



Understanding impact

Using your theory of change to develop a measurement and evaluation framework

James Noble, Peter O'Flynn, Anne Kazimirski

February 2020

Contents

Introduction: Why should charities evaluate their work?	3
Incentives: Why charities want to understand their impact	3
What questions should we be asking?	4
Why do we need measurement? What's wrong with our own judgement?	4
What makes measurement so difficult?	5
Should we just give up and go home?	6
Part One: Translating your theory of change into a plan for measurement	7
A theory of change highlights what is relevant to measure	7
Five types of data you can use with your theory of change	8
Bringing this all together	23
Summary	26
Part Two: Prioritising what you measure	27
Introduction	27
Identify the critical elements of your theory of change	27
Use your 'assumptions' to highlight a list of research questions	27
Don't measure things you know already with enough confidence	28
Consider what information will be most helpful for your decisions	29
Develop an evaluation plan you can share with others	29
Considering the impact question	29
Summary	32
Our approach for service delivery charities	33
Conclusion	34
Glossary	35
Theory of change terms	35
Common measurement terms	35
Further reading	37
NPC and Inspiring Impact Resources	37
External resources	38
References	41

We are grateful to Esmée Fairbairn Foundation for their generous support for this work.

Introduction: Why should charities evaluate their work?

This guide is for charitable service providers who seek to evaluate their work or measure their impact. When we talk about 'measuring' we mean using data and evidence, both quantitative and qualitative, to gain useful insights about a charity's work. At heart, measurement is about understanding whether your project or service has made a difference in people's lives. If yes, how so? If not, why not?

We begin by outlining how to turn your theory of change into a plan for measurement, the five types of data you will need from service users, and the additional information that will be helpful to you. Finally, we explore how to prioritise what to measure. We've also included seven factsheets on our website offering a closer look into specific aspects of measurement.

We've written this guide for smaller charities and charitable services, especially those who are new to this or have limited budgets or capacity for data collection. We focus less on campaigning activities, systems change, or shared evaluation approaches, but much of what we say here will be relevant to any evaluation of social impact.

This guide follows on from [Theory of change in ten steps](#), which we want you to use as the basis for any measurement process. As in our [ten steps](#) guide, we bring together a wide range of knowledge into a comprehensive approach to help you achieve something worthwhile without getting lost in the process.

Helping you to maximise the difference you make is central to our work at NPC, with measurement being one of our core consultancy services for charities and funders. In our 2014 guidance, [Building your measurement framework: NPC's four pillar approach](#), we described the fundamentals of measurement as:

- Mapping your theory of change.
- Prioritising what you measure.
- Choosing your level of evidence.
- Selecting your sources and tools.

We argued that practical decisions about *what* you measure and *how* to measure it should be informed by your priorities, who you want to be interested in the results, and what is actually achievable. Since our 2014 guidance was published, we've helped hundreds of organisations to understand their impact, through direct support or training. By working with so many charities, we're continually uncovering new ideas and techniques. This updated guide builds upon everything we've learnt, and draws on our other NPC papers on this topic. We include new advice on how data should be collected, analysed and used, and discuss what an evidence-based organisational culture could look like.

Incentives: Why charities want to understand their impact

In our experience, charities have two main drivers for wanting to do 'measurement':

- **Learning** from information to improve effectiveness and achieve more for beneficiaries.
- **Showing** funders and supporters what they have achieved.

Learning is undoubtedly the right motivation. Focussing on 'learning' rather than 'showing' helps charities to:

- **Improve day to day practice and efficiency** by being open to new information and change.
- **Make better decisions** by learning which activities are likely to be most effective.
- **Respond more intelligently** to the complexity of the social issues they are trying to influence.
- **Become more accountable and credible** through recognising that things can always be better and being open to feedback.
- **Be more open** about success and failure.

Conversely, too much focus on 'showing' can lead you to eschew these benefits because they don't suit your narrative. Only presenting results that put you in the best light will cloud your thinking of what could be better.

What questions should we be asking?

The [Results Based Accountability approach](#) (RBA) condenses the range of things we *could* ask to three core questions:

- How much did we do?
- How well did we do it?
- Is anyone better off?

We think this is a helpful overview so we will refer to these questions regularly. We add to them in Part Two when we discuss how to define specific questions applicable to your circumstances.

Why do we need measurement? What's wrong with our own judgement?

Whatever you're doing it's likely you'll have a sense of whether it's working or not, whether you are making a difference and what could be better. But there are good reasons to choose a more systematic approach:

- You won't always be able to see or understand all the effects of what your organisation or programme is doing, especially if you are a larger organisation or running a longer-term programme. You need insights across programmes and from everyone involved.
- You can't assume things based on what you happen to see or hear. You have to talk to people. You can't generalise from one or two cases, so you should try to talk to a range of people, including those who drop out or do not attend. It's not enough to rely solely on those who volunteer their views.
- People will leave or move on from services. So, unless you put effort into research, you will not see how influential your project has been long term.
- Your judgements might be wrong. There are charities who, while well-intentioned, have delivered projects which are ineffectual, even harmful. Evaluation data should give you day to day insights to help you adapt your programme or service to meet changing needs.
- It is hard to be impartial if you are involved in designing and delivering something. Good evaluation is more objective because there are standards against which data collection and analysis can be assessed.
- Your stakeholders are unlikely to be persuaded by your own judgement. They need a stronger and more diverse selection of evidence. Good evaluation supports accountability, so donors, taxpayers and beneficiaries can be confident that money is being spent wisely.

What makes measurement so difficult?

The root of the measurement problem is that the three RBA questions above get progressively more difficult to answer. Asking whether anyone is any better off thanks to your intervention gives charities a real headache. Why is this? There are two inherent difficulties:

1. The data collection problem

It's hard to collect reliable data from people about things that have changed. To be 'better off' means change needs to be sustained. This is what we mean by 'outcomes', and to a greater extent 'impact'. The challenges to collecting this kind of data are:

- The people experiencing outcomes and impact are probably no longer available to tell us what changed. They are busy living their lives. Collecting data therefore demands a lot of effort to get hold of people and persuade them to answer your questions. This is particularly hard with transient groups like students or homeless people, who quickly move on, or with people whose employment or family life is too busy for them to respond.
- Charities often aim to achieve change that is personal and therefore less obviously observable. For instance, changes in knowledge, attitudes, or feelings of wellbeing and security. Measuring these things demands self-reflection by service users which they might be unwilling or unable to do. You often need quite sophisticated methods to collect this kind of information.

Some service users, such as very elderly people, very young children, or people with severe disabilities, will struggle to tell you about the outcomes they are experiencing.

These factors mean collecting any data about outcomes and impact is normally difficult and often costly. Moreover, to actually measure these things quantitatively we need to collect data on a large enough scale to draw statistically significant conclusions. There are some fortuitous circumstances, such as accessing data collected by schools or hospitals, but for most charities this is a challenge. Our [data labs](#) project was set up to solve this, and you can find out more on our website.

2. The attribution problem

Even if you can measure it, how can you determine how much of the change which took place was down to you? People are complex and ever-changing, as are the societies in which we live. By contrast, the effects of social programmes are generally fleeting and highly variable across individuals and contexts. Even if you are able to collect some outcomes and impact data, you still have the problem of identifying the effect of your own work alongside everything else going on.

This is the attribution problem. It is the most difficult and contentious aspect of evaluation. Charities often encounter it without realising. For example:

- A charity is teaching maths to children. Are the improvements it measures higher than it would normally be expected as children get older and learn more?
- A charity is raising awareness of a health condition. How much of the improvements it measures are down to their efforts rather than the work of others?
- A charity is working with people leaving prison to prevent reoffending. Its results compare well to the national average for reoffending, but were they working with people less likely to reoffend in the first place?

There are multiple ways to respond to the attribution challenge, which we cover in Part One in our discussion of the five types of data. These include experimental or quasi-experimental approaches in which a group of service users (the treatment group) are compared to another group as similar as possible except that they didn't receive the service (control group). If the treatment group has a different outcome to the control group

then this is likely due to the impact of the service. This is the basis for our [data labs](#) initiative which uses government data to create virtual control groups.

Unfortunately, for most charitable programmes, getting to the point where you have enough data from both treatment and control groups to demonstrate a statistically significant difference is neither financially nor practically viable. Even if you could achieve it, your results would still only be specific to that context.

Given the complexity of the world, most evaluators now recognise that merely asking 'what works' is oversimplistic. At the very least, we should be asking 'what works, for whom, in what circumstances?'¹.

Should we just give up and go home?

At this point you might be thinking that pondering these problems are a pessimistic way to begin. But the fact is that any organisation working on a measurement framework needs to face up to them, and to appreciate that they are burdens all organisations share.

The good news is that these challenges are only applicable to the 'is anyone better off?' question. Moreover, we've found there are usually ways to tackle this modestly and manageably. Meanwhile, the questions 'how much did we do?' and 'how well did we do it?' are eminently answerable and will deliver plenty of useful information.

So, a positive note upon which to begin is that charitable organisations can get this right and develop measurement frameworks that have a positive effect on their work and promote greater impact for beneficiaries. This is not a farfetched aim.



Part One: Translating your theory of change into a plan for measurement

The best place to start thinking about your measurement framework is your theory of change. Theory of change is a process for how charities think about and describe what they aim to achieve and how they hope to achieve it. At NPC we think of theory of change as the foundation of strategy. Our [Theory of change in ten steps](#) guide is a great place to start.

