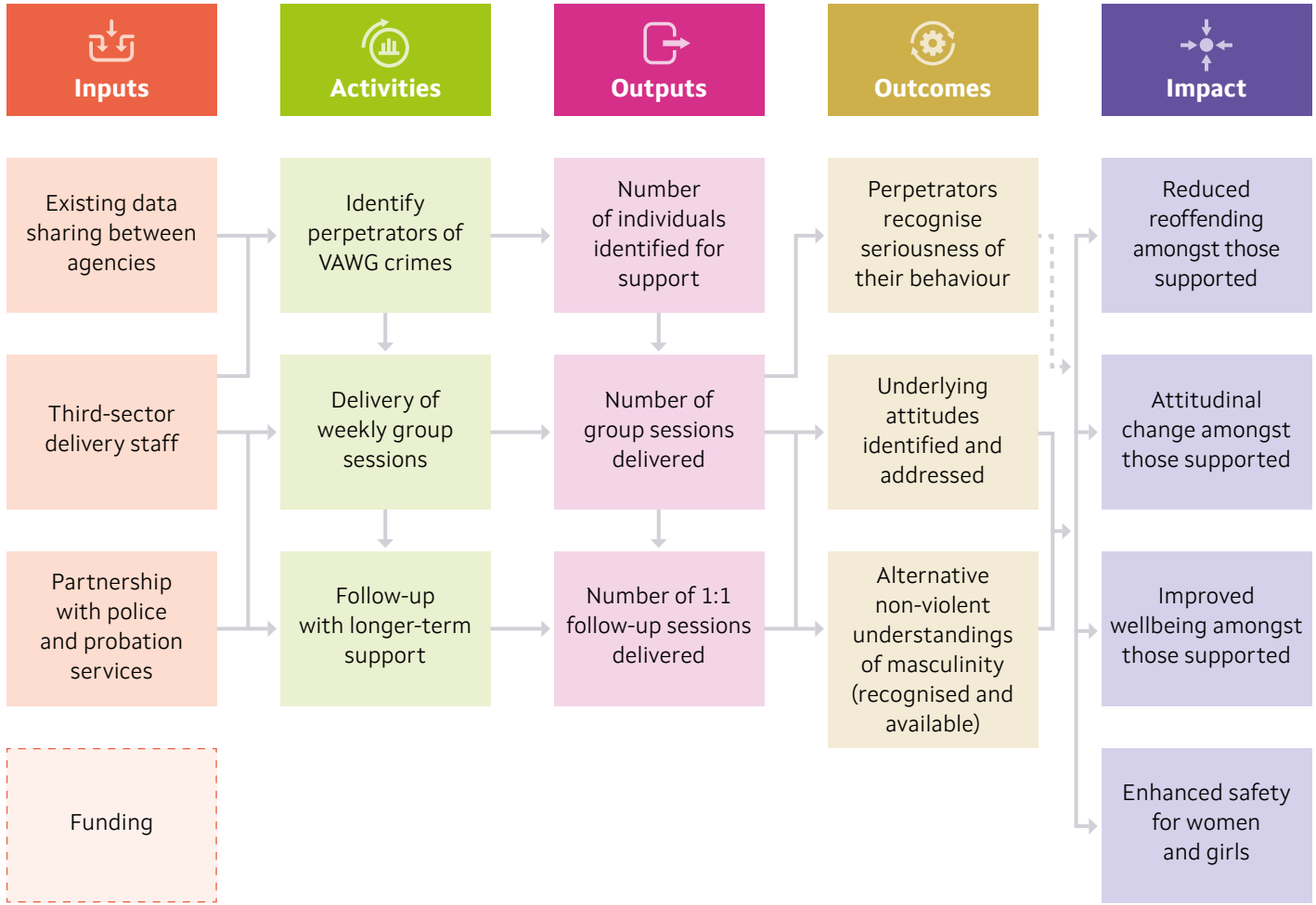
Perpetrators of VAWG crimes hold toxic attitudes towards women and masculinity. There are a lack of perpetrator-focused interventions which aim to hold perpetrators to account, while educating them and offering them opportunities to change their attitudes and behaviours.

Assumptions

Identifiability of gendered and sexual violence perpetrators (who may also be victims). Prompt availability (and take-up) of group sessions. Availability and take-up of tailored follow-up support/good relationships with probation services.

Risks

Low take-up, difficulty getting access to perpetrators, difficulty getting data sharing agreements in place due to sensitive nature of the data.



2.2 When and how to use Theories of Change

ToCs are an important tool for all stages of project delivery – they are flexible documents that can be used before, during and after an intervention. They can provide a great opportunity to ensure that your thinking is consistent from the design stage to the post-implementation stages, and in all decision-making processes. As such, you should use ToCs during:

(Co) Design of the project

- Developing activities to achieve the intended outcomes/impacts – it will support you in designing your service in a way that will optimise achieving the outcome you are looking for.
- Identifying the inputs required to deliver the activities effectively.

The development of a business case

- Ensuring the intervention meets identified needs, for example, does it link to the VRU's Strategic Needs Assessment (SNA)?
- Demonstrating to partners and funders that the intervention is logical and supported by evidence.

Delivery

- Effectively communicating the approach to delivery teams and referrers.
- Identifying when an intervention is not being delivered as intended and/or isn't leading to desired outcomes.

Evaluation

- Knowing what you are evaluating from the outset, helping to define your research questions.
- Focused and effective monitoring and evaluation approaches can be put in place from the outset of delivery.

2.3 How to create a Theory of Change

Developing a ToC should be an iterative process. By going through the development stages detailed below, we are forced to focus on the key elements and assumptions of the intervention to arrive at a ToC that is logical and accessible. This requires some critical thinking about why you are doing what you are doing.

There are four key stages involved in developing a ToC, which we'll explain in this section. They are:

- Desk review
- Initial draft
- Stakeholder consultation
- Re-draft/re-consult/finalise.

As we noted earlier, you should also be prepared to revisit your ToC throughout your project lifecycle, and update and amend it as necessary.

Desk review

When you are developing a ToC, it is very important to understand the problem you are trying to address with your intervention or project. Begin with an analysis of the context and situation, focussing on the background to the project, and outlining the problem in society you are trying to address, your target group, and their needs and characteristics.[4] Your ToC should be based on evidence, which can come from:

- **Wider literature:** Many VRU interventions are based on/adapted from tried and tested interventions, which likely have a ToC and/or supporting evidence. Whilst it is important to develop a ToC based on your specific intervention (e.g. within your local context), existing evidence is there to be used.
- **VRU/intervention documents:** For example, VRU Strategic Needs Assessments should provide the rationale and intended outcomes (that activities should be developed to meet), and business cases for existing interventions will detail the proposed activities/outcomes. This can also include evidence that your organisation has already collected, and any relevant published research.
- **The experiences of stakeholders:** Knowledge about violence and crime prevention from internal and external stakeholders (see page 20).

You might find some evidence that contradicts your theory. It is important to think this through and, if necessary, modify your approach or activities to reflect what the evidence tells you.[5] During these early stages of the ToC development process, you should ask the following questions and use research/evidence to determine answers:

1. What is the change that your programme is trying to achieve?

2. Why do you think this change is needed?

3. How will you achieve this change?

4. What is the context that might affect the ToC?

5. Who will benefit?

6. What is the most appropriate intervention for this process of change?

7. Are there alternative strategies to meet your aims?

8. What does existing evidence say about what works?[6]

The evidence can be used to establish the **rationale** (problem statement) of your ToC.

4 Harries. E., Hodgson. L., and Noble. J., 2014. Creating your Theory of Change, p8. Available at: https://golab.bsg.ox.ac.uk/documents/Creating-your-theory-of-change1.pdf

5 Harries. E., Hodgson. L., and Noble. J., 2014. Creating your Theory of Change, p8. Available at: https://golab.bsg.ox.ac.uk/documents/Creating-your-theory-of-change1.pdf

6 Government Analysis Function,2023, The Theory of Change Process – Guidance for Outcome Delivery Plans. Policy details. https://analysisfunction.civilservice.gov.uk/policy-store/the-analysis-function-theory-of-change-toolkit/#resources-to-help-you-assess-evidence

Initial draft

Once you have collected evidence and established the rationale for your intervention, you can begin to draft your ToC. You might find it useful to start by hand-writing your first draft, using sticky notes and a blank ToC template (see section 2.4.) to start with. Your initial draft should be based on the desk review to provide a starting point for stakeholder consultations. Here we provide some guidance on identifying what should be included in the different elements of your ToC; distinguishing between outputs and outcomes, understanding what outcomes should be, and distinguishing between outcomes and impact.

Distinguishing between outputs and outcomes

When developing a ToC, it is vital to know the difference between outputs and outcomes. Outputs are **a measure of activity**. In the context of youth violence prevention, common outputs include the number of young people engaged, the number of sessions delivered, the number of knives seized or the creation of a partnership. Outcomes refer to **the effect of those activities**. This might include improved educational attainment or wellbeing, or reduced knife carrying amongst those young people supported. Put simply, the outcomes are the change you wish to see, and the outputs are the activities or actions that will contribute to achieving that change.

Distinguishing outputs from outcomes is important because outputs may not necessarily lead to positive outcomes, and therefore do not accurately convey the impact of an intervention on their own. Outputs are important; for example, engaging 10 young people in youth work sessions is a good thing. However, this alone does not tell us anything about the effect of that work; it may not lead to discernible reductions in the overall levels of knife crime. Think about the outputs as the 'what' of your project. The outcomes are the 'so what?' in terms of what the outputs will achieve. It is important to commit to being outcomes-focussed- that is, to understand the 'so what?' of outputs. This involves focussing on measurable changes which can be attributed to your outputs.

What should outcomes be?

When you are defining your outcomes, it can be useful to consider the changes in behaviour, wellbeing, attitudes, skills, and abilities you expect to see. Use words like 'increased', 'improved', 'enhanced' or 'reduced' to describe the change you would like your programme to have. Ensure that your outcomes are SMART:

Specific: Components of the ToC, and how they connect, should be specific.

Measurable: Elements of the ToC should be measurable.

Achievable: The proposed activities and outputs, and anticipated outcomes/impacts should be, theoretically, achievable.

Relevant: Focus on what is most relevant (and attributable) to the intervention.

Time-bound: Consider the timeframes to deliver activities and achieve the anticipated outcome/impacts.

It is important that your outcomes are easy to understand by staff, beneficiaries, and stakeholders.

You can use existing outcomes frameworks to select outcomes you want to work towards. Some resources:

- We recommend the South Wales Violence Prevention Unit Outcomes Indicators – See 2. Data Indicators from page 6 onwards. The first two columns include indicators and measures for a given outcome. This resource also includes data sources for measuring these outcomes – including secondary datasets and primary methods.
- Lancashire Violence Reduction Network Outcomes Framework 2020.
- 'Home Office Indicators of Integration framework 2019 third edition' guidance on GOV.UK – select the 'Home Office Indicators of Integration framework 2019: Theory of Change Guide notes Part B' link to download the PDF and go to page 22 to read the section called 'Annex 2: The bank of Outcome Indicators of Integration.
- For measuring outcomes, you can also use the Outcomes STAR (which you will need training and a license for). This can be tailored to specific services, and is widely used across Europe, USA, and Australia.

Distinguishing between outcomes and impact

The difference between outcomes and impact lies in the fact that outcomes are typically observable during, or shortly after, the delivery of your intervention. For example, you may be able to measure an improvement in a young person's educational attainment during several weeks or months of working with them. However, the broader impact on knife crime or violence will probably not be observable for several years. Think of the impact as the 'bigger picture', long-term change you wish to create, and the outcomes as the short or medium-term, measurable precursors.

'Backwards mapping'

Backwards mapping is one useful approach to completing your draft ToC. This involves beginning your draft ToC with your long-term goal and working 'backwards' toward the earliest changes that need to occur.

First, ask yourself 'what is the overall change that this program is trying to achieve?'. A final goal or **impact** should be relatively long-term, obviously beneficial, and something that funders, commissioners or supporters would be interested in funding. It must be relevant to the needs of the target population, and plausibly linked to your project's activities.[7] Impacts can relate to the individual young people you work with, for example, reduced criminality amongst those supported, or an improvement in their wellbeing or future prospects. You can also consider broader societal impacts, for example, reduced knife crime, or enhanced safety for women and girls, wider systemic impacts, or even impacts for your organisation.

7 Harries. E., Hodgson. L., and Noble. J., 2014. Creating your Theory of Change, p7. Available at: https://golab.bsg.ox.ac.uk/documents/Creating-your-theory-of-change1.pdf

Once you have defined the impact you wish to achieve, you can then work backwards to plot the preceding stages and understand how this change will happen. This requires defining the **outcomes** which will support your planned impact (see section 2.3 on what outcomes should be).

You are now ready to plot which **outputs** and **activities** will help you to bring about the outcomes. There is some overlap between these two aspects of the ToC so it makes sense to work on these areas together; remember, the activities component of the ToC describes what you will do, and the outputs quantify those activities. Work backwards from your outcomes and think about what outputs will be needed to bring them about.

When you are defining your activities and outputs, it can be useful to consider questions such as:

1. What is the direct result of these activities?

2. What outputs need to be produced to achieve your outcomes?

This will involve thinking creatively about the outputs that will be most effective in bringing about your desired outcomes, and when they are best delivered. It is important to refer back to your desk research here, to consider carefully if your planned outputs will lead directly to your desired outcomes. For example, ask yourself *why* and *how* workshops will lead to improved social and emotional skills for young people, and whether there is evidence or past experience to support this.

At this point, also think more broadly about which activities to include in your ToC to support the achievement of your outputs. As well as your direct delivery (i.e., workshops, group or 1:1 sessions, sports-based activities), this might also include staff training, identification of young people, or the establishment of referral pathways. It is important to have a clear sense of the activities which could help you to reach your desired goals. This will help you work out what is achievable and where there might be gaps.[8] Some outputs will involve collaboration with other agencies and some outcomes may only be achieved if other services are also involved in some way, so these contact points or joint activities will need to be charted as part of your ToC.

If you are creating a ToC for an existing project, plot your existing activities and outputs. This is a good chance to discuss to what extent outputs are delivering your desired outcomes and may lead to some revision of what your project does. At this point, it is also helpful to explore which outcomes your work directly contributes to, and which ones are beyond the scope of your work. When using your ToC for evaluation, this step is essential to ensure you evaluate the right things (see section 5).[9]

8 Government Analysis Function, 2023, The Theory of Change Process – Guidance for Outcome Delivery Plans. Policy details. https://analysisfunction.civilservice.gov.uk/policy-store/the-analysis-function-theory-of-change-toolkit/#resources-to-help-you-assess-evidence

9 NCVO, How to build a Theory of Change. Available at: https://www.ncvo.org.uk/help-and-guidance/strategy-and-impact/strategy-and-business-planning/theory-of-change/#how-to-build-a-theory-of-change

Next, identify your **inputs**. Inputs could include resources such as funding, existing operational structures and intelligence (i.e., data sharing), people and partners. Consider which resources you have available and any limitations on them.