In summary, the ten steps are:

- **Step 1: Situation analysis:** What is the issue you hope to tackle? What does your organisation offer?
- **Step 2: Target groups:** Who are the people you can help or influence the most?
- **Step 3: Impact:** What are the sustained or long-term changes you want to see?
- **Step 4: Outcomes:** What shorter-term changes for your target group might contribute to impact?
- **Step 5: Activities:** What are you going to do?
- **Step 6: Change mechanisms:** How will your activities cause the outcomes you want to see?
- **Step 7: Sequencing:** In what order do outcomes and impact might occur?
- **Step 8: Your theory of change diagram:** How do you reflect and communicate your intentions to others?
- **Step 9: Stakeholders and 'enabling factors':** Who (and what) are the amplifiers and inhibitors to you achieving your desired impact?
- **Step 10: Assumptions:** Does your theory of change stand up to scrutiny?

One of the main reasons for doing a theory of change is to inform your evaluation and data collection activities. In this chapter we show you how to do this by considering the data you could collect around your theory of change.

A theory of change highlights what is relevant to measure

Your theory of change represents an agreed picture of what you want to deliver and how you want your target groups to benefit. It narrows your focus from potentially measuring anything towards a clear list of what is relevant. For example, a youth charity might have agreed priority outcomes for service users that include 'greater awareness of what is going on in the community' and 'increased willingness to take part in things', but not 'a better understanding of how to use social media'. The first two would therefore be relevant to measure, whilst the third would not be.

A subtle trap to avoid is thinking that your only 'measurement' task is to *test* your theory of change. This risks limiting your understanding because you cannot anticipate everything. Crucially:

- Context for people is rich and varied. While you might think you have a good grasp of your 'problem' (step one in the theory of change), there will always be more to learn, barriers to overcome, stories to hear etc.
- People will experience your work uniquely and may benefit in ways you couldn't expect.
- You need to ask people how things could be different.

In the example above, it might turn out that some beneficiaries really need to know how to use social media better or that your service helps them do that. It would therefore become a relevant point to measure, despite staff initially not thinking it would be. If you don't give users the opportunity to give feedback in their own terms you will not find this out. So, you need ways to hear peoples' stories and experiences in their own words through qualitative approaches and open-ended questions. Asking these kinds of questions gives people a proper voice.

This goes back to our introductory argument that measurement should be driven by 'learning' rather than 'showing'. A charity driven by 'showing' might only ask about the things in the theory of change because their motivation is to show the theory of change works. It's all they care about. A charity driven by learning will want to learn about the things in the theory of change and keep an open mind about what else might be going on.

Five types of data you can use with your theory of change

It helps to appreciate that there are only a few types of data that any charity *could* collect for performance management or evaluation. From our experience, we've found the most helpful information from service users tends to fall into one of [five types of data](#):

- **User data:** Is your service effective at reaching the intended target group?
- **Engagement data:** How effective is your service at continuing to engage your target service users?
- **Feedback data:** What do people think about the service?
- **Outcomes data:** How have people been influenced or helped by your service in the short-term?
- **Impact data:** Have the outcomes achieved (above) helped people to change their lives for the better?

In addition, there are two types of internal data that you should collect, which link closely to engagement data:

- **Output data:** How much have you produced or delivered?
- **Quality data:** What is your own assessment of how well you do things?

You should also consider what is already known about the area in which you are working:

- **Background evidence:** What can existing knowledge tell you about the problem and people's needs? This could draw upon all types of data.

Helpfully, each type of data corresponds to a different step in theory of change, as shown in the table below. Note though the difference in sequence. Our [ten steps](#) guide is written in the order in which you answer the different questions, whereas the table below is arranged according to the order in which people will experience them.

Whenever NPC supports a charity to do a measurement framework, we work through these types of data and background evidence to ensure we cover all the right questions and don't miss anything out. Our examples have service delivery in mind, but the framework can also be applied to campaigning and advocacy.

Step in the theory of change process	Corresponding type of data
Step 1: Situation analysis	<p>Background evidence: Existing knowledge of the problem and people's needs</p> <ul style="list-style-type: none"> • Any kind of existing information about the scale and causes of the problem • Previous research into possible solutions and existing evidence on similar programmes • Consultation with stakeholders and those affected to help develop your ideas
Step 2: Target group	<p>User data: Background information about the people you reach:</p> <ul style="list-style-type: none"> • Their characteristics • Issues they face or experience
Step 5: Activities	<p>Engagement data: The extent to which people engage with your activities:</p> <ul style="list-style-type: none"> • Frequency, timing, duration of engagement and mix of activities • The quality of engagement <p>Output data: The quantity of what you produce or deliver</p> <p>Quality data: Your own assessment of how well you do things</p>
Step 6: Mechanisms and quality	<p>Feedback data: What people thought of the project</p> <ul style="list-style-type: none"> • Whether activities are experienced in the way intended? What they thought and felt while using the service? • What the 'quality' was like? • How did they rate it overall? Would they recommend it to a friend? • Specific likes / dislikes. How could it be better?
Step 4: Outcomes	<p>Outcomes data: Short-term changes or benefits your target groups may get from the project (e.g. changes in knowledge, attitudes and behaviour)</p> <ul style="list-style-type: none"> • Have people started to make the positive changes you intend? • How else have people benefited from or been changed by the project? • Which parts of the service have helped, for whom in what circumstances
Step 5: Impact	<p>Impact data: Sustained, meaningful change</p> <ul style="list-style-type: none"> • Has any lasting or sustained change been achieved? • Has the wider community benefited in some way?

Over the following pages we explain each type of data in detail. In each case we discuss the following questions:

- What is it?
- What are the key questions?
- Why is it important?
- What are the challenges?
- How do we collect it?

Then in Part Two we go on to discuss ways to prioritise different questions and types of data.

Background evidence

What is it?

Background evidence can draw from any of the five types of data. It refers to any existing external data, research, evidence or information that might inform, support or improve your work. No matter what you are working on there will always be something relevant you can find and use.

What are the key questions?

The main types of background evidence you can look for are:

- **Data from official or other sources that can tell you about the scale of the problem, levels of need and which communities it might be best to target.** There are nearly always official statistics available to help you look at the scale and prevalence of an issue. This data is much easier to access than it used to be. Look at government research, statistics, publications, and community data produced by local authorities.
- **Information about your target beneficiaries, the challenges they face and what they think will be most valuable.** You might find existing research or the results of a previous consultation, but you could also conduct your own primary research with potential beneficiaries. This could range from surveys and focus groups to full involvement in your theory of change process and [co-production](#). Consultation is always valuable, and in situations where there are lots of unanswered questions it will be vital. This is separate from user data in that it is about understanding the situation rather than evaluating your intervention.
- **Academic insight, theories, or hypotheses.** It's good to be familiar with the latest academic thinking in your area and draw upon it. For example, in criminal justice [desistance theory](#)² is a prevailing model so a lot of charities seek to align themselves with it. Academic insight is especially valuable because it tends to relate to the connection between outcomes and impact, which is the hardest part of a theory of change to evidence yourself (see below).
- **Previous research or evaluations of similar interventions.** You might be able to find published evaluations of similar projects to your own. If not, there might yet be research from different contexts or settings that help you think about your own work. This could include international research and evaluations of projects that are different but share some common elements.

Why is it important?

The more background evidence you have underpinning your theory of change, the stronger it will be and the more confidence you and others will have in it. Familiarity with background evidence saves time and effort, because any aspect of your theory of change that is well supported by research does not need to be tested again. For example, if your intervention aims to get people exercising more, you do not need to show how exercise improves health and wellbeing, because that is already well-evidenced by research.

How do I collect it?

Doing a theory of change is a good prompt to explore background data because you can use it to inform the process through a better understanding of your 'situation' and the most likely solutions. Beyond this, you should routinely look for relevant publications that might support your work or encourage you to reflect.

Your first port of call should be online searches, which should help you find the data that is most relevant. If you are lucky you might find that someone has summarised or synthesised evidence for you. In many policy areas in the UK we now have '[What Works Centres](#)' that have taken on this task. These would be a good place to start.

How long it takes ranges from commissioning a systematic literature review through to an afternoon spent searching online and looking at research summaries. The first option is better, and we have seen some charities do this, but either option is better than nothing.

What are the challenges?

It can be daunting to engage with research if you don't have an academic background. Much of it is not even accessible because journals want you to pay for it, although working papers can sometimes be available through websites such as [ResearchGate](#).

Alternatively, you may be able to get help from academics themselves. Some charities have developed good links with universities and secured help from students and academics at little or no cost. For academics there may be an opportunity to further their own interests through the work you are doing. It doesn't hurt to ask.

But make sure you don't cherry-pick

When looking for external evidence it can be tempting to select only the sources which support your model. This will give a distorted picture of the external evidence base and the extent to which your programme aligns with it. The most robust way to prevent cherry-picking is to conduct a systematic literature review of what has already been said on the topic. This is a lengthy and costly process, out of reach of most charities. If this isn't an option, it is still important to look widely for potentially relevant evidence, including drawing on systematic reviews that have already been conducted, and present a balanced picture of what you find.

In reviewing external evidence, attention to context is key

Knowing that a model has delivered successful results in other settings cannot guarantee that it will do the same in yours. To understand how changing the context could influence the strength of supporting external evidence, you should pay close attention to enabling factors within your theory of change. These are the elements which need to be in place for you to achieve outcomes, such as the policy context, levels of need, delivery partners, and infrastructure.

User data

What is it?

User data is information about the characteristics of the people who use or sign-up to your service. User data can be collected early on in your relationship with people, ideally at referral or sign-up.

What are the key questions?