When you are thinking about the limitations on resources you might consider questions like:

1. Is your budget ring-fenced for a specific purpose?

2. Do the people you can deploy have the correct skills?

3. Is there a time-constraint on the availability of facilities?[10]

Once you have plotted out your ToC (using the steps above), you can reflect on the **assumptions** that underpin it. Assumptions are the conditions that need to be in place to make the theory work. Thinking through assumptions is helpful to do as a team, to ensure that you agree on why one outcome will lead to another. This can draw upon evidence, common sense, and intuition. Consider the different roles of involved stakeholders, and whether they have access to the required resources (i.e., training, time, Data Sharing Agreements), and the logic behind your theory. When you are defining the assumptions, it can be useful to consider questions such as:

1. What assumptions, tested and untested, are you making for this project to work?

2. What beliefs, attitudes, behaviours, and values have shaped your ToC?

3. What evidence is there to support the links you are including in your plan?

4. What is the quality of this evidence and how reliable is it?

5. What has worked well in previous projects?

6. What do you expect to remain the same to enable your project to be carried out as planned?[11]

As you identify assumptions, consider which are critical – that is, those on which your theory rests, and which may derail your project if they turn out to be false.[12]

This leads us to identifying **risks**, which will help you assess the limitations of your programme and develop appropriate plans to minimise negative consequences. Consider what might disrupt the flow of your ToC, and how you can mitigate risks within your control. Identifying assumptions is linked to risks, for example, you may assume that young people will engage throughout the programme, but a risk is that they will disengage, which will affect the outcomes.

10 Government Analysis Function, 2023, The Theory of Change Process – Guidance for Outcome Delivery Plans. Policy details. https://analysisfunction.civilservice.gov.uk/policy-store/the-analysis-function-theory-of-change-toolkit/#resources-to-help-you-identify-assumptions

11 Government Analysis Function, 2023, The Theory of Change Process – Guidance for Outcome Delivery Plans. Policy details. https://analysisfunction.civilservice.gov.uk/policy-store/the-analysis-function-theory-of-change-toolkit/#resources-to-help-you-identify-assumptions

12 NCVO, How to build a Theory of Change. Available at: https://www.ncvo.org.uk/help-and-guidance/strategy-and-impact/strategy-and-business-planning/theory-of-change/#how-to-build-a-theory-of-change

Finally, you may want to apply **causal links** (arrows) between different components of your ToC. Examine how each of the components connect and lead to each other; this will help you to show which activities lead to particular outcomes, and which outcomes lead to each impact. This process will help you to visualise and demonstrate how your program will have an effect, justify spending, and monitor important steps. You can either do this at the end of your initial draft, or as you go. You could go through your ToC and link:

- inputs to activities
- activities to outputs
- outputs to outcomes
- outcomes to impacts.

You do not *have* to create your draft in this reverse order. Rather, do so in a way that makes sense to you and your team. We do, however, highly recommend that you begin with rationale and impact, and then go from there. When considering the content and design of your ToC, take care that is it clearly labelled and laid-out, with colour-coding and clear pathways. Be prepared to adapt and rewrite your draft as you consult with wider stakeholders.

Stakeholder consultation

An effective ToC should take a broad view of the topic by considering the views of a diverse range of stakeholders. You should use stakeholder consultation to discuss and gather feedback on the initial draft, but it is also helpful to involve colleagues – and even other relevant stakeholders – in the drafting stage itself. Effective consultation and engagement in creating the ToC supports ownership of the document and can help with securing buy in to the project more widely.

> *Stakeholders include everyone who is directly or indirectly affected by, or has influence on, the outcomes of the project or programme.*
>
> **Government Analysis Function**

Who to involve and how to involve them

The example ToCs presented in section 2.1. are general templates; you could use them as a basis for developing your own ToC, but they should be revised and refined to reflect the specifics of your own local context and identified problem. Working with partners and affected parties can help you think through whether the response you are considering makes sense in your circumstances.[13]

Before you begin the process of identifying your stakeholders it is helpful to consider the following questions:

1. Who is affected by this project?

2. Who may have influence on decisions?

3. Who knows about issues affecting the local area of the project?

4. Who has an interest in the project?

5. Who contributes to the programme?

6. Who can enhance the scope of the issues being considered?

7. Who can help identify other stakeholders?

8. Who can adequately represent those with different viewpoints?[14]

You could include staff, partners, trustees, funders, other local organisations. In line with the public health approach, consulting with young people and communities (affected by violence) at this stage is valuable. As a ToC is a live document, it is helpful to review whether any stakeholders have been missed over the course of project delivery, especially if priorities change, and ask for their input to review your ToC at a later date.

When asking for input to the ToC, it's important to first make sure stakeholders understand what a ToC is and what purpose it serves. A facilitated group discussion, or a workshop, is an important aspect of consultation; this approach means consensus can be more easily gained and is an effective use of time.

Consultation should focus on testing the logic of the intervention with those with expert and local knowledge. It's helpful for a facilitator to have some pre-prepared questions for workshop participants to discuss; these can include views on the long-term goals of the project (the rationale and the impact), whether the assumptions the ToC is based on are correct, whether there are any gaps in the activities/outputs/outcomes/impacts, and whether any identified causal links are valid.

13 College of policing, 2021, Knife crime A problem solving guide, p5. Available at: https://assets.college.police.uk/s3fs-public/2021-11/Knife-crime-a-problem-solving-guide.pdf

14 Government Analysis Function, 2023, The Theory of Change Process – Guidance for Outcome Delivery Plans. Policy details. https://analysisfunction.civilservice.gov.uk/policy-store/the-analysis-function-theory-of-change-toolkit/#resources-to-help-you-identify-assumptions

Stakeholders are likely to come to consultations with different perspectives and views, and the workshop facilitator should be prepared to guide the discussion to help achieve a consensus.[15]

Here is a resource which provides more detail on ensuring effective stakeholder engagement.

Refinement and revisions

Developing a ToC is an iterative process. Once you have completed your first draft, you should re-engage stakeholders and tweak it based on their views. A ToC should be a living document – it should be revisited as new evidence and perspectives emerge or as you develop new ways of working. Evaluation can help provide evidence for whether your ToC works, and also suggest where your theory needs to be modified.[16]

Check that your ToC makes sense by asking yourself if it is:

- **Meaningful:** Does it describe the project accurately in ways that staff, trustees, volunteers and stakeholders agree with?
- **Well-defined:** Is it clear what your project will do and who your target groups are?
- **Comprehensible:** Does it enable you to give someone the 'two-minute story' of the service? Would a member of the public understand the theory?
- **Do-able:** Are the services and activities likely to contribute to the desired outcomes and impact?
- **Plausible:** Is it realistic? Does it take into account your organisation's capacity? It should be something that the programme, project or organisation could really do, not just wish it could.
- **Credible:** Are people outside your organisation likely to believe it? Is the secondary evidence you include credible with your stakeholders?
- **Testable:** Can you test the theory through a series of testable hypotheses? All elements should theoretically be able to be assessed using research and observation (even though you may not have the resources to assess this yourself).[17]

15 Brown. AM., 2019, How to Develop a Theory of Change. Available at: https://www.linkedin.com/pulse/how-develop-theory-change-ann-murray-brown/

16 NCVO, How to build a Theory of Change. Available at: https://www.ncvo.org.uk/help-and-guidance/strategy-and-impact/strategy-and-business-planning/theory-of-change/#how-to-build-a-theory-of-change

17 Harries. E., Hodgson. L., and Noble. J., 2014. Creating your Theory of Change, p21. Available at: https://golab.bsg.ox.ac.uk/documents/Creating-your-theory-of-change1.pdf

2.4 Theory of Change templates

There is a range of software you can use to create a ToC; MS PowerPoint or MS Word are popular as they are widely accessible and available. However, we also recommend Draw.io, which is free online software for creating diagrams. This software is especially helpful if you want to create a more complex ToC including causal links, as it is easier to create pathways between the boxes.

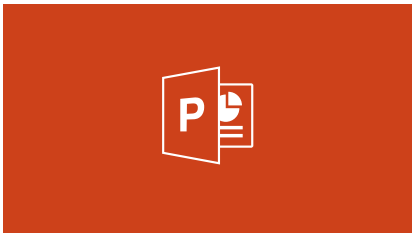
In this toolkit, we have provided a selection of blank ToC templates you can use to form the basis of your own document. These are fully editable – do feel free to delete or add boxes and arrows to fit the context of your project.

To edit the draw.io file simply **click on the icon on the left**.

You can view the ToC templates by opening the attachments pane:

view > show/hide > navigation panes > attachments

and then double click on the relevant template.

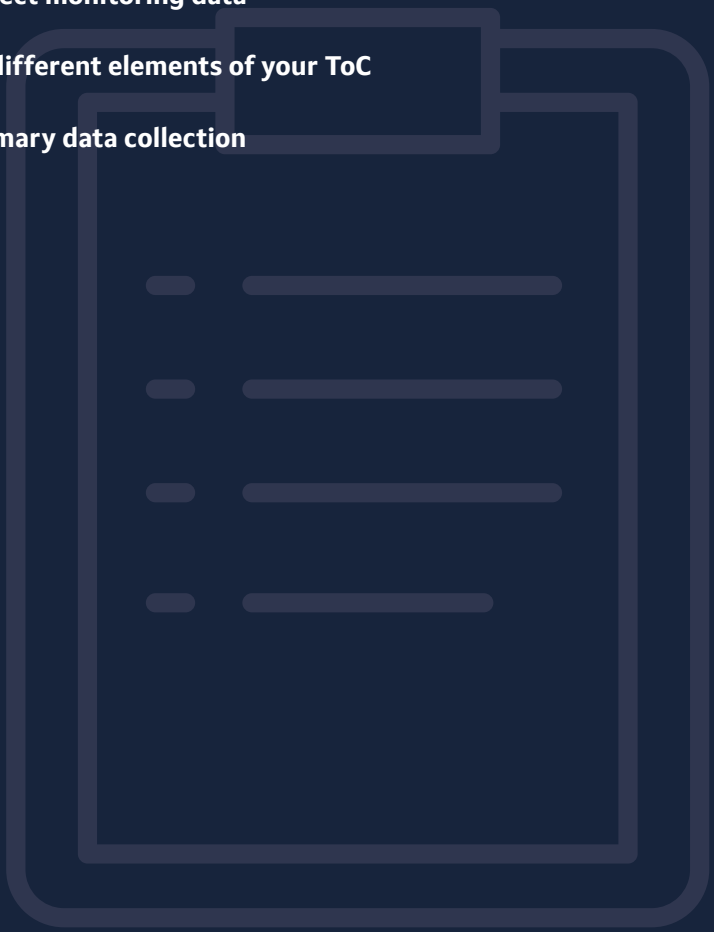


Blank simple ToC template

Blank casual ToC template

3.0

Monitoring:
Understanding if your project is working



Delivered by

West Yorkshire Combined Authority | Tracy Brabin Mayor of West Yorkshire

West Yorkshire Violence Reduction Partnership

ECORYS | Answering tomorrow's challenges today

It's likely that your organisation is well experienced in collecting monitoring data, especially the routine information asked for by funders. However, most routine monitoring focuses on outputs, and many organisations are less used to collecting data on outcomes. This means there is often a gap in understanding how outputs lead to outcomes, and with that, a gap in knowledge about what works. Getting your ToC in place (and with it, a clear understanding of what your project aims to achieve) can help you figure out what aspects of your intervention you should be collecting evidence about to prove what works. Once you know this, you can set up your monitoring processes at the outset of project delivery, meaning the right data is in place to answer questions about efficacy when you get to the point of evaluating your work.

When to collect monitoring data

Through consultation interviews, delivery partners told us it can be difficult to engage participants in data collection, and that there can be a lack of time to collect data. It can be useful to identify key touch-points for data collection, and how the data at each stage of delivery can help you understand whether your intervention is meeting it's aims. Figure 5 provides some examples of the types of data you could collect at different points of delivery.

Figure 5: Useful points for collecting monitoring data aligned to your ToC

1 Eligibility

- Depending on the intervention focus, there might be pre-existing data about people who are potentially eligible for support, and some of their characteristics.
- This data could be used to support recruitment into project activities (depending on data usage permissions) and/or support the identification of comparator groups for impact evaluation.

2 Referral

- If young people are referred from other agencies to the intervention, (standardised) referral forms can provide key demographic information, including an initial outline of needs.

3 Assessment

- Your intervention might include an initial assessment to confirm eligibility and/or identify support needs. Again, existing data and/or standardised forms can facilitate efficient monitoring data collection.
- This data will help you to understand who is accessing support, and whether those using services are your intended target groups.

4 Support

- During support, you should aim to capture intervention intensity (e.g. number of sessions attended/not attended through a register) and potentially distance travelled through reassessment.
- This data will tell you whether you have attained your outputs as planned.

5 Exit

- Data collected at exit can tell you if the young person completed support and, ideally, whether their identified needs were met. What was their destination after support?
- This data can help you understand whether your intervention is achieving its planned outcomes, and potentially, impact.

Monitoring different elements of your Theory of Change

When planning your monitoring approach, it can be helpful to take those time points for data collection one step further and map sources of information against the different elements of your ToC. Doing this will assure you that once you get to the stage of evaluating your project, you have evidence in place for all elements of your ToC and can demonstrate whether your project has done what it set out to do. The elements or metrics in your ToC should, where possible, have been developed to be SMART – specific, measurable, achievable, relevant, and time-bound (as explained in section 2.3). This means that monitoring each of these elements should be easier.

While the monitoring of inputs, activities, and outputs is usually routine, monitoring outcomes can be more challenging. This usually involves some sort of consultation with services users to understand the effects of the intervention, and this can be difficult within time and budget constraints. We believe that approaches to monitoring (and evaluating) your ToC should be proportionate to the scale of your intervention and the types of activities you deliver.

Table 1 provides examples of different data sources that could be used to evidence different elements of your ToC, while Table 2 explores the pros and cons associated with different primary data collection methods (that is, data you collect directly from participants). You might use some of these methods as part of your routine monitoring of outcomes, and then use others during your evaluation to supplement your monitoring data (see section 5).

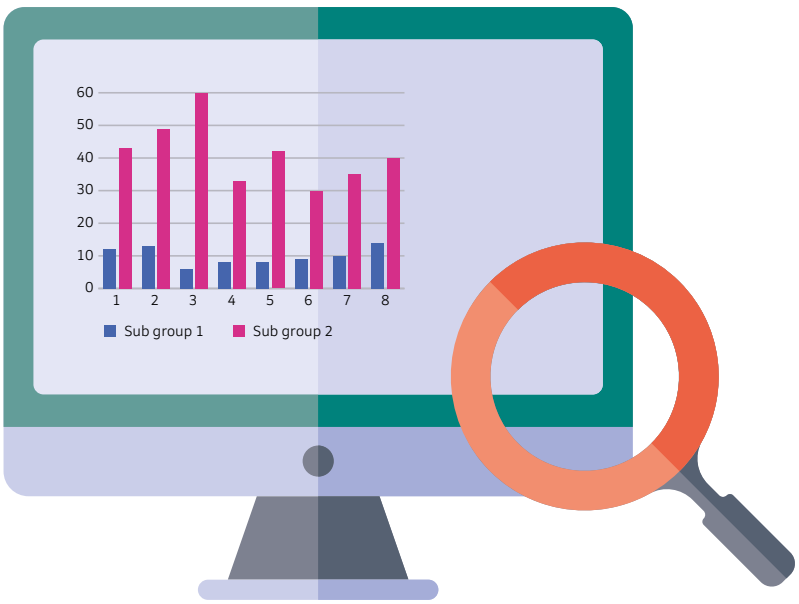


Table 1: Monitoring different elements of your ToC

ToC component	Source of monitoring data
Inputs	
Funding	Organisational financial records and accounts
Partnership meetings	Meeting notes and minutes
Use of school premises	Session records, booking confirmations
Staffing	Timesheets
Activities/outputs	
Number of people attending group sessions in schools	Session registers, school data
Number of people attending 1-2-1 mentoring sessions	Mentor's session notes
Number of training sessions provided to professionals	Session registers, trainer session notes
Outcomes	
School pupils have an improved understanding of the impact of violent crime	Session feedback forms
Participants develop improved social and emotional skills	Completion of standardised wellbeing scales at the start and end of support, observations of participants.
Professionals are better equipped to support young people at risk of violence	Session feedback forms, follow up interviews with sample of professionals trained.
Impact	
Reduced reoffending amongst those supported	Follow up interviews with participants at multiple time points after leaving the project.
Analysis of secondary data on offending rates.	