The user data you need should be directly informed by the characteristics of your intended target group defined in step two of the theory of change. Remember that target groups can either be the ultimate beneficiaries of your work or others through whom you want to work to reach beneficiaries indirectly (such as teachers or doctors). It

should be straightforward to translate this directly into the information you collect from people when they first encounter your project. The key question is:

- What are the characteristics of the people you reach?
 - **Objective characteristics:** Gender, age, ethnicity profile, educational level, housing status, income.
 - **Subjective characteristics:** Current attitudes and aspirations.

Capturing these characteristics lets you find out whether you are reaching the people in your target group and identify who you are not reaching.

In addition, you might also want to ask wider questions like:

- What are the experiences of the people you are reaching? What issues do they face? What is their story? What else is going on that might affect their engagement, outcomes and impact? This may overlap with data collection on baseline outcomes (see section on outcomes data).
- What other services are they engaging with? What have they engaged with in the past?

These questions will help you to understand your service users and their needs better, as well as providing contextual data to help you understand some of the other information you collect. You may have already considered some of these questions in step nine of the theory of change, which asked you to consider 'enabling factors' and things that might affect how people could benefit from your work.

Why is it important?

- **User data encourages us to check we are reaching the right people.** For example, if your service aims to reduce loneliness you need to know that you reach people who are lonely. This is sometimes overlooked. Often there's plenty of demand for a service and it's tempting to be satisfied with that. But it's always important to challenge yourself on whether you are reaching those whose needs are greatest or whom you can help the most. When charities consider this they usually think about types or groups of people they feel they could reach more. Very often these are the most vulnerable people from the most disadvantaged and marginalised parts of society. By collecting user data, we *challenge ourselves* on whether we are reaching the right people.
- **User data is the first meaningful indicator of success you have.** Showing that you can effectively reach the right people is an important goal in its own right, particularly in areas like youth outreach, prison work, or health interventions with disadvantaged groups.
- **User data enhances analysis.** By collecting user data, we position ourselves to identify any variations in how different types of people engage with services later. It lets us compare demographics, feedback, and factors associated with success and failure.
- **User data helps us measure and understand 'need'.** As highlighted above you can use the opportunity of collecting user data to ask wider questions about people's experiences. This information can inform how you deliver services or campaigns.

For example, we worked with a parenting charity that collected extensive information about the issues parents were facing, like debt and housing concerns, and used this to influence their local authority. Frontline organisations are uniquely placed to collect this information. When measured over time it can be a powerful resource revealing how things are changing.

What are the challenges?

There are two opposing challenges around collecting user data:

- There are always many relevant questions you could ask.

- Collecting user data is time-consuming and may involve intrusive questions, especially in the early stage when trust needs to be built. Aggressive data demands can discourage people from returning.

You need to find the right balance between the information you need and not harming the service itself. Sometimes it is necessary to delay asking people questions until you have established a stronger relationship.

How do I collect user data?

You should collect user data routinely from everyone you encounter as early as possible in your relationship with them. This is because:

- You are trying to understand reach. If some people drop out almost immediately you will not get an accurate picture unless you do it quickly.
- You have the opportunity to collect it, because people are accustomed to being asked questions when they sign-up with a new organisation.

The obvious way to collect this data is to have a sign-up form with relevant questions. For example, 'How old are you?', 'What is your ethnic group?', 'How often do you feel lonely?'. In some situations, it might be possible to also get user data from partners or referrers who are already working with people and have asked those questions before. This is always worth pursuing if it seems possible—whilst still complying with data protection legislation.

Ideally, all user data should be entered into a database or client management system against which all further data for that individual can be recorded.

Outputs and quality data

What are they?

Alongside engagement data, outputs and quality data look at your processes and what you deliver. We think of them as separate from the five types of data because they are collected internally rather than from your service users.

- **Outputs** refer to the quantity of what you produce, for example how many sessions you deliver, how many leaflets are produced, how many services are provided.
- **Quality** refers to judgments of how good your work is based on information on service delivery.

What are the key questions?

To count your outputs, simply tally the number of things you produced, irrespective of who engaged. For example, the number of sessions you delivered.

For judging quality, there are both quantitative and qualitative measures you can use. Quantitative measures could include:

- Time to answer calls.
- Time spent per individual.
- Clients seen within a day of referral.
- Availability of different options.
- Proportion delivered by trained volunteers.

Meanwhile, qualitative assessment of how good things are could include:

- Peer review of processes.
- Self-reflection of practice by staff. This is more difficult as it is subjective.

Why are they important?

Counting outputs usually isn't very important or enlightening. Indeed, a common criticism of charitable data collection and evaluation in the past has been the tendency to rely too heavily on outputs. The only real value of output data is:

- To provide a simple understanding of what you have done. For example, to inform a funder about the quantity of things taking place.
- To combine with other types of data to highlight the likely scale of achievements.

Judging quality is useful to check you are meeting your own standards and aims, or those which apply to your sector, and identify ways in which you could improve.

What are the challenges?

Output and quantitative quality data is mostly easy to collect. Qualitative quality data will take more effort as that will entail some resources to undertake.

How do we collect them?

You will probably need to spend a bit of time making sure you have a system in place to count outputs, especially if you are working across multiple sites.

Appropriate systems for quality data will vary a lot depending on what type of service you are providing.

Engagement data

What is it?

Engagement data follows closely from user data. It relates to how the people you reach use your services, in particular whether those you reach continue to engage, and engage in the ways you hope. Engagement data relates most closely to step five (activities) in your theory of change process.

What are the key questions?

Like user data, engagement data can be divided into:

- **Objective:** Frequency, timing and duration of engagement and the mix of activities they engage in.
- **Subjective:** The quality of engagement:
 - How do people seem to respond? Are they actively listening and taking part or are they just present?
 - What kind of relationship do you establish with people?

Engagement data addresses the key question of whether a project is effective at retaining people and delivering the intended activities with them.

Why is it important?

The importance of engagement data is obvious. Logically people need to engage with your work for it to have any chance of success, so you need to collect data to check and show that this happens. Like user data, engagement

data provides an early indication of success. If the data says people are not engaging as much as you expect then something is going wrong and needs attention.

Good engagement data will enhance your analysis possibilities. When linked to user data you can start exploring what types of people come back and continue to work with you. If it is linked to feedback and outcomes data (see below) then you can begin investigating the relationship between duration and quality of engagement and likelihood of success.

What are the challenges?

As with all data collection, time and effort is an issue. Some charities, especially those delivering light touch interventions like events or advice services, decide that a simple register or record of contact is all they can manage. But any service that works with people for longer should track engagement by individual users using a case management system.

How do we collect it?

Objective engagement data is relatively easy to collect. The main method will be staff or volunteers entering data during or soon after their interactions with service users. For example, recording who attends a particular session.

As with user data, you should collect engagement routinely for everyone you work with. It should be part of your basic monitoring processes and ideally entered into a case management system.

It gets more complicated when you want to go beyond recording whether someone has shown up. For example, how long they stay and what they do. You will need to decide how important this is to your theory of change.

For good data collection we suggest you:

- Make data entry as easy as possible.
- Ask staff or volunteers how they think data should be collected.
- Encourage data to be entered as-and-when rather than storing it and entering it in bulk.
- Ensure staff are consistent in the data they enter.
- Try to move to an electronic rather than a paper approach to minimise data entry.

Your approach to both user data and engagement data should be closely linked. Ideally it will be done in the same system. This will increase efficiency and help you to look at how reach and engagement are related, such as who is coming back and who isn't, and what strategies you can put in place to retain those who you think will benefit the most.

Feedback data

What is it?

Feedback is what people think about your services. Feedback lets you test whether your service is experienced as you intended, which should be clear from the mechanisms and quality section of your theory of change (step six), but it's also about giving people a voice. Feedback data is an opportunity for people to share what they think *in their own terms*. Which is the right thing to do **and** how you get the best information on how to improve.

Feedback does not only come from your users. You can also collect it from staff, volunteers and other stakeholders. Don't forget non-users or those who drop out. You can learn a lot from them about how to improve.

What are the key questions?

You will want to check that your services are experienced in the way you intended. Your feedback questions will therefore be partly determined by the mechanism and quality section of your theory of change (step six). To do this, start by asking 'closed' questions directly about those things, as opposed to open questions. For example, in a mentoring project it will be things like feeling comfortable, a sense of trust, feeling listened to and so on. For an advice service you will want to know if people felt the service was quick, polite, relevant, and useful.

We also recommend that you ask some general questions to rate the service overall using a 0-10 scale:

- How satisfied or dissatisfied are they?
- Would they recommend it to someone else?

You should also ask open-ended questions that allow people to answer in their own words and give you suggestions for improvement.

- How did they find out about your service? What attracted them to it?
- What aspects do people rate the best and worst?
- How do people think the service could be improved?

Why is it important?

Feedback data helps you test your service model to establish whether your service gets the reaction you want and whether it is beginning to work in the way you intended, which will help you think about ways to improve.

What are the challenges?

We have the usual concerns about overburdening staff and beneficiaries, but with feedback we can worry about this a bit less because people usually like to be asked their opinion on things.

The key challenge is to get **representative** feedback. It's often the most positive and most negative who respond. The main job in designing a good feedback system is to figure out how to get the views of people who *don't normally volunteer to be consulted*, including those who drop out or disengage. This comes down to putting effort into reaching out to these people and designing feedback forms that are short and feel relevant.

How do we collect it?

Feedback data can be informal or formal:

	Informal feedback	Formal feedback
Description	Encouraging users to share their views over time and in different ways.	Using structured methods like questionnaires or interviews.
Why it's useful	You will get suggestions for improvement and an early indication of where things could be better. Users might feel listened to and more engaged.	You can direct people towards the specific questions you are interested in, encourage more people to take part and collect feedback from a more representative group. Formal feedback will also be more persuasive for stakeholders and funders.