Table 2: Pros and cons of different methods for collecting primary data

	Pros	Cons
Surveys	▮ Online surveys can collect lots of data quickly and cheaply. ▮ Can identify subgroup patterns.	▮ Not always accessible. ▮ Risk of low response rate.
Interviews and focus groups	▮ Collects in-depth insight into an intervention and can complement quantitative data collection.	▮ Time consuming. ▮ Biased views. ▮ No numerical estimates.
Monitoring information	▮ Low cost and quick.	▮ Can have a high admin burden on staff who collect it.
Observations	▮ Allows for a deeper understanding of individual experience. ▮ Reduces bias from self-reporting.	▮ Participants may act differently if they know they are being observed. ▮ More costly.
Case studies	▮ Can capture real life situations. ▮ Complements other data collection methods such as surveys.	▮ Difficult to generalise findings to other settings.

Based on: Quick guide to research methods commonly used for evaluation, Magenta Book (2023).

Ethics in primary data collection

When collecting primary data, it is vital to take an ethical approach. This is particularly important considering that the target groups for many VRU-funded interventions will face multiple vulnerabilities. The core principals of social research ethics include:

- **Social responsibility:** researchers should aim to maximise the benefits of the research and minimise the potential risk or harm to participants and researchers. This is often referred to as a '**Do No Harm**' approach.
- **Independence:** mitigate conflict of interest or partiality on behalf of the researchers, funding or commissioning body.
- **Informed consent and voluntary participation:** research participants should be given sufficient information about the research and how their data will be used and offered the opportunity to consent or refuse participation without negative consequences.
- **Anonymity and confidentiality:** ensure that there is no way to identify a person from the information provided (i.e., by keeping personal details separate from survey responses, interviews, or focus groups) unless they explicitly agree. Also, ensure that only the researchers collecting or analysing the data have access to respondents' personal information.
- **Integrity and transparency:** All stages of research design and data collection, cleaning, coding and analysis should be documented appropriately so the research process is transparent and there is an audit trail.[18]

More information can be found here:

- Ethics in evaluation: Why it is important
- Understanding the Ethics of Data Collection
- Ethical considerations in research and evaluation with Children and Young People

18 Halej, J. 2017. Ethics in primary research (focus groups, interviews and surveys). Available at: https://forms.docstore.port.ac.uk/A816773.pdf

4.0

Evaluation:
Laying the foundations

Delivered by

West Yorkshire Combined Authority | Tracy Brabin Mayor of West Yorkshire

West Yorkshire Violence Reduction Partnership

ECORYS | Answering tomorrow's challenges today

What is evaluation?

Evaluation is the process of assessing the quality and value of something and why it has that value or quality. It should use relevant evidence to inform decisions about the value of policies and programmes and support the allocation of resources.

Evaluation is the why, which differs from monitoring as the what. Evaluation involves reflecting on the data collected as part of monitoring and considering what it tells us about the impact of the project or programme.

VRU delivery partners might experience evaluation in a number of different ways: as the subject of someone else's evaluation, by doing self-evaluation or by commissioning an evaluation from a professional researcher.

Why is evaluation important?

Evaluation can help us understand **what works** to deliver the maximum impact on projects and programmes. It also helps to highlight any difficulties and uncertainties that might impact on what the project achieves.

Evaluation seeks to uncover the following questions:

- 'What works', for who and why?
- What difficulties and uncertainties lie within an approach.
- What evidence can be used to inform public debate or policy.
- What learning and insights can be gained to inform future project decisions.
- How much a project really costs to deliver.

There are several practical reasons to conduct evaluation. These include:

- To demonstrate project value, creating a business case for future funding.
- To fulfil a funding requirement.
- To support replication of a project.
- To get an independent view on any processes.
- Contribute to the wider evidence-base about what works in a public health approach to violence reduction.

Ultimately, analysis and reflection on the information you gather during the delivery of your project can help you to form value judgements on whether the project's aims have been met.

When should you use evaluation in your work?

You can use evaluation at any point in your project, including during project design, during project delivery and after project completion. Evaluation can also be conducted in an ongoing cycle – at all three stages – to continuously provide information on the successes and challenges of the project or programme. It doesn't need to only happen at the end of delivery.

Figure 6: Evaluation throughout a project lifecycle

Before
Use evaluations from previous similar projects to inform your approach and understand any risks.

It can be really helpful to set up your evaluation before project implementation or during the project design phase, ensuring you have mechanisms in place to collect the right data right from the start. However, you might also want to refer to evaluations and insights from previous similar projects. This will help to inform your approach and understand any risks. It will also help you understand whether the wider policy landscape could learn something new from your context and approach.

The YEF website provides an archive of evaluations which can be sorted by approach to identify similar projects.

During
Formative evaluation allows learnings to feed into delivery for continual improvement.

Conducting evaluation during project delivery is called **formative evaluation** and can allow learning to feed directly into continual improvement of the project or programme as it is being rolled out. This will help you to understand what is working, what is not working and to begin to interrogate why this is the case.

After
An evaluation of a completed project or programme can conclude on the impact and provide lessons for the future.

An evaluation of a completed project or programme can conclude on the impact and provide lessons for the future. Conducting evaluation after the project has been completed allows learning to feed into future projects or programmes. This approach is known as a **summative evaluation**.

Who conducts evaluation?

Some funders will request that an evaluation be conducted as a condition of funding being awarded, but even where this is not the case, we have seen how important evaluation is for contributing to an evidence base of what works. However, it can be difficult to find the capacity, skills, and knowledge to conduct a thorough evaluation in-house, particularly for smaller organisations. You might also want an objective view of your intervention from someone outside your organisation. In either case, it can be valuable to commission an evaluator or researcher to support the process (known as an external or independent evaluation).

When deciding who will conduct your evaluation, consider:

- **The skills required** – do we have a staff member with appropriate expertise and analytical skills?
- **Capacity** – is it possible for staff to take time away from their normal duties to conduct the evaluation?
- **Budget** – how much do you have to spend on resourcing the evaluation? Is it enough to commission an external partner?

If an evaluation is being conducted internally, it's helpful to develop a separate workplan and delivery plan for the evaluation, as well as separation of responsibilities in the team to support objectivity, where possible. The size of the team and the level of involvement of team members will depend on timescales, and the size, complexity, and purpose of the evaluation. It is helpful to assign different aspects of your evaluation to different team members, and keep track of deliverables and deadlines using something like the evaluation and planning table in Annex C of this resource.

Commissioning an external evaluation

If you decide to commission an external partner, such as a freelance consultant, a research agency or an academic partner, there are several steps involved in the commissioning process.

1. Write a specification for your evaluation. This acts as a brief for those who tender for the work, although the details of research methods can be ironed out later in the process, with the evaluators. You should stipulate the timescales for the work, and any deliverables you would like at the end (such as summary reports, full reports and presentations).
2. Set a budget for the amount you want to spend. You might have already included this in the overall budget for your funding.
3. Share the specification and invitation to tender with relevant networks. You could ask colleagues in other organisations for recommendations or share the invitation to tender through professional bodies.
4. Develop criteria for scoring the bids you receive. This should include considerations of the quality of the tender and the price quoted. You can also ask the highest scoring bidders to 'pitch' for the work in a meeting, where you can ask clarification questions about their proposals.
5. Once a successful bidder has been identified, agree terms of reference with them and put a contract in place. An inception meeting with the supplier can be useful to help finalise the aims, objectives and approach for the evaluation.

There is more detailed information about each of these steps in this good practice guide[19] to commissioning an evaluation.

4.1 An overview of different types of evaluation

Although there are many different approaches to conducting evaluation activity (see section 5), broadly there are three main **types of evaluation**. Each type focuses on a different aspect of an intervention:

- **Process evaluation:** What can be learned from how the intervention was delivered?
- **Outcome or Impact evaluation:** What difference did the intervention make?
- **Value for money evaluation:** Was this a good use of resource?

19 https://golab.bsg.ox.ac.uk/documents/Good_Practice_Guide_-_Commissioning_Evaluations_Final_feb_14.pdf

A **process evaluation** gathers evidence on how the intervention has been implemented in practice and can offer insights into why an impact has or has not been realised. This type of evaluation would answer questions related to inputs, activities, and outputs from the ToC, as well as giving insight into what has helped or hindered delivery and what are the lessons being learnt. There is more information on methods for process evaluation in section 5.1.

An **impact or outcome evaluation** is usually conducted after an initiative is completed (though preparation should begin while the project is being delivered, with appropriate data collection systems being put in place). Both these types of evaluation aim to understand the extent to which the outcomes or impacts set out in the ToC have been realised. The key difference between them is whether they have some causal attribution or not: an impact evaluation uses a **counterfactual** to assess what would have likely happened in the absence of the intervention, in order to better attribute changes to the intervention. An outcomes evaluation does not usually look at attribution, instead measuring the outcomes of the intervention without as assessment of what would have happened in the absence of the programme. There is more information on methods for outcomes and impact evaluation in section 5

A **value for money evaluation** weighs up the relationship between the costs of an intervention and the benefits or effects the intervention achieves. This is often done by allocating 'proxy' monetary values to outcomes and impacts and weighing those values against the costs associated with the intervention. For more information on conducting value for money evaluation, see section 5.5.

These different evaluation types are not mutually exclusive but complimentary. Often, evaluations will use a combination of these different approaches to gain a broad understanding of whether an intervention achieved the outputs and outcomes laid out in a ToC.

4.2 Evaluation frameworks and research questions

This section will help you to put the foundations in place to conduct an evaluation, through the development of an evaluation framework.

Put simply, an evaluation framework is a plan that clearly sets out how you will measure the success of the different elements of your ToC, through collecting and analysing data, to evidence and learn from your story of change.

Ideally, you should try to write your framework before your project starts so that you can make sure you are collecting appropriate data from the beginning; this can reduce the costs of data collection by building evaluation activities into project delivery. Planning an evaluation also requires consideration of both the design and the project management of the evaluation.[20] Although this will require some thought at the beginning, which can be difficult when you are short on time and resource, it is an opportunity to keep stakeholders focussed on the ultimate goal of the project.

20 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879438/HMT_Magenta_Book.pdf

Developing an Evaluation Framework is a process that looks at:

- Defining the intervention to be evaluated (i.e. the scope).
- Defining the evaluation aims and research questions.
- Selecting appropriate evaluation approaches and methods.

Defining the intervention

The first step in any evaluation is to define the scope, which means considering what the intervention to be evaluated is. Developing your ToC first will allow you to clearly understand the rationale for your project, what it aims to achieve, and how. At this stage, you should be clear on the overall purpose of the project and be able to specify how its activities will contribute to a chain of effects that bring about its intended outcomes and impact.[21] **Your ToC will underpin your evaluation design.**

Defining the evaluation aims and research questions

At the outset, you must be clear on what is to be evaluated (your intervention) and what you (and your audience, or stakeholders) want to learn. Defining the aims of an evaluation depends on its purpose and use, whether that be to:

- Identify risks and modify your project in response.
- Monitor progress towards outcomes and take action when necessary.
- Respond to external scrutiny.
- Communicate impact and win future funding.
- Create internal learning and influence the design of future provision.

The Magenta Book includes more information on key evaluation uses in Table 2.1, on page 27.

By setting out the aims of your evaluation (in other words, what you and your stakeholders hope to get out of it), you can be clear about what these questions are and how the findings from them are expected to be used, by whom and when. This will inform the evaluation approach to be used, help focus the evaluation, and ensure the findings stand the strongest chance of having an impact and being used.[22]

When considering the purpose of your evaluation, think about whether you would like to understand the **process** by which your project was implemented; the **impact** of the project; or the **value-for-money** of the project. This will influence the types of research questions your evaluation will aim to answer. Table 2.2 in the Magenta Book, page 31, includes useful examples of evaluation questions for each type of evaluation.

Finally, the scope of your evaluation will also be defined by what resources (budget, staff, and time) are available to conduct the work.

21 https://www.gov.uk/government/publications/evaluation-in-health-and-well-being-overview/planning-an-evaluation

22 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879438/HMT_Magenta_Book.pdf

Developing research questions

Evaluations can be designed to answer a range of questions about whether and how the aims of your intervention (set out in your ToC) were achieved. Initially, it can be helpful to write down all the potential questions you would like your evaluation to answer. This is likely to result in a long list – you will probably not be able to answer every question posed, typically because of the time and resources that would be required to answer them all, as well as methodological limits.[23] It is useful to narrow the list down to around 5- 6 high-level research questions. This will keep your evaluation manageable and focussed. Your research questions should be:

- Clear, specific, and well-defined
- Focus on a program or program component
- Measurable by the evaluation
- Aligned with your Theory of Change.

It is helpful to think of the research questions as ways to 'test' the 'theory' behind your ToC. Ultimately, your research questions should be centred on whether or not your intervention met its intended objectives. You can work through your ToC and consider whether the project went as expected and if your assumptions were correct, thereby mapping your questions to the different ToC elements. For example, did our inputs help support the delivery of activities as we predicted? Did our activities lead to the outputs we planned?

The decisions you make about your research questions (directed by what you want to gain from the evaluation), and the evidence needed to answer those questions, will directly influence the types of evaluation approaches and methods you choose. More specifically, your evaluation focus will depend on which elements of your ToC the research questions are linked to. For example, if you decided that you want to learn lessons on whether the identification of young people to take part in the program was effective, you would be assessing how the 'outputs' and 'activities' on your ToC played out in reality. This would require a process evaluation (see section 5.1).

23 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879438/HMT_Magenta_Book.pdf

Figure 7 below illustrates how research questions about inputs, activities, and outputs will fall into process evaluation, and those about outcomes and impact will fall into an outcome or impact evaluation.