Example methods	<ul style="list-style-type: none"> • Pop-up surveys • 'Show of hands' • Satisfaction tokens • Suggestion boxes • Social media comments • Just asking people 	<ul style="list-style-type: none"> • Surveys • The data that staff or volunteers collect about peoples' progress • Focus groups • In-depth interviews • Observation/ethnography • Interviews with other stakeholders
Suggested frequency	All the time. Users should feel that you are interested in their views and have a range of opportunities to share them.	Formal feedback can be more infrequent because no-one wants to answer surveys all the time. Different options for delivering a feedback questionnaire are discussed in the outcomes section below. You will probably want to ask for feedback and outcomes data at the same time.

Introducing outcomes and impact data

The final two types of data will give you the greatest headaches. These are where the two challenges of collection and attribution mentioned in the introduction really start to matter. It's important to understand the distinction between them, particularly when doing a theory of change. Hence, when it comes to discussing the challenges and approaches, we address the data collection challenge by talking about outcomes, and the attribution challenge by talking about impact.

Given the challenges and costs involved in collecting outcomes and impact data, charities should always ask themselves whether they need to do it. We would go as far as to say that for smaller individual organisations, especially those providing light touch interventions, outcomes and impact data should be seen as the exception rather than the norm. The next chapter on prioritising discusses this further.

Outcomes data

What is it?

Outcomes data is about how people have been influenced or helped more immediately in the short-term by your service or campaign. It is directly linked to the outcomes section in your [theory of change](#) (step four). The best way to think about outcomes is as changes in attributes (knowledge, attitudes and behaviours) that people retain within a short time of your service. For example, a few days or weeks.

What are the key questions?

As with feedback, you will want to ask 'closed' questions directly about the outcomes that you said you wanted to achieve in your theory of change. For example:

- If you are working with young people, are they feeling more confident about their future or less so? Are they more engaged in their local community or less so?
- If you are working with people with a health condition, do they feel more knowledgeable about that condition, more able to cope? Do they engage with health services and manage their condition more effectively?

- If you are campaigning for a change in policy, do more people understand the issue, agree with your perspective, and take whatever actions they can to support that change?

As with feedback data, you should also ask more open-ended questions that allow people to answer in their own words. These questions are subtly different to feedback questions, as they are specifically about what people have done, rather than what they think of your service.

- What is different now? Have you made changes?
- How, if at all, has our service helped?
- How would things have been different without our service?
- Which aspects of our service have helped?

Why is it important?

Outcomes data starts to indicate whether you are really making a difference for people, whether they are gaining resources, benefits or assets from the service. Funders and donors are likely to be interested in outcomes because it tells them that their money is making a difference.

If there is good existing evidence supporting the link between outcomes and impact in your theory of change then evidence of outcomes will give you confidence that your service users will enjoy positive long-term change (impact). Outcomes are signs that your service is having an effect, so it is always appealing to measure them.

What are the challenges?

Collecting outcomes data is harder than collecting user, engagement and feedback data. This is because the people experiencing the effects of your work may no longer be in contact with you. Even if you still see people, it can be hard to persuade them to spend time answering your questions.

The changes you hope to achieve are often internal (like a new way of thinking) or unseen (like new behaviours or social groups). These are sometimes referred to as soft outcomes. People may not realise how they have changed or may not want to tell you. Therefore, as well as being time consuming and potentially intrusive, it is difficult to find the right questions to get a good quality response.

As with feedback, trying to get representative data on outcomes is an important challenge. Charities will be able to get outcomes information in the form of case studies about some 'success stories', which is nice for communications, but less valid for performance management or evaluation because you don't know how representative these stories are. Respondents to charitable data collection will typically be the most favourable people, or those who have been helped the most (perhaps alongside a small number of people at the other extreme). It means you always need to work hard to get people to answer questions and to engage those who don't normally volunteer to be consulted.

Finally, there are other common sources of bias to be aware of and to try to avoid:

- **Confidence bias:** Something we hear repeatedly is how people report high levels of confidence, wellbeing and aspiration at the start of an intervention and then reduce their scores as they start to trust services more or because of self-reflection and learning how far they could improve. There's no easy answer to this, other than really stressing how much you want honest answers from the start or making data collection a joint exercise between a worker and a beneficiary, such as the [Outcomes Star](#)³ approach. Good training is essential to ensure consistency.
- **Courtesy bias:** Many people will give the answers they think you want to hear, out of politeness or because they think negative answers might have bad effects for staff or for themselves. This is very common. As

above, the only solution is to try hard to make sure people understand the purpose of the data collection and to be as honest and candid as possible.

- **Impatience bias:** Longer questionnaires are particularly vulnerable to people getting bored and not completing them or answering them in a slapdash way. Obvious solutions are to keep things short and give questionnaires to people at the right times. You can also change the order in which you ask questions so it's not always the same questions at the end which get missed.

How do I collect outcomes data?

Outcomes data can be collected through the same methods as formal feedback (and possibly at the same time). Beyond this we have six top tips:

1. Use approaches that have been designed before

Helpfully, in most areas of charitable work you will find there have been previous attempts at measuring what you are interested in. There are often established outcomes tools or frameworks. Over the last few years, many service providers have come to realise that they are working towards the same outcomes and impact. For example, most youth organisations have similar outcomes and there is consensus about what is important. This commonality is the basis for [shared measurement initiatives](#) like the [Youth Investment Fund](#) learning project, which reflects the fact that there are a many good reasons for collecting outcomes data in a consistent way.

Your first port of call should be to review what's already available to see if any are suitable. Ideally you will find an outcome tool that you can adopt wholesale. Alternatively, you can also take questions from different tools and build your own. However, a 'pick and choose' approach will mean you won't benefit as much from any testing or validation of those tools, which will reduce the validity and credibility of your data.

A downside of many of the most respected tools is that they have many questions and take a long time to complete. They are designed for use in academic studies so charities often struggle to administer them. They can feel like overkill and the quality of data you end up with might be poor because it is so hard to get people to fill them in. Hence, we would argue that unless your intervention has a reasonable prospect of getting a representative sample, you should be careful about using longer outcomes tools.

2. A before-and-after approach

The best approach to using any outcomes tool, either off-the-shelf or one you have made yourself, is to ask the questions of your service users **'before-and-after'** the intervention. The first time to ask is when collecting user data. You should then repeat the same questions over time to measure change for individuals.

If you can't take this approach, or if it seems like overkill, then you will have to rely on questions that ask people to say what they believe has changed as a result of your intervention. This approach is more open to error. For instance, a health based dietary intervention may ask what someone had for dinner the night before, but it is very difficult for people to remember what they had for dinner a month ago.

3. Think about what will work in practice

Deciding on the right way to administer questionnaires or tools depends a lot on the strength of relationship you have with your service users:

- In situations where you have an ongoing and strong relationship with people, you might be able to build outcomes measurement into your routines. For example, a housing charity might use an outcome measurement tool, like the Homelessness Star, to start useful conversations with people. Similarly, a charity working in schools might be able to give out questionnaires termly in lessons.

- Alternatively, you could think about occasional surveys. For example, we worked with a charity that concentrated all of its efforts to collect outcomes data into one week in every year.
- Some services may only meet people once. For example, those that just provide information. Here the options are very limited. These organisations will find collecting outcomes data especially hard and may only be able to collect it there-and-then with one or two questions at most.

4. Use staff or volunteer assessments or proxy information

Aside from questionnaires or tools, you might be able to collect other observable indicators of change that can act as proxies. For example, do people interact more with others? Do they manage their health better? Do they seem happier?

The best way to measure outcomes is often for staff or volunteers to make judgements about how people are progressing. It's not a particularly strong methodology because it's based on interpretations, and it's a challenge to make these consistent. But it has the advantage of being feasible, because staff and volunteers should want to help you and will probably understand the kinds of changes you are looking to identify.

5. Think about qualitative research and smaller studies

Another good alternative is to conduct smaller scale studies and/or interviews or focus groups with samples of beneficiaries. This is especially useful if you can talk to people at a few different times to better understand how things change. We discuss this further in the following section on impact data.

6. Think about sampling and question frequency

It's important to consider whether you need to measure all beneficiaries' outcomes. Ideally yes, but this is likely to require lots of time and money, including good case management software and a well organised team who can collect and enter the data. In practice, it is usually a better option to concentrate your efforts on collecting better quality outcomes data from a sample of service users rather trying and failing to get it from everyone.

Similarly, it's important not to try to collect outcomes data too often because:

- Outcomes usually take time. You don't need to keep checking-in on people.
- Service users and staff will not want to spend all their time collecting this type of data.

Again, the right answer will depend on the type of relationship you have with your service users. Our overall advice is to ask for outcomes data as infrequently as possible and avoid asking too early.

Impact data

What is it?

Impact data is the long-term sustained change that people achieve as result of your services. In our [ten steps](#) guide we discuss impact at step three. It is worth reviewing the criteria we applied to it there to see how it is distinguished from outcomes (step four).

What are the key questions?

The key question is: What lasting or sustained change has occurred?

This question implies a positive change, and indeed this is what you will particularly want to know. But you should also be interested in negative changes and collect data in such a way as not to miss these. In some situations, it

is not so much a change as an achievement. For example, passing an exam, recovering from a health problem, recovering from substance misuse. It can also mean preventing things from getting worse, delaying the onset of problems or maximising wellbeing, or other benefits, in the face of difficulties.

The usual next question is: To what extent has our work contributed to this lasting change?

This introduces the attribution challenge that we explained in the introduction. We present some possible responses to it below.

Finally, you can also consider: Has the wider community benefited in some way? Although it is usually very difficult to see the effects of your work at the community level, unless your service is particularly large or you are working in policy or systems change.

Why is it important?

Data on impact is highly desirable. We all want to understand when we've made a difference and we want to tell people about it.

What are the challenges?

In one respect, the **data collection challenge** is harder for impact data than it is for outcomes data. This is simply because impact is further into the future so it will be even harder to stay in touch with people to ask them questions. By implication, there are many situations where impact measurement is simply impossible. For example, charities conducting open access youth work have no way to learn the long-term education or employment impacts of the people they work with other than anecdotally.

But in other ways collecting impact data is *sometimes* easier:

- Impact tends to be more observable, as by definition, impact is a substantial positive change. It is easy to see major changes like getting a qualification or job, staying out of hospital, or making a change in the law. It is more straightforward to ask someone about these things than it is for outcomes like self-esteem.
- Sometimes you will find relevant government or official data about impact. For example, the NHS collects data about patient outcomes, schools measure attendance and results, government conducts a wide range of different surveys about employment, wellbeing, health etc. Theoretically, all of this data could be useful to charities, but the challenge is getting access to this data in the first place, and then identifying your service users in it. Data protection means this is usually impossible, but there are exceptions. Charities working in schools might be able to get good data on attendance and exam results, while charities in the UK justice sector can get data on reoffending through the [Ministry of Justice Data Lab](#). At a national level, NPC is working hard to establish better processes for all charities to access this kind of data.

Notwithstanding these points, the advice we gave above about collecting outcomes data is equally applicable to impact data. This leaves us with the **attribution** problem. How do we know it was our services that made the difference? Would change have happened anyway? We discuss some possible responses to this below.

How do I collect impact data?

There are different ways to respond to the challenges, which we list below. The first three options listed below are probably the best options for smaller charities to do by themselves, while options four and five should only be attempted with expert help. You may also have the opportunity to take part in shared evaluation initiatives through large grant programmes such as the [Youth Investment Fund](#) or [Building Connections Fund](#).

1. Draw upon the existing research

You should have found some support for your project's model through your review of the existing research (background evidence), particularly around the links between short-term outcomes and long-term impact. For example, there is already strong evidence that:

- Exercise improves health.
- Support and recognition for carers prevents emotional strain.
- Encouraging parents to engage in their children's schooling has a positive effect on exam results.

You should not feel compelled to demonstrate such links all over again. You can use the existing research to do that for you. In an ideal scenario, you may have collected some good outcomes data and will use this research to assume or assert that some positive impact is a likely result. Keep in mind the context though. Knowing that a model has delivered successful results in other settings cannot guarantee that it will do the same in yours.

2. Smaller-scale research

Smaller-scale quantitative or qualitative research is typically easier and less costly to do than quantitative impact evaluation. It can provide good indicative information about potential outcomes and impact. For example, you should be able to collect some evidence of people's journeys and possibly the part your project has played in improving outcomes and impact.

This can be done for a sample of service users rather than everyone. But if you are using a sampling approach it is important not to only talk to those beneficiaries who engage or benefit the most, as this will create overly positive results that lack credibility. Rather, we suggest selecting people randomly, and keeping in touch with those who do not engage as well.

The attribution challenge can also be partly addressed through qualitative research:

- You can simply ask people what they think would have happened without the service in question, and whether there are other reasons for improvements in their lives, such as other support or significant life changes. For example, people leaving prison can reflect on what would have happened without any help. This is obviously quite speculative and will not provide 'proof' of impact but the answers are often quite revealing.
- More sophisticated qualitative approaches⁴, like The Qualitative Impact Assessment Protocol (QuIP)⁵, Process Tracing⁶ and Qualitative Comparative Analysis⁷, can be used to look for patterns in fairly large qualitative datasets, to understand the factors associated with change. Again, these approaches fall short of 'proof' but are increasingly respected as a meaningful and relatively robust response to a difficult question.

3. Accessing data from other sources

There may be data about your wider target population that you can use as a benchmark. For example, we know the national average levels of wellbeing for different population groups. How does the group you are working with compare? Can you see change against this average as your project continues?

Finding data that helps you contextualise your results might be an option. For example, a health charity might find data from a national survey about the awareness of a condition to compare service users to.

As noted above, there may already be data about your beneficiaries that you can use to study outcomes and impact. For example, the NHS have data like hospital admissions data which could tell you whether your service users have been able to make a positive change. Although this data can be hard to access, it is worth speaking with your partners or commissioners to see what is feasible.

Another type of secondary data that charities sometimes use to talk about impact is population level data. For example, if you are working in a school to raise attainment you might look at the exam results for the whole school, or if you are working on wellbeing you might look at data from the Office for National Statistics about happiness in a local authority area. This is useful, but it's very unlikely to help a charity understand its impact because any effect of its work will be lost amongst all the other things happening in that wider population. To make a claim about impact in population data, you will either need to demonstrate that your service is transformational or that you have reached the majority of people in that community.

4. Identifying opportunities to find control groups

There will be people in your target group who are not receiving your service. They can be a 'control' group to compare your service users to. Within this option there is a big range, from weaker approaches that find control groups amongst people who haven't volunteered for a service, through to the supposed 'gold standard' of research, the Randomised Control Trial, in which people who volunteer for a service are assigned to a 'treatment' or 'control' group and compared in the same way that medical researchers test new treatments. The more likely options available to charities are as follows:

- If you have a waiting list for your project, you could collect a small amount of data from those on it and look at how this data changes compared to those who receive your service (this is known as a wait-list design).
- Similarly, if you can't deliver to an area or community yet, you could collect data from this community and compare it to a community where you do deliver (known as a 'natural experiment'). It is difficult to collect data from non-users, so this is generally only an option when an organisation is already delivering some services to a population and your project is an additional service.

5. Commission an impact evaluation

The fifth option is to commission an impact evaluation from independent researchers. This costs money so is usually only necessary or feasible when a project is reaching enough people and has sufficient budgets for robust research. When the time comes, we strongly recommend getting expert help to do it properly.

Bringing this all together

Having described the different types of data, let us conclude this chapter by giving you some templates for bringing all this information together. The following table summarises the discussion above in terms of the different methods you might deploy for each type of data we discuss above. We hope you find this to be a useful checklist and that it helps prompt ideas.

Types of data you could collect from service users	Quantitative	Qualitative
User data Background information about the people you reach.	Sign-up / registration / referral forms.	Sign-up interviews with service users, open-ended questions on referral forms, conversations with referral partners.

<p>Engagement data</p> <p>The extent to which people engage with your activities.</p>	<p>Counting outputs (sessions delivered, guides produced, invitations sent).</p> <p>Recording attendance and reattendance by individuals.</p> <p>Usage monitoring (attendance, website hits, downloads, ‘click throughs’, calls to helplines etc.).</p> <p>Quality checking—time spent per individual, clients seen within a day of referral, availability of different options / choices, proportion delivered by trained volunteers.</p>	<p>Staff / volunteer processes for assessing the quality of delivery and engagement by individuals (e.g. simple scales or qualitative notes on how well someone engaged).</p>
<p>Feedback data</p> <p>What people thought of the project.</p>	<p>Informal approaches:</p> <p>Pop-up surveys. ‘Show of hands’. Smiley face / satisfaction tokens.</p> <p>Formal approaches:</p> <p>Surveys / questionnaires. Key worker assessments of people’s progress.</p>	<p>Informal approaches:</p> <p>Suggestion boxes. Social media. Case studies. Just asking people.</p> <p>Formal approaches:</p> <p>Focus groups. In-depth interviews. Ethnography. Interviews with other stakeholders.</p>
<p>Outcomes data</p> <p>Short-term changes or benefits that your target groups may get from the project.</p>	<p>Informal approaches:</p> <p>Pop-up surveys. ‘Show of hands’. Smiley face / satisfaction tokens.</p> <p>Formal approaches:</p> <p>Surveys / questionnaires. Key worker assessments of people’s progress.</p>	<p>Informal approaches:</p> <p>Suggestion boxes. Social media. Case studies. Just asking people.</p> <p>Formal approaches:</p> <p>Focus groups. In-depth interviews. Ethnography. Interviews with other stakeholders.</p>
<p>Impact data</p> <p>Sustained, meaningful change.</p>	<p>Secondary data</p> <p>Follow-up / longitudinal research.</p>	<p>Follow-up / longitudinal research with small samples.</p>



How does this apply in practice? Below, we outline some of the indicators that may be applicable to a youth charity. Here we have assumed that the charity focuses on providing activities related to improving educational outcomes for young people at risk.