Figure 7: Aligning research questions to your ToC

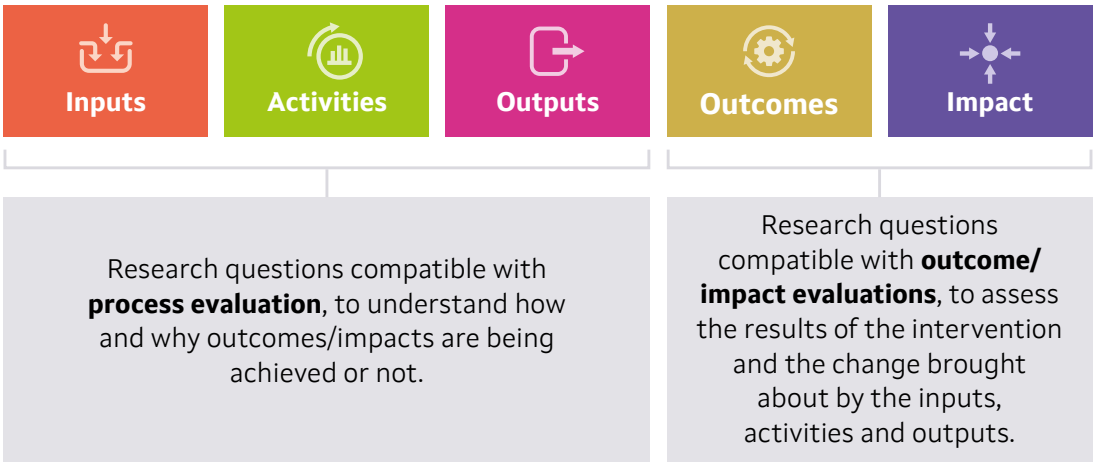


Figure 7 highlights that there is crossover between process and outcomes evaluations. The process evaluation is concerned with the **processes** by which **outcomes** and impact are (or are not) realised, and the **outcome/impact** evaluation is concerned with the outcomes and impact brought about by those **processes** (inputs, activities and outputs). Whilst they are different, process and outcome-focussed evaluations are often done at the same time, their findings being combined to give a deep, well-rounded understanding. There is information on how to combine findings from different evaluation approaches in section 5.4.

You may want to include more detailed sub-questions under some of your research questions. For example:

- **Did our activities lead to the outputs we planned?**
- **Were there enough resources?**
- **Were there any unexpected or unintended issues in the delivery of the activities?**
- **To what extent has the intervention reached all the people that it was intended to?**

Figure 8 shows examples of research questions developed for each element of the example ToCs set out in section 2. See more on process and outcomes evaluation approaches in section 5.

Figure 8: Research questions for different types of evaluation and their associated ToC element

Process evaluation	
▮ To what extent and how: – Did existing data sharing between agencies support the identification of perpetrators of VAWG crimes (potential beneficiaries of support)? – Were police and probation available to support the identification of perpetrators, and provide timely initial and longer-term support to the project? – Was the right staffing in place to deliver the intervention? ▮ Was the funding for intervention sufficient and how was it used?	Inputs to activities
▮ How were offenders identified? ▮ How were the initial group interventions delivered? ▮ How were the longer-term interventions delivered?	Activities
▮ Was the identification of offenders timely and accurate? ▮ Did those identified engage with the initial intervention? ▮ Did those identified engage with longer-term interventions?	Activities to outputs

Figure 8: Research questions for different types of evaluation and their associated ToC element continued

Outcome/impact evaluation questions	
▮ Did participants recognise the seriousness of the situation through the initial group intervention? ▮ Were underlying risk factors identified (at initial intervention) and subsequently supported (through longer-term) intervention? ▮ Were alternative life choices recognised by and available to those supported? ▮ Were there any unintended (positive or negative) outcomes?	Outcomes
▮ Were underlying risk factors addressed? ▮ Were alternative behaviour choices addressed/taken up? ▮ Was there a direct/independent link between participant recognition of their offences and impacts achieved by the project?	Outcomes to impact
▮ (Relative to what would have likely happened in the absence of intervention) was there a reduction in reoffending, arrest rates, criminality and risky behaviours for those supported?	Impact

Monitoring	
▮ How many perpetrators were identified and what were their characteristics? ▮ How many of the identified perpetrators were supported through initial and subsequent longer-term support?	Outputs

It is worth noting that whilst testing the ToC should be the primary purpose of your evaluation, there may be other research questions which are not captured in the ToC. For example, you might want to consider the experiences and outcomes of different groups (taking into account gender, ethnicity, neurodiversity and other characteristics) with a view to understanding how you could tailor your intervention for different barriers or needs.

Mapping research questions to data sources

Once you have set-out your research questions, you can begin to consider exactly how you might answer them. This involves considering the data sources and evidence needed to show whether you achieved the outputs and outcomes detailed in your ToC.

For each measure of progress, it will be important to find out what information you have already, what additional information you will need to collect, what methods you will use to gather it, who will take responsibility for collecting the information, and the timescale.[24]

Ask yourself:

- What relevant data do you already collect?
- Are you missing any data to answer this question?
- How could you collect the data in a way that's proportionate to your evaluation budget and the sort of project you have delivered? Can it be incorporated into project delivery?
- Who can provide this data? Consider whether you can collect it yourself from participants (for information primary data, see sections 3.0 and 5.1.) or if you will access existing data collected by others (for information on secondary data, see below).
- Are there any barriers to collecting this data?

It is particularly important to plan early what data and evidence should be collected before implementation and during the lifetime of your intervention. If data collection is left until the end or after the lifetime of your project, it may limit your ability to conduct appropriate evaluation (for example, it may be necessary to collect baseline data before delivery starts).

Top tip: Set up systems to gather data on a regular basis. Clarify what information you will need to help you to answer or explore your research questions. Think about the records you will be collecting anyway as part of your project monitoring that will provide some of the information you need.

Then: Decide what additional information you will gather specifically for the evaluation, and how you will gather it. The following table shows how some different data sources can be aligned to specific research questions. You should choose different methods in order to get a 'full picture'. Also, different ways of gathering information will suit different projects and research questions. Which you choose to use, and how much time you spend gathering data, should be tailored to the capacity of your organisation.[25]

24 https://www.jrf.org.uk/evaluating-community-projects-a-practical-guide
25 https://www.jrf.org.uk/report/evaluating-community-projects-practical-guide p7

Figure 9: Mapping research questions to data sources

Research questions	Desk review	Management information	Interviews	Survey	Secondary data
To what extent and how did existing data sharing between agencies support the identification of those at risk of involvement/ already involved in crime?	●		●		
Did young people recognise their vulnerability through the initial intervention?			●	●	
(Relative to what would have likely happened in the absence of intervention) was there a reduction in arrest rates, criminality and risky behaviours for those supported?		●			●
How many young people (potential perpetrators) were identified and what were their characteristics?		●			

Secondary data and data sharing

If you are commissioning an evaluation, it is unlikely you will need to source secondary data yourself. However, if you are designing an internal evaluation, you will most likely need to access secondary data. Secondary data is information that already exists, collected by other people or organisations for a different purpose. This can come from various sources, including national and local crime statistics, previous evaluation reports, or data collected by other organisations. It may be publicly available, although you might need permission to access it. It is necessary to consider whether the secondary data you plan to use for your evaluation is:

- Available – readily accessible
- Reliable and trustworthy
- Relevant to your work and appropriate for evaluation use

One of the main reasons for collecting secondary data is to avoid duplicating work that has already been done. If you can use secondary data sources, you may be able to save both time and expense. There are other reasons for reviewing or collecting secondary data:

- It will show the gaps in existing information and the quality of evidence already available.
- It can provide a context in which to place your analysis of the primary data that you are collecting (for example, if there is a general trend nationwide of a reduction in young people knife crime, this should be acknowledged by the evaluation).
- It can give you a greater understanding and insight into the problems, issues and practice related to the field in which you are evaluating (and often has greater scope).
- It can provide a basis for comparison for the data that you are collecting.

Note that using secondary data comes with a risk of misinterpreting how the data was sourced and how it should be analysed. Most public sources will have guides of how you can use the data.

Quantitative secondary data sources

You can freely access national and local statistics through government websites, such as www.statistics.gov.uk. The Office for National Statistics (www.ons.gov.uk) is the largest independent source of national statistics. There are also a number of sites where you can get more specific data on different topics, for example you can also find data about reported incidents in a specific locality on www.crime-statistics.co.uk and www.police.uk. The outcomes framework[26] developed by South Wales Violence Prevention Unit includes links to other relevant datasets.

Furthermore, there is a wealth of data held by public bodies such as police, health and local authorities. Access to this data can be key to supporting prevention activity, though you may need to develop data sharing agreements in partnership with these agencies, although some information could be obtained through Freedom of Information Requests.[27]

Qualitative secondary data sources

You can undertake a literature review (a summary of relevant literature on a topic, or of research findings which relate to the project or programme being carried out) or a document review (including organisational documents, client reports or meeting minutes). You can also analyse media or online resources. In each case, it is important to verify the dependability and reputation of the data source,[28] and add weight to your findings accordingly.

Now that you have clearly defined your intervention (through your ToC) and established concise research questions and their aligned data sources, you can complete your Evaluation Framework by deciding on appropriate evaluation approaches and information collection methods.

26 https://www.violencepreventionwales.co.uk/cms-assets/global/Violence-Prevention-Indicators_Wales-VPU_2021.pdf

27 https://www.violencepreventionwales.co.uk/cms-assets/research/Violence-Prevention-Evaluation-Toolkit.pdf

28 NCVO, Using Secondary Data, Available at: https://www.ncvo.org.uk/help-and-guidance/strategy-and-impact/impact-evaluation/planning-your-impact-and-evaluation/choosing-evaluation-methods/using-secondary-data/

5.0

Evaluation methods

5.1 Process evaluation

Generally, process evaluation seeks to capture the experiences of those delivering the intervention as well as those experiencing the intervention (i.e., young people, communities) to identify what is working well and any aspects that are, or may prove to be, problematic. Various methods are used to assess how an intervention is working and why it might work better in a different context. The task at this stage is to select the research methods most appropriate for answering your research questions. Process evaluations predominantly use qualitative primary methods,[29] including:

- Interviews
- Focus groups
- Observations
- Open-ended questionnaires
- Creative methods (i.e., photographs, videos, diary entries).

You will most likely also collect secondary data as part of a process evaluation, for example, conducting a document review or analysing project monitoring documents. It is important to uphold the ethical considerations laid out in Section 3 when conducting all of these research methods.

Sampling for qualitative methods

To begin with, it is important to consider who will be involved in your research. Depending on your research questions, you may wish to speak to staff at different levels within your organisation or project, intervention participants, members of the community, or funders. It is often not practical to involve everyone in your project within the research, so it is necessary to select a group of people (a 'sample'). Sampling methods which tend to be used in qualitative research include:

- **Purposive sampling**, whereby you choose respondents because they have a particular characteristic that you are interested in.
- **Snowball sampling**, whereby you ask one contact to put you in touch with others, and then they may also put you in touch with others (the sample size therefore 'snowballs' and becomes larger through word of mouth).
- **Convenience sampling**, which involves selecting people who are available.[30]

For qualitative data collection, the goal is to ensure that you capture the richness of views and opinions of the people involved; or to illustrate specific stories or contexts through in-depth case studies. There are no guidelines for how many respondents should be selected – it depends on the scale of your project and evaluation. It is important to understand that the objective with qualitative research is not to be able to produce findings that are statistically representative, but rather, to understand the range of views on a particular topic and to provide deeper insights which cannot be quantified.

29 It is possible to include quantitative methods such as surveys as part of a process evaluation.
30 https://golab.bsg.ox.ac.uk/documents/Good_Practice_Guide_-_Commissioning_Evaluations_Final_feb_14.pdf

Reaching **saturation point** (the point at which the addition of new data doesn't add anything new to the findings) is a good indication that the breadth of views and opinions has been captured.[31]

Interviews

Interviews are conversations between an interviewer (researcher) and an interviewee, (sometimes called a respondent or participant) in which questions are asked to gain information. Your research questions will guide who you will aim to speak to and the questions you will ask during the interview. The goal of interviews is to collect data and narrative information to better understand the respondents' unique perspectives, opinions, and world-views; the **lived experience** of other people is at the heart of this approach. Interviews invite the participant to make sense of their own experiences and to share these experiences with the researcher.[32] There are some important practical considerations for interviews here.

There are three types of interview:

- A **structured** interview is one with a set list of pre-planned questions. There is no scope for probes, and the conversation follows a specific pattern.
- A **semi-structured** interview, in which there are some guidelines around questions to ask (in the form of a list or 'topic guide'), but there is also flexibility around where the conversation can go. Researchers can ask follow-up questions and steer the conversation depending on the interviewee's answers to the pre-set questions.
- **Unstructured** interview, which can be described as 'conversations with purpose in mind' but without a set plan.[33] Whilst the researcher should keep the conversation on-track with the subject matter in mind, there are no set questions.

Interview questions can be **open-ended**, which encourage the interviewee to give their own answer, or **closed**, in which there are a limited set of possible answers (including 'yes' or 'no' answers). Open-ended questions are favourable in interviews, in order to gain more detail.

To begin with, use your research questions to think about what you would like to ask during the interview. It is not the case that research questions and interview questions are the same, however, there will be some overlap. The goal is to 'tap into' the respondents' experiences and perspectives, which requires asking questions that encourage the interviewee to talk and 'open up'. It is useful to formulate interview questions with the view that the interviewee should speak more than the interviewer during the conversation.

31 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879438/HMT_Magenta_Book.pdf

32 https://www.imperial.ac.uk/education-research/evaluation/tools-and-resources-for-evaluation/interviews/best-practice-for-interviews/

33 https://www.questionpro.com/blog/types-of-interviews/

These pointers may be useful when developing interview questions:

- **Word the questions so that respondents will answer completely and honestly.** Do not use 'leading' questions, for example, 'It is more difficult to engage young men than young women, isn't it? What do you think?'. Instead, ask 'Have you experienced any differences in engagement levels between different groups?'.
- **Develop probes that will give a more detailed response.** Simple probes might be 'would you like to say any more about that?' or 'could you please elaborate?', or they could relate directly to the question (see some examples in Figure 10 below).
- **Begin the interview with 'warm-up' questions.** This can be anything from 'Please tell me about your background and how you got involved with the program?' to 'Could you tell me about your interests and hobbies?'.
- **Think about the flow of the conversation** – what topics should come first, and what flows on 'naturally'. It is useful to consider embedding sensitive topics in the middle of the interview, rather than beginning or ending with them (although note that sensitive issues may arise at any stage, because the point of interviews is to encourage people to 'open up'. This is dependent on the topic area and respondent).
- **End the interview on a positive note**, leaving the respondent feeling empowered, listened to, and glad they took part.[34] You could ask if they have any further comments and let them know how helpful and insightful their comments have been. It is helpful to 'cool off' by asking if they have plans for the rest of the day, or even to talk about the weather.

Your interview questions will depend on what type of interview you are doing. Below is an example of the thought process (rationale) when developing interview questions with a research question in mind.

34 https://sociology.fas.harvard.edu/files/sociology/files/interview_strategies.pdf

This example is from a semi-structured interview with project delivery staff about VRU-funded interventions:

Figure 10: An example of the thought process behind developing interview questions with a specific research question

Research question	
Did the project reach the intended population and, if not, why?	
Interview question	**Rationale**
How would you describe the intended participant group for your intervention?	This will allow you to explore what project staff understood the target group to be, and some of the context around the choice of target group. It may also encourage them to express challenges or successes in engaging certain groups
How does the intervention reach and engage the intended target group? Probe: ▮ Is this effective? In what ways? ▮ To what extent does it reach those considered most at-risk?	Here you can begin to understand the ways in which the target population is reached and, from the project staffs' perspectives, the strengths and weaknesses of this approach. You could encourage them to elaborate on whether those most at-risk have been reached, and whether the approach was effective for all. This will help you to understand why outputs (and in turn, outcomes) have or have not been met for certain groups.
What has enabled/inhibited reaching and engaging with the intended target group?	This question will allow to you gain an understanding of barriers to reaching and engaging the target population. You could explore external factors, differences amongst groups, and how engagement could have been approached differently.

Focus groups

Focus groups are facilitated group discussions. Whilst interviews are often with one or two interviewees, focus groups use group dynamics to encourage people with shared experiences to share their views. Whilst there is no set size for the groups, having around 6-8 participants works well.

Focus groups are particularly useful if you want to want to use the group dynamic to develop new ideas which may not be possible through one-to-one interviews or other research methods, and if you want to consult with a range of people about a limited number of topics. They are, however, less appropriate if you want to explore confidential matters. There is a useful guide on setting up, conducting and analysing focus groups here.

Open-ended questionnaires

Questionnaires offer a systematic way to collect information from individuals or groups. Whilst surveys are the main way of collecting information in quantitative research, including qualitative, open-ended questions can allow you to gather data that is not easily quantified. This includes attitudes, beliefs, habits, and challenges. Open-ended questions can allow respondents to express the thought process behind survey responses and add context to quantitative data. Examples of open-ended questionnaire questions include:

- Please describe any challenges your project faced in relation to engaging young people.
- How would you describe the key activities offered to the young people involved?

Observation

Observations involve turning what you see and hear into evidence for research purposes. It allows you to assess a project or situation in real-time, seeing the activities and outputs laid out in a ToC in practice. This can provide a good understanding of the intervention being evaluated, from first-hand experience. Often, observational data is used in conjunction with other approaches as part of an evaluation design. For example, interviews, focus groups, or questionnaires might be used together with observations to build a more complete assessment of a certain activity.

The key to using observational data as evidence in an evaluation is to take a systematic and consistent approach as you collect, organise, and analyse what is observed. The steps to using observational study for evaluation (explained in detail here) include:

- **Setting the scope.** Establish the focus and approach- are you interested in a particular type of activity? Will you undertake a structured or unstructured approach?
- **Prepare the tools and the people.** This includes organising whether you will take notes or record, and organising the team and participants.
- **Conduct the observation.**
- Organise and analyse the data.[35]

You may wish to combine the above methods and data to create **case studies**, which focus on a particular aspect of your intervention, such as a person, place, or element of delivery. With this approach, you can gain insights into how certain factors such as policy change, local context, or life experiences have impacted individuals and communities.[36] A case study approach might include conducting 5 interviews with practitioners and 5 interviews with beneficiaries of one activity or adopting various methods with one beneficiary to create a case study of their individual story.

There is a range of other approaches you might use to collect qualitative data for your evaluation. Here is a useful resource with details on creative methods for qualitative research.

35 https://education.nsw.gov.au/teaching-and-learning/professional-learning/pl-resources/evaluation-resource-hub/collecting-data/observation/step-4-organise-and-analyse-the-data

36 https://www.evalcommunity.com/career-center/case-study-evaluation-approach/

Analysing qualitative data

The analysis stage of your process evaluation is the point that you begin to answer your research questions. It involves finding patterns and themes in the data that you have collected, which will help you report on it effectively and use it to make decisions. Qualitative analysis relies heavily on interpretation, so it is helpful to run the findings by multiple people in order to make sure that interpretation is not biased. If resource allows, you could test findings with a 'steering group' of participants or other stakeholders.

It is important to use a systematic approach to analysis, following these steps:

- **Prepare the data:** Firstly, prepare the raw data for analysis by transcribing audio recordings of interviews or by writing up interview notes. You may want to organise the data by research question, or stakeholder type. Here is a useful guide on transcribing interviews. You may want to analyse the data manually or using software (i.e., Nvivo or excel).
- **Get to know the data:** Read and re-read the data. Start to jot down thoughts and interesting points.
- **Identify common codes from an initial review:** A 'code' is a brief description of what is being said in the interview or focus group data. Think of it as a word or a short phrase that captures the meaning of specific quotes- a description, not an interpretation. It might be something like 'trust between young people and youth worker' or 'improved self-esteem'.
- **Categorise codes under themes:** A theme is broader than a code and captures something important about the data in relation to the research purpose. Read through the data and identify recurring themes, by applying the same code to sections of the text that represent the same meaning (i.e., if a young person said 'I feel more confident in myself', that would be coded to 'improved self-esteem'). You can use sticky notes and coloured highlighters to visually separate themes if you are doing your analysis manually rather than using software.
- **Review and refine your themes:** Ensure that there are identifiable differences between your themes and that they are useful and an accurate representation of the data. Also, consider whether anything is surprising and identify new themes you might have missed.
- **Report on the findings:** Reporting needs to go beyond just describing your data and should include your own analysis to make an argument about the story you present. Think of this as telling a story about your data, using quotes to back up points.

There are two methods of coding when conducting thematic analysis:

- **Deductive**, through which you come to the data with pre-determined themes that you expect to find based on knowledge or evaluation questions (this is more common in evaluation).
- **Inductive**, which allows the data to determine your themes as they emerge (this takes longer but is more exploratory).

These steps are summarised from this guide to thematic analysis.

5.2 Designing surveys for your evaluation

Define the objectives and research questions

If you have followed the steps set out in section 4.2 (on mapping research questions to data sources) then you should have identified which, if any, research questions you will answer through a survey. You will need to have a clear idea of what you are trying to answer through the survey to avoid making it too long, getting poor quality data or never using the results.

Determine the sample and administration method.

The first steps in gathering data with a survey is to decide who your target audience is – that is, your sample – and what method you're going to use to ask your survey questions. Your sample and the method of administering the survey will likely be interlinked; your intended audience may affect how you choose to distribute the survey. Both your sample and your method should be proportionate to your evaluation budget and the sort of project you have delivered.

Sampling

When determining who your survey respondents should be, firstly you will need to choose your **population** and then choose a smaller **sample** of this population to distribute the survey to.

Population: The entire group you want to draw conclusions from.

Sample: A specific group from your population that you will collect data from.

Some sampling methods have already been discussed in this toolkit (see section 5.1). However, you will most likely survey a larger group than you would for qualitative interviews, so there are more sampling methods that could be used. Ideally your sample should be:

- **Representative** – your sample should represent the core demographics of your population – that is the total number of people whose views you are interested in. For example, if you are interested in people's outcomes from engaging in your project, then your population would likely be the number of people who have engaged with your project in a certain period of time. If you're interested in young people's experiences of crime, then your population might be all those aged 16-24 in your local authority area).
- **Chosen at random** – this helps to ensure a representative sample and eliminate any bias in selection. For example, those who are most engaged in a programme are more likely to respond to a survey but may respond in a different way to individuals who were less engaged. If it is possible to select a random group from your population, then you should.

In some cases, your population may be the same as your sample. This would be the case if you are only interested in those who receive support from your project, and it is feasible for you to survey everyone who has taken part. Otherwise, you will need to consider how many people should be in your sample. The size of your sample will vary according to the size of your population and how accurate you would like your findings to be – that is, how much confidence you will have that your data reflects the views of your whole population. You can find more information about confidence intervals in section 5.3 (statistical significance), or here.

Once you have determined these two factors, you can use a simple online sample size calculator to work out how many responses you need.

Remember, it's unlikely that everyone you send your survey to will respond. This means you will need to send out more surveys than the number of responses you need. There are many variables that can affect your response rate, but a good survey response rate tends to fall between 5% and 30%, and an excellent response rate would be 50%.

Choosing an administration method

The way you choose to share your survey will depend on the way you interact with your chosen group and the resources you have available. For example, if you have regular face to face contact with your chosen group this may be the most effective for getting a high response rate. Some options are:

Type	Advantages	Disadvantages
Face to face	▮ Likely to get better responses to open ended, more qualitative questions. ▮ Inclusive for those with low literacy levels. ▮ Can support e.g., younger respondents, or those with learning difficulties, by helping them to interpret questions. ▮ Higher response rate if sampling the public.	▮ High cost. ▮ Respondents might be less likely to express their true feelings.
Telephone	▮ Likely to get better responses to open ended, more qualitative questions. ▮ Inclusive for those with low literacy levels.	▮ High cost. ▮ Might not have current contact details. ▮ Hard to reach people who won't answer unknown phone numbers.

Type	Advantages	Disadvantages
Text message/ SMS	▮ Quick to deploy, meaning you receive fast feedback. ▮ Wide reach to a large number of people. ▮ Cheaper than telephone surveys. ▮ High response rates due to easy participation.	▮ Limited survey length. ▮ Lack of detail in both questions and responses; limited numbers of characters to convey information. ▮ Might not have current contact details.
Online	▮ Low cost. ▮ Responses can be automatically collated ready for analysis.	▮ Excludes those who can't use technology well. ▮ Assumes a certain level of literacy in your sample. ▮ Can be exclusive of those with English as an additional language. ▮ Risk of low response rate.
Paper based	▮ Potentially more accessible, particularly if posted or given directly to your sample.	▮ Less environmentally friendly. ▮ Assumes a certain level of literacy in your sample. ▮ Can be exclusive of those with English as an additional language. ▮ More time consuming (and potentially expensive) to distribute. ▮ More time consuming to collate responses for analysis.

Resources for online surveys:

- Microsoft forms – Free. The software will automatically analyse your data for you, but you can't download the graphs produced.
- Google forms – Free. The software will automatically analyse your data for you, and you can download the graphs easily.
- Survey monkey – Mainly free, but some features are paid. Allows for more sophisticated question types.
- Smart Survey – Free for surveys of up to 15 questions, and 100 responses per month. Provides template surveys to personalise.

Ethical considerations

The data protection act covers how any organisation uses personal information. It gives individuals the right to know how their data is being used or have data erased amongst other things. It also stipulates that information collected is:

- used fairly, lawfully, and transparently.
- used for specified, explicit purposes.
- used in a way that is adequate, relevant, and limited to only what is necessary.
- accurate and, where necessary, kept up to date.
- kept for no longer than is necessary.
- handled in a way that ensures appropriate security, including protection against unlawful or unauthorised processing, access, loss, destruction, or damage.
- You should also consider whether using a survey is the appropriate data collection method. Sometimes when you are collecting sensitive data it may be more appropriate to conduct an in-person interview where you can provide support to respondents if they need it.

Design the survey questions

The next step is to design the survey, ensuring the questions you ask align with your research objectives. The questionnaire can include a mix of closed questions (multiple choice, rating scales) and open-ended questions. You might also want to include some questions on personal characteristics or demographics, so you can compare results for different groups.

Type of questions	
Question type	**Description**
Multiple choice	Lots of types of multiple-choice ques-tions including Likert scales, multiple statement multiple choices and more.
Likert scale	Type of multiple choice for rating how much an individual agrees or disa-grees with a statement. Either a 5- or 7-point scale; it's better to keep it simple.
Open ended	Free text boxes to get richer answers or to follow up a closed question.
Dichotomous	Questions with two responses, normal-ly Yes/No or Agree/Disagree.

Top tips when designing survey questions	
Tip	**Description**
Don't use leading questions.	**Bad:** 'Do you agree that this project is the most effective initiative in the community?' **Better:** 'How effective do you think the service has been?'
Don't use loaded questions.	**Bad:** 'Don't you think it's unbelievable that other projects are failing while our violence prevention project is making a real difference?' **Better:** 'Do you think this project is making a difference?'
Use plain English and use simple and short language where possible.	**Bad:** 'What are your perspectives on the efficaciousness of this project in mitigating societal harm from violence?' **Better:** 'How effective do you think this project has been in preventing the negative impact of violence on society?'
Use consistent phrases.	Just stick to one phrase or the other, don't mix and match or duplicate meanings of questions. Example: 'Do you think this project is successful?' and 'Do you feel the project has achieved its objectives?'
Keep the survey as short as possible.	You will be less likely to get responses if you produce a long survey, and many survey technologies will tell respondents roughly how long it takes to fill out. If you need a long survey, you may want to consider an incentive to complete it, such as prize draws or vouchers.
Research other surveys.	There may be other surveys done by your organisation, or others available online, which provide model questions you can utilise (for example, for demographics).

Informed participation

Regardless of your method of distribution, you should make sure that respondents understand why they are being sent the survey and what the purpose of the data collection is. You can do this verbally in telephone or face to face interviews, and with a cover page or letter in online or written surveys.

As a minimum you should include:

- The organisation behind the study, including a contact name and details for any queries.
- Why and how the respondent was selected.
- The aims of the study.
- What will happen with the information they provide.
- Data protection measures.
- Where relevant, any potential benefits or harms that could result from the study (and signposting to further support if you're surveying on a sensitive topic).
- That taking part assumes consent.

Pilot the survey

Once the sample and method of administration are chosen, and the survey is designed, it is important to conduct a pilot test of the survey with a small group of respondents. This help to identify any flaws, such as unclear questions or technical issues and can avoid problems with data quality later down the line.

Data collection

Once the survey is finalised, you can go ahead with data collection. You should monitor the data collection process while the survey is live to address any issues or concerns that arise – for example, if you're getting a poor response rate, why might that be? What can you do to improve it?

Data analysis

The next section of the toolkit, and especially page 58, has steps you can follow for analysing and interpreting the results of your survey.

You should also give extra consideration to consent if any of your respondents are under the age of 16. If this is the case, it is important to gain consent from parents or carers, especially if any of your questions might be considered sensitive or potentially upsetting. Where parental consent is needed, you should make sure parents are also given sufficient information to help them make an informed decision. This article[37] explores some of the issues around gaining consent from young people in online research and sets out some suggestions for best practice approaches

37 Harris, J and Porcellato, LA Opt-Out Parental Consent in Online Surveys: Ethical Considerations. http://researchonline.ljmu.ac.uk/id/eprint/8588/

5.3 Outcomes evaluation

Outcome evaluation is a tool for assessing the effectiveness and impact of your intervention. It is an objective process that involves collecting and analysing data to determine whether the intervention achieved its intended goals and objectives, and whether the outcomes are meaningful and beneficial.[38] It could be the first step in an impact evaluation (see Section 5.4.) and/or combined with a process evaluation (see section 5.1.) to create a mixed-method evaluation.

Specific research questions to be addressed by an outcome evaluation could be:

- To what extent did the project reach its outcomes?
- How did beneficiaries' outcomes change over time?
- Did the types/scale of outcomes differ by group characteristics?
- Did those that engaged more experience greater outcomes?
- Were there any associations between outcomes?

It should be noted here that an **outcomes evaluation cannot attribute a change in outcomes to your project**. It is for this reason that an outcomes evaluation is generally combined with process evaluation, so that qualitative findings can be used to support and interpret findings from the analysis of outcomes data.

Collating data on outcomes

The first step for conducting an outcomes evaluation is to begin with a review of the outcomes laid-out in your ToC. Selecting the appropriate data collection methods and outcomes measures is best done with your ToC at the forefront of your mind.

As covered earlier, your outcomes should be all be SMART (see section 2.3) and you should have established how you are going to measure them. Ideally, you will have been monitoring and collecting outcome data or have an identified secondary data source from the start of the intervention. Ideally, quantitative data should be collated into an electronic data set to make analysis easier – Microsoft Excel is useful enough for this, and most survey software will allow data to be exported to Excel.

Example: In Section 2.1, we shared an example ToC for a Hospital Navigators project. In that example, two primary outcomes were:

- 'Victim recognises vulnerability/seriousness of current situation'.
- 'Alternative non-violent life choices (recognised and available)'.

Each of these outcomes could be measured through a survey of participants when they join the intervention which is then repeated when they leave (called a pre- and post- intervention survey). Results from the survey can be collated into a data set as shown in Table 3.