Types of data you could collect from service users	Methods chosen	Indicators
<p>User data</p> <p>Background information about the people you reach.</p>	<p>User forms during sign-up or core information sent from referrals.</p>	<p>Date of birth; Sex / Gender; Location; Race, Ethnicity; Household status; Schooling (location); Schooling (achievement to date).</p>
<p>Engagement data</p> <p>The extent to which people engage with your activities.</p>	<p>Electronic case management system.</p>	<p>Assessment of how often young people come to activities / sessions (and how long for); time spent with case workers per individual; average referral time until first meeting or service provision.</p> <p>Assessment of interest / boredom by type of activity.</p>
<p>Feedback data</p> <p>What people thought of the project.</p>	<p>Surveys or simple satisfaction rating systems.</p> <p>Informal methods—like social media, one to one conversations or suggestion boxes.</p>	<p>Extent of enjoyment of activities / service. Whether they would recommend the service to someone else.</p> <p>Qualitative feedback on facilitator of activities; suggestion box on how sessions could be better applied, or games for the session to make more fun.</p>
<p>Outcomes data</p> <p>Short-term changes or benefits that your target groups may get from the project.</p>	<p>Focus Groups.</p> <p>Surveys.</p>	<p>Changes in knowledge, perception, attitudes and behaviours with regards to education.</p> <p>Young people's views of the role of the services in generating that change (attribution) as opposed to teachers, parents, or other actors.</p> <p>Which aspects of the service have helped which types of service users in which circumstances? And which have not?</p>

<p>Impact data</p> <p>Sustained, meaningful change.</p>	<p>Control group design for sample.</p>	<p>Increased educational attainment of a group of young people who used the service compared to similar young people on the waiting list.</p>
--	---	---

Summary

In this section we have described the five types of data you could collect, the benefits of each type of data and some of the issues each type raises. As we said at the beginning of the chapter, we use this framework whenever we are working with a charity on measurement framework because it ensures we cover all the necessary ground.

As shown, the challenges increase as we go through the five types of data until we get to impact at the end, with impact typically the most difficult type of data to collect. Collecting user, engagement and feedback data should be routine. They are the easiest to collect and offer you actionable insight on the questions of ‘how much did we do?’ and ‘how well did we do it?’, which will help you improve the services you offer.

Meanwhile, outcomes and impact offer more fundamental insights that will help you question your underlying theory of change, but this is not something you need to be doing all the time. Collecting this data can therefore be more occasional and exceptional.

In Part Two, we suggest an approach to balancing these considerations, which we think will work for most service delivery charities.



Part Two: Prioritising what you measure

Introduction

In Part One we discussed how your theory of change highlights what data might be relevant, and how by applying your theory of change to the different types of data you can start to think about your measurement framework. But there is a further level of prioritising to do before finalising your plans.

Why is this needed?

- It's unlikely you will be able to afford the time or money to collect data about everything. Nor should you collect more data than you have the resources to analyse or use.
- You will have some important information needs, some less important and some things you know already with enough confidence.
- Your different stakeholders and audiences will want to know different things.
- You will face some practical barriers as some things are harder to measure than others.

In this chapter, we share some ideas on prioritising within your theory of change, to help you identify a more specific set of questions beyond just trying to measure everything. We conclude this guide by suggesting an approach that we think should work for most charitable organisations.

Identify the critical elements of your theory of change

It is sometimes possible to simply look at your [theory of change](#) and identify the most important things. This might feel strange, because if something is less important you might wonder why it's in your theory of change at all. In reality, there is usually a hierarchy of importance, particularly amongst outcomes and mechanisms.

For example, on a youth employment programme you might have a range of outcomes like self-esteem, determination, autonomy, empathy, positive attitudes to work etc. All are potentially relevant, but you can identify that what all participants really must gain from the programme is a plan for what they need to do next and to have positive aspirations around that plan.

In this way, prioritising is just a simple exercise of asking yourself what is most important. Better still, you could ask volunteers or former service users what they feel are the most important things.

Use your 'assumptions' to highlight a list of research questions

A more sophisticated way to use your theory of change is to identify the key questions it raises, both for yourself and others (including funders). You should have already started to do this in steps nine and ten of your [theory of change](#), in which you will have written a list of your biggest concerns or risks. Helpfully, these are also your main research questions.

A simple illustration of this in action is to imagine a simple theory of change for a school breakfast club:

Serve healthy breakfasts → Children eat healthy breakfasts → Children concentrate better in lessons → Children's attainment increases.

If we didn't prioritise, we'd measure all of these things. But if we also developed some 'assumptions', based on our thinking about where the theory of change is weak, at risk or untested, this would identify which are more important.

For example, for the school breakfast illustration, the assumptions are:

1. That we can produce a 'healthy breakfast' with the small kitchen and limited budget we have.
2. That targeted parents will bring children to the breakfast club in the first place.
3. That targeted parents will not feed them an unhealthy breakfast beforehand.
4. That we can get the children to eat our healthy breakfasts.
5. That the lack of healthy breakfasts is the cause of the concentration problems we have been seeing amongst some of our children.
6. That better concentration will lead to increased attainment for the children.

This represents a more refined set of questions than trying to measure everything that is relevant, so this gives us a better steer towards what types of data we need to collect. For example, we now know we'll need to observe who eats the breakfasts and how much they eat, and we need a way to measure concentration levels.

Don't measure things you know already with enough confidence

You almost certainly have some information already that can help narrow your focus, whether it is your own or from other people's research. To learn how to collect this, see the section on **background evidence** in Part One.

How much information you have available will depend on how new or innovative your service or campaign is. The development of your service or campaign will heavily influence your research questions. Our [proportionate evaluation](#) paper provides more detailed advice on this.

Taking stock of what you already have

When thinking about performance management and evaluation we are usually doing so for existing services and campaigns. There is usually some evaluation history and existing data collection to consider:

- What data do I have and where is it held? Who is responsible for collecting and managing that data?
- How, if at all, is the data currently being used? Is it useful?
- Is the data accurate and consistent? Is there anything missing?
- Which data collection exercises have worked well? Which have not?

Consider what information will be most helpful for your decisions

Possibly the best way to prioritise is to anticipate the most important decisions you and your stakeholders will need to make, and therefore what information will be most helpful for making better decisions.

For example, in the breakfast club illustration above:

- We will need to make a quick decision about whether the breakfast club is appealing to and engaging the right children, whether parents bring their children, and whether they eat their breakfasts. So, we need to keep detailed information about take-up.
- Assuming we do not reach everyone in our target audience right away, how are we going to increase its appeal? We will need information from non-users about their perceptions of the breakfast club and the barriers to engagement.
- Ultimately, both our funders and us will need to decide whether to keep running the service based on what it costs and whether it seems to be making a difference. So, we will need information that tells us the level of take-up and any changes in concentration levels and attainment amongst targeted pupils.

Develop an evaluation plan you can share with others

From all of the above, you should be able to start developing a measurement framework. This will list the main questions that you have decided are your priorities and build on the work you did in Part One, using the different types of data. This is a useful way to show others which methods and indicators you plan to use to answer the questions above.

Considering the impact question

The biggest decision charities need to make is how much to prioritise the ‘is anyone better off?’ question. In other words: ‘impact measurement’.

We would obviously all like to know whether our work makes a difference, whether it achieves impact. But because of the inherent challenges discussed in Part One, in most circumstances getting this data will prove illusive. This point is still not widely accepted, and consequently the charity sector spends a lot of time agonising over this. The power dynamic between charities and their funders can fuel the problem.

In this section we run through the main arguments that determine how important impact measurement is in different circumstances, so you can weigh them up and figure out whether it is a priority for you. We conclude with a suggested approach to bringing these arguments together, in a way that we think is the best bet for most smaller charitable services.

1. Impact measurement is *morally* the right thing to do

There is a moral argument for answering the impact question. Charities use money and time given voluntarily by donors and service users, so we should feel obligated to test whether their resources are being put to best use, that they are making a positive difference, and not causing harm. Even in situations where charities are completely voluntary, it’s hard to argue that we shouldn’t try to make charitable work as effective as possible by studying what is achieved.

The difficulty is that because measuring impact is difficult and takes time and resources, we quickly run into the counterargument that these resources would still be better spent on helping people. So, in a sense, while it's right to pose the question, it's wrong to spend all our time and money trying to answer it.

The solution is to find the right balance between what we would ideally like to know and what is possible and proportionate. So, for instance:

- We should only try to answer the impact question selectively and occasionally, when it is most pressing. And, once we have answered it with enough confidence, we do not have to keep asking it again.
- We should draw on all the information that already exists to make sure that we are only trying to answer questions that need answering. So, if there is already evidence around the connections in one part of the theory of change, this is not a priority for further measurement.
- When we attempt to measure impact, we should answer our questions well enough so that we can trust the results and so that the information can be useful to someone else.

In other words, do it occasionally but do it well and do it to support a growth in overall knowledge rather than to meet the needs of your organisation alone.

2. To raise your profile and increase your funding

This is a common motivation for impact measurement, especially for charities who rely on trusts and foundations, philanthropists, commissioners, or impact investors. They believe that having good impact data represents valuable currency in the ongoing battle to secure funds. Sometimes funders themselves make specific requests for impact measurement, they may even fund charities to do it.

Conversely, charities who rely on public donations or paid services tend to feel under slightly less pressure, because smaller donations come with less scrutiny.

We tend to think that it's wrong to pursue impact measurement only because you think it's what your funders want. Doing so will encourage you to want to 'show' rather than 'learn', which has negative consequences. In this worst-case scenario:

- You only choose approaches and questions that have the potential to demonstrate impact. Meanwhile, approaches that might help you learn or do better, like qualitative research, get ignored or deprioritised.
- You take on the most difficult questions, like trying to show impact statistically, and risk spending a lot of time and money but still ending up disappointed.
- You will struggle to report your findings objectively or honestly.
- It puts you in a competitive mindset. Rather than sharing what you have learned with others so that they might benefit, you will only release what puts you in a favourable light.

We would also question whether pursuing impact measurement for funding purposes will even have the effect you want. We talk to a lot of charities who *perceive* that funders are interested in impact but actually will be happy with other kinds of data. We recommend talking to funders and other stakeholders about what information they need and what questions they want answering. Streamline your data collection so it meets your internal needs as well as those of your external stakeholders. It is useful to explain to funders what your data collection decisions are based on, so that you can negotiate what to report on. Share your theory of change, evidence gaps, and evaluation questions, as well as any existing frameworks for your sector.