38 https://www.evalcommunity.com/career-center/outcome-evaluation/

Table 3: Collating outcomes data

	Demographic information			Outcome data	
	Age	**Gender**	**Previous conviction**	**Outcome 1** Victim recognises their situation	**Outcome 2** Alternative non-violent life choices (recognised and available)
Individual 1	25	F	Yes	Yes	No
Individual 2	23	M	Yes	No	No
Individual 3	19	M	No	Yes	Yes

Analysing quantitative data

Dealing with a large dataset can be overwhelming, and it can be difficult to pin-point what the numbers are telling us. The following steps will enable you to demonstrate the effectiveness of your program in a manageable and tangible way.

Data cleaning

First, you must ensure that your data is ready to use, which means ensuring that it is accurate and free from any errors. Consider:

- Correcting spelling mistakes and inconsistent entries.
  - Imagine that you are using software such as Excel to analyse your data. Survey respondents may have inputted various answers to a question on their sex, i.e., 'Female', 'F' or 'Fmeale'. The software will not recognise that these entries mean the same thing, so you must go through the data and ensure that entries are consistent and free from typos. This can be a time-consuming task, but it cannot be avoided.
- Check for missing data and decide how to approach this, e.g. will a record be removed if enough data is missing? It is especially important to create a consistent approach if you are going to use summary statistics (see page 61).

Data matching

If your project has managed to secure data sharing agreements and you have access to data from a third party such as the police, you may need to match this to your own data. Matching data means combining two or more data sets based on a **matching variable**. In the case of individual level data, the matching variable will be an individual participant. To match individuals, you could allocate a **unique code** for each one, anonymising them whilst still making it easy to match records together. Combining data sets in this way can help you piece together a more complete story for your participants.

Example: You have two data sets: one contains demographic data for your sample and another the outcomes. Under GDPR, you may be required to remove any data that could identify a person which is not required for analysis, like names. However, allocating ID codes can also help with any issues arising from typing errors (see page 58). You may assume that Sam and Samuel are the same person, but with common names it is possible there are duplicates, and you may not have enough information to distinguish between them.

Data set 1			
Name	**Unique ID**	**Age**	**Sex**
Samuel Smith	S9867	25	M
Jane Jones	P0987	19	F

Data set 2			
Name	**Unique ID**	**Outcome 1**	**Outcome 2**
Samuel Smith	S9867	Yes	Yes
Jane Jones	P0987	Yes	No

Once your data set is cleaned, matched and complete, you can move on to analysing the data.

Summarising your data

To perform statistical analysis, you would ideally have a 'large' data set. However, there is no set definition of small or large data sets. An arbitrary minimum of 30 rows has been set by statisticians, however the general rule for analysis is the **bigger the data set the better**. The larger the number of people you have outcomes for, the more confident you can be that your analysis represents your target population.

There is also more you can do with a larger data set in terms of presenting analysis. Table 4 outlines the different statistical measures you can use to summarise your data.

Table 4: statistical measures for summarising data

Measure	Description	Type of data[39]	Example
Frequency	The 'count' or number of responses in each category.	Numeric, categoric, ordinal	25 individuals achieved outcome 1.
Percentage	The proportion of people who give a particular response.	Numeric, categoric, ordinal	50% of individuals in the programme achieved outcome 1.
Mean	The arithmetic average.	Numeric	On average individuals spent 14 days a term in exclusion.
Standard deviation	The average spread of scores around the mean.	Numeric	The standard deviation for number of days excluded was 10 days, indicating there was a lot of variation in experiences.
Median	The 'middle value' (another measure of average).	Numeric	Individuals spent a median of 10 days a term in exclusion.
Mode	The most common response to a question (another measure of average).	Numeric, categoric, ordinal	The outcome most commonly achieved was outcome 1.
Range	The difference between the highest and lowest response value.	Numeric	The number of days individuals in the programme were excluded in a term varied from 0 to 25.

Tools for data analysis

These calculations can be done easily in a range of computer programmes.

- Excel: The analyse tab and pivot tables provide quick and easy ways to analyse data. There are many free courses to help you use excel such as on LinkedIn, or you can find some cheat sheets here.
- Other data coding programmes are available, such as R (free) or Stata (paid).

39 You can read more about the four types of data here: https://www.upgrad.com/blog/types-of-data/

Issues to consider in your analysis

Small data

For small data sets/samples, take caution when presenting your data. Percentages and means can be misleading. For example, if you had
5 individuals and one individual thought the programme was 'great' you could say:

- **Option 1:** 20% of individuals we supported thought the service was 'Great'.
- **Option 2:** One individual we supported thought the service was 'Great'.

Option 1 sounds much more positive than option 2, however it is misleading as it doesn't tell the reader this was only out of a sample of 5.

There may be valid reasons why you have a small sample (for example, you might be looking at one particular aspect of your intervention). If this is the case you can still report outcomes, but the richness of your evaluation is more likely to come from qualitative data.

Summary statistics

It's important when you present summary data that you do so without misrepresenting your results. For example, of 100 individuals supported by an intervention, we have data that tells us that all 100 went on to secure a job. However, we only have prison records for 70 of them. Out of the 70, 40 had been in prison. Consider the following statistics we could pull from this:

- **Option 1:** 40% of individuals who got a job at the end of the programme had been in prison. This would be incorrect, as we do not know whether the 30 people for whom we do not have records were in prison or not.
- **Option 2:** 57% of the individuals who got a job at the end of the programme had been in prison. While this is factually correct, it doesn't offer the reader details of what sample was used. In other sections you may wish to refer to the full sample of 100.
- **Option 3:** 57% of individuals (of the 70 individuals we have records for) who got a job at the end of the programme had been in prison. This option is the most accurate, providing the reader with sufficient detail (and thus, transparency) of where the figure has come from. If you feel this level of detail is confusing, you could add it as a footnote or in a methodology note in the annex of a report or presentation.

Looking for data trends

By now you should have calculated some key statistics for your data. You may decide this is sufficient for your evaluation (if so, you can move on to page 64 – presenting your data). However, you could also take your analysis a step further and consider looking into trends in your data. This could be trends over time, or it could be trends in different subgroups in your sample. By digging further into the data, you will be better placed to answer questions linked to your ToC, particularly in relation to how your activities impacted on different groups, or whether different activities led to different outcomes.

Consider the following questions:

- How do trends vary between different subgroups?
- What are the key differences since you last analysed this data?
- How have trends changes over time?

Sub-group analysis

Subgroup analysis splits individuals into subgroups based on a variable (for example, age group or gender). Subgroup analysis can be used to compare differences within or between groups, and is important to pick up any effects on particular groups that might be hidden or less visible in the wider data set.

Example: Highlighting differences in experience through sub-group analysis.

Table 5 shows a sample of satisfaction ratings for five participants, including some demographic data.

Table 5: Satisfaction by gender and age

	Gender	Age	Satisfaction with service
Individual 1	F	20	Great
Individual 2	F	17	Ok
Individual 3	M	30	Ok
Individual 4	M	33	Poor
Individual 5	M	31	Poor

For this data, the median age for the whole group is 30, and the mean is 26.2.

Whilst those figures give an overview of the individuals receiving support, it would be much more illustrative to say that females who received support are generally younger than males, with an average age of 19 compared to 32.

Differences between subgroups may be more difficult to pick out when you have lots of data, but using software that creates quick summaries can quickly break your data down. In Excel, you could try using pivot tables to create cross tabulations of your results.

Example: In exactly 10 clicks we were able to create a pivot table (Table 6) using data from the previous example. From this table we can quickly see that only females thought the service was great and that only males though the service was poor. We could also say that generally older respondents thought the service was less good than younger respondents. By analysing data in this way, it raises questions about why there is a disparity in experience between sexes; answering that question could form vital learning for future delivery.

Table 6: Pivot table of satisfaction, by age and gender

Row Labels	Average of Age
Great	**20**
F	20
Ok	**23.5**
F	17
M	30
Poor	**32**
M	32
Grand Total	**26.2**

Choosing subgroups to explore

The Wales VPU outcomes framework resource includes suggestions of what sub groups would be useful to look at for different outcomes. For example, for the indicator (or outcomes associated with) 'violence related injuries', it is suggested that disaggregation could usefully be done by age; sex; ethnicity; injury type; day/time; relationship to perpetrator; violence type; deprivation quintile.

You might be interested to **compare your data over time** to build up a bigger picture of how the programme has evolved, or to understand the length of support needed to achieve outcomes. This could be done in a few different ways depending on what your outcomes are and how they are recorded.

Example: You are evaluating your programme after two years of delivery. There are three outcomes we are interested in looking at. If we didn't look at trends over time, we could say something like: '*By Q8 of the programme Outcome 1 was achieved by four individuals, Outcome 2 was achieved by 5 individuals and Outcome 3 was achieved by 3 individuals.*'

	Outcome 1	Date achieved	Outcome 2	Date achieved	Outcome 3	Date achieved
Individual 1	Y	Q3	Y	Q1	N	-
Individual 2	N	-	Y	Q2	Y	Q7
Individual 3	Y	Q3	N	Q1	Y	Q8
Individual 4	Y	Q4	Y	Q2	Y	Q8
Individual 5	Y	Q4	Y	Q1	Y	-

However, we are missing an opportunity to comment on time varying trends within the data. By looking at the data over time, you could say something with more insight like '*By Q8 of the programme Outcome 1 was achieved by four individuals, Outcome 2 was achieved by 5 individuals and Outcome 3 was achieved by 3 individuals. There was also variation in when the outcomes were achieved; all individuals had completed Outcome 2 by Q2 whereas no participants achieved Outcome 3 until Q7.*'

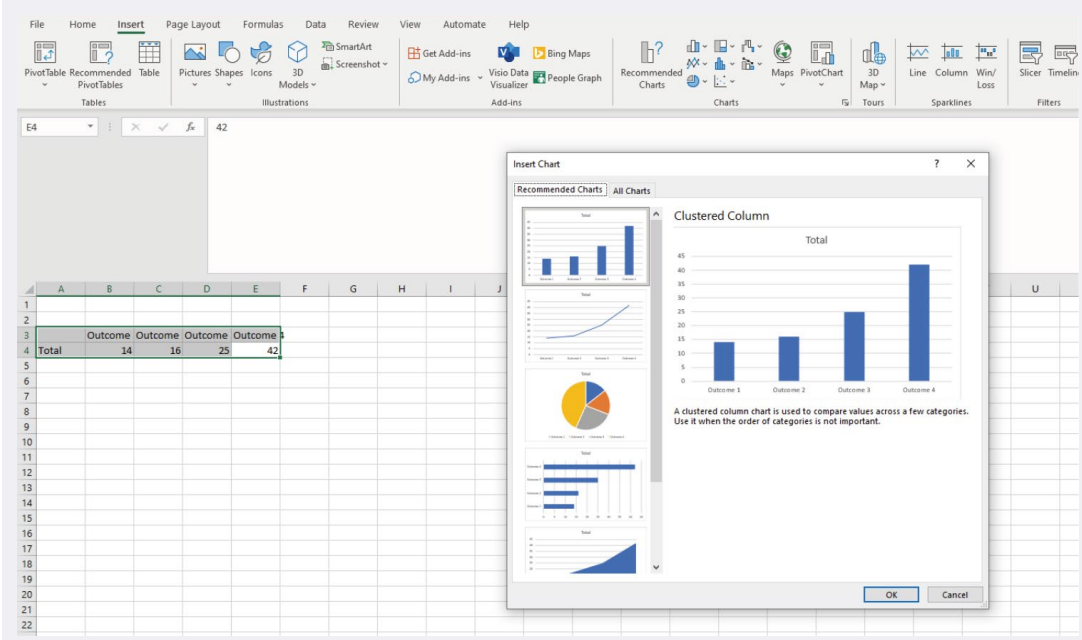
The next section will explore presenting your data graphically which is the easiest way to show any trends over time.

Presenting the data

From the processes outlined above, you should now have some outcome results, whether they are just summaries or whether you have identified more complex patterns in your data. The next step is deciding how to present your findings. Tables and charts are very useful to help present your results quickly and easily to another audience. If you are using basic data on excel, it will select suitable charts for you. It also gives tips on under what conditions to use the charts.

For example: if you wanted to show how many of each outcome your sample achieved, it would likely suggest a clustered bar chart. It also gives tips on under what conditions to use the charts. You can see an example of this in Figure 11.

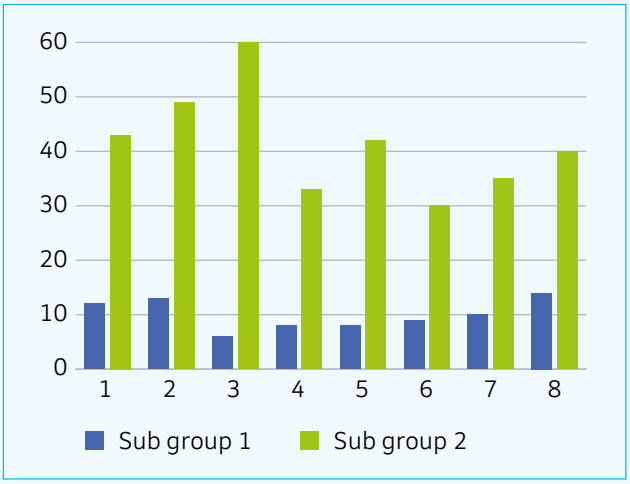
Figure 11: Choosing charts in Excel



You may want to present your analysis in graphs too, for example findings for trends over time or within subgroups. This may require some manipulation of your data to get it in the correct format to produce a graph. You can also make graphs using pivot tables.

Example 1: Outcomes achieved by different subgroups

Outcome	Sub group 1	Sub group 2
1	12	43
2	13	49
3	6	60
4	8	33
5	8	42
6	9	30
7	10	35
8	14	40



From the chart in Example 1, the reader can clearly see:

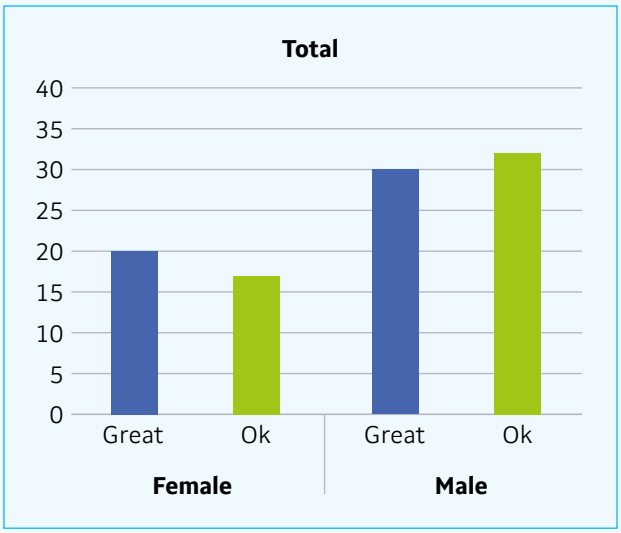
a) Subgroup 2 achieved more of every outcome than subgroup 1.

b) Outcome 8 was the most common outcome achieved by subgroup 1.

c) Outcome 3 was the most common outcome achieved by subgroup 2.

Example 2: Charts from pivot tables

Row Labels	Average of Age
Great	**20**
F	20
Ok	**23.5**
F	17
M	30
Poor	**32**
M	32
Grand Total	**26.2**



From the chart in Example 2, the reader can see the same conclusions we drew from this pivot table on levels of satisfaction by age and gender (see page 63, Table 6), but more easily and quickly.

There are several online tools and blogs to help you select graphs. One example is here.