We are happy to see that more funders are getting better at having realistic expectations and are valuing charities who focus on learning and having a good learning culture, rather than those who boast about their achievements.

A word to funders and commissioners

Funders should recognise the role they play and their potential to drive both good and bad practice. The important thing is to be conscious of how challenging the impact question is. What might seem like a simple information request might translate into a world of difficulty, stress and wasted resources.

In our view, funders and commissioners should make sure they understand what impact measurement is, only make data requests if they are prepared to help fund the data collection and are clear on how they are going to use the data. Moreover, it is best to focus on collaborative research projects that test big questions of general interest, rather than solely whether organisation A or B achieves impact, which may add very little to the evidence base.

3. Impact measurement for learning

The best argument for impact measurement is to learn something new that will clearly influence your work and the work of others providing similar services. To understand what information would be genuinely new and useful, you need to be up to date with the existing evidence base, so you don't repeat other research but can take an informed view about the questions you could usefully answer.

What this means in practice will vary very much depending on the sector or area you are working in and how novel your work is, as discussed in our guide to [proportionate evaluation](#). As an illustration, we discuss some different scenarios below:

- Many charities provide mentoring for people leaving prison. It is something which has been researched and refined over many years. There isn't much new to be learned about good practice in mentoring, but there are enduring questions about its long-term impact on reducing reoffending that are hard for small charities to answer. These charities should try to measure impact, but they should do it together so that they can invest their combined resources into the highest quality research. Our [criminal justice research](#) identifies the key areas where charities are making a difference.
- Many charities are trying to confront online bullying amongst teenagers. This is a relatively new phenomenon and research is still in its infancy. Lots of things are being tried, so it is important to look for evidence around outcomes and impact to help us quickly understand what the promising approaches are.

4. There is an opportunity to do it

Finally, it's worth noting that sometimes a fortuitous opportunity to measure impact arises because someone is already collecting data that you can use. This usually only applies in institutions like schools and hospitals where this kind of data is collected routinely, but even if the data exists it usually remains hard to get hold of and difficult to analyse.

The best thing charities themselves can do is be opportunistic and to engage with those who hold useful data to see if they will work with you to access it. Very often, data owners are also those people commissioning services, like the NHS and local authorities. We have seen situations in which commissioners ask charities for impact data that the commissioner already has. We want to see many more conversations between commissioners and providers about making data available and useful.

Summary

The aim of this discussion has been three-fold:

- Firstly, we want you to appreciate that it's important not to jump into 'impact measurement'. Don't see it as something you must do all the time.
- Secondly, that the right motivation for impact measurement is to learn something new. This needs careful thought and a solid understanding of what the existing evidence base says before moving forward.
- Thirdly, that if you decide that you do need to measure impact then you should aim to do it well by combining your efforts with others, doing it infrequently but to the highest quality possible, seeking expert help and raising funds to do it properly.

Think about stakeholders and understand their evidence needs

Different stakeholders will have different priorities for the questions they want answers to, and different views about what counts as good evidence. Discussing evidence with them helps you to ask the right questions and prevents you from collecting data that isn't needed or useful. Funders and service deliverers may need to communicate and negotiate on evaluation plans, based on existing evidence and mutually agreed priority outcomes. No stakeholder will be totally predictable, but typical values are suggested below:

- **Frontline staff and volunteers** may see data collection as a burden, and will need convincing that it makes a difference. Measurement should provide useful learning for your staff, though this may vary depending on how well you understand your programmes already.
- **Beneficiaries'** perspectives are crucial and must not be overlooked. Research findings are important for keeping organisations accountable to beneficiaries, and provide helpful examples of how others have achieved their goals.
- **Commissioners** must balance budgets with needs. Indicating that your work supports outcomes that align with their strategic priorities is important, but the required standard of evidence is key. Commissioners are likely to value how outcomes can be supported within specific costs. This does not necessarily entail expensive data collection. Commissioners might be just as happy with forecasts or estimates using social value calculators.
- **Grant-makers and donors'** evidence requirements will depend on the kinds of organisation they're looking to fund and the outcomes that they're hoping to support. Certain approaches and metrics will be commonly recognised and align with funder interests. Strong, experimentally derived data might attract some funders, but many will be just as interested in programme theory and how change mechanisms might work (which could include qualitative data on outcomes). Funders often want grantees to develop a mutually beneficial approach to measurement and reporting. A good understanding of evidence needs and a plan for meeting them can significantly influence reporting criteria. NPC's research into funder impact practice found that evidence of impact plays a bigger role in renewing funding than initial grants.⁸
- **Policymakers** will often be interested in evaluations that help answer questions related to programme efficacy ('what works') and will be influenced by government guidance prioritising quantitative data derived from a high-standard (counterfactual) research design⁹. Depending on the stage of policy development, policymakers will be interested in promising findings and answering broader questions, such as 'understanding behaviour change, value for money, unintended outcomes and longer-term impacts'.¹⁰ If you're looking to influence policymakers, how you use your evidence may be just as important as the kind of evidence you produce.¹¹

Our approach for service delivery charities

Let us wrap up the discussion of the last two chapters with a suggested approach that we think works for many service delivery charities and reflects the advice we have given above. It can be summarised through the maxim:

'A little from a lot, a lot from a few'.

This means collecting a little bit of user, engagement and feedback data from as many service users as possible, ideally from everyone, while concentrating your outcomes and impact data collection on smaller samples of representative service users. The idea is that:

- Collecting user and engagement data from everyone will tell you the scale of your reach, how many people you helped.
- Collecting feedback from everyone gives them a voice and maximises your opportunities to improve.
- Collecting outcomes and impact from smaller groups or samples can tell you the type and amounts of change your service is capable of, as well as further insights about how it works for people.
- So long as your sample is genuinely representative, you have the basis for combining user, engagement, outcomes and impact data to generalise about what your service has achieved. Whilst you cannot prove anything outright, you can have confidence that you are achieving what you intended. Seeking 'proof' of impact may not be proportionate or feasible for your intervention.
- Consider using a counterfactual or control group but use expert advice if it's an option for you.

Some important points to note about this suggested approach:

- The routine collection of user, engagement and feedback data should not be seen as especially taxing, it is just part-and-parcel of running a good quality service.
- The most challenging aspect is that data must be linked at the individual level, meaning that the feedback you get from someone can be linked back to data you collected about them at the outset along with data on their engagement in the service. The ideal scenario is that every piece of data from an individual can be brought together. This greatly enhances the analysis you can do, such as tracking individual changes over time, and the story you can tell. This is difficult because it raises data security and confidentiality issues. Furthermore, as your service grows, you will need a good electronic case management system.
- The occasional collection of outcomes and impact data can be done periodically, perhaps once every couple of years, or regularly but only for a sub-sample of service users.
- We cannot emphasise the need for representativeness strongly enough. It means that your research on outcomes and impact will have to include people who dropped out of the service, didn't really engage, and even those who disliked you. You need to cover all these different responses before you can generalise and say that this is what is typical. Unfortunately, the tendency in charities is to focus on success stories. Select randomly to achieve representativeness.
- Combine quantitative data, like engagement, with qualitative data, like people's stories. This is often called triangulation. The best attitude to adopt is seeing all data and information as potentially contributing to your learning and to the narrative that you want to convey about your service.

Conclusion

We hope that you find our advice helpful. Our mission is to support charities and funders to maximise their social impact. As a charity ourselves, we understand the challenges of making something happen with limited time and resources, and have taken this into account in this guidance. As we are on our own journey to improve, we would very much like to hear your views about our advice, and what more we could do. Please contact info@thinkNPC.org with your comments.

Remember that you do not need to address everything we have recommended that you think about in order to improve your approach to understanding your impact, and the best way to start is through testing approaches and working out what works best for you and your organisation. We wish you luck on your learning journey.



Glossary

Theory of change terms

Theory of change: A process for thinking about and describing; the change you want to see; and your plans for achieving that change.

Problem: A summary of the issue or challenge your work is aimed to tackle.

Context: The nature, scale, causes and consequences of the problem you are addressing.

Target group(s): The group(s) of people you believe you can help or influence the most.

Impact: The sustained change you want to see in your target group(s), which they will achieve themselves.

Outcomes: Changes in your target group(s) that you believe will contribute to impact.

Activities: Whatever actions you take.

Mechanisms: What you want people to be thinking, feeling or doing whilst they are experiencing your service or campaign.

Quality: How we plan to deliver our services or campaigns, so that people experience them in the way we want.

Enabling factors: Things outside your immediate control which might help or hinder the theory of change.

Assumptions: Challenges or weaknesses, that reveal the main underlying beliefs on which your theory of change is based.

Evidence: Existing research and data that has informed your thinking.

Common measurement terms

Attribution: Isolating and accurately estimating the contribution of an intervention to an outcome.

Baseline: A minimum or starting point in an intervention used for comparisons.

Causality: The relation between an event or events (cause or causes) and a second event or events (effect or effects), where it is understood the second is a consequence of the first.

Contribution analysis: A technique generating evidence on the role and influence of an intervention on an outcome, in the absence of experimental data.

Counterfactual: An estimate of what would have happened in the absence of the intervention or organisation.

Efficacy: The ability to produce a desired or intended result.

Evaluation: The use of information, from monitoring and elsewhere, to judge and understand the performance of an organisation or project.

Experimental approach: An evaluation that compares outcomes of recipients of an intervention to those of a control group who did not receive the intervention.

Fidelity: Delivery being faithful to the original model in terms of its core components, the duration of the intervention, and the quality of the service.

Impact measurement: The set of practices through which an organisation attempts to learn something about the sustained difference it has achieved. Can be used interchangeably with impact evaluation.