After you have selected your charts and have your key figures, you should write up your analysis. This should include describing the results from all the steps described so far. Key figures should be summarised, and descriptions of graphs should be included picking out any interesting results. This can then be included in your evaluation report (see section 6).

Optional: Statistical significance

Statistical significance confirms that the result of your calculations did not happen 'by chance'. This can be done through statistical tests that provide us with a '**p value**' – a value that tells us the probability that the results we have seen have happened by chance or do in fact show a significant difference between groups.

T-tests

T-tests are used to compare the mean (average) of two or more populations or samples. It can confirm that two groups are statistically significantly different from each other. This could be used to test whether the cohort has significantly increased on an outcome. However, it's important to remember that this will not prove that your programme caused this change.

For example: A cohort of 1,000 young people took part in a sports programme where one outcome was to reduce the truancy rate amongst the group. This is measured by the number of unauthorised absences in a school term. The average number of absences across the group was 30 days before the intervention, and 25 after. A t-test compares these two values using the variation and sample sizes to confirm whether this change is significant or not.

There are several online calculators that you can copy your data into and it will perform the test for you. However, there are pre-conditions that the data must fulfil to use these tests effectively. Try this resource for more information and to access an online calculator.

Confidence intervals

Confidence intervals tell you how 'good' or accurate an estimation is, or how sure we are about the validity of our findings. This can be done as part of a T-test. For example, if there is a lot of variation in a data set, using a standard deviation test you can calculate a set of values which 95% of the expected observations should fall within. There is a more detailed explanation of confidence intervals here.

Should you wish to delve deeper into the statistical significance of your results, you can find more information in Graph Pad.

5.4 Combining findings from process and outcome evaluations

As explained in section 4, evaluation approaches should be thought of as complimentary. Process and Outcomes evaluations are rarely done in isolation and when paired together they can offer important insight into whether your project achieved change, and if so, how and why.

> *"An outcome evaluation tells you whether a program achieved its goals. A process evaluation tells you how and why."*[40]

An evaluation which combines different methods and approaches to collect data is called a **mixed-method evaluation**. For example, you could combine interviews and observations from a process evaluation with pre- and post-intervention surveys, and reviews of government statistics from an outcome evaluation. Whilst the two types of evaluation are geared towards different research questions, it is really useful to look at their findings together.

For example, imagine that your outcome evaluation research question is 'were alternative, non-violent life choices recognised by the young people involved? What impact has this had on their lives?'. Your process evaluation might explore which types of activity were particularly helpful for the young people, and why. You can combine the findings to gain an understanding of exactly how and why the intervention arrived at the observed outcomes, which is vital learning to take forward into future provision.

This approach can also be described as a **triangulation** of your findings across your different research methods, using a range of data sources and methods to ask the same question. You can begin to triangulate your data by thinking about trends in findings from different sets of data to see if findings converge or diverge. Where answers corroborate, you can be more confident in the validity of your findings.

Here is a useful resource on conducting mixed methods evaluations.

5.5 Methods for conducting impact evaluation

An impact evaluation, if done correctly, can attribute the impact on your desired target group and wider society to your project. In comparison, an outcome evaluation can only detail the change in outcomes in your target groups – it cannot attribute this change to your project.

Example: Outcome analysis: 25 individuals who took part in the programme did not commit an offence in the two years following the programme.

Impact evaluation: The programme *caused* 25 individuals to not commit an offence in the two years following the programme.

40 https://www.jbassoc.com/resource/whats-difference-understanding-process-outcome-evaluation/

5.0 Evaluation methods

Impact evaluation builds knowledge about what works, by assessing the change in outcomes which are attributable to the particular programme or policy. However, it is a difficult task attributing impact to your project. **Impact evaluation often involves using complex statistical methods and should only be used if you or your team have received training on the subject.** For most VRP projects where staff are conducting evaluation themselves, process and outcomes evaluation should be adequate to evidence if change has happened, and to understand why.

This toolkit will give a brief overview of impact methods, with links to other resources that will provide more detail on how to undertake one.

Why use impact evaluation?

- It is needed to establish whether programmes and policies have worked to improve outcomes.
- There is a growing trend to shift focus from inputs and outputs to outcomes, for better accountability and decision-making.

In his book *The Effect*,[41] Nick Huntington-Klein provides some examples of why it is important to be able to attribute a change to a policy or programme:

> *"We don't want to know if countries with higher minimum wages have less poverty, we want to know if raising the minimum wage reduces poverty. We don't want to know if people who take a popular common-cold-shortening medicine get better, we want to know if the medicine made them get better more quickly. We don't want to know if the central bank cutting interest rates was shortly followed by a recession, we want to know if the interest rate cut caused the recession."*

What research questions does impact evaluation answer?

There are many research questions that an impact evaluation can help to answer. Some examples of this are below:

- **What is the effect due only as a result of the intervention (and not other interventions)?**
- **To what extent can a specific (net) impact be attributed to the intervention?**
- **What would have happened without the intervention?**
- **How much of the impact can be attributed to the intervention?**

Ultimately impact evaluation will estimate the **causal effect** (impact) of:

- an intervention '**P**' (programme, treatment, policy) on
- an outcome '**Y**' (indicator, measure of success).

41 https://theeffectbook.net/

Unfortunately, this can be really challenging to do in social research. In an ideal scenario we would have a **counterfactual** to tell us what would have happened in the absence of P (programme, treatment, policy). This is normally in the form of a control group – that is, a group that did not receive P. The classic example in a medical trial would be those individuals who receive a placebo rather than treatment.

If we have a counterfactual and can 'control' for all other variables (referred to 'closing the back doors'; see more information on that here) then we could prove that the impact is due to P (programme, treatment, policy). However, it is difficult to form a counterfactual group for a social policy programme; it would be unethical to offer support to some people with identified vulnerabilities but not offer it to others. However, we will briefly explore different techniques to use to show impact by either inferring or creating a counterfactual.

Frequently asked questions about counterfactuals

Why don't we just select those who didn't take part in the project?

Selection bias: The reason a person did not engage might ultimately change the effect. For example, an optional course on how to get the best grades was offered to a classroom of school children. The children who took the course ended up with higher grades than those who didn't. However, it is likely that those children who took the course were more interested in getting higher grades and therefore might work harder, which would likely have an impact on their final grade. A correct approach here (though not necessarily ethical) would be to only allow half of the children who applied for the course to take it (at random) and then you could compare these groups.

Why don't we just compare before and after of the same group?

Contextual change: Other things may have happened in this time frame that we cannot control for and wouldn't be identifiable if you only study the treated group. For example, some of the children who took the course on how to get the best grades fell ill in the period before their exams. This impacted on their final grades and skewed the results for the cohort.

Data and project requirements

Not all programmes or policies are suitable for impact evaluation. Certain elements must be present to consider an impact evaluation:

- **Clear target outcomes/impacts** which should be SMART.
- **Intervention or project is consistently delivered** with no variation in approach.
- **Strong theoretical links between the project and desired impact:** there should be existing research and evidence that this type of programme would achieve the desired impact.
- **The *potential* for a counterfactual:** this can be hard to ascertain, but in some situations it might be particularly difficult (for example if a policy was rolled out across a whole nation).
- **Sufficient sample size:** as you need to use statistical tests for an impact evaluation, a minimum sample size is required. See here for more details on how to work out what this could be.

Impact evaluation methods

Generally, there are five key steps in the process of conducting an impact evaluation:

1. Feasibility study to establish whether the relevant data is available.

2. Research design, including sampling and establishing an analysis plan.

3. Randomisation (allocation to counterfactual groups) and data collection.

4. Analysis and testing assumptions.

5. Reporting.

The next few sub-chapters will explore some methods for impact evaluation design and analysis approaches.

Randomised Control Trials – The 'Gold standard' impact evaluation.

RCTs are **experiments** made of 3 main features: (1) control groups, (2) randomization, and (3) blinding.

1. **Control group:** a group not receiving treatment to be compared with the treatment group.
2. **Randomisation:** process of allocating participants randomly in treatment and control groups.
3. **Blinding:** procedures that prevent study participants, caregivers, or outcome assessors from knowing which participants were in treatment or control groups.

A typical example is a clinical trial, where one group is given regular treatment, while the other is given a placebo. The difference in outcomes is the net effect of the treatment.

RCTs are increasingly being applied in the field of social sciences despite the complexity of the approach. In the Youth Endowment Fund (YEF) toolkit, you can see examples of VRU projects that have been evaluated by RCTs.

As noted, not all programme or interventions are suitable for impact evaluation, and that includes RCTs. Table 7 explains when an RCT might be relevant and appropriate.

Table 7: Factors defining when to use an RCT

When is an RCT useful?	When is an RCT not useful?
▮ Interest in net impact. ▮ Intervention consistently delivered. ▮ RCTs planned from the start. ▮ Clear, measurable impacts.	▮ Interest in causal factors. ▮ High variation in how intervention is delivered. ▮ Quick answers are needed. ▮ Hard to measure or diverse results.
When is an RCT feasible?	**When is an RCT not feasible?**
▮ Possible to compare groups affected and not affected by the intervention. ▮ Possible to assign participants randomly. ▮ Large sample size.	▮ Not possible to compare groups affected and not affected by the intervention. ▮ Not possible to assign participants randomly. ▮ Small sample size.

Other Impact Evaluation methods (quasi experimental)

If an RCT is not possible or appropriate, there may be other impact evaluation methods you could apply in your evaluation. Below is an overview of other impact evaluation methods suited to different availability of data. You can follow the hyperlinks to resources providing more information on each.

Q: Do you have information on both treatment and control groups before and after the intervention?

- Difference-in-difference: Estimates impact by taking the difference between groups and before/after, controls for selection bias.

Q: Do you have individual-level data on known factors including participation for both groups?

- Propensity Score Matching (PSM): Matches the characteristics of treatment and comparator groups, allowing for better comparisons.

Q: Are individuals assigned to treatment and control group based on a cut-off point?

- Regression Discontinuity Design (RDD): Takes advantage of a programme 'threshold' (usually eligibility criteria to compare similar groups.

Q: Do you have data on trends before and after your intervention, but no concurrent control group?

- Interrupted Time-Series (ITS) analysis: when there is an expected 'interruption' due to a policy change.

Q: Can you use historical data to construct a 'clone' of a group receiving an intervention?

- Synthetic Control groups Method (SCM): Constructs artificial control groups based on a weighted average of existing comparator groups.

Theory-based impact evaluation

Theory-based impact evaluation can be used when using a counterfactual of any form is not possible. Instead, it tests whether the causal chains – the assumptions you made in your Theory of Change about the elements of your intervention that should lead to effects – are supported by strong evidence, and that other explanations for change can be ruled out. However, theory-based evaluation is less robust than using a counterfactual and is not able to conclude that change can be directly attributed to an intervention, but rather establish whether an intervention contributed to change.

Theory-based impact evaluation methods are recommended for use in very complex settings, for complex interventions, or when the intervention/policy/outcomes are designed to be adaptive (and so, aren't delivered in a consistent way with every participant).

Theory-based impact evaluation should broadly follow three main stages:

1

Develop a Theory of Change

Follow the process described in section 5 to develop your Theory of Change alongside stakeholders to accurately map any causal mechanisms. In a theory-based evaluation, alternative theories of how the change could have occurred should also be mapped.

2

Gather evidence

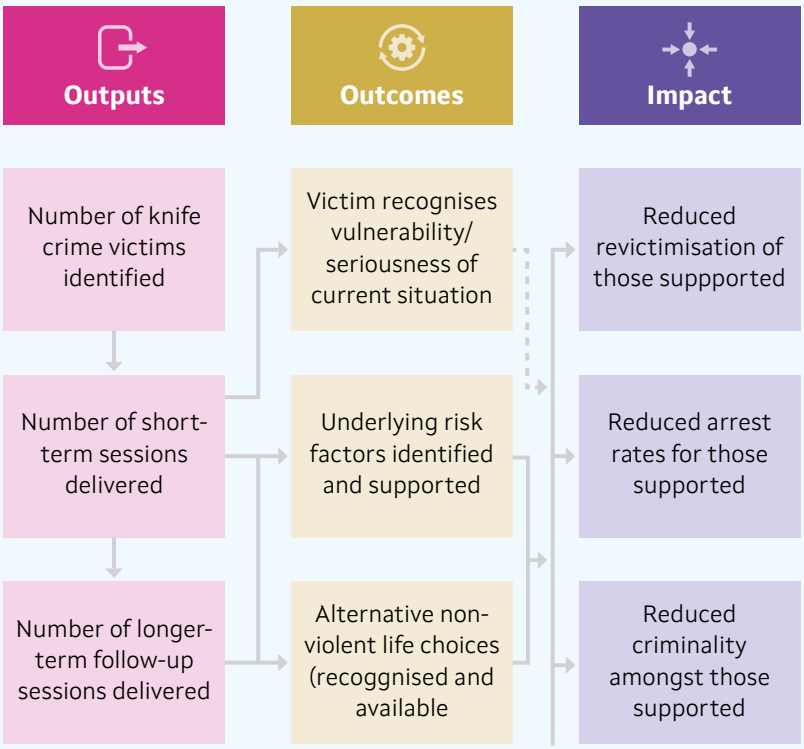
Evidence could be qualitative or quantitative and should be gathered with the purpose of testing causal pathways. The research questions should be developed accordingly.

3

Test causal pathways

Triangulate the evidence (see section 5.4.) and assess if, how, why, and when the intervention works.

Example: Looking back at the hospital navigators example ToC in section 2.1., we can see there is a causal chain between the output 'number of short term and long-term sessions delivered' and outcome 'non-violent life choices (recognised and available)'. To test whether this output led to the outcome, we could look at evaluation evidence which shows that sessions contributed to alternative non-violent life choices (such as exit interviews with participants, or follow up surveys), but also explore what other factors could explain changes in individuals' life choices. We could also test causal links from planned outcomes to impact and explore whether the take up of alternative life choices did lead to reduced criminality in the cohort.



There are different methods available to help weigh up whether your evidence is sufficient to conclude that your intervention contributed to the change. These are detailed in Table 3.2 of the Magenta Book.

Some other useful resources for theory-based approaches include:

- Contribution Analysis: an approach to understand the contribution of your program has made (or is currently making) to particular outcomes (as laid out in your ToC).
- Collaborative Outcomes Reporting.

5.6 Value-for-Money evaluation

You may have heard of value for money or 'cost benefit analysis' being used to explore the feasibility of capital or infrastructure projects. However, it can also be used in the evaluation of social projects or policies to express exactly what it says: how much value for money a programme or intervention provides. It does this by assessing the value of the outcomes or impact (not necessarily with a monetary value) and comparing this to the cost of the programme inputs. It may also consider what costs or benefits could have been incurred if another scenario took place (e.g. a different policy or programme). It can be used to answer research questions regarding whether the intervention was delivered efficiently or how much it cost to provide the outcomes delivered by the policy/project/programme.

Often, socially-focused projects aim to avoid or prevent a negative incident (such as a violent crime) taking place. Placing a monetary value on such outcomes can be complex, and generally involves assembling 'proxy' values of the cost to the public purse; in the case of violent crime, this may include costs of a hospital admission, an arrest, the criminal justice system and imprisonment.

Economic evaluation can be a powerful tool to communicate the value of your project. However, as with impact evaluation, **a value for money calculation should only be attempted if you or your team has received training**. The approach can be complex, and inadvertently creating misleading financial figures is likely to problematic for your organisation. This toolkit will briefly explain the different types of value for money evaluations, however it will not go into detail of how to undertake these methods. However, there are a number of resources that can help to assign a monetary value to outcomes if you do have access to an economic researcher:

- The Green Book: Government issued guidance on how to appraise projects.
- Manchester Combined Authority CBA tool: This includes a cost benefit analysis tool to input your project details into and also a generic data base that you can extract 'use values' from.

Requirements for assessing VfM

For a value for money assessment, you generally need:

- Outcomes or impact data for your cohort.
- Input cost data e.g. management information on staff costs, buildings costs, materials.

The following examples shows what kind of data we may need to conduct an economic evaluation of a programme.

Example: We have the following data for the individuals involved in the project.

Individual	Outcome 1: Not arrested violence for 12 months
1	Yes
2	Yes
3	Yes
4	No

Costs for the programme:

- **Staffing:** £15,000
- **Other overheads:** £5,000
- **Proxy costs for resources provided in kind (e.g. partner staffing, use of venues):** £5,000

We should also have a robust methodology in place for calculating the proxy cost to society of an individual avoiding an arrest for a violent crime.

Value-for-Money methods

There are different types of value for money assessments that can answer different questions:

- **Cost effectiveness analysis:** cost per outcome compares the cost of alternative ways of creating the same outputs. This is often used when cost benefit analysis cannot be undertaken as some valuations cannot be made but an outcome could still be counted (e.g. lives saved).

  Example: The programme in the example above cost £25,000 and achieved 3 crime-related outcomes at a cost of 25,000/3 = £8,333. This could be done in advance of delivery, with predicted outcomes, and compared to other programmes. Or it can be done after a period of delivery to assess the actual cost-effectiveness, and compare it to other programmes with the same outcomes.

- **Cost benefit analysis** assesses the impact of different interventions valued in monetary terms (where possible) by looking at all the costs and benefits of an intervention. Sometimes, the results are presented as a ratio of the value of benefits for every £1 spent on delivery. This type of approach is sometimes referred to as a Social Return on Investment analysis.

  Example: The cost to the 'public purse' of the crime 'violence without injury' is calculated as £6,480 and 'violence with injury' £15,353. The national VRU evaluation report published in 2022 applied these costs to outcome data generated by analysis of outcomes in a quasi-experimental impact evaluation,[42] and found the total estimated benefits of the programme were valued at £883 million. In this case, this represented a cost benefit ratio of 4.1 – that is, for every £1 invested in the programme, there was an estimated return of £4.10.

42 https://www.gov.uk/government/publications/violence-reduction-units-year-ending-march-2022-evaluation-report/violence-reduction-units-year-ending-march-2022-evaluation-report#section2

6.0

What next?

Delivered by

West Yorkshire Combined Authority | Tracy Brabin Mayor of West Yorkshire

West Yorkshire Violence Reduction Partnership

ECORYS | Answering tomorrow's challenges today

6.1 Disseminating evaluation findings

Dissemination is the final stage of an evaluation; it refers to the process of sharing findings with stakeholders and wider audiences. Whilst it is important to let your funders and staff know about the findings from your evaluation, it is also crucial to communicate findings (whether positive or negative) to the broader sector so that future provision can be improved.

When developing your research outputs, the first thing to think about is your audience. It is important to identify your target audience and tailor your message accordingly, because different audiences have different needs, interests, and expectations. These are things to consider when choosing the format, language, and tone of your communication. Ask yourself:

- Who are the audience(s), and what do they want to know from the evaluation?
- How might the audience best understand what we are trying to tell them?

Your audience(s) might include:

- The organisation delivering the intervention (e.g. delivery partners).
- Funders of the project, programme, or service (e.g. The VRU, The Home Office).
- Policy makers (e.g. the Home Office).
- Other organisations delivering relevant or similar interventions.
- Academics or researchers.
- Primary recipients of your programme, project, or service (the target population).
- Secondary recipients of your programme, project, or service (e.g. the local community).
- The public.

Depending on who the output is for, you will likely need to take a different approach to its preparation. While your findings should include information about all aspects of the evaluation, different parts of your audience are likely to be particularly interested in different elements. For example, funders may want a detailed understanding of outcomes and Value for Money, but practitioners may prefer practical information about good practice and what worked for whom (in other words, how the findings might affect how they work). Concise and engaging infographics, blogs, or press releases may be more appropriate if your audience consists of young people and the community.

6.0 What next?

Choosing the right format for your communication is essential. When planning your dissemination activity, you should consider your audience, and the resources and networks you have access to. There are a number of ways you can present and share your findings, including:

- Reports are comprehensive documents that outlines the processes, data, and findings of an evaluation. They can include recommendations and implications for practice or policy.
- Briefs are shorter, focused documents that highlight the main findings and implications of research.
- Infographics are visual representations of data or key messages and can be effective for capturing attention and conveying complex information in a simple way.
- Presentations can be delivered at conferences, workshops, or seminars, and can help network and get feedback from peers or stakeholders.
- Webinars are online presentations that can be delivered to a wider and more diverse audience and allow for interaction and discussion through chat or Q&A features.
- Blogs are short and informal articles that can be written or contributed to, and can help share insights and opinions, and invite comments and feedback from readers.
- Social media platforms can be used to promote and share research findings, engage with an audience, and create a buzz around a topic.[43]

Here is a useful resource to help you structure your evaluation outputs and inform your approach to disseminating your findings.

43 https://www.linkedin.com/advice/0/what-best-practices-disseminating-your-research-findings

6.2 List of resources

Although we have linked to various sources of information throughout this toolkit, the resources below are particularly helpful for developing your Theory of Change, sourcing evidence relating to other interventions, and developing your own evaluation.

Resource	Description	Link
Better Evaluation	A knowledge platform with info about how to plan, manage, conduct and use evaluation.	https://www.betterevaluation.org/frameworks-guides/managers-guide-evaluation
Magenta Book	Central Government guidance on evaluation.	https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879438/HMT_Magenta_Book.pdf
The Step-by-Step Guide to Evaluation: How to Become Savvy Evaluation Consumers	A step-by-step guide to evaluation.	https://search.issuelab.org/resource/the-step-by-step-guide-to-evaluation-how-to-become-savvy-evaluation-consumers-4.html
A Short Guide to Evaluation	An NHS-commissioned brief guide to evaluation. Includes a useful evaluation checklist.	https://www.nhsdg.co.uk/wp-content/uploads/2019/09/A_Short_Guide_to_Evaluation.pdf
Youth Endowment Fund – Guide to evidence and evaluations	Video explainers on YEF evidence and evaluation.	https://youthendowmentfund.org.uk/guide-to-evidence-and-evaluation/
Youth Endowment Fund – Resources for evaluators	List of resources for evaluators including guidance, templates, and policies.	https://youthendowmentfund.org.uk/resources-for-evaluators/
Education Endowment Foundation- Evaluation Design	Resources for impact and process evaluation, and outcomes measurement.	https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluation/evaluation-guidance-and-resources/evaluation-design
UNICEF Theory of Change	UNICEF brief on Theory of Change.	https://www.betterevaluation.org/sites/default/files/Theory_of_Change_ENG.pdf

Resource	Description	Link
Early Intervention Foundation – 10 steps to evaluation success	EIF Evaluation guide (including video on developing a ToC).	https://www.eif.org.uk/resource/10-steps-for-evaluation-success
Knife crime: a problem solving guide	Good practice guide on tackling knife crime from college of policing.	https://assets.college.police.uk/s3fs-public/2021-11/Knife-crime-a-problem-solving-guide.pdf
Youth Endowment Fund toolkit	An overview of existing research on approaches to preventing serious youth violence.	https://youthendowmentfund.org.uk/toolkit/
South Wales VPU outcome indicators	A comprehensive list of indicators and potential data sources with which to measure them.	https://www.violencepreventionwales.co.uk/cms-assets/global/Violence-Prevention-Indicators_Wales-VPU_2021.pdf
Draw.io Diagrams	Open source (free) and available through the Microsoft Store, useful for developing ToCs.	https://apps.microsoft.com/store/detail/drawio-diagrams/9MVVSZK43QQW?hl=en-us&gl=us
Microsoft Visio	Fee attached for but this software is usually included as part of Office 365 subscriptions.	https://www.office.com/
Qualitative guide: Ritchie, J., Lewis, J., Nicholls, C.M. and Ormston, R. eds., 2013. Qualitative research practice: A guide for social science students and researchers. sage.	Textbook on qualitative research – design, sampling, data collection, analysis and reporting.	https://uk.sagepub.com/en-gb/eur/qualitative-research-practice/book237434
The Effect: An Introduction to Research Design and Causality	Covers the key concepts in a very accessible way and includes analysis code. Available as a hard copy or for free online.	https://theeffectbook.net/index.html
Youth Endowment Fund Outcomes Indicators 2022	Outcomes framework for YEF projects.	https://youthendowmentfund.org.uk/wp-content/uploads/2022/08/YEF-Outcomes-Framework-August-2022.pdf

7.0

Glossary

Delivered by
West Yorkshire Combined Authority | Tracy Brabin Mayor of West Yorkshire
West Yorkshire Violence Reduction Partnership
ECORYS | Answering tomorrow's challenges today

7.0 Glossary

Glossary	
Activity	The actions taken or work performed through which inputs (resources, time, funding, etc.) are turned into outputs.
Analysis	The act of studying or examining something in detail to discover or understand more about it, or your opinion and judgment after doing this.
Assumptions	Conditions that must be true for the ToC to be valid and for the desired change to occur.
Baseline	Information collected before or at the start of a project or program that provides a basis for planning and/or assessing subsequent progress and impact.
Beneficiaries	The individuals, groups, or organisations, whether targeted or not, that benefit directly or indirectly from the intervention.
Case study	A systematic description and analysis of a single project, program, or activity.
Causal pathways	Any links between a cause and an effect that is part of the mechanism for change. For example, there would be a causal pathway from: ▮ lack of regular exercise; ▮ weight gain; ▮ increased risk of heart disease.
Causality	The relationship between one event (the cause) and another event (the effect) which is the direct consequence (result) of the first.
Conclusion	The opinion you have after considering all the information about something for example after reviewing evaluation evidence.
Data	Information collected; this could be either qualitative or quantitative (see other definitions).
Data collection	The collection of information to use in evaluation; this can be quantitative or qualitative.
Evaluation	The rigorous collection and analysis of information about program/intervention activities, characteristics, and outcomes that determine the merit or worth of the program/intervention. Evaluation studies provide credible information for use in improving programs/interventions, identifying lessons learned, and informing decisions about future resource allocation.
Evaluation framework	An overall framework for evaluations across different programs or different evaluations of a single program (e.g. process evaluation; impact evaluation). For example, this could include guidance on data, management or an underlying theory applied to the evaluation.

7.0 Glossary

Glossary continued	
External evaluation	When an evaluation is completed by someone outside of the organisation that delivered the intervention/project/programme.
Findings	Information that is discovered during an official examination of a problem, situation, or object for example in an evaluation.
Goal	An aim or purpose.
Impact	The long-term, cumulative effect of programs/interventions over time on what they ultimately aim to change, such as a change in crime rates, knife-related injuries or killings, and gang violence. Note: Impacts at a population-level are rarely attributable to a single program/intervention, but a specific program/intervention may, together with other programs/interventions, contribute to impacts on a population.
Impact evaluation	An objective test of what changes have occurred, the scale of those changes and an assessment of the extent to which they can be attributed to the intervention.
Indicator	A measure of progress of the Theory of Change.
Inputs	Resources provided for program implementation. Examples are money, staff, time, facilities, equipment, etc.
Internal evaluation	An evaluation conducted by the same organisation or team that delivered the intervention/project/programme.
Intervention	A specific activity or set of activities intended to bring about change in some aspect(s) of the status of the target population.
Monitoring	Routine tracking and reporting of priority information about a program/project, its inputs and intended outputs, outcomes and impacts.
Outcome	A change or result that a program or intervention aims to achieve in order to reach its final impact goals.
Outcome evaluation	A type of evaluation that determines if, and by how much, intervention activities or services achieved their intended outcomes.
Outcome monitoring	The process of systematically observing and assessing the outcomes of a particular activity, project, program, or policy.
Outputs	The results of program/intervention activities; the direct products or deliverables of program/intervention activities, such as the number of sessions completed, the number of people served, the number of knives seized.
Participatory evaluation	An evaluation approach that involves the stakeholders of a programme or policy in the evaluation process.
Pre- and Post-	Pre is before an intervention, and post is after an intervention is delivered.

Glossary continued	
Process evaluation	Process evaluation is designed to understand the following: ▮ whether an intervention is being implemented as intended; ▮ whether the design is working; ▮ what is working more or less well and why.
Programme	A set of interventions, activities or projects that are typically implemented by several parties over a specified period of time and may cut across sectors, themes and/or geographic areas.
Programme evaluation	Evaluation of a set of interventions designed to attain specific global, regional, country, or sector development objectives.
Project	An individual endeavour delivered to achieve some output.
Qualitative data	Non-numeric data to provide in-depth understanding collected through methods such as interviews, focus groups, or case studies.
Quantitative data	Numerical data (e.g., counts) collected through methods such as surveys. It can be used for statistical analysis and measurements.
Reliability	Consistency or dependability of data with reference to the quality of the instruments, procedures and used.
Representative	Typical of, or the same as, others in a larger group of people or things
Sample	A group of people or things that is chosen out of a larger number and is asked questions or tested in order to get information about the larger group.
SMART outcome	A statement of what one expects to achieve that is Specific, Measurable, Achievable, Realistic, and Time Bound
Stakeholder	Individuals, groups, or organizations that have an interest or are affected by a program or intervention, including program staff, beneficiaries, partners, and policymakers
Thematic Analysis	A way of producing themes from texts such as interview or focus group transcripts.
Theory of Change (ToC)	A Theory of Change describes how and why a program is expected to work and explains the underlying causal chain or theory that connects the building blocks and ultimate goal. It is often depicted as a map or pathway of change considering inputs, activities, outputs, outcomes, and impact.
Triangulation	The process whereby evaluation findings are integrated around each of the research questions.
Validity	The quality of being based on truth or reason, or of being able to be accepted.
Control Group	A randomly selected group that does not receive the services, products or activities of the program being evaluated.

7.0 Glossary

Glossary sources
UNAIDS: https://www.unaids.org/sites/default/files/sub_landing/files/7_1-Basic-Terminology-and-Frameworks-MEF.pdf
Girls Not Brides: https://www.girlsnotbrides.org/documents/1035/Session-3e-ME-training-handout-final.pdf
Glossary of Evaluation Terms: https://pdf.usaid.gov/pdf_docs/Pnado820.pdf
Magenta Book: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879438/HMT_Magenta_Book.pdf
Cambridge Dictionary: https://dictionary.cambridge.org/
Better evaluation: https://www.betterevaluation.org/methods-approaches/approaches/participatory-evaluation#:~:text=Participatory%20evaluation%20is%20an%20approach,the%20reporting%20of%20the%20study
Evaluation academy: https://www.evalacademy.com/articles/interpreting-themes-from-qualitative-data-thematic-analysis

Delivered by

West Yorkshire Combined Authority | Tracy Brabin Mayor of West Yorkshire

West Yorkshire Violence Reduction Partnership

ECORYS | Answering tomorrow's challenges today

For further information, please email:

WYVRU.KnowledgeHub@westyorks-ca.gov.uk