Indicator: A marker of accomplishment/progress. It is a specific, observable, and measurable accomplishment or change that shows the progress made...Common examples of indicators include participation rates, attitudes, individual behaviours, incidence and prevalence. The indicators you select should answer your evaluation questions and help you determine whether or not your program objectives have been achieved. ([Source](#))

Meta-analysis: A study which looks at evidence across multiple studies.

Monitoring: A systematic way to collect and record information to check progress and enable evaluation.

Outputs: Products, services or facilities that result from an organisation's or a project's activities.

Quasi-experimental approach: A form of experimental evaluation where the control group is identified by matching the characteristics of people who are receiving the intervention using statistical techniques.

Randomised controlled trial (RCT): A study in which people are randomly allocated to one of several interventions.

Validity: The method which most accurately measures the issue you want it to.

Systematic review: A type of literature review that collects and critically analyses multiple studies or papers.

Qualitative: Research used to gain an understanding of underlying reasons, opinions, and motivations.

Qualitative Comparative Analysis: An analytic approach and set of research tools that combines detailed within-case analysis and cross-case comparisons.

Quantitative: Information or data based on quantities obtained using a quantifiable measurement process.

Pre and post: Research into beneficiaries and outcomes before and after the receipt of an intervention.

Triangulation: The use of multiple data sources and types of data collection to test a particular research question.



Further reading

NPC and Inspiring Impact Resources

- Our guidance for smaller charities starting to think about their Impact. **Theory of Change in Ten Steps:** <https://www.thinknpc.org/resource-hub/ten-steps/>
- **Building your measurement framework:** NPC's four pillar approach: <https://www.thinknpc.org/resource-hub/npcs-four-pillar-approach/>
- NPC's classification blog on the **Five Types of Data:** <https://www.thinknpc.org/blog/5-types-of-data-for-assessing-your-work-an-explainer/>
- **Inspiring Impact:** the Inspiring Impact partnership's high level advice on the full range of topics related to 'impact practice', including a simple diagnostic tool that offers tailored advice based on the five types of data: <https://www.inspiringimpact.org/>
- NPC's report **Balancing act: A guide to proportionate evaluation** goes into more detail on how smaller organisations might think about approaching evaluation issues: <http://www.thinknpc.org/publications/balancing-act-a-guide-to-proportionate-evaluation/>
- NPC's sectoral contributions to ensure that the **social sector is evidence-led:** <https://www.thinknpc.org/resource-hub/towards-an-evidence-led-social-sector/>
- **Data with destiny: How to turn your charity's data into meaningful action:** <http://www.thinknpc.org/publications/data-with-destiny/>
- **Listen and learn: How charities can use qualitative research:** <https://www.thinknpc.org/resource-hub/listen-and-learn-how-charities-can-use-qualitative-research/>
- Our guide on when it is useful to do an **economic analysis:** <https://www.thinknpc.org/resource-hub/economic-analysis-what-is-it-good-for/>
- **User voice: Putting people at the heart of impact practice:** <https://www.thinknpc.org/resource-hub/user-voice-putting-people-at-the-heart-of-impact-practice/>
- **User involvement: Why impact matters:** <https://www.thinknpc.org/wp-content/uploads/2018/10/Make-it-count-Why-impact-matters-in-user-involvement-1.pdf>
- **Learning culture: how to develop a learning organisation:** <https://www.thinknpc.org/resource-hub/developing-a-learning-organisation/>
- **Choosing an electronic case management system for your organisation:** <https://www.thinknpc.org/resource-hub/choosing-an-electronic-case-management-system-for-your-organisation/>
- **The benefits of shared evaluation:** <https://www.thinknpc.org/resource-hub/shared-measurement-greater-than-the-sum-of-its-parts/>
- **The latest developments in evaluation:** <https://www.thinknpc.org/resource-hub/global-innovations-in-measurement-and-evaluation/>

External resources

Elsewhere, we recommend the following resources, broken down by themes in the report:

Guidance around evaluation and data collection

- [Better Evaluation](#) has a vast array of guidance, and notes regarding different tools and approaches to evaluation.
- [Evaluation Support Scotland](#) has a great array of support guides and evaluation explained videos for beginners.
- Nesta have a great [guide](#) on ensuring evidence informed decision making.
- The [Centre of Excellence for Development Impact and Learning](#) and the [Centre for Development Impact](#) both come out with very detailed papers suggesting how to go about evaluations in unique contexts.

Evidence standards

- Nesta's [standards of evidence](#) sets out a very clear framework for what evidence quality is required.
- [EEF Security of Findings](#) has a padlock rating based off how strong the evidence base is.
- NPC's scale from our publication '[Untapped potential](#)' shows how evidence standards can be applied in the field of health and care transformation.

Conducting surveys

- This [guide](#) is useful if you are wondering about how to present different Likert scales, based off the number of points that you need.
- We often recommend that charities adopt the standard [Net Promoter Score approach](#) of 'How likely is it that you would recommend our company/product/service to a friend or colleague?' Using a 0 to 10 scale.
- For demographic questions you can use the [census questionnaire](#).
- For questions on wellbeing you can use the [ONS standard questions](#).

Sampling

- [Better Evaluation](#) has a very clear list of the different types of sampling approaches that may be useful to your specific research question.
- Social Research Methods has a [language buster](#) for all sample related terminology.
- Qualtrics (one of the more common online tools for surveys) has some [recommendations](#) on how to avoid sampling errors.

Sample sizes/margin of error

- Two very useful sample size/margin of error calculators can be found at [Survey System](#).

Experimental designs

- A simplified guide to randomised control trials can be found on [Better Evaluation](#).
- Regression Discontinuity Design: A type of a quasi-experimental research design where participants are assigned to the treatment and control groups on the basis of whether they meet a certain threshold. This is best addressed [here](#).

Qualitative approaches

- NCVO [here](#) provides an overview of why qualitative research is used, and what expectations we should have for qualitative methods.
- Inspiring Impact has a good intro to the benefits of [conducting interviews](#), and advice for setting up [focus groups](#).
- Regarding more specialised qualitative tools that attempt to address causality, one can look at Qualitative Comparative Analysis ([QCA](#)), or tools such as the qualitative impact protocol ([QUIP](#)).
- The Cabinet Office has produced a [framework](#) for assessing qualitative evidence which looks at practical ways to conduct, improve and use qualitative evidence to a very high standard.

Theory based approaches

- Theory based evaluation seeks to look at how an intervention is expected to produce its results. Examples of theory based approaches include [Realist Evaluation](#), [Process Tracing](#), and [Contribution Analysis](#). Each of which have their own relative strengths and weaknesses.

Economic analysis

- [Social Value UK](#) have a number of resources on how to start to monitor the [social value](#) that a charity is creating, and how to [maximise impact](#).
- [Pro Bono Economics](#) have a number of resources and case studies on how to measure economic impact.
- Finally, there are tools such as Social Return on Investment ([SROI](#)), that seek to calculate the economic benefit of social actions conducted by a charity.

Ethics

- The Social Research Association's (SEA) [ethical guidelines](#) are useful for understanding the first principles of research
- Similarly, the UK Evaluation Society has produced useful [guidelines](#) of good practice.
- For further advice, the [Market Research Society's](#) code of conduct is a very useful framework.

Data protection

- We highly recommend that you take some time to read the [ICO's guidance](#) on GDPR.

Questionnaire design

These resources say many of the same things that we say here but are worth reviewing if you have time and interest:

- <https://www.inspiringimpact.org/learn-to-measure/do/surveys/>
- http://psr.iq.harvard.edu/files/psr/files/PSRQuestionnaireTipSheet_0.pdf
- <http://www.surveystar.com/startips/jan2010.pdf>
- <http://www.wireuk.org/ten-steps-towards-designing-a-questionnaire.html>

Other interesting links

- [Cyenfin framework](#).
- If you only have three minutes to learn about a theory of change, [watch this video](#).
- [Youth Impact's](#) compendium of tools to evaluate impact.
- A Centre for Youth Impact blog on the issue of ['The Holy Grail' of evaluation and evidence](#).
- The American Evaluation Association has a [daily blog](#), which features a lot of the issues that evaluators and social researchers can face on a day to day basis.
- Bond's [evidence principles](#).

References

-
- ¹ Better Evaluation https://www.betterevaluation.org/en/approach/realist_evaluation
 - ² Clinks <https://www.clinks.org/publication/introducing-desistance>
 - ³ Outcome Star <http://www.outcomesstar.org.uk/about-the-star/what-is-the-outcomes-star/>
 - ⁴ NCVO <https://blogs.ncvo.org.uk/wp-content/uploads/sally-cupitt/qualitative-methods-of-assessing-impact.pdf>
 - ⁵ Better Evaluation <https://www.betterevaluation.org/en/plan/approach/QUIP>
 - ⁶ Itad <https://itad.com/process-tracing-in-impact-evaluation-potentials-and-pitfalls/>
 - ⁷ Better Evaluation https://www.betterevaluation.org/en/evaluation-options/qualitative_comparative_analysis
 - ⁸ Kail, A., Van Vliet, A. and Baumgartner, L. (2012) [*Funding Impact: impact measurement practices among funders in the UK*](#). New Philanthropy Capital.
 - ⁹ HM Treasury (2011) [*The Magenta Book: Guidance for evaluation*](#)
 - ¹⁰ Hills, D. and Junge, K. (2010) [*Guidance for transport impact evaluations: Choosing an evaluation approach to achieve better attribution*](#). Tavistock Institute.
 - ¹¹ Breckon, J. and Dodson, J. (2016) [*Using evidence: what works?*](#) Alliance for Useful Evidence